Bayesian Uncertainty Directed Trial Designs

Steffen Ventz[1,2,⋆*], Matteo Cellamare[1,2,⋆†], Sergio Bacallado[3‡], and Lorenzo Trippa[1,2§]

[1]Dana-Farber Cancer Institute, US, [2]Harvard T.H. Chan School of Public Health, US,

[3]University of Cambridge, UK

⋆co-first authorship

June 28, 2018

Abstract

Most Bayesian response-adaptive designs unbalance randomization rates towards the most promising arms with the goal of increasing the number of positive treatment outcomes during the study, even though the primary aim of the trial is different. We discuss Bayesian uncertainty directed designs (BUD), a class of Bayesian designs in which the investigator specifies an information measure tailored to the experiment. All decisions during the trial are selected to optimize the available information at the end of the study. The approach can be applied to several designs, ranging from early stage multi-arm trials to biomarker-driven and multi-endpoint studies. We discuss the asymptotic limit of the patient allocation

*steffen@jimmy.harvard.edu
†matteoc@jimmy.harvard.edu
‡sb2116@cam.ac.uk
§ltrippa@jimmy.harvard.edu

proportion to treatments, and illustrate the finite-sample operating characteristics of BUD designs through examples, including multi-arm trials, biomarker-stratified trials, and trials with multiple co-primary endpoints.

Keywords: Multi-arm clinical trials, Information theory, Decision theory, Response-adaptive designs.

1 Introduction

We discuss a class of Bayesian adaptive designs for randomized clinical trials that seek to maximize the acquisition of information on the effectiveness of new experimental treatments. We call this class of designs Bayesian uncertainty directed (BUD) designs. For a BUD design, the investigator specifies a Bayesian model that will be continuously updated during the trial and a metric to quantify the accumulated information on experimental treatments. This information measure is a summary of the posterior distribution, and all decisions during the trial are selected to approximately optimize the available information on experimental treatments at the end of the study. We illustrate the approach through examples, including controlled multi-arm trials, biomarker-defined subgroup trials [40], and trials with multiple co-primary endpoints [25].

Adaptive designs for clinical experiments use the accumulating data during the study to modify characteristics of the ongoing trial [8, 33]. For instance patient eligibility can be modified during the study, the overall sample size may be re-estimated at interim analyses, or randomization probabilities can be unbalanced towards the most promising arms [21, 34, 38, 40]. Adaptive designs have been applied in several settings, ranging from dose finding studies to Phase III trials [1, 12, 13, 20, 29]. These designs have been developed to save resources, to protect patients from ineffective treatments, and to gain efficiency in the development of new treatments.

For most clinical studies one can consider multiple designs, including adaptive randomization [8]. Practitioners can compare and select from competing candidate methodologies. Selecting according to interpretable performance criteria is desirable [39, 41], and it forces the investigator to explicitly specify the primary aims of the trial. The Bayesian decision-theoretic framework identifies the design that maximizes the performance criteria in expectation. The approach enables the use of utility functions u to represent investigator preferences and it incorporates *a priori* information into the design through Bayesian modeling [2]. The utility $u = u(Y, \theta, d)$ is a random function of the unknown parameters θ, for instance response probabilities that one seeks to estimate during the trial, and of the data Y collected during the study using the design d. Bayesian designs maximize the utility in expectation.

Bandit problems have been studied extensively in the context of clinical trials [5, 6, 14, 23, 42, 46]. A multi-arm Bandit problem describes a sequential decision process. At each stage of the process, a decision maker chooses to "play" one of K arms, this arm then generates a random outcome and a corresponding payoff. In a clinical trial, the arms typically correspond to K treatments, and the outcome summarizes the response of a patient to the assigned treatment. In a Bandit problem the decision maker seeks to select arms during the trial to maximize the total payoff of the sequential experiment [5]. This class of decision-theoretic designs can be used to maximize an explicit utility in expectation, for instance the expected number of responders at the end of the trial. The solution of a Bandit problem is the sequence of decisions that maximizes the expected payoff, and in some cases it can be computed with backward induction (BI) [2].

It has been well documented that BI is often computationally prohibitive [2], and several approximation strategies have been suggested. For instance, Carlin et al. [11] introduced a forward sampling algorithm. Alternative algorithms [9, 19, 24] approximate Bandit solutions by combining

forward sampling with backward induction. M-steps myopic look-ahead procedures, which select at each state the action that maximizes the expected utility at completion of a fictitious M-stages horizon, have been discussed in [10]. These procedures can be extended to allow for randomized actions [28]. Indeed randomization is often a requirement in clinical trials. In these cases, for each patient an arm a is randomly assigned, with randomization probabilities chosen within a fixed and pre-specified subset of the simplex, to optimize the payoff at M-stages.

Popular Bayesian designs, including Bayesian adaptive randomization (BAR) [34], unbalance randomization probabilities towards the most promising arms. These designs appear consistent with the goal of maximizing the number of positive treatment outcomes during the study, although the primary aim of the trial could be different. Here we focus on uncertainty directed designs for clinical trials. Figure 1 illustrates relevant differences between the randomization probabilities of BUD and BAR designs. In this figure each panel summarizes the posterior distribution for a 3-arm trial (without control) with the primary goal of selecting the most effective treatment.

BUD designs use information measures that are convex functionals of the posterior distribution $p(\theta \mid Y)$. Examples will include the entropy of the posterior probability of a positive treatment effect $H(I(\gamma_a(\theta) > 0) \mid Y)$ for an experimental arm a, and the posterior variance of treatment effects $\mathrm{Var}(\gamma_a(\theta) \mid Y)$. The use of a convex measure ensures — by straightforward arguments based on Jensen's inequality — that, on average, information increases with every additional observation. In a BUD design decisions during the trial are selected to maximize these information increments. Using simulations, we show that in multi-arm studies, the operating characteristics of BUD designs are comparable to decision-theoretic designs derived through BI when the corresponding utility function u and information metric coincide.

The paper is structured as follows. Section 2 introduces BUD designs and illustrates their

asymptotic behavior with two specific examples. Section 3 examines the operating characteristics of BUD designs for multi-arm trials. Section 4 discusses BUD designs for biomarker-defined subpopulation trials. The aim here is to match targeted treatments with subgroups of patients that benefit from them. Section 5 deals with an application of the BUD approach to a multi-arm trial with several co-primary endpoints [25]. We conclude the paper with a discussion in Section 6. Details on the computational implementation of the methods proposed and R functions are provided in the supplementary material.

2 Bayesian Uncertainty Directed Designs

In the decision-theoretic framework for sequential experiments the investigator specifies a probability model for unknown parameters and observations during the trial. At pre-planned stages, the experiment can be modified by selecting actions, which may correspond for example to the decision to close the enrollment for a patient subgroup or to modify the randomization probabilities.

The decision maker selects, at stages $t = 1, \ldots, T$, an action A_t from a set $\mathcal{A}$ in a sequential manner. Each action A_t, in turn, generates a random outcome $Y_t \sim p(Y_t \mid A_t, \theta)$ with value in $\mathcal{Y}$. We use a prior probability for the parameters $\theta \in \Theta$ of the outcome distribution. The action taken at time $t+1$ is a random variable A_{t+1}, whose distribution depends on the history of previous actions and outcomes $\Sigma_t = \{(A_\ell, Y_\ell), \ell \leq t\}$. We write $A_{t+1} = d(\Sigma_t)$ to make this dependence explicit. In different words, previous actions and outcomes Σ_t are translated either into the selection of a point in $\mathcal{A}$, for instance the selection of a specific arm, or into a distribution over $\mathcal{A}$, for example the randomization probabilities for patients that will be enrolled.

Several authors proposed and justified multi-stage designs using utility functions [18]. In a two-stage single-arm trial with binary endpoints with n_t patients for stages $t = 1, 2$, the actions

$(A_2, A_3) \in \{0,1\}^2$ represent the decisions (i) to continue the trial after the first interim analysis if the number of responses is sufficiently promising $A_2 = d(\Sigma_1) = I(Y_1 > y_1)$, and subsequently (ii) to recommend the treatment for a confirmatory trial if more than $y_2 > 0$ responses have been observed during the trial, $A_3 = d(\Sigma_2) = I(Y_1 > y_1) \times I(Y_1 + Y_2 > y_2)$. In the example Y_t is a binomial random variable with size n_t and probability θ.

For a given experiment, we let $\mathcal{D}$ denote the set of decision functions that map the data into actions. For the two-stage design $\mathcal{D}$ can be identified by varying the thresholds $(y_1, y_2) \in \{0, \ldots, n_1\} \times \{0, \ldots, n_1 + n_2\}$. Throughout the article $u(d, Y, \theta)$ indicates the utility function. It is a random quantity, which is a function of the parameter θ and the data Y generated under design $d \in \mathcal{D}$. A rational investigator selects the design d from $\mathcal{D}$ that maximizes the expected utility $U(d) = \mathbb{E}[u(d, Y, \theta)]$ of the experiment [2],

$$d^\star = \arg\max_{d \in \mathcal{D}} U(d). \tag{1}$$

In some cases $d^\star$ can be computed exactly using backward induction (BI) [2], but this is often infeasible. The two-arm bandit problem with binary outcomes, as described in [5, 14], is a good example to describe how quickly the computational burden of the BI algorithm increases with the sample size. For a trial with 50 observations, the BI algorithm requires $\binom{54}{4} = 316,251$ operations dedicated to each possible configuration of the sufficient statistics, these are the combinations of positive and negative outcomes for each arm after $t = 0, 1, \cdots, T$ assignments. The number of operations increases quickly with the sample size, for instance, by doubling the sample size the number of operations increases to $\binom{104}{4} = 4,598,126$. Strictly related considerations have been discussed for biomarker trials in [46].

We focus therefore on a myopic approximation of $d^\star$ [2], and proceed in three steps:

Step (1) - Action space: We first specify the set of actions $\mathcal{A}$ that can be selected at each of the T stages of the experiment. For example, in a controlled two-arm trial, which randomizes n patients during each stage t, the action A_t is the number of patients assigned to the experimental arm. Hence, $\mathcal{A} = \{0, 1, \ldots, n\}$. The randomization probability $r_t \in (0, 1)$ toward the experimental arm is a function of the history of actions and outcomes Σ_{t-1}, and $A_t \mid \Sigma_{t-1} \sim \text{Binomial}(n, r_t)$.

Step (2) - Information metric: Next we quantify the acquired information through the accumulated data Σ_t until stage t. Our utility function $u(\cdot)$ will quantify the information accrued by the experiment. Large values of $u(\Sigma_t)$ correspond to low uncertainty levels. We use functions $\tilde{u}$ that translate the posterior $p(\theta \in \cdot \mid \Sigma_t)$ into utilities, i.e. $u(\Sigma_t) = \tilde{u}(p(\theta \in \cdot \mid \Sigma_t))$. The utility functions u can have negative values. In the next paragraphs we provide some examples:

(i) *Multi-arm trials:* Consider a multi-arm study with binary endpoints and primary aim of providing accurate estimates of the treatment effects γ_a for the effective experimental treatments a with $\gamma_a > 0$. Here $\gamma_a = \theta_a - \theta_0$ is the difference between the response probability θ_a for treatment a and the control arm. We can define

$$u(\Sigma_t) = \sum_{a=1}^{K} \Big(v_a - \text{Var}(\gamma_a \times I(\gamma_a > 0) \mid \Sigma_t) \Big) \tag{2}$$

to measure information up to stage t, with $v_a = \text{Var}(\gamma_a \times I(\gamma_a > 0))$ denoting the prior variance.

(ii) *Biomarker-stratified trial:* We test K treatments for patients with and without a genomic alteration. The trial will measure binary outcomes with parameters $\theta_{x,a}$ for subgroups $x = 0, 1$ and treatments $a = 0, \ldots, K$. The primary aim is to test the presence of effects within

subgroups $E_{x,a} = I(\theta_{x,a} > \theta_{x,0})$ and in the overall population $E_a = I(\theta_a > \theta_0)$. Here $\theta_a = \beta\theta_{1,a} + (1-\beta)\theta_{0,a}$ and $\beta \in [0,1]$ is the prevalence of the biomarker. Let $H[p(X)] = -\mathbb{E}[\log p(X)]$ indicate the entropy of a random variable X. We can use a summary u that weights the entropy values associated to interpretable posterior probabilities,

$$u(\Sigma_t) = -\sum_{a=1}^{K} \left\{ H[p(E_a \mid \Sigma_t)] + w \times \Big(H[p(E_{1,a} \mid \Sigma_t)] + H[p(E_{0,a} \mid \Sigma_t)] \Big) \right\}, \text{ with } w \geq 0.$$

(iii) *Dose-finding trial:* We select one of K candidate dose levels $\mathcal{A} = \{1, 2, \ldots, K\}$ using binary efficacy and toxicity outcomes. We let $\theta_{E,a}$ and $\theta_{T,a}$ denote the probabilities of response and toxicity at dose level a. For each dose level a a score weights efficacy $\theta_{E,a}$ and toxicity $\theta_{T,a}$, say $\mathcal{S}_a(\theta) = w\theta_{E,a} + (1-w)(1-\theta_{T,a})$ with $0 \leq w \leq 1$ and dose level $A^\star = \arg\max_a \mathcal{S}_a(\theta)$ has the highest score. In Bayesian modeling $(\theta_{E,a}, \theta_{T,a})_a$, as well as $\mathcal{S}_a(\theta)$ and $A^\star$, are random variables. The posterior distribution $p(A^\star = a|\Sigma_t) = p\big(\cap_{a'} \{\mathcal{S}_a(\theta) \geq \mathcal{S}_{a'}(\theta)\}|\Sigma_t \big), a = 1, \cdots, K$, changes over time as more information becomes available. We can use the (negative) entropy of the posterior and specify $u(\Sigma_t) = \sum_a p(A^\star = a|\Sigma_t) \log p(A^\star = a|\Sigma_t)$.

For each example we only mention one information metric tailored to the aim of the trial. Several alternative measures u could be used. The unifying element is the use of functionals of the posterior to quantify information.

Step (3) - Myopic approximation: As we mentioned the information metric $\tilde{u}$ is specified by a convex functional over the convex space of distributions on Θ. In different words

$$\tilde{u}(w \times p_1 + (1-w) \times p_2) \leq w \times \tilde{u}(p_1) + (1-w) \times \tilde{u}(p_2)$$

for every pair of probability measures p_1 and p_2, when $w \in [0,1]$. Let $u(\Sigma_t) = \widetilde{u}(p(\theta \in \cdot \mid \Sigma_t))$ be the current value of the information function. By Jensen's inequality, given the action selected at the next step $A_{t+1} = a$, the information on average increases

$$\begin{aligned}\mathbb{E}[u(\Sigma_{t+1}) \mid \Sigma_t, A_{t+1} = a] &= \mathbb{E}[\tilde{u}(p(\cdot \mid \Sigma_{t+1})) \mid \Sigma_t, A_{t+1} = a] \\ &\geq \tilde{u}(\mathbb{E}[p(\cdot \mid \Sigma_{t+1}) \mid \Sigma_t, A_{t+1} = a]) = u(\Sigma_t). \end{aligned} \tag{3}$$

It is desirable for $\tilde{u}$ to satisfy this inequality, as any additional observation should tend to reduce uncertainty.

BUD designs select actions that generate large information increments. The myopic decision rule $\tilde{d} \in \mathcal{D}$ selects at each stage the action that maximizes the gain of information $A_{t+1} = \tilde{d}(\Sigma_t) = \arg\max_{a \in \mathcal{A}} \Delta_t(a)$, where $\Delta_t(a) = \mathbb{E}[u(\Sigma_{t+1}) \mid A_{t+1} = a, \Sigma_t] - u(\Sigma_t) \geq 0$. In most clinical trials, non-randomized policies like the myopic decision rule described above, are inappropriate [7]. For this reason we use the BUD design d_{BUD}, which is a randomized version of the myopic design. The design translates Σ_t into a distribution on $\mathcal{A}$. In particular, if A_{t+1} is the treatment assigned to the next patient, then

$$p(A_{t+1} = a \mid \Sigma_t) \propto \Delta_t(a)^{h(t)}. \tag{4}$$

Here $h(\cdot)$ is a non negative function. With large values of $h(t)$ the the BUD design d_{BUD} and the myopic design $\tilde{d}$ become nearly identical. On the other extreme, with $h(t) = 0$, the randomization probabilities become identical across arms. The choice of $h(\cdot)$ has the goals of (i) satisfying the requirement to randomize patients and (ii) the approximate optimization of utility criteria u.

2.1 Asymptotic Allocation Proportions in BUD Designs

Our goal is the study of BUD designs with realistic sample sizes (Sections 3–5). However, understanding the asymptotic behavior of BUD policies is useful to interpret simulation results and to ascertain if the proportions of patients allocated to different arms converges to a nearly optimal limit. Here we discuss with examples asymptotic characteristics of BUD designs. For simplicity we consider a constant parameter $h(\cdot) = h \geq 0$ in (4).

Lemma 1 is a technical result that we will later use to derive asymptotic allocation proportions in two examples. A proof following stochastic approximation arguments can be found in the appendix. Let $\widehat{p}_{a,t}$ be the proportion of samples allocated to arm a by time t.

Lemma 1. *Consider the allocation of patients to arms $\mathcal{A} = \{0, 1\}$ with a BUD design. Define $F_t = -\widehat{p}_{0,t} + \Delta_t^h(0)/(\Delta_t^h(0) + \Delta_t^h(1))$. Suppose that, on a set of probability 1, for any $\varepsilon > 0$ there is a random time T, and number $c > 0$ such that $F_t < -c$ wherever $\widehat{p}_{0,t} > \rho_0 + \varepsilon$, and $F_t > c$ wherever $\widehat{p}_{0,t} < \rho_0 - \varepsilon$ for all $t > T$. Then $\widehat{p}_{0,t} \to \rho_0$ a.s.*

Remark 1. Before applying this result to examples, we observe that for some utility criteria u it is straightforward to extend the lemma to experiments with multiple arms $\mathcal{A} = \{0, \ldots, K\}$. In particular, when the information metric $u(\Sigma_t)$ is an additive functional of the posterior of each parameter θ_a, $a \in \mathcal{A}$, for example $u(\Sigma_t) = -\sum_a \mathrm{Var}(\theta_a \mid \Sigma_t)$, the information gain $\Delta_t(a)$ depends only on the prediction of outcomes from a single arm a. For any pair of arms (a_1, a_2), the subsequence of samples assigned to these two arms is equivalent to a two-arm BUD design. Therefore, if the conditions of Lemma 1 hold,

$$\frac{\widehat{p}_{a_1,t}}{\widehat{p}_{a_1,t} + \widehat{p}_{a_2,t}} \xrightarrow{\text{a.s.}} \tilde{\rho}_{a_1,a_2}$$

for every $a_1, a_2 \in \mathcal{A}$. Then, the allocation proportions $(\widehat{p}_{0,t}, \ldots, \widehat{p}_{K,t})$ converge to a limit $\rho = (\rho_0, \ldots, \rho_K)$, which is the unique solution to the linear system

$$\sum_{a=0}^{K} \rho_a = 1 \,, \qquad \rho_{a_1} = \tilde{\rho}_{a_1,a_2}(\rho_{a_1} + \rho_{a_2}) \quad \text{for all } \{a_1, a_2\} \subset \{0, \ldots, K\}.$$

Example 1. Multi-arm trial with normal outcomes. Assume the outcome $Y_t \mid A_t = a \sim N(\theta_a, \sigma_a^2)$ is normal with unknown mean θ_a and known variance σ_a^2. We use independent $N(0, \sigma^2)$ prior distributions for θ_a, $a \in \mathcal{A} = \{0, \ldots, K\}$ and show that in a BUD design, driven by the information measure

$$u(\Sigma_t) = -\sum_{a=0}^{K} \mathrm{Var}(\theta_a \mid \Sigma_t) = -\sum_{a=0}^{K} \frac{1}{\sigma^{-2} + t\widehat{p}_{a,t} \times \sigma_a^{-2}},$$

the arm-specific sample size proportions converge a.s. to

$$\rho_a = \frac{\sigma_a^{\frac{2h}{1+2h}}}{\sum_{\ell=0}^{K} \sigma_\ell^{\frac{2h}{1+2h}}} \quad \text{for all } a = 0, \ldots, K. \tag{5}$$

For large h the BUD design becomes nearly identical to the myopic design, and the limit (5) coincides with the Neyman allocation $\rho_a \propto \sigma_a$ [17].

To show the convergence to (5), we note that the information gain $\Delta_t(a)$ is equal to

$$\begin{aligned} \Delta_t(a) &= \frac{1}{\sigma^{-2} + t\widehat{p}_{a,t}\sigma_a^{-2}} - \frac{1}{\sigma^{-2} + (t\widehat{p}_{a,t} + 1)\sigma_a^{-2}} \\ &= \frac{\sigma_a^2}{(\sigma_a^2/\sigma^2 + t\widehat{p}_{a,t})(\sigma_a^2/\sigma^2 + 1 + t\widehat{p}_{a,t})} = \frac{\sigma_a^2}{t^2\widehat{p}_{a,t}^2} + \mathcal{O}((\widehat{p}_{a,t}t)^{-3}). \end{aligned} \tag{6}$$

Using (6), we can write

$$F_t = -\widehat{p}_{0,t} + \frac{\Delta_t^h(0)}{\Delta_t^h(0) + \Delta_t^h(1)} = -\widehat{p}_{0,t} + \frac{\left(\widehat{p}_{0,t}^{-2}\sigma_0^2 + \mathcal{O}(t^{-1}\widehat{p}_{0,t}^{-3})\right)^h}{\left(\widehat{p}_{0,t}^{-2}\sigma_0^2 + \mathcal{O}(t^{-1}\widehat{p}_{0,t}^{-3})\right)^h + \left(\widehat{p}_{1,t}^{-2}\sigma_1^2 + \mathcal{O}(t^{-1}\widehat{p}_{1,t}^{-3})\right)^h}. \tag{7}$$

Consider a similar sequence

$$\tilde{F}_t = -\widehat{p}_{0,t} + \frac{\widehat{p}_{0,t}^{-2h}\sigma_0^{2h}}{\widehat{p}_{0,t}^{-2h}\sigma_0^{2h} + (1-\widehat{p}_{0,t})^{-2h}\sigma_1^{2h}}. \tag{8}$$

Note that $\tilde{F}_t$ is strictly decreasing in $\widehat{p}_{0,t} \in [0,1]$ with a zero at ρ_0, which implies that for any $\varepsilon > 0$ there is a $c > 0$ satisfying $\tilde{F}_t < -c$ wherever $\widehat{p}_{0,t} > \rho_0 + \varepsilon$, and $\tilde{F}_t > c$ wherever $\widehat{p}_{0,t} < \rho_0 - \varepsilon$. These are the assumptions of Lemma 1.

We need to show that $F_t - \tilde{F}_t \to 0$ as $t \to \infty$ to apply the lemma and derive (5). First, consider the F_t subsequence when $t^{-1/3+\delta} < \widehat{p}_{0,t} < 1 - t^{-1/3+\delta}$ for some $\delta \in (0, 1/3)$. In this case $\widehat{p}_{0,t}$ does not approach zero or one too quickly, and the difference between (7) and (8) vanishes for large t. Second, when $\widehat{p}_{0,t} \leq t^{-1/3+\delta}$, it can be verified that $\Delta_t(1)/\Delta_t(0) \to 0$ as $t \to \infty$, and that $F_t \to 1$. Conversely, when $\widehat{p}_{0,t} > 1 - t^{-1/3+\delta}$, $F_t \to -1$. We can therefore apply the lemma to conclude that $\widehat{p}_{0,t} \to \rho_0$. For a multi-arm BUD design, $K > 1$, it is sufficient to apply Remark 1.

Example 2. Multi-arm trial with binary outcomes. Now consider the case where Y_t is binary with $p(Y_t = 1 \mid A_t = a) = \theta_a$, and a prior $\theta_a \sim \text{Beta}(\alpha, \beta)$ for $a \in \mathcal{A}$. We consider the same information metric of the previous example and prove that the allocation proportions converge with probability 1 to the limit in (5) with $\sigma_a^2 = \theta_a(1-\theta_a)$. In this example, the information gain $\Delta_t(a)$ has a closed-form expression which simplifies the derivation of the asymptotic result. The proof follows the same arguments of the previous example, and it is deferred to the appendix.

The examples above are by no means exhaustive, but their simplicity permits a self-contained analysis, which points to mathematical techniques. An important case not covered here is that of BUD designs in which the utility function u is not additive on the treatment arms. The information metric (2) in Section 2 is an example where Lemma 1 is not directly applicable. Such information metics are of practical relevance, for example when some of the experimental agents are potentially significantly inferior to the standard of care. We believe that the extensive literature on stochastic approximation can be used for additional asymptotic analyses of BUD designs, including non additive utilities u. The monograph [26] provides a useful guide. The general strategy would consist of writing the vector of allocation proportions $(\widehat{\boldsymbol{p}}_{0:K,t})_{t\geq 0}$ as a stochastic approximation process; namely, one would express the increments $\widehat{\boldsymbol{p}}_{0:K,t+1} - \widehat{\boldsymbol{p}}_{0:K,t} = \gamma_t(F(\widehat{\boldsymbol{p}}_{0:K,t}) + \epsilon_{t+1} + \mathcal{O}(t^{-1}))$ for some sequences $(\epsilon_t)_{t\geq 0}$ and $(\gamma_t)_{t\geq 0}$ with $\mathbb{E}[\epsilon_{t+1} \mid \Sigma_t] = 0$, $\sum_t \gamma_t = \infty$, and $\sum_t \gamma_t/t < \infty$. Then the arguments used to prove Lemma 1 can be adapted to establish that $\{\widehat{\boldsymbol{p}}_{0:K,t}\}_{t\in\mathbb{N}}$ approaches the solution of a differential equation.

3 BUD Designs for Phase II Multi-Arm Trials

We discuss BUD designs for two multi-arm Phase II trials with distinct aims: (i) to identify all experimental arms with a positive treatment effect [43], and (ii) to select the treatment with the most favorable outcome distribution [30].

3.1 Estimation of Treatment Effects in a Controlled Multi-Arm Study

We first considered a study with K experimental arms compared to a control therapy. For each patient t, action $A_t = a$, with $a \in \mathcal{A} = \{0, \ldots, K\}$, indicates the assignment of patient t to arm a, where $a = 0$ denotes the control arm. The primary outcome is the binary response to treatment

with probability of response $p(Y_t = 1 \mid A_t = a, \theta) = \theta_a, a = 0, \ldots, K$. We use independent beta random variables $\theta_a \sim \text{Beta}(\theta_a; v_1, v_2)$ for $a = 1, \cdots, K$ to define the prior for $\theta = (\theta_0, \ldots, \theta_K)$.

The trial is designed to generate estimates of the treatment effects, γ_a, with low posterior variance at the end of the trial. We use an information measure which is consistent with this goal,

$$u(\Sigma_t) = \sum_{a=1}^{K} \Big(v_a - \text{Var}(\gamma_a \mid \Sigma_t) \Big), \tag{9}$$

where $v_a = \text{Var}(\gamma_a)$. When the primary goal is to test the null hypothesis $\mathcal{H}_{0,a} : \gamma_a \leq 0$, a slightly different utility function $-\text{Var}(g(\gamma_a) \mid \Sigma_t)$ can be considered, where $g : [-1, 1] \to [0, 1]$ is a monotone function with plateaus. Each patient is assigned to treatment $a \in \mathcal{A}$ with probability

$$p(A_{t+1} = a \mid \Sigma_t) \propto \left(\sum_{a'=1}^{K} \mathbb{E}\Big[v_{a'} - \text{Var}(\gamma_{a'} \mid \Sigma_{t+1}) \;\Big|\; A_{t+1} = a, \Sigma_t \Big] - u(\Sigma_t) \right)^{h(t)}. \tag{10}$$

Simulation Study: We discuss results of a simulation study for a 4-arm trial using the information measure (9) to estimate treatment effects. We considered four scenarios with constant response rate of 0.4 for the control arm (see Table 1) and use uniform prior $v_1 = v_2 = 1$ for θ_a, $a = 0, \cdots, 3$. In Scenario 1 all experimental arms are ineffective with response rates equal to 0.4. Whereas in the other scenarios some of the arms have positive treatment effects. To simplify comparison to balanced randomization (BR) we set $T = 336$, because a BR design with one-sided Fisher's exact test and type I and II error rates of 0.05 and 0.20 for the alternative 0.6 vs 0.4 requires 84 patients per arm. We compare the BUD design to three alternative randomization methods: BR, Bayesian adaptive randomization (BAR) as described in [38], and the doubly adaptive coin design (DBCD) [16] targeting assignment frequencies equal to the Neyman allocation $\rho_a \propto \sqrt{\theta_a(1-\theta_a)}$ (DBCD1) or equal to $\rho_a \propto \sqrt{\theta_a}$ (DBCD2) [16].

Table 1 shows the average number of patients randomized to each arm across 5,000 simulations and the mean squared error (MSE) of the treatment effect estimates. We also show the power of Fisher's exact test (without correcting for adaptivity [3]) for the null hypothesis $\mathcal{H}_{0,a} : \gamma_a \leq 0$. Most notable, among all five designs, the BUD design has the lowest MSEs for all experimental arms across all four scenarios. When all experimental arms are ineffective BR, DBCD1 and DBCD2 randomize on average 84 patients to each arm with standard deviations (SDs) of 8, 4 and 6. BAR and the BUD design randomize more patients to the control arm compared to the remaining designs (118, 97). For BAR this is expected by construction of the randomization probabilities [38]. In the BUD design, the assignment of a patient to the control arm reduces the uncertainty on θ_0 and therefore leads to uncertainty reductions for all treatment effects $\gamma_a = \theta_a - \theta_0$, $a = 1, \ldots, K$, while the assignment of a patient to arm $a, a > 0$, reduces the expected variance only of θ_a and γ_a.

When only arm 1 has a treatment effect (Scenario 2), BAR assigns on average 103 and 100 patients to arm 1 and the control arm (SD 10 and 14) and has 87.5% power. The BUD, DBCD1 and DBCD1 designs have 82.2%, 79.8% and 81.2% power respectively, and for all three designs the standard deviation of the number of enrolments to arm $a = 0, \ldots, K$ is substantially smaller than for BAR. The utility function (9) targets treatment effect estimates with low uncertainty. One can therefore expect similar numbers of patients assigned to each experimental arm. In contrast with BAR the arm-specific sample sizes tend to be markedly different across experimental treatments. In the sections 3.2, 4 and 5 we discuss BUD designs driven by different information measures that, similar to BAR, tend to increase the sample size of the best experimental arms.

We also compare the BUD design to an alternative Bayesian design [34] (BAR2) which assigns patients to arms with probability proportional to $p(\theta_a > \theta_{a'} \text{ for } a' \neq a|\Sigma_t)^{t/(2T)}$. Compared to the BUD and BAR designs, BAR2 assigns more patients on average to the most effective arm,

$(84, 161, 180, 130)$ assignments for BAR2 in scenarios 1 to 4 compared to $(80, 103, 117, 83)$ for BAR and $(73, 73, 75, 70)$ for the BUD design, respectively. BAR2 tends to have larger MSEs for treatment effect estimates compared to the BUD and BAR designs. For example, the MSEs of arm 3 in Scenarios 1 to 4 equal 7.3, 9.7, 10 and 9.9 for BAR2 compared to 5.5, 5.4, 4.7 and 5.1 for the BUD design.

We further investigated the behavior of BUD design by repeating the simulations over a range of sample sizes T. In this comparison we also included, for each scenario, a hypothetical oracle design. For a fixed total sample size T, the oracle design is defined by the number of patients that should be assigned to each experimental arm (with sum equal to T) in order to maximize $u(\Sigma_T)$ in expectation, assuming that the oracle knows θ. We computed for each scenario and $T = 1, \ldots, 200$, the combination of sample sizes $\{T_a\}_{a=0}^K$, $T = \sum_{0 \leq a \leq K} T_a$, that minimizes the expected value of $\sum_{a=1}^K \text{Var}(\gamma_a \mid \Sigma_T)$. The top row of Figure 2 shows, for each scenario, the difference between the expected information $\mathbb{E}_d[u(\Sigma_T)]$ of the four randomization methods and the expected information of the oracle design for different values of T. For the BUD design we observe the lowest regret values across all sample sizes $T = 1, \ldots, 200$. This is expected because the BUD design, unlike the other designs, is driven by the utility u, which is also used to define regret values. We observe that the regret of the BUD design is considerably closer to the benchmark utilities of the oracle design compared to the other designs, and that the discrepancy between the BUD design and the benchmark vanishes quickly as the sample size increases.

We also explore the sensitivity of BUD designs to different information measures that are consistent with the study aims. We used the same scenarios as in the previous paragraphs and consider BUD designs defined by the entropy, $-\sum_{a=1}^K \mathbb{E}_{p(\gamma_a|\Sigma_t)}[\log p(\gamma_a|\Sigma_t)]$ (BUD-E), the sum of the mean absolute error of treatment effects estimates (BUD-MAD), or the posterior variance

and entropy of the discretized treatment effect γ_a^c (BUD-DV and BUD-DE), which is defined as $\gamma_a^c = 0, 1,$ or 2 when the treatment effect γ_a falls into the intervals $(-1, 0]$, $(0, 0.25]$ or $[0.25, 1]$. Table S3 in the supplementary material shows the mean squared error (MSE), the average number of enrollments per arm and power for these BUD designs. The designs BUD-E and BUD-MAD have similar MSEs for the treatment effect estimates across scenarios and assign on average a similar number of patients to each treatment arm. By contrast, the truncation of the treatment effect in BUD-DE and BUD-DV increases the variability of treatment effect estimates across scenarios.

In the supplementary material we add to these comparisons the evaluation of alternative BUD designs for controlled multi-arm trials where the utility u is representative of different study aims. The supplementary material includes a sensitivity analysis (Table S2) of the BUD design to variations of the prior model.

3.2 Selection of the Best Experimental Treatment

Companies and investigators often explore multiple experimental treatments within a single study to select the best treatments [30]. We consider a K-arm study without control arm to identify treatment $a^\star = \underset{1 \le a \le K}{\arg\max}\, \theta_a$.

The trial estimates the response rate $\theta_{a^\star}$ of the most effective therapy $a^\star$ with posterior distribution $\theta_{a^\star}$

$$p_{\theta_{a^\star}}(x \mid \Sigma_t) = \sum_{a=1}^{K} \left[p_{\theta_a}(x \mid \Sigma_t) \prod_{j \neq a} p(\theta_j \le x \mid \Sigma_t) \right] \quad \text{for } x \in (0, 1).$$

Patients are randomized with the aim to minimize uncertainty on $\theta_{a^\star}$. Note that a BUD design can minimize uncertainty on $a^\star$, or it can minimize uncertainty on the unknown response rate $\theta_{a^\star}$.

We use a utility function that is consistent with the aim of minimizing uncertainty on $\theta_{a^\star}$ and measure information using the (negative) posterior entropy of $\theta_{a^\star}$, $H[p(\theta_{a^\star} \mid \Sigma_t)] = \int_0^1 p_{\theta_{a^\star}}(z \mid$

$\Sigma_t) \log p_{\theta_{a^\star}}(z \mid \Sigma_t) dz$. Other measures, such as the posterior variance of $\theta_{a^\star}$ can be considered. When the primary goal is to select arm $a^\star$ one may specify u equal to the entropy of the posterior distribution $p(a^\star|\Sigma_t)$.

Using the randomization scheme of section 2,

$$p(A_{t+1} = a \mid \Sigma_t) \propto \left(\mathbb{E}\Big[H[p(\theta_{a^\star} \mid \Sigma_{t+1}] \Big| \Sigma_t, A_{t+1} = a \Big] - H[p(\theta_{a^\star} \mid \Sigma_t)] \right)^{h(t)}.$$

Simulation Study: We conducted a simulation study for a 4-arm trial that selects the most effective treatment. We consider three scenarios (see Table 2). Arm 4 is the most effective treatment in all scenarios. We compare the BUD design to BR, BAR (using Thomson's rule) and the randomized-play-the-winner (RPW) design [44]. For all randomization schemes, at completion of each simulation we compute the probabilities $p(\theta_a = \max_{a'=1,\ldots,4} \theta_{a'} \mid \Sigma_T)$ using a uniform prior for $\theta = (\theta_1, \ldots, \theta_4)$ and select the arm $a^\star$ with the highest posterior probability of being the best available treatment.

Table 2 shows the average number of patients randomized to each arm across 10,000 simulated trials with $T = 30, 50$ or 70 patients, and provides additional operating characteristics. The BUD design selects the most effective arm more frequently across simulations compared to the three alternative designs. In scenario 1 after $T = 30$ assignments, for example, in 57% of the simulated trials the BUD design selects the right arm, compared to 52%, 56% and 53% for BR, BAR and RPW, respectively. The BUD design assigns on average the same number of patients to each arm as BAR, but has lower assignment variability. BAR is also similar to the BUD design when we consider the probabilities of selecting the best available treatment, but it has the largest assignment variability among all randomization schemes. As shown in the last column of Table 2, across all

scenarios the BUD design has the lowest mean squared error for the estimate of $\max_a \theta_a$.

We also derived optimal sequential Bayesian designs $d^\star$ in (1) using BI for comparisons. The entropy $H(\theta_{a^\star} \mid \Sigma_T)$ defines the utility function and the expected utility $\mathbb{E}_{d^\star}[u(\Sigma_T)]$. We considered small sample sizes T, up to 30 patients, for which BI is feasible. These computations were followed by a comparison based on 10^5 simulated trials for BUD, BR BAR, and the RPW designs. In each simulation, patients respond to treatments with probabilities $\theta = (\theta_1, \ldots, \theta_4) \sim p(\theta)$ from the prior (independent beta) that we used to compute $d^\star$. We estimate the expected utility of each design d and the regret $\mathbb{E}_{d^\star}[u(\Sigma_T)] - \mathbb{E}_d[u(\Sigma_T)]$, see Figure S1. The regret of the BUD design remains close to zero across all sample sizes T that we considered, whereas the regret of BR and RPW designs are larger; for instance the regret for BUD, BAR, BR and RPW designs is equal to $(0.014, 0.035, 0.15, 0.12)$ and $(0.017, 0.047, 0.30, 0.22)$, with $T = 10$ and 30 patients, respectively.

We evaluate the sensitivity of BUD designs to the choice of the randomization parameter h. We conducted simulations under the three scenarios in Table 2 using either $h(t) = 1, 2, 6, 12$ or 20. The parameter h in our simulations has minor influence on the resulting operating characteristics, see supplementary Table S5. Across scenarios the average number of treatment assignments to each arm and the proportion of simulated trials that recommend the best arm for further investigation remains nearly identical when we vary h.

4 Biomarker-Stratified Clinical Trials

The development of anti-cancer treatments focuses increasingly on therapies that target specific genetic alterations. Consequently, there has been an increasing interest in biomarker-driven studies [21, 37, 40, 45]. These studies enroll patients with multiple genomic abnormalities in a single multi-arm trial, and identify subgroups of patients that respond to experimental treatments.

We consider a design that evaluates K experimental therapies in biomarker subgroups. For each patient t the vector $X_t \in \{0,1\}^B$ defines the patients' biomarker profile, here $X_{t,\ell} = 1$ if patient t has a positive marker status for biomarker $\ell = 1, \ldots, B$ and zero otherwise. In our examples B will be equal to 2,3 or 4. Patients can be partitioned into 2^B groups. We assume binary endpoints with probability of response $p(Y_t = 1 \mid X_t = x, A_t = a, \theta) = \theta_{x,a}$ for patients with profile $x \in \{0,1\}^B$. Here $\theta_{x,0}$ is the subgroup-specific response probability for the control arm $a = 0$.

In oncology, treatments are commonly developed for patients with specific genomic abnormalities [27]. Pre-clinical studies typically suggest effects of a treatment a in a targeted biomarker subgroup. But in many cases positive effects in other groups of patients are hypothesized, and treatments are evaluated in a population larger that the target groups [27]. We assume that treatment $a > 1$ targets biomarker $1 \leq b_a \leq B$, and it is therefore *a priori* more likely to be effective for patients with positive biomarker $X_{t,b_a} = 1$ status than for other subgroups.

The indicator $E_{\ell,a} = \mathbf{1}\Big(\bigcup_{x \in \{0,1\}^B : x_{b_a} = \ell} \{\theta_{x,a} > \theta_{x,0}\} \Big)$ corresponds, for $\ell = 1$ and 0, to treatment effects for biomarker b_a positive and negative patients, respectively. We first specify the prior probabilty $p(E_{1,a} = 1) = \pi$. Since treatment effects for patients without genomic alteration b_a can not be ruled out we specify $p(E_{0,a} = 1 \mid E_{1,a} = 1) = \lambda$. Lastly, since drug a was developed primarily for biomarker b_a we assume $p(E_{0,a} = 1 \mid E_{1,a} = 0) = 0$.

The prior distribution for $\theta = (\theta_{x,a}; x \in \{0,1\}^B, 0 \leq a \leq K)$ conditional on the indicators $E_{\ell,a}$ is a product of Beta$(1,1)$ distributions restricted to the subset $R \subset [0,1]^B$ consistent with these indicators,

$$p\Big(\theta \,\Big|\, E_{\ell,a}; \ell = 0,1, a = 0, \ldots, K\Big) \propto I(\theta \in R) \prod_{\substack{x \in \{0,1\}^B, \\ a = 0, \ldots, K}} \text{Beta}(\theta_{x,a}\,;\, 1,1). \tag{11}$$

Alternative prior distributions, different from the simple Bayesian model (11) that we specified, could be applied within the BUD framework, see for instance [45].

We define a BUD design that evaluates treatments in the targeted subgroups, and explores possible additional treatment effects in the biomarker negative subgroups. The information measure is

$$u(\Sigma_t) = -\sum_{a=1}^{K} \Big\{ H_{\text{as}}\big[p(\{E_{1,a}=1\}\cup\{E_{0,a}=1\} \mid \Sigma_t)\big] + w \times \Big(H_{\text{as}}\big[p(E_{1,a}=1 \mid \Sigma_t)\big] + H_{\text{as}}\big[p(E_{0,a}=1 \mid \Sigma_t)\big] \Big) \Big\}. \tag{12}$$

Here $H_{\text{as}}[p] = p - p^{\beta}$, with $\beta > 1$, is the asymmetric entropy [22] and $w > 0$ weights the information gains in the subgroups and in the overall population. Allocations in the BUD design using (12) are driven by the expected reduction of the posterior entropy of positive treatment effects. The parameter β controls the curvature $H''_{\text{as}}[p] \propto -p^{\beta-2}$ of the entropy, and, using a Taylor approximation, the expected variation in the entropy of $p(E_{\ell,a}=1|\Sigma_t)$ given $A_{t+1}=a$ equals approximately

$$H''_{\text{as}}(p(E_{\ell,a}=1|\Sigma_t)) \times \text{Var}[p(E_{\ell,a}=1|\Sigma_{t+1})|\Sigma_t, A_{t+1}=a]/2.$$

Values of $\beta > 2$ tend to reinforce the randomization probabilities of the arms that are more promising accordingly to the posterior probabilities $p(E_{\ell,a}=1|\Sigma_t)$ and $p(\{E_{1,a}=1\}\cup\{E_{0,a}=1\} \mid \Sigma_t)$.

In contrast to previous BUD designs in Section 3, which focus on treatment effects estimation, the metric (12) quantifies uncertainty on the presence or absence of treatment effects in biomarker targeted subgroups and in the overall population. This is consistent with the study aim of testing efficacy.

The BUD design assigns patients to treatment a with probability

$$p\big(A_{t+1}=a \mid X_{t+1}=x,\Sigma_t\big) \propto \left[\mathbb{E}\Big[u(\Sigma_{t+1}) \;\Big|\; X_{t+1}=x, A_t=a, \Sigma_t\Big] - u\big(\Sigma_t\big)\right]^{h(t)}, \text{ for } a=0,\ldots,K.$$

Simulation Study: Table 3 summarizes the results of a simulation study for a controlled biomarker trial with four experimental treatments and an overall sample size of $T = 500$ patients.

We considered five scenarios with four (Scenarios 1-3), three (Scenario 4) or two (Scenario 5) biomarkers B. The true response probability for the control equals $\theta_{x,0} = 0.35$ across all five scenarios and subgroups $x \in \{0,1\}^B$. Treatments $a = 1,\ldots,4$ target biomarker $b_a = a$ in the first three scenarios. In Scenario 1 there are no positive treatment effects (PTEs), and for all therapies $a > 0$ the probabilities $\theta_{x,a}, x \in \{0,1\}^B$ are identical to the control. Whereas in Scenarios 2-3, treatment $a = 1$, and treatment $a = 2$ (only in Scenario 3), have PTEs with $\theta_{x,a} = 0.55$ for all patients with positive biomarker status for $b_a = a$, i.e. $X_{t,b_a} = 1$. In Scenario 4, treatments $a = 1, 2$ both target biomarker $b_a = 1$ and have a PTE on this subgroup with $(\theta_{x,1}, \theta_{x,2}) = (0.55, 0.65)$. The remaining treatments $a = 3, 4$, which target biomarker 2 and 3, have no PTEs. Lastly, in Scenario 5, therapies $a = 2, 3, 4$ have PTEs in their target population $b_a = 1, 2, 2$ with $\theta_{x,a} = 0.55, 0.55$ and 0.65 for all biomarker positive patients $X_{t,b_a} = 1$. Additionally, the PTE of treatment $a = 1$ ($b_1 = 1$) extends to all patients irrespectively of the profile $X_t = x$, with $\theta_{x,1} = 0.5$.

We used $(w, \beta) = (5, 6)$ to define the utility function of the BUD design, and compared the design to a balanced design (BR) that randomizes patients to the five arms with equal probabilities. For each treatment $a > 0$, we tested treatment effects in the biomarker-targeted group (patients with $X_{i,b_a} = 1$) and in the biomarker negative group ($X_{i,b_a} = 0$) using a bootstrap test similar to [12, 38, 41], accounting for adaptive randomization.

Table 3 shows mean and variability for the allocation of patients to treatments, and the power of detecting treatment effects. The targeted type I error rate is set at 10%. Similar to the multi-arm BUD designs in section 3, the biomarker BUD design allocates substantially more patients to the control arm than the BR design to reduce uncertainty on treatment effects. This translates into a higher power across all scenarios compared to the BR design. In Scenario 1, without treatment effects, the BUD design assigns on average a higher number of patients with targeted biomarker profile $X_{b_a} = 1$ to the corresponding experimental treatment $a = 1, \ldots, 4$ than the BR design. It tends to match targeted treatments and biomarker profiles. This is due to the specification of a higher a priori probability of PTEs for the targeted biomarker-subgroups. Interestingly, in Scenario 4, where arms $a = 1$ and 2 have PTEs in the first biomarker subgroup, the BUD design on average assigns more patients to the arm with the smaller treatment effect, which is more difficult to identify. This translates into a gain in power of 11% compared to BR. Similarly, in Scenario 5, where all treatments $a > 1$ have PTEs the BUD design randomizes on average more patients to experimental arms with small treatment effects. This translates into substantial gains in power compared to BR.

In contrast to BAR, under a BUD design with utility (12), the average arm-specific sample size is not expected to increase with the magnitude of the treatment effect. Indeed with large treatment effects, uncertainty on the indicators $E_{1,a}$ and $E_{0,a}$ in the targeted and non-targeted subgroups vanishes rapidly with a limited number of observations.

We evaluated the sensitivity of the BUD design to different choices of the parameters (β, w) which define the uncertainty function (12). We repeated the simulations under scenarios 2 and 4 using either $w = 5$ with $\beta = 2, 6, 8$ or using $w = 2, 5, 8$ with $\beta = 6$ (Tables S6 and S7). A BUD design with symmetric information measure $\beta = 2$ assigns on average more patients to ineffective

experimental arms than using $\beta = 6$ or $\beta = 8$. The latter two choices have similar operating characteristic; compared to $\beta = 2$, a BUD design with $\beta \geq 6$ assigns on average fewer patients to ineffective agents and increases the power by 11% for arm 1 in scenario 2, and by 5% to 9% for arms 1 and 2 in scenario 4 (scenario 2: 74%, 86% and 85% power for $\beta = 2, 6, 8$ using Fisher exact test). As shown in Tables S6 and S7 the BUD design is relatively insensitive to the choice of w.

5 Trials with Co-Primary Endpoints

In several contexts, such as Alzheimer's disease, a single endpoint has been shown to be insufficient to capture patients' response to treatments adequately [15]. Several authors recommend to evaluate new treatments using multiple clinical endpoints [15]. Trial designs that evaluate treatments using multiple outcomes have been proposed in [31, 35]. We consider a BUD design for a controlled multi-arm trial that evaluates K experimental treatments using two binary endpoints $Y_t = (Y_{t,1}, Y_{t,2})$, for example, in Alzheimer's disease, the cognitive performance and the physical status after treatment. Regulatory guidelines [15] recommend the investigator to demonstrate superiority of the experimental treatment $a > 0$ compared to the control arm $a = 0$ on both endpoints.

For each treatment a, $p(Y_t = y \mid A_t = a) = \theta_{y,a} \geq 0$, with $y \in \{0,1\}^2$, and $\sum_{y\in\{0,1\}^2} \theta_{y,a} = 1$. We use independent Dirichlet prior distributions for the arm-specific parameters $\theta_a = \big(\theta_{(1,1),a}, \theta_{(1,0),a}, \theta_{(0,1),a}, \theta_{(0,0),a}\big)$. The parameters θ_a specify the marginal probabilities $(\nu_{1,a}, \nu_{2,a})$ of the two endpoints. We use $\gamma_{\ell,a} = \nu_{\ell,a} - \nu_{\ell,0}$, $\ell = 1, 2$, to indicate treatment effects for both endpoints, and define the indicators $E_{\ell,a} = \mathbf{1}\{\gamma_{\ell,a} > 0\}$, $\ell = 1, 2$, and $E_a = E_{1,a} \times E_{2,a}$.

The posterior probability $p(E_a = 1 \mid \Sigma_t)$ summarizes the available evidence of therapy a as suitable alternative to the control therapy. Although the investigator's focus is on testing $E_a = 0$

versus $E_a = 1$, improvements for one of the two outcomes are also relevant. We therefore consider a composite uncertainty measure

$$u_1(\Sigma_t) = -\sum_{a=1}^{K} \left\{ H_{\text{as}}[p(E_a = 1|\Sigma_t)] + w \times \left(H_{\text{as}}[p(E_{1,a} = 1|\Sigma_t)] + H_{\text{as}}[p(E_{2,a} = 1|\Sigma_t)] \right) \right\}, \quad (13)$$

where as before $H_{\text{as}}[p] = p - p^{\beta}$, $\beta > 1$ and $w > 0$, which is consistent with the study goals of testing E_a and improvements for single endpoints $E_{a,\ell}, \ell = 1, 2$.

To explore variations of the BUD design's operating characteristics with different information measures we also considered an alternative utility function

$$u(\Sigma_t) = -\sum_{a=1}^{K} \left\{ \text{Var}(\gamma_a|\Sigma_t) + w \times \left(\text{Var}(\gamma_{1,a}|\Sigma_t) + \text{Var}(\gamma_{2,a}|\Sigma_t) \right) \right\}, \quad (14)$$

where γ_a is the difference between arm $a > 0$ and the control arm in the probability of an individual's positive response on both endpoints.

Simulation Study: We discuss a simulation study for a multi-arm trial with four experimental arms and an overall sample size of $T = 348$ patients. For the arms $a = 0, 2, 3, 4$ the parameters $\theta_a = (0.15, 0.25, 0.4, 0.2)$ are identical across the four scenarios that we considered. In Scenario 1, arm $a = 1$ has no PTE while in Scenarios 2–4 $(\gamma_{1,1}, \gamma_{2,1}) = (0.2, 0.2), (0.2, 0.05)$, and $(0.05, 0.2)$ respectively.

Table S8 in the supplementary material, illustrates the MSE of the effects estimates $\widehat{\gamma}_{\ell,a}$, power for null hypotheses $\mathcal{H}_{0,a}^{\ell} : \gamma_{\ell,a} \leq 0$, $\ell = 1, 2$ and the proportion of simulations in which both hypotheses are rejected for BUD designs with utility function (13) (BUD_1) or utility function (14) (BUD_2). We also compared BUD designs to a balanced randomized (BR) design that assigns patients to treatments with equal probabilities. As expected, the BUD_2 design, which seeks to

minimize the posterior variance of the treatment effects, attains the lowest MSEs. The BUD designs have nearly identical power and BR has across all alternative scenarios the lowest power.

We also computed the Bayesian optimal design $d^\star$ using BI with utility defined by (13) for a trial with $K = 2$ experimental arms. We compare the expected utility of the optimal design $d^\star$ to the expected utility of the BUD and BR designs $d = d_{BUD}, d_{BR}$. Expected utilities are computed by integrating with respect to the prior. With co-primary endpoints dynamic programming becomes quickly computationally challenging for moderate sample sizes $T \approx 25$, and realistic sample sizes $T \approx 378$ are infeasible. For values of $T = 5, 10, 15$, we observe a reduction of expected utility equal to (0.59%, 1.21%, 2.04%) for BUD and (36.92%, 61.39%, 77.4%) for BR.

6 Discussion

There are competing objectives in the design of clinical trials. These include identification of treatment effects, accurate estimates, and limiting the number of patients exposed to suboptimal treatments. Bayesian adaptive designs usually target the latter aim, or attempt to balance competing aims, by unbalancing randomization toward more promising arms. In contrast, BUD designs target an information measure u consistent with the goal of testing or estimating treatment effects.

Bayesian adaptive randomization (BAR) has been applied in various clinical studies and, in several settings, it has better operating characteristics than alternative designs [4, 12, 21, 34, 38, 40]. Limitations of BAR have also been discussed in the literature. It is known, for example, that in two-arm trials BAR may increase the overall sample size compared to balanced randomization (Korn and Freindlin, 2011). In multi-arm trials BAR can be highly variable in the arm-specific enrollment proportions and in some cases assigns with non negligible probability more patients to

inferior arms than balanced randomization [32, 34]. This suggests that investigating alternative approaches to Bayesian adaptive trials, such as BUD designs, remains a relevant problem.

Both, BAR and BUD designs translate posterior distributions into randomization probabilities. These two approaches can also be described as randomized myopic decision rules. BAR approximately optimizes the number of positive outcomes [4], while BUD designs approximately optimizes posterior summaries, such as the variance or entropy, that quantify uncertainty on key parameters. In the two-arm case, [4] showed that BAR has, in several scenarios, operating characteristics similar to the Bayesian optimal design when the utility function coincides with the number of positive outcomes. Unlike BAR, BUD designs focus on information measures based on the primary aims of the trial, such as testing efficacy or estimating treatment effects.

We propose several information measures which reflect the aims of multi-arm trials with biomarker subpopulations and multiple endpoints, and show that in a number of simulation scenarios, frequentist operating characteristics are improved with respect to other strategies. In fact, the utility at the end of the experiment is often near the optimal value produced by intractable backward induction algorithms. The simplicity of the randomization probabilities makes it possible to derive asymptotic limits in certain examples.

The computing time to simulate a BUD trial increases linearly with respect to the sample size T and with respect to the number of possible actions. In our examples the computing time for BUD simulations, as expected, is very similar to BAR, and generally significantly faster than computing the optimal Bayesian design obtained with dynamic programming.

A crucial aspect of BUD designs is the selection of information metrics u that represent the primary aims of the clinical study. We can classify studies into broad categories based on their primary aims; (i) estimating treatment effects, (ii) identifying therapies with positive effects, and

(iii) studies with multiple objectives, for instance identifying relevant treatment effects on multiple endpoints or within subgroups.

In case (i), the focus is on key parameters of interest and a symmetric uncertainty metric, such as the entropy or posterior variance, can be used to represent estimation accuracy. The sensitivity analyses in Section 3.2 shows that BUD designs with these information metrics have similar operating characteristics.

In case (ii), as in hypothesis testing, the parameter space is partitioned into regions with and without treatment effects. We can use discrete random variables $E_a \in \{0, 1\}$ or monotone functions $g(\gamma_a)$ representative of this partition. The posterior variance or (asymmetric) entropy of these random variables can be used to define uncertainty measures. For example, in a multi-arm study with two experimental drugs and binary response to treatment, $\gamma_a = \theta_a - \theta_0, a = 1, 2$ indicate the difference between the response rate of experimental arm a and the control. The entropy of $E_a = I(\gamma_a > 0)$ or variance of $g(\gamma_a) = \gamma_a/(1 + \gamma_a^2)^{1/2}$ can be adopted as uncertainty metrics.

In case (iii), when the study has multiple related goals, a composite information measure $u = \sum_j w_j u_j$ can be used to weight different criteria u_j. For example in Section 5, the BUD design for a trial with two endpoints is based on a composite utility function.

References

[1] A. Barker, C. Sigman, G. Kelloff, N. Hylton, D. Berry, and L. Esserman. I-spy 2: An adaptive breast cancer trial design in the setting of neoadjuvant chemotherapy. Clinical Pharmacology & Therapeutics, 86(1):97–100, 2009.

[2] J. Berger. Statistical decision theory and Bayesian analysis. Springer, 1985.

[3] D. A. Berry. Interim analysis in clinical trials: The role of the likelihood principle. The American Statistician, 41(2):117–122, 1987.

[4] D. A. Berry and S. G. Eick. Adaptive assignment versus balanced randomization in clinical trials: a decision analysis. Statistics in medicine, 14(3):231–246, 1995.

[5] D. A. Berry and B. Fristedt. Bandit problems: sequential allocation of experiments. Springer, 1985.

[6] D. A. Berry and C.-H. Ho. One-sided sequential stopping boundaries for clinical trials: A decision-theoretic approach. Biometrics, pages 219–227, 1988.

[7] S. M. Berry and J. B. Kadane. Optimal bayesian randomization. Journal of the Royal Statistical Society: Series B, 59(4):813–819, 1997.

[8] S. M. Berry, B. P. Carlin, J. J. Lee, and P. Muller. Bayesian adaptive methods for clinical trials. CRC press, 2010.

[9] A. E. Brockwell and J. B. Kadane. A gridding method for bayesian sequential decision problems. Journal of Computational and Graphical Statistics, 12(3):566–584, 2003.

[10] S. Bubeck, N. Cesa-Bianchi, et al. Regret analysis of stochastic and nonstochastic multi-armed bandit problems. Foundations and Trends® in Machine Learning, 5(1):1–122, 2012.

[11] B. P. Carlin, J. B. Kadane, and A. E. Gelfand. Approaches for optimal sequential decision analysis in clinical trials. Biometrics, 54(3):964–975, 1998.

[12] M. Cellamare, M. Milstein, S. Ventz, E. Baudin, L. Trippa, and C. Mitnick. Bayesian adaptive randomization in a clinical trial to identify new regimens for mdr-tb: the endtb trial. The International Journal of Tuberculosis and Lung Disease, 20(12):S8–S12, 2016.

[13] M. Cellamare, S. Ventz, E. Baudin, C. D. Mitnick, and L. Trippa. A bayesian response-adaptive trial in tuberculosis: The endtb trial. Clinical Trials, 14(1):17–28, 2017.

[14] Y. Cheng and D. A. Berry. Optimal adaptive randomized designs for clinical trials. Biometrika, 94(3):673–689, 2007.

[15] B. G. Druss, R. M. Rohrbaugh, C. M. Levinson, and R. A. Rosenheck. Integrated medical care for patients with serious psychiatric illness: a randomized trial. Archives of general Psychiatry, 58(9):861–868, 2001.

[16] F. Hu and W. F. Rosenberger. The theory of response-adaptive randomization in clinical trials. John Wiley & Sons, 2006.

[17] C. Jennison and B. W. Turnbull. Group sequential methods with applications to clinical trials. CRC Press, 1999.

[18] S.-H. Jung, T. Lee, K. Kim, and S. L. George. Admissible two-stage designs for phase ii cancer clinical trials. Statistics in medicine, 23(4):561–569, 2004.

[19] J. B. Kadane and P. K. Vlachos. Hybrid methods for calculating optimal few-stage sequential strategies: Data monitoring for a clinical trial. Statistics and Computing, 12(2):147–152, 2002.

[20] E. S. Kim, R. S. Herbst, I. I. Wistuba, J. J. Lee, G. R. Blumenschein, A. Tsao, D. J. Stewart, M. E. Hicks, J. Erasmus, S. Gupta, et al. The battle trial: personalizing therapy for lung cancer. Cancer discovery, 1(1):44–53, 2011.

[21] J. J. Lee, X. Gu, and S. Liu. Bayesian adaptive randomization designs for targeted agent development. Clinical Trials, 7(5):584–596, 2010.

[22] P. Lenca, S. Lallich, and B. Vaillant. Construction of an off-centered entropy for the supervised learning of imbalanced classes: Some first results. Communications in Statistics—Theory and Methods, 39(3):493–507, 2010.

[23] R. J. Lewis and D. A. Berry. Group sequential clinical trials: a classical evaluation of bayesian decision-theoretic designs. Journal of the American Statistical Association, 89(428):1528–1534, 1994.

[24] P. Müller, D. A. Berry, A. P. Grieve, M. Smith, and M. Krams. Simulation-based sequential bayesian design. Journal of Statistical Planning and Inference, 137(10):3140–3150, 2007.

[25] W. Offen, C. Chuang-Stein, A. Dmitrienko, G. Littman, J. Maca, L. Meyerson, R. Muirhead, P. Stryszak, A. Baddy, K. Chen, K. Copley-Merriman, W. Dere, S. Givens, D. Hall, D. Henry, J. D. Jackson, A. Krishen, T. Liu, S. Ryder, A. J. Sankoh, J. Wang, and C.-H. Yeh. Multiple co-primary endpoints: Medical and statistical solutions: A report from the multiple endpoints expert team of the pharmaceutical research and manufacturers of america. Drug Information Journal, 41(1):31–46, 2007.

[26] R. Pemantle. A survey of random processes with reinforcement. Probability surveys, 4:1–79, 2007.

[27] J. Polivka and F. Janku. Molecular targets for cancer therapy in the pi3k/akt/mtor pathway. Pharmacology & therapeutics, 142(2):164–175, 2014.

[28] D. Russo and B. Van Roy. Learning to optimize via information-directed sampling. In Advances in Neural Information Processing Systems, pages 1583–1591, 2014.

[29] P. Schöffski, S. Riggert, P. Fumoleau, M. Campone, O. Bolte, S. Marreaud, D. Lacombe,

B. Baron, M. Herold, H. Zwierzina, et al. Phase i trial of intravenous aviscumine (rviscumin) in patients with solid tumors: a study of the european organization for research and treatment of cancer new drug development group. Annals of Oncology, 15(12):1816–1824, 2004.

[30] N. Stallard and T. Friede. A group-sequential design for clinical trials with treatment selection. Statistics in Medicine, 27(29):6209–6227, 2008.

[31] A. Teixeira-Pinto, J. Siddique, R. Gibbons, and S. L. Normand. Statistical approaches to modeling multiple outcomes in psychiatric studies. Psychiatric annals, 39(7):729–735, 2009.

[32] P. Thall, P. Fox, and J. Wathen. Statistical controversies in clinical research: scientific and ethical problems with adaptive randomization in comparative clinical trials. Annals of Oncology, 26(8):1621–1628, 2015.

[33] P. F. Thall. Bayesian models and decision algorithms for complex early phase clinical trials. Statistical Science, 25(2):227, 2010.

[34] P. F. Thall and J. K. Wathen. Practical bayesian adaptive randomisation in clinical trials. European Journal of Cancer, 43(5):859–866, 2007.

[35] P. F. Thall, R. M. Simon, and E. H. Estey. Bayesian sequential monitoring designs for single-arm clinical trials with multiple outcomes. Statistics in medicine, 14(4):357–379, 1995.

[36] W. R. Thompson. On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. Biometrika, 25(3/4):285–294, 1933.

[37] L. Trippa and B. M. Alexander. Bayesian baskets: A novel design for biomarker-based clinical trials. Journal of Clinical Oncology, 35:681–687, 2017.

[38] L. Trippa, E. Q. Lee, P. Y. Wen, T. T. Batchelor, T. Cloughesy, G. Parmigiani, and B. M. Alexander. Bayesian adaptive randomized trial design for patients with recurrent glioblastoma. Journal of Clinical Oncology, 30(26):3258–3263, 2012.

[39] S. Ventz and L. Trippa. Bayesian designs and the control of frequentist characteristics: A practical solution. Biometrics, 71(1):218–226, 2015.

[40] S. Ventz, W. T. Barry, G. Parmigiani, and L. Trippa. Bayesian response-adaptive designs for basket trials. Biometrics, 73(3):905–915, 2017.

[41] S. Ventz, G. Parmigiani, and L. Trippa. Combining bayesian experimental designs and frequentist data analyses: motivations and examples. Applied Stochastic Models in Business and Industry, 33(3):302–313, 2017.

[42] S. Ventz, M. Cellamare, G. Parmigiani, and L. Trippa. Adding experimental arms to platform clinical trials: randomization procedures and interim analyses. Biostatistics, 19(2):199–215, 2018.

[43] J. Wason and T. Jaki. Optimal design of multi-arm multi-stage trials. Statistics in medicine, 31(30):4269–4279, 2012.

[44] L. Wei. The generalized polya's urn design for sequential medical trials. The Annals of Statistics, pages 291–296, 1979.

[45] Y. Xu, L. Trippa, P. Müller, and Y. Ji. Subgroup-based adaptive (suba) designs for multi-arm biomarker trials. Statistics in Biosciences, 8(1):159–180, 2016.

[46] Y. Zhang, L. Trippa, and G. Parmigiani. Optimal bayesian adaptive trials when treatment efficacy depends on biomarkers. Biometrics, 72(2):414–421, 2016.

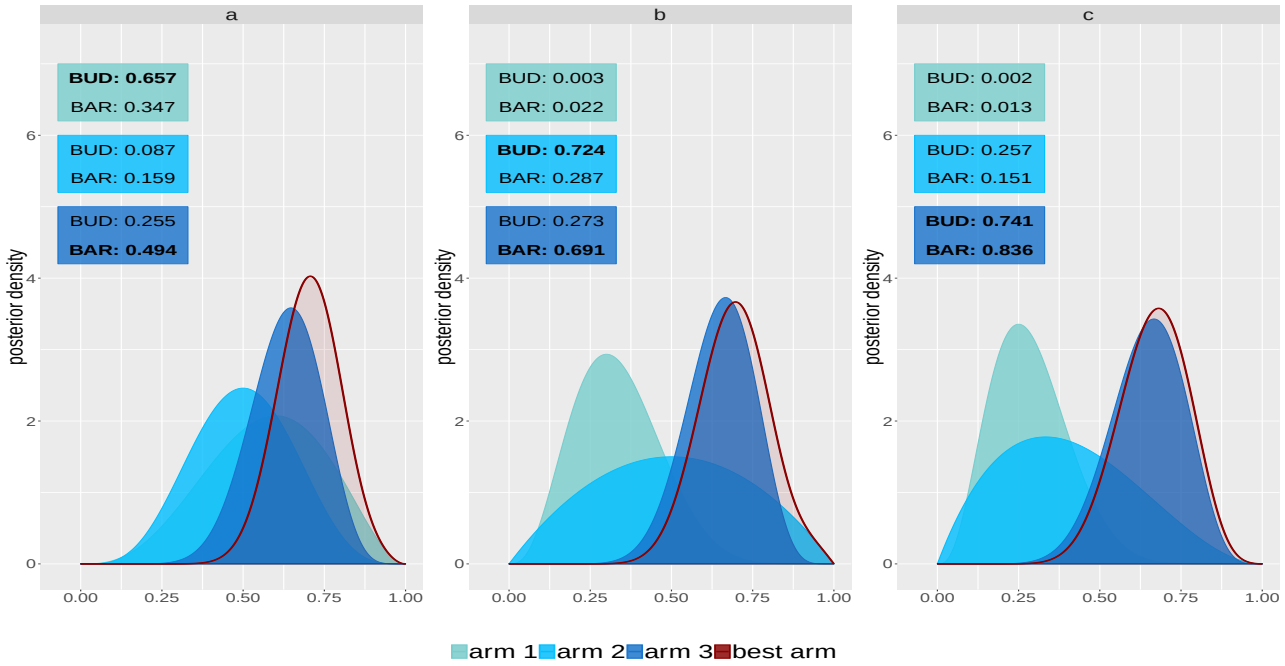

Figure 1: Comparison of a Bayesian uncertainty directed (BUD) design and Bayesian adaptive randomization (BAR). We consider an early stage 3-arm trial without control, with the primary goal of selecting the best experimental arm $a^\star = \arg\max_{a=1,2,3} \theta_a$. The total sample size is equal to $T = 100$. BAR defines randomization probabilities to treatments $a = 1, 2, 3$ that mirror the posterior distribution of a^* [36]. The BUD design randomizes patients with the goal of approximately minimizing the posterior entropy of $a^\star$. Panels (a), (b) and (c) show three different simulations of the 3-arm trial after the enrollment of 50 patients. The blue curves show the posterior densities of the response probabilities θ_a. The red curves indicate the posterior distribution of the maximum $\theta_{a^\star} = \max_{a=1,2,3} \theta_a$. In each panel we also indicate for BAR and BUD designs the randomization probabilities of the next patient $t = 51$ that will be enrolled in the trial.

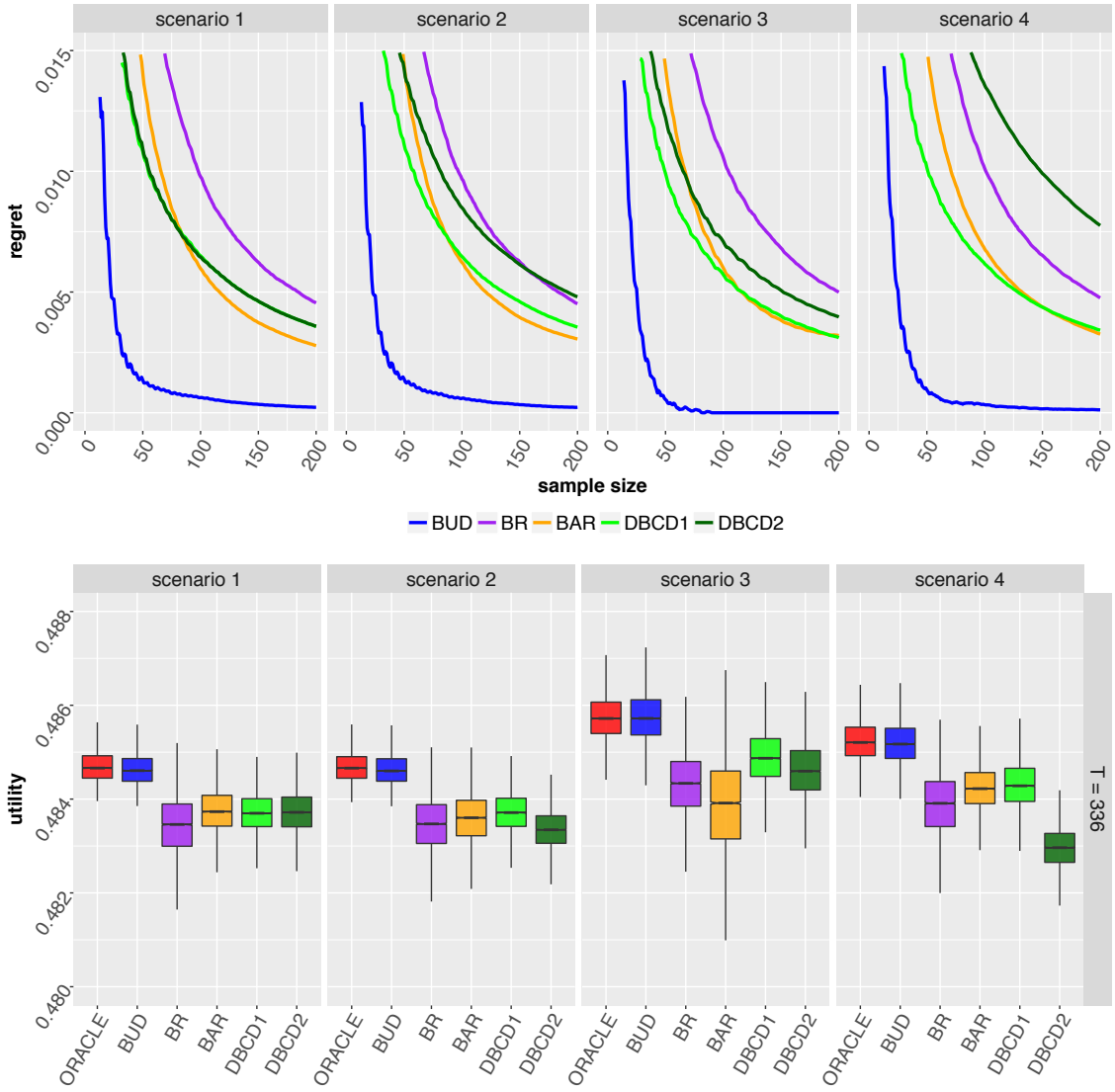

Figure 2: Information collected at the end of a multi-arm clinical trial with four experimental arms. Results are based on 5,000 simulations of a clinical study with four experimental arms using balanced randomization (BR), Bayesian uncertainty directed (BUD) design, Bayesian adaptive randomization (BAR) and two versions of the doubly adaptive coin design (DBCD1 and DBCD2). The top panel shown for each of the four designs the regret function, which is defined as the difference between the expected information $\mathbb{E}[u(\Sigma_T)]$ of the design and the expected information of the oracle design for different values of $T = 1, 2, \ldots, 200$. The bottom panel shows boxplots of the utility $u(\Sigma_t) = \sum_{k=1}^{K} \Big(v_k - \text{Var}(\gamma_k \mid \Sigma_t)\Big)$, with prior variance v_k, across the simulated trials for a sample size of $T = 336$.

	Control	Arm 1			Arm 2			Arm 3		
Design	ESS (SD)	ESS (SD)	Power	MSE	ESS (SD)	Power	MSE	ESS (SD)	Power	MSE
	Scenario 1: no effective arm $(\theta_0, \theta_1, \theta_2, \theta_3) = (0.4, 0.4, 0.4, 0.4)$									
BR	84 (8)	84 (08)	04.0	5.89	84 (8)	04.1	5.86	84 (8)	03.2	5.74
BUD	118 (3)	73 (03)	03.8	**5.44**	73 (3)	03.8	**5.43**	73 (3)	04.0	**5.52**
BAR	97 (8)	80 (21)	03.9	7.33	79 (21)	03.6	7.30	80 (21)	03.8	7.28
BAR2	84 (23)	84 (22)	03.9	7.26	84 (22)	04.0	7.17	84 (23)	04.2	7.37
DBCD1	84 (4)	84 (4)	03.8	5.93	84 (4)	03.8	5.85	84 (4)	03.6	5.82
DBCD2	84 (6)	84 (6)	03.6	5.89	84 (6)	03.9	5.92	84 (6)	03.5	5.82
	Scenario 2: one superior arm $(\theta_0, \theta_1, \theta_2, \theta_3) = (\mathbf{0.4}, \mathbf{0.6}, 0.4, 0.4)$									
BR	85 (8)	84 (8)	78.6	5.75	83 (8)	04.2	5.75	84 (8)	03.3	5.80
BUD	118 (3)	73 (3)	82.2	5.48	73 (3)	03.6	**5.46**	73 (3)	03.8	**5.40**
BAR	100 (10)	103 (14)	**87.5**	**4.89**	67 (18)	03.2	7.65	66 (18)	03.6	7.96
BAR2	58 (17)	161 (28)	85.1	6.80	58 (17)	3.4	9.80	58 (17)	3.5	9.65
DBCD1	84 (4)	84 (4)	79.8	5.99	84 (4)	03.9	6.07	84 (4)	04.0	6.02
DBCD2	80 (6)	97 (6)	81.2	5.63	80 (6)	03.4	6.16	80 (6)	03.7	6.09
	Scenario 3: one superior and one inferior arm $(\theta_0, \theta_1, \theta_2, \theta_3) = (\mathbf{0.4}, \mathbf{0.6}, 0.4, 0.2)$									
BR	84 (8)	84 (8)	78.9	5.92	84 (8)	03.6	5.70	84 (8)	0.00	4.72
BUD	122 (4)	75 (3)	84.3	5.10	75 (3)	03.6	**5.22**	63 (5)	0.00	**4.68**
BAR	111 (10)	117 (13)	**90.6**	**4.40**	75 (20)	04.0	7.22	32 (11)	0.00	8.12
BAR2	62(18)	180 (27)	88.2	6.24	62 (18)	3.4	9.04	31 (8) 0.0	9.95	
DBCD1	88 (4)	88 (4)	81.1	5.69	88 (4)	03.7	5.66	72 (7)	0.00	5.20
DBCD2	85 (7)	104 (6)	83.8	5.25	85 (7)	03.8	5.95	61 (8)	0.00	5.76
	Scenario 4: three superior arms $(\theta_0, \theta_1, \theta_2, \theta_3) = (\mathbf{0.4}, \mathbf{0.6}, \mathbf{0.65}, \mathbf{0.7})$									
BR	84 (8)	85 (8)	79.4	5.89	84 (8)	92.5	5.86	84 (8)	98.6	5.74
BUD	120 (3)	74 (3)	**83.6**	**5.34**	72 (3)	**95.1**	**5.18**	70 (4)	**99.1**	**5.14**
BAR	90 (4)	80 (9)	79.2	7.33	82 (8)	93.4	7.30	83 (8)	98.6	7.28
BAR2	35 (10)	75 (22)	55.1	11.94	96 (26)	78.8	10.77	130 (29)	95.3	9.87
DBCD1	86 (4)	86 (4)	80.4	5.93	84 (5)	94.0	5.85	80 (5)	98.6	5.82
DBCD2	70 (6)	85 (5)	75.8	5.89	89 (5)	91.7	5.92	92 (5)	98.4	5.82

Table 1: Expected sample size (ESS), standard deviation (SD), power and mean squared error (MSE), for each experimental arm in a 4-arm trial with 336 patients using either balanced randomization (BR), Bayesian uncertainty directed (BUD) design, Bayesian adaptive randomization (BAR) as described in [38] and the version described in [34] (BAR2), and two versions of the doubly adaptive biased coin design, DBCB1 and DBCD2. The MSE is scaled by a factor of 10^3.

T	Design	Arm 1 ESS (SD)	P_1	Arm 2 ESS (SD)	P_2	Arm 3 ESS (SD)	P_3	Arm 4 ESS (SD)	P_4	MSE
	Scenario 1			$(\theta_1, \theta_2, \theta_3, \theta_4) = (0.3, 0.4, 0.5, \mathbf{0.6})$						
30	BR	7 (3)	0.059	8 (3)	0.155	8 (2)	0.269	8 (4)	0.517	13.63
	BUD	4 (10)	0.035	6 (8)	0.117	8 (10)	0.274	12 (12)	**0.573**	**8.83**
	BAR	4 (12)	0.044	6 (7)	0.123	8 (7)	0.273	13 (13)	0.560	9.16
	RPW	7 (4)	0.056	7 (3)	0.133	8 (3)	0.280	8 (4)	0.531	11.9
50	BR	12 (3)	0.033	13 (3)	0.101	12 (3)	0.276	13 (5)	0.590	9.22
	BUD	6 (15)	0.018	9 (12)	0.075	13 (16)	0.247	22 (15)	**0.659**	**6.35**
	BAR	5 (16)	0.024	8 (15)	0.086	13 (10)	0.251	24 (21)	0.640	6.80
	RPW	11 (4)	0.031	12 (4)	0.091	13 (3)	0.2592	14 (6)	0.619	8.23
70	BR	17 (3)	0.015	18 (4)	0.079	17 (4)	0.263	18 (5)	0.643	7.65
	BUD	7 (23)	0.008	11 (18)	0.054	18 (20)	0.222	34 (15)	**0.715**	**5.10**
	BAR	6 (23)	0.011	10 (25)	0.060	18 (16)	0.226	36 (30)	0.703	5.31
	RPW	15 (6)	0.013	17 (5)	0.067	18 (4)	0.240	20 (9)	0.680	6.77
	Scenario 2			$(\theta_1, \theta_2, \theta_3, \theta_4) = (0.4, 0.4, 0.4, \mathbf{0.8})$						
30	BR	7 (3)	0.049	8 (3)	0.058	8 (2)	0.040	8 (4)	0.853	13.95
	BUD	3 (10)	0.024	4 (11)	0.026	3 (13)	0.028	20 (12)	**0.921**	**11.98**
	BAR	3 (12)	0.035	3 (8)	0.036	3 (13)	0.033	20 (13)	0.894	13.79
	RPW	7 (4)	0.041	7 (2)	0.039	7 (7)	0.041	10 (2)	0.878	13.47
50	BR	12 (3)	0.022	13 (3)	0.022	12 (3)	0.013	13 (5)	0.943	10.94
	BUD	4 (17)	0.006	4 (20)	0.008	4 (23)	0.006	37 (20)	**0.979**	**6.33**
	BAR	4 (20)	0.014	4 (18)	0.016	4 (23)	0.014	38 (22)	0.955	8.11
	RPW	11 (4)	0.014	11 (5)	0.013	11 (11)	0.012	17 (4)	0.960	10.03
70	BR	17 (3)	0.007	18 (4)	0.008	17 (4)	0.003	18 (5)	0.982	8.61
	BUD	5 (26)	0.002	5 (30)	0.003	5 (33)	0.003	55 (30)	**0.992**	**4.01**
	BAR	5 (30)	0.006	5 (28)	0.008	4 (33)	0.008	57 (32)	0.977	5.29
	RPW	15 (4)	0.005	16 (5)	0.004	15 (15)	0.006	24 (7)	0.984	7.59
	Scenario 3			$(\theta_1, \theta_2, \theta_3, \theta_4) = (0.35, 0.45, 0.7, \mathbf{0.8})$						
30	BR	7 (3)	0.012	8 (3)	0.056	8 (2)	0.314	8 (4)	0.618	8.12
	BUD	3 (10)	0.006	3 (8)	0.024	9 (11)	0.290	15 (8)	**0.679**	**8.01**
	BAR	2 (12)	0.012	3 (10)	0.029	9 (12)	0.306	15 (11)	0.652	8.99
	RPW	6 (6)	0.016	6 (3)	0.034	8 (7)	0.315	9 (5)	0.635	8.37
50	BR	12 (3)	0.005	13 (3)	0.021	12 (3)	0.277	13 (5)	0.697	6.57
	BUD	3 (13)	0.002	4 (11)	0.006	14 (20)	0.241	28 (15)	**0.751**	**5.11**
	BAR	3 (22)	0.003	4 (18)	0.011	15 (22)	0.272	28 (19)	0.714	5.86
	RPW	10 (9)	0.003	11 (2)	0.014	14 (10)	0.273	16 (7)	0.711	6.43
70	BR	17 (3)	0.001	18 (4)	0.011	17 (4)	0.242	18 (5)	0.747	5.57
	BUD	4 (18)	0.001	5 (17)	0.001	19 (26)	0.203	42 (25)	**0.795**	**3.70**
	BAR	3 (32)	0.001	4 (28)	0.005	20 (32)	0.241	42 (28)	0.752	4.61
	RPW	13 (10)	0.001	15 (2)	0.003	20 (14)	0.239	22 (11)	0.757	5.03

Table 2: Expected sample size (ESS), standard deviation (SD), proportion of simulations that selected therapy $a = 1, \cdots, 4$ as the most effective treatment (P_a) and mean squared error (MSE) of the response rate of the best arm for 4-arm trial using balanced randomization (BR), Bayesian uncertainty directed (BUD) design, Bayesian adaptive randomization (BAR) and the randomized play the winner (RPW) design. The overall sample size equals either $T = 30, 50$ or 70. Results are based on 10,000 simulated trials. The MSE is scaled by a factor of 10^3.

	Expected Sample Size in BMK-positive groups (SD)					Power	
Biomarker (BMK)		1	2	3	4	Po_+	Po_-
Scenario 1	arms $a = 1, \cdots, 4$ target BMKs a, BMK prevalence (0.5,0.5,0.5,0.5)						
BR	each arm	50(7)	50(7)	50(7)	50(7)		
BUD	control	91 (11)	90(11)	91(11)	91(11)		
	arm 1	**49(30)**	37(26)	37(26)	36(26)	10.15	9.96
Scenario 2	arms $a = 1, \cdots, 4$ target BMKs a, BMK prevalence (0.5,0.5,0.5,0.5)						
BR	control	50(7)	50(7)	50(7)	50(7)		
	arm 1	**50(7)**	50(7)	50(7)	50(7)	**77.09**	9.97
	arm 2	50(7)	**50(7)**	50(7)	50(7)	9.98	9.80
BUD	control	90(12)	90(11)	90(10)	90(10)		
	arm 1	**47(29)**	38(24)	38(24)	38(23)	**86.92**	10.00
	arm 2	37(26)	**49(30)**	37(26)	38(26)	10.35	10.92
Scenario 3	arms $a = 1, \cdots, 4$ target BMKs a, BMK prevalence (0.7,0.3,0.5,0.5)						
BR	control	70(8)	30(5)	50(7)	50(7)		
	arm 1	**70(8)**	30(5)	50(7)	50(7)	**86.47**	10.44
	arm 2	70(8)	**30(5)**	50(7)	50(7)	**61.29**	10.40
	arm 3	70(8)	30(5)	**50(7)**	50(7)	9.86	10.24
BUD	control	120(13)	57(8)	89(10)	89(10)		
	arm 1	**52(33)**	16(11)	34(21)	34(21)	**92.52**	10.81
	arm 2	59(32)	**35(18)**	38(22)	38(22)	**75.82**	10.28
	arm 3	60(38)	21(16)	**51(30)**	39(26)	9.87	9.44
Scenario 4	arms $1, 2, 3, 4$ target BMKs $1, 1, 2, 3$, BMK prevalence (0.5,0.5,0.5)						
BR	control	50(7)	50(7)	50(7)			
	arm 1	**50(7)**	50(7)	50(7)		**76.90**	10.41
	arm 2	**50(7)**	50(7)	50(7)		**96.15**	10.53
	arm 3	50(7)	50(7)	**50(7)**		10.08	10.61
BUD	control	89(11)	90(11)	90(11)			
	arm 1	**47(27)**	38(22)	38(22)		**87.45**	10.30
	arm 2	**28(18)**	26(16)	26(16)		**97.17**	10.37
	arm 3	43(30)	**54(30)**	41(28)		10.43	9.13
Scenario 5	arms 1,2,3,4 target BMKs 1,1,2,2, BMK prevalence (0.5,0.6)						
BR	control	50(7)	60(7)				
	arm 1	**50(7)**	60(7)			**59.41**	**59.69**
	arm 2	**50(7)**	60(7)			**76.51**	10.54
	arm 3	50(7)	**60(7)**			**81.93**	10.07
	arm 4	50(7)	**60(7)**			**97.76**	10.57
BUD	control	85(10)	102(11)				
	arm 1	**38(23)**	52(27)			**61.78**	**67.99**
	arm 2	**52(24)**	62(30)			**89.04**	9.87
	arm 3	45(23)	**54(27)**			**92.48**	9.98
	arm 4	30(17)	**29(17)**			**98.52**	9.71

Table 3: Average sample size, standard deviation (SD) and power in the biomarker positive (targeted) and negative subgroups (Po_+ and Po_-) for a 5-arm trial with an overall sample size of $T = 500$. Results are based on 5,000 simulations of a trial using either BUD or BR designs. For each scenario, the patients' biomarker status has been generated independently for each marker according to the specified prevalences. The response rate for the control therapy is constant across scenarios and equal to 0.35. The first three scenarios refer to a trial with four biomarkers, and therapy $a = 1$ (scenarios 2,3) and $a = 2$ (scenario 3) have a positive treatment effect (PTE) for patients with biomarkers $X_{i,a} = 1$. Scenarios 4 and 5 correspond to a trial with three and two biomarkers respectively. Therapies $a = 1, 2$ in Scenario 4 have PTEs for patients with $X_{i,a} = 1$. In scenario 5 all therapies $a = 2, 3, 4$ have a PTE in their target population, and arm $a = 1$ PTEs extend to all patients.