Journal of the Royal Statistical Society Series B: Statistical Methodology, 2025, **87**, 1055–1084
https://doi.org/10.1093/jrsssb/qkaf006
Advance access publication 24 February 2025
Original Article

ROYAL STATISTICAL SOCIETY
DATA | EVIDENCE | DECISIONS
Journal of the Statistics Society
Series B
Statistical Methodology

Adaptive experiments toward learning treatment effect heterogeneity

Waverly Wei[1], Xinwei Ma[2] and Jingshen Wang[3]

[1]Department of Data Sciences and Operations, University of Southern California, Los Angeles, CA, USA
[2]Department of Economics, University of California, San Diego, CA, USA
[3]Division of Biostatistics, University of California, Berkeley, CA, USA

Address for correspondence: Jingshen Wang, Division of Biostatistics, University of California, 5408 Berkeley Way West, Berkeley, CA 94720, USA. Email: jingshenwang@berkeley.edu

Abstract

Understanding treatment effect heterogeneity has become an increasingly popular task in various fields, as it helps design personalized advertisements in e-commerce or targeted treatment in biomedical studies. However, most of the existing work in this research area focused on either analysing observational data based on strong causal assumptions or conducting post hoc analyses of randomized controlled trial data, and there has been limited effort dedicated to the design of randomized experiments specifically for uncovering treatment effect heterogeneity. In the manuscript, we develop a framework for designing and analysing response adaptive experiments toward better learning treatment effect heterogeneity. Concretely, we provide response adaptive experimental design frameworks that sequentially revise the data collection mechanism according to the accrued evidence during the experiment. Such design strategies allow for the identification of subgroups with the largest treatment effects with enhanced statistical efficiency. The proposed frameworks not only unify adaptive enrichment designs and response-adaptive randomization designs but also complement A/B test designs in e-commerce and randomized trial designs in clinical settings. We demonstrate the merit of our design with theoretical justifications and in simulation studies with synthetic e-commerce and clinical trial data.

Keywords: covariate-adjusted response-adaptive designs, design of experiments, frequentist adaptive design, subgroup analysis

1 Introduction

1.1 Motivation

Understanding and characterizing treatment effect heterogeneity has become increasingly important in many scientific fields. For example, identifying differential treatment effects is an important step toward materializing the benefits of precision health, as it provides evidence regarding how groups of patients with specific characteristics respond to a given treatment either in efficacy or in adverse effects (He et al., 2019). As another example, individuals from different socioeconomic backgrounds may benefit differently from government programmes, meaning that a careful evaluation of the programme's possibly heterogeneous impacts is crucial for effective policy-making (Karlan & Zinman, 2008; Kharitonov et al., 2015).

The existing literature in this research area mostly focuses on conducting retrospective analyses that employ observational or randomized experiment data. Even with large-scale observational data or carefully collected randomized experiment data, statistical bias in these analyses cannot be overlooked. On the one hand, in observational studies, statistical bias can arise due to potential violations of untestable causal assumptions. For example, one of the commonly imposed causal assumptions in practice is the unconfoundedness assumption (Athey et al., 2018; Cattaneo et al., 2019; Djebbari & Smith, 2008; Hill, 2011; Huang et al., 2012; Ma & Wang, 2020;

Received: December 11, 2023. **Revised:** January 10, 2025. **Accepted:** January 15, 2025

Raudenbush & Bloom, 2015), which states that conditional on measured confounders, the treatment assignment is as good as random. Due to the untestable nature of the unconfoundedness assumption and the possibility of unmeasured confounders in observational data, the validity of established causal conclusions under this assumption cannot be guaranteed. On the other hand, when analysing classical randomized experiment data, although carrying out valid causal conclusion does not require imposing untestable causal assumptions, exploring treatment effect heterogeneity could still be susceptible to the winner's curse bias if seemingly promising heterogeneous treatment effects are selected from the data in an ad hoc fashion (Andrews et al., 2024; Guo & He, 2021; Ma et al., 2023; Stallard et al., 2008).

In this manuscript, we tackle the problem of uncovering treatment effect heterogeneity from a fresh perspective by proposing an adaptive data collection mechanism called 'adaptive randomized experiments.' This adaptive experiment approach enables the collection of reliable causal evidence that is specifically focused on understanding treatment effect heterogeneity. By adaptive, we mean that experimenters have the flexibility to sequentially allocate and modify experimental efforts, such as adjusting the treatment allocation probability and proportions of sequentially enrolled (sub)groups of participants. This adaptability allows experimenters to respond and make adjustments based on the evidence accrued during the experiment (Follmann, 1997; F. Hu & Rosenberger, 2006; Rosenberger & Lachin, 1993; Rosenblum et al., 2020; D. J. Russo et al., 2018; Simchi-Levi & Wang, 2023a, 2023b; Villar et al., 2015; Xu et al., 2016); see Section 1.3 for literature review and Section 2 for a detailed introduction. To collect robust evidence towards learning treatment effect heterogeneity, we formalize our experimental design goal as maximizing the probability of correctly selecting subpopulations (or subgroups) who respond favourably to the treatment motivated by the large/moderate deviation principles (Section 3). Without loss of generality, we refer to the subpopulation with the highest treatment level as the best subpopulation or best subgroup throughout this manuscript.

1.2 Our contribution

In what follows, we break down our contributions from three perspectives:

Our proposed adaptive experiment strategy offers two potential benefits compared with the classical post hoc analysis approaches. First, because treatments are randomly assigned in adaptive experiments, they are independent of any potential unmeasured confounding variables, which means the proposed adaptive experiment design strategy generates samples enabling valid causal conclusions without imposing any untestable causal assumptions (see Section 2 for our experimental setup). This is in stark contrast to analyses using observational data. Second, compared with conducting post hoc analyses with randomized experiment data, our design is equipped with the flexibility to revise the experimental strategy sequentially. Thus, the proposed adaptive experimental design strategy can detect individuals who respond favourably to the treatment and then optimize experimental effort spending based on the inferred context. As a part of this endeavour, compared to analysing data collected from completely randomized experiments, this design feature offers advantages not only in improving the statistical efficiency of detecting treatment heterogeneity (Proposition 2) but also in reducing the necessary sample size to correctly identify the best subgroup (Proposition 3).

Next, on the theoretical side, we first leverage the large and moderate deviation principle to characterize an 'oracle' allocation strategy, which maximizes the asymptotic probability of detecting the best subgroup when the underlying data generation process is unknown (Section 3). Because this oracle allocation strategy depends on the unknown data-generating process, the empirically feasible design strategy must be sequentially revised using adaptively collected data. As our second theoretical contribution, we demonstrate the oracle allocation is attainable using our proposed design strategies (Theorem 1). Unlike classical RAR designs, we do not restrict the potential outcomes to following any parametric form, hence alleviating the burden of choosing what type of parametric assumptions should be used in practice. Our third contribution concerns the large-sample properties of the estimated treatment effects. Under mild moment restrictions, we show that the proposed design delivers asymptotically normally distributed estimators for subgroup treatment effects, and in particular, for the effect size of the best subgroup. In addition, we provide consistent asymptotic variance estimators and hence offer valid statistical inference procedures (Theorem 2).

We tackle three major challenges in our theoretical investigation, which arise as the data are sequentially collected, and hence, they are not statistically independent. First, leveraging upon martingale methods, we develop a general framework and show that consistent estimation of features of the potential outcome distributions is possible for a large class of adaptive experimental designs (online supplementary material, Lemma E.1 and its corollary), provided that the allocation probability and the enrichment proportion are bounded away from zero. Importantly, this result does not require independence and allows the data to be collected in an adaptive manner. More broadly, this technical lemma not only unifies adaptive enrichment (AE) designs and RAR but also applies to other adaptive design settings with objective functions different from ours. We expect the general consistency result to be of independent interest. The second challenge in our theoretical investigation is that the optimized treatment assignment rule, both in finite samples and asymptotically, may not be unique, rendering standard M-estimation proof techniques inapplicable to our setting. Nevertheless, building on the general consistency result (online supplementary material, Lemma E.1), we are able to show that with probability approaching one, any empirically optimized treatment assignment rule will be arbitrarily close to one of the oracle allocations (online supplementary material, Lemma E.2). Such 'set consistency' result, seems new in the literature on adaptive experimental design. Finally, we establish the asymptotic normality of our treatment effect estimators. This result is established with a martingale central limit theorem (Hall & Heyde, 1980), and hence, it takes into account the adaptive nature of the proposed experimental design and the induced dependence. In this process, a key step is to show that the cumulative empirical treatment probability (or the enrichment proportion) also converges to its oracle counterpart, which in turn helps to verify that a conditional variance will stabilize asymptotically. See online supplementary material, Section C.7 for details and additional discussions.

From a practical point of view, our proposed adaptive experiment strategy presents a unified framework incorporating both classical enrichment designs and response-adaptive randomization (RAR) designs. Furthermore, our framework can be applied in both multistage and fully adaptive settings. See Table 2 for a summary of our design strategies. Thanks to its versatility, our designs can be applied to online experiments conducted in e-commerce platforms, clinical trials conducted in health industries, and policy evaluation experiments conducted for social science research.

1.3 Existing literature

Adaptive experiments have been frequently adopted in clinical trials where patients are enrolled sequentially based on certain eligibility criteria. In recent years, they also gained considerable popularity among online platforms for conducting A/B tests or digital randomized experiments. Our design strategy, proposed from a frequentist perspective, does not impose any parametric assumptions on the underlying data-generating process. Our setting is distinctly apart from Bayesian adaptive designs (Atkinson & Biswas, 2005; Cheng & Shen, 2005; Gsponer et al., 2014; Park et al., 2022; Thall & Wathen, 2007).

Our framework falls into the realm of frequentist response adaptive designs. Such design strategies can be roughly divided into RAR design and AE design in the existing literature. Response-adaptive randomization design often refers to the design strategy in which the treatment assignment probabilities are adapted during the experiment based on the accrued evidence in the outcomes, with the goal of simultaneously achieving the experimental objectives and preserving statistical inference validity (F. Hu & Rosenberger, 2006; Robertson et al., 2023; Rosenberger, 2002). The classical RAR framework often revises the treatment assignment probabilities at infinitely many stages, a design strategy we refer to as fully adaptive settings. In fully adaptive settings, a popular class of RAR designs is the doubly adaptive biased coin design (DBCD). The early DBCD can be found in (Eisele, 1994), which has its root in Efron's biased coin design (Efron, 1971). The asymptotic properties of DBCD are studied in various works (F. Hu & Zhang, 2004b; F. Hu et al., 2009; Tymofyeyev et al., 2007; Zhu et al., 2023). Our work shares some connections with (Tymofyeyev et al., 2007) in that we both incorporate an optimization perspective into the problem of finding optimal treatment allocation, although our design objectives and the implementation of the optimal treatment allocation differ. Other than the fully adaptive settings, existing RAR designs also accommodate multistage settings (Pocock, 1977; van der Laan, 2008; J. Zhao, 2023). Such multistage designs propose to revise the treatment assignment probability by minimizing the

asymptotic variance of the average treatment effect estimator. Nevertheless, this design is carried out in two stages and is not designed to identify treatment effect heterogeneity.

Instead of relying solely on the outcome variable to optimize for the experimental goals in RAR, one may further incorporate covariate information. Response-adaptive randomization that further incorporate covariate information is known as covariate-adjusted response-adaptive (CARA) designs (Bandyopadhyay & Biswas, 1999; Rosenberger et al., 2001; van der Laan, 2008; L.-X. Zhang et al., 2007). Early work in Zelen (1994) proposes to balance covariates based on the biased coin design. J. Hu et al. (2015) propose a family of CARA designs that could account for both efficiency and ethics. Zhu and Zhu (2023) generalizes CARA to incorporate semiparametric estimates. Some related CARA designs are also discussed in Y. Lin et al. (2015), Villar and Rosenberger (2018), W. Zhao et al. (2022), and J. Zhao (2023). We note that much of the existing work on RAR designs and CARA aims to optimize the estimation efficiency of the overall treatment effect but is not tailored to study treatment effect heterogeneity (F. Hu & Rosenberger, 2003; Rosenberger & Hu, 2004).

On top of RAR, AE designs are often adopted in clinical trials, and interim data is used to identify treatment-sensitive patient subgroups by changing patient enrolment criteria. In these designs, experimenters often partition the population into predefined subgroups based on biomarkers measured at baseline and enrol patients in multiple stages (Burnett et al., 2020; Rosenblum et al., 2014; Rosenblum & van der Laan, 2011; Wang et al., 2007). For example, the early work in Follmann (1997) considers revising the enrolment proportions of two discrete patient subgroups defined by a single biomarker and provides conditions under which the Type I error rate is controlled. Stallard (2022) considers overlapping subgroups defined by a continuous biomarker. To our knowledge, different from our goal of identifying the best subgroup with high probability, much of the existing work on AE designs aims to preserve the Type I error rate of the estimated subgroup treatment effect.

As the data are sequentially collected using our design strategy, the toolkit we adopted to ensure the validity of statistical inference (martingale limit theories in particular) has also been used in the existing literature (Glynn & Juneja, 2004). For example, Luedtke and Van Der Laan (2016), Hadad et al. (2021), and Zhan et al. (2021) focus on analysing adaptively collected data either from adaptive randomized experiments or online policy learning, different from our goal in designing an adaptive data collection mechanism.

While our work is connected to the classical design of experiments literature, we believe that the adaptive nature of our framework sets it apart. The early experimental designs can be traced back to Fisher (1936), which introduces the design principles such as blocking, randomization, and replication. The seminal work by Jeff Wu and Hamada (2011) lays down the foundation for diverse techniques and theories in experiment designs. For example, the orthogonal designs are a way to ensure that experiments yield clear, independent insights about each factor, thereby maximizing the information return on the experimental efforts and ensuring more reliable conclusions (Butler, 2001; C. D. Lin et al., 2010; Sun & Tang, 2017). As another example, with the advancement of computational capacity in modern designs, (Wu & Xu, 2001) introduces a generalized minimum aberration criterion for evaluating asymmetrical factorial designs. A thorough review of factorial designs can be found in Mukerjee and Wu (2006) and the reference therein. Lastly, our work is also connected to the board class of bandit literature and the literature on learning optimal policy (Audibert et al., 2010; Dudik et al., 2011; Garivier & Kaufmann, 2016; Kasy & Sautmann, 2021; Kaufmann et al., 2016; D. Russo, 2020; Simchi-Levi & Wang, 2023a, 2023b; Simchi-Levi et al., 2023). We provide a more detailed review in online supplementary material, Section A.

2 A synthesized adaptive experiment framework

In this section, we introduce a unified design framework in a two-arm (a treatment arm and a control arm) experiment. Our design framework operates within the frequentist framework. It encompasses classical RAR design with AE design, both widely used in practice.

Suppose experiment participants are sequentially enrolled in T stages. The total number of enrolled subjects is $N = \sum_{t=1}^{T} n_t$, where n_t denotes the number of subjects in Stage t, for $t = 1, \ldots, T$. In Stage t, we denote the treatment assignment status of subject i as $D_{it} \in \{0, 1\}$, where i ranges

from 1 to n_t. Here, $D_{it} = 1$ corresponds to the treatment arm, while $D_{it} = 0$ corresponds to the control arm. Denote subject i's covariate information as $X_{it} \in \mathbb{R}^p$ and the observed outcome as $Y_{it} \in \mathbb{R}$.

To formally introduce treatment (or causal) effects, we follow the potential outcomes model. Define $Y_{it}(d)$ as the potential outcome we would have observed if subject i receives treatment d at Stage t, for $d \in \{0, 1\}$. The observed outcome can then be written as

$$Y_{it} = D_{it}Y_{it}(1) + (1 - D_{it})Y_{it}(0), \quad i = 1, \ldots, n_t, \ t = 1, \ldots, T.$$

Consistent with the existing literature on adaptive experiments, we assume that the outcomes are observed without delay, and their underlying distributions do not shift over time (F. Hu & Rosenberger, 2006). Furthermore, we define the history, which represents the collected data up to Stage t,

$$\mathcal{H}_t = \{\mathcal{H}_s\}_{s=1}^t \triangleq \{(Y_{is}, D_{is}, X_{is}), i = 1, \ldots, n_s\}_{s=1}^t.$$

To investigate treatment effect heterogeneity, we partition the covariate sample space $\mathcal{X}$ into m prespecified nonoverlapping regions, denoted as $\{\mathcal{S}_j\}_{j=1}^m$ (an extension of an overlapping division shall be discussed in online supplementary material, Section I). In clinical settings, each partition of the sample space is commonly referred to as a subgroup (Assmann et al., 2000; Kubota et al., 2014; Xu et al., 2016), where each subgroup comprises subjects with distinct characteristics. To evaluate the effectiveness of the treatment within each subgroup, we measure the mean difference between potential outcomes in the treated and control arms:

$$\tau_j = \mathbb{E}[Y_{it}(1) - Y_{it}(0) | X_{it} \in \mathcal{S}_j], \quad t = 1, \ldots, T, \ j = 1, \ldots, m.$$

Furthermore, we denote the total number of subjects enrolled in subgroup j as $N_j = \sum_{t=1}^T n_{tj}$, where $n_{tj} = \sum_{i=1}^{n_t} \mathbb{1}_{(X_{it} \in \mathcal{S}_j)}$.

In adaptive experiments, practitioners have the flexibility of sequentially allocating experimental efforts to reach certain prespecified design goals. Such efforts include actively recruiting subjects of different characteristics in multiple stages and revising treatment assignment (or allocation) probabilities based on accrued evidence during the experiment. Within the existing literature, two commonly employed design strategies have emerged to distribute these experimental efforts differently, which we will discuss in detail below.

The first strategy is called RAR design or CARA design. In these designs, experimenters can sequentially revise the treatment assignment strategies based on responses accumulated during the experiment but, unlike enrichment designs, often do not change the enrolment criteria across multiple stages. Response-adaptive randomization designs incorporating additional covariate information are more frequently referred to as CARA designs. The design goals of RAR designs tend to vary in different application areas, and we refer interested readers to Robertson et al. (2023) for a comprehensive review. Formally, by defining the treatment assignment probability (or propensity scores) for subjects in subgroup j as

$$e_{tj} = \mathbb{P}(D_{it} = 1 | X_{it} \in \mathcal{S}_j), \quad t = 1, \ldots, T, \ j = 1, \ldots, m.$$

Response-adaptive randomization and CARA design aim to dynamically revise e_{tj} to reach desired design goals.

The second strategy is called (adaptive) enrichment design, which has been frequently carried out in clinical settings to identify patient subgroups that benefit the most from a given treatment (Follmann, 1997; Leung Lai et al., 2019; Rosenblum et al., 2020; Simon & Simon, 2013). In these designs, experimenters often fix the treatment allocation probability during the entire experiment, but they sequentially enrol different subgroups of participants over different stages. Here, the word 'enrichment' spells out the action of actively recruiting a new batch of subjects who may have characteristics different from the previous stage, and the word 'adaptive' indicates that the

enrolment proportions of subjects with different characteristics can be adaptively revised based on the current understanding of treatment effect heterogeneity. Formally, by defining an auxiliary variable $Z_{it} \in \{1, 0\}$ that indicates if subject i is enrolled at Stage t, we introduce the enrichment proportion of subjects falling into region $\mathcal{S}_j$ in Stage t as

$$p_{tj} = \mathbb{P}(X_{it} \in \mathcal{S}_j | Z_{it} = 1), \quad t = 1, \ldots, T, \ j = 1, \ldots, m.$$

Enrichment designs sequentially revise p_{tj} across multiple stages to reach their design objectives.

Our proposed adaptive experimental design framework unifies RAR designs and enrichment designs by formalizing them as a sequential policy learning problem (see Table 1 for a summary). We hope that this unified framework broadens the practicability of the proposed design framework under various practical constraints. In particular, we define a sequential policy $\boldsymbol{\pi}$ consisting of a sequence of policies $\pi_1, \ldots, \pi_{T-1}$, and each π_t is a mapping from the historical data $\mathcal{H}_t = \{\mathcal{H}_s\}_{s=1}^t$ accumulated up to Stage t to either the subgroup enrichment proportions $\boldsymbol{p}_{t+1} \triangleq (p_{t+1,1}, \ldots, p_{t+1,m})$, or to the treatment assignment probabilities $\boldsymbol{e}_{t+1} \triangleq (e_{t+1,1}, \ldots, e_{t+1,m})$, that is

$$\begin{aligned} \pi_t &: \mathcal{H}_t \to \boldsymbol{e}_{t+1} \triangleq (e_{t+1,1}, \ldots, e_{t+1,m}) \quad \text{Response-adaptive randomization design,} \\ \pi_t &: \mathcal{H}_t \to \boldsymbol{p}_{t+1} \triangleq (p_{t+1,1}, \ldots, p_{t+1,m}) \quad \text{Adaptive enrichment design.} \end{aligned}$$

Other than dispensing different experimental strategies, practitioners can also flexibly choose the number of stages T and the number of participants n_t in each stage of the experiment. We refer to experimental design strategies with large n_t and finite T as multistage designs, and we refer to designs with small n_t and large T as fully adaptive designs. While both designs tend to share similar large-sample properties, they have different strengths and can often be applied in scenarios with different practical constraints. On the one hand, multistage designs can be preferable in clinical settings or social experiments where experimenters often have a limited number of opportunities to revise the experimental effort allocated during the experiment (see Gertler et al., 2012; Karlan & Zinman, 2008 for example). Fully adaptive designs are more readily integrated into digital experiments such as online A/B testing or digital clinical trials in which sequentially allocating experimental efforts in a large number of stages is more practical and less costly (see Kharitonov et al., 2015; Robertson et al., 2023 for example). On the other hand, as seen in our simulation studies in Section 6, benefiting from frequently updated experimental strategy, fully adaptive designs tend to have superior finite sample performance compared to multistage designs when the sample size N is rather small.

Benefiting from the above framework, while existing adaptive experiments normally target one of the experimental schemes listed in Table 1, the design strategies we shall propose can be applied in all four settings. This demonstrates that the proposed design strategy is flexible and completes existing frequentist adaptive design strategies, suggesting our designs can be potentially applied to online experiments conducted in e-commerce platforms, clinical trials conducted in health industries, and policy evaluation experiments conducted for social science research. In what follows, we introduce the general goal of our design strategy.

3 Design objectives and oracle allocation strategies: a large deviation perspective

Adaptive experiments are frequently designed with specific predetermined goals in mind. Our adaptive experiment is designed with the goal of gathering strong evidence for learning treatment effect heterogeneity by identifying specific subgroups of participants who are more likely to benefit from the treatment.

Accurately identifying the best-performing subgroups provides several practical advantages, particularly in cases where one treatment is not universally beneficial for the entire population and treatment effects vary across different subpopulations. In clinical research, identifying the beneficial subgroup contributes to the development of personalized medicine, allowing treatments to be tailored to individual patients based on their unique characteristics or predictive markers. By

Table 1. Examples of frequentist data collection mechanisms in response adaptive experiments

π_t	(Regime 1) Small n_t with large T 'Fully adaptive'	(Regime 2) Large n_t with finite T 'Multistage'
Response-adaptive design $\mathcal{H}_t \to e_{t+1}$	Response-adaptive randomization (Eisele, 1994; F. Hu & Rosenberger, 2006; F. Hu & Zhang, 2004b; F. Hu et al., 2009; Robertson et al., 2023; Rosenberger, 2002; Tymofyeyev et al., 2007; Zhu et al., 2023)	Adaptive propensity score (Pocock, 1977; J. Zhao, 2023)
	Covariate-adjusted response-adaptive (Bandyopadhyay & Biswas, 1999; J. Hu et al., 2015; Rosenberger et al., 2001; Villar & Rosenberger, 2018; Zelen, 1994; L.-X. Zhang et al., 2007; Zhu & Zhu, 2023)	Sequential rerandomization (Morgan & Rubin, 2012, 2015; Q. Zhou et al., 2018)
Enrichment design $\mathcal{H}_t \to p_{t+1}$	Not available	Frequentist enrichment design (Burnett et al., 2020; Follmann, 1997; Rosenblum et al., 2014; Rosenblum & van der Laan, 2011; Stallard, 2022; Wang et al., 2007)

identifying subgroups most likely to benefit, our trial design establishes the groundwork for targeted and individualized interventions. In social economics research, accurately identifying the best-performing subgroups enables policymakers and practitioners to understand which specific subpopulations are most positively affected by certain interventions or policies. This knowledge allows for more targeted and effective interventions to address social and economic challenges. By focusing resources and efforts on the subgroups that stand to benefit the most, policymakers can maximize the impact of their initiatives and improve overall societal well-being.

In statistical languages, our design goal is to construct reliable estimators of the subgroup average treatment effect so that the probability of correctly identifying the subgroups with the most beneficial (or harmful) effects is maximized when the experiment ends. Formally, without loss of generality, we assume that the population subgroup average treatment effects satisfy $\tau_1 > \tau_2 > \cdots > \tau_m$ (generalizations to other possible effect orders are provided in online supplementary material, Section I), and suppose we have constructed consistent estimators $\hat{\tau}_1, \ldots, \hat{\tau}_m$ of $\tau_1, \ldots, \tau_m$ based on the collected data at the end of the experiment. Because the joint distribution of $\hat{\tau}_1, \ldots, \hat{\tau}_m$ not only depends on the underlying data distribution of the potential outcome and covariates but also crucially relies on the treatment assignment mechanism and subgroup enrolment proportions, these estimators can be viewed as a function of the historical data and the corresponding policy adopted in the adaptive experiment. Then, in a simple case where we aim to find the best subgroup with the largest treatment effect in the population (i.e. the first subgroup $\mathcal{S}_1$), our design objective is to find a sequential policy $\boldsymbol{\pi}$ belonging to a set of feasible policies $\mathbf{\Pi}$, so that the probability of the estimated first subgroup treatment effect margins out the others is maximized. As in this simple case, the first subgroup has the largest treatment effect in the population; the correction selection probability can be written as $\mathbb{P}(\hat{\tau}_1 \geq \max_{2 \leq j \leq m} \hat{\tau}_j)$.

Unfortunately, without imposing additional parametric distributional assumptions on the historical data, directly searching for a policy that maximizes the correct selection probability results in an intractable optimization problem, as deriving a general analytic form of the correct selection probability is nearly impossible. One seemingly natural alternative is to consider solving this optimization problem in an asymptotic sense. By letting the total sample size N go to infinity, it is possible to approximate the distribution of $\hat{\tau}_j$ with a Gaussian distribution under mild conditions. However, even in this asymptotic framework, given $\tau_1 > \tau_2$ and for any policy π, the correct selection probability $\mathbb{P}(\hat{\tau}_1 \geq \max_{2 \leq j \leq m} \hat{\tau}_j)$ grows exponentially fast to one as $N \to \infty$. Consequently, it

is no longer a function of $\boldsymbol{\pi}$, implying that directly searching for a sequential policy that maximizes the correct selection probability in an asymptotic sense is infeasible.

To address the challenges mentioned above, we temporarily shift our focus from studying a sequential policy that maximizes correct selection probability. Instead, we consider an idealized 'oracle' scenario in which we possess complete knowledge about the underlying data distribution. With this oracle in hand, we can explore the best strategy to allocate experimental efforts and design the experiment to achieve the highest possible correct selection probability.

While acknowledging that this idealized scenario is not practically attainable, studying it can offer valuable insights and serve as a benchmark for evaluating the performance of more realistic strategies and policies in real-world adaptive experiments. However, even in the oracle scenario with perfect knowledge of the data distribution, the correct selection probability can still exhibit complex behaviour with finite samples or tend to 1 as the sample size tends to infinity. Consequently, searching for the optimal allocation strategy remains a challenging task. In light of this, we are motivated to magnify the correction selection probability through the lens of the large and moderate deviation principle (Dembo, 2009; Eichelsbacher & Löwe, 2003; Glynn & Juneja, 2004; Hollander, 2000; Petrov, 1975).

In essence, the large and moderate deviation principles provide a precise characterization of the correction selection probability using a set of rate functions. Specifically, under appropriate conditions with some $a_N \to \infty$ (as $N \to \infty$), the correct selection probability satisfies:

$$\begin{aligned} \lim_{N\to\infty} \frac{1}{a_N} \log\left(1 - \mathbb{P}\left(\hat{\tau}_1 \geq \max_{2\leq j\leq m} \hat{\tau}_j\right)\right) &= -\min_{2\leq j\leq m} G(\mathcal{S}_1, \mathcal{S}_j; e_1, p_1, e_j, p_j), \\ G(\mathcal{S}_1, \mathcal{S}_j; e_1, p_1, e_j, p_j) &= \frac{(\tau_j - \tau_1)^2}{2\left(\mathbb{V}_1(e_1, p_1) + \mathbb{V}_j(e_j, p_j)\right)}, \end{aligned} \tag{1}$$

where $\mathbb{V}_j(e_j, p_j)$ is the variance of $\hat{\tau}_j$, for $j = 1, \ldots, p$. The rate function $G(\mathcal{S}_1, \mathcal{S}_j; e_1, p_1, e_j, p_j)$ thus captures the exponential decay rate of the probability of the rare event where the estimated treatment effect in the best subgroup $\hat{\tau}_1$ is smaller than the estimated treatment effect in subgroup $\hat{\tau}_j$, as the sample size $N \to \infty$. The derivation of this result is explained in detail in online supplementary material, Section J for mathematical clarity under both large and moderate deviation principles. Furthermore, depending on the design strategy, the rate function $G(\mathcal{S}_1, \mathcal{S}_j; e_1, p_1, e_j, p_j)$ typically has a closed-form expression that depends on the treatment allocations (e_1 and e_j) and subgroup enrichment proportions (p_1 and p_j) in the best subgroup $\mathcal{S}_1$ and subgroup $\mathcal{S}_j$; see (3) and online supplementary material, Section C Eq (1) for their closed-form expressions.

We are now ready to define *oracle allocation strategies* in the RAR designs and the enrichment designs. In RAR designs, when the enrolment criteria are fixed, and the subgroup proportions cannot be modified, we define the oracle treatment allocation probabilities $\boldsymbol{e}^* \triangleq (e_1^*, \ldots, e_m^*)$ as the solution to the following constraint optimization problem:

$$\max_{\boldsymbol{e}} \left\{ \min_{2\leq j\leq m} G(\mathcal{S}_1, \mathcal{S}_j; e_1, e_j) \ : \ \sum_{j=1}^{m} p_j e_j \leq c_1, \ c_2 \leq e_j \leq 1 - c_2 \right\},$$

where $c_1 \in (0, 1)$ and $c_2 \in (0, 1/2)$. Similarly, in enrichment designs, when the treatment assignment probabilities in different subgroups are fixed, and the propensity scores $\boldsymbol{e} = (e_1, \ldots, e_m)$ cannot be modified, we define the oracle subgroup enrichment proportions $\boldsymbol{p}^* \triangleq (p_1^*, \ldots, p_m^*)$ as the solution to the following constraint optimization problem:

$$\max_{\boldsymbol{p}} \left\{ \min_{2\leq j\leq m} G(\mathcal{S}_1, \mathcal{S}_j; p_1, p_j) \ : \ \sum_{j=1}^{m} p_j = 1, \ p_j \geq 0 \right\}.$$

The closed-form solution of the above optimization problems relies on the choice of the subgroup treatment effect estimators. We thus leave more detailed discussions of the oracle allocation strategies for the RAR design in Section 4 and the enrichment design in online supplementary material, Section C. As shall be made clear in later sections, the oracle allocation strategies offer

considerable advantages over traditional randomized experiments, including improving the efficiency in estimating the best subgroup treatment effect (Proposition 2) and allowing the population treatment effect of the second-best subgroup to stay closer to that of the best subgroup (Proposition 3).

In practice, when experimenters have no prior knowledge about the joint distribution of the subgroup treatment effect estimators, adaptive experiments offer a natural environment to sequentially learn the unknown parameters in each subgroup and adjust the allocation of experimental efforts during the experiment. In the following sections, we aim to answer the following two research questions: When we have no prior information about the data-generating process, is it possible to carry out adaptive experimental design strategies that sequentially study the joint distribution of the underlying data and meanwhile use learned information to allocate experimental efforts better as the experiment progresses? When the experiment is finished, can such designs produce subgroup treatment effect estimators that have competing performances with the ones under the oracle allocation strategies?

4 Response-adaptive randomization design with adaptive treatment allocation

In this section, we present the oracle treatment allocation strategy for RAR designs. Subsequently, we propose two design strategies for fully adaptive and multistage settings (refer to Table 1), both of which address the questions raised at the end of the previous section.

4.1 Oracle treatment allocation in response-adaptive randomization designs

As the rate function depends on the choice of the subgroup treatment effect estimators, we adopt the inverse propensity score weighting (IPW) estimator with estimated propensity scores to estimate the subgroup treatment effects, that is

$$\hat{\tau}_j = \hat{\tau}_{Tj} = \frac{\sum_{t=1}^{T}\sum_{i=1}^{n_t} \mathbb{1}_{(X_{it}\in\mathcal{S}_j)} D_{it} Y_{it}}{\sum_{t=1}^{T}\sum_{i=1}^{n_t} \mathbb{1}_{(X_{it}\in\mathcal{S}_j)} D_{it}} - \frac{\sum_{t=1}^{T}\sum_{i=1}^{n_t} \mathbb{1}_{(X_{it}\in\mathcal{S}_j)} (1-D_{it}) Y_{it}}{\sum_{t=1}^{T}\sum_{i=1}^{n_t} \mathbb{1}_{(X_{it}\in\mathcal{S}_j)} (1-D_{it})}, \quad j = 1, \ldots, m. \tag{2}$$

We adopt this particular estimator as it is semiparametrically efficient, following results documented in Hirano et al. (2003). We leave a discussion on the augmented IPW estimator (Robins et al., 1994) to online supplementary material, Section I. When the IPW estimator is adopted, we are able to derive a closed-form expression of the rate function:

$$G(\mathcal{S}_1, \mathcal{S}_j; e_1, e_j) = \frac{(\tau_j - \tau_1)^2}{2\big(\mathbb{V}_1(e_1) + \mathbb{V}_j(e_j)\big)}, \quad \mathbb{V}_j(e_j) = \frac{\sigma_j(1)^2}{p_j e_j} + \frac{\sigma_j(0)^2}{p_j(1-e_j)}, \tag{3}$$

where $\mathbb{V}_j(e_j)$ is the asymptotic variance of the estimator $\hat{\tau}_j$, and $\sigma_j(d)^2 = \mathbb{V}[Y(d)|X \in \mathcal{S}_j]$, $d = 0, 1$. We note that in RAR designs, the subgroup enrolment proportions p_j's remain fixed throughout the experiment. Consequently, we denote the rate function and the asymptotic variance solely as functions of the subgroup treatment assignment probabilities e_j's.

With the closed-form expression of the rate function in hand, we are now ready to explore the oracle treatment allocation $\boldsymbol{e}^* \triangleq (e_1^*, \ldots, e_m^*)$, which solves the following optimization problem:

$$\textbf{Problem } \textit{A}: \quad \max_{\boldsymbol{e}} \min_{2\le j\le m} \frac{(\tau_j - \tau_1)^2}{2\big(\mathbb{V}_1(e_1) + \mathbb{V}_j(e_j)\big)}, \quad \leftarrow \text{Maximize correct selection probability}$$

$$\text{s.t.} \sum_{j=1}^{m} p_j e_j \le c_1, \quad \leftarrow \text{'Cost' / practical constraint}$$

$$c_2 \le e_j \le 1 - c_2, \ j = 1, \ldots, m, \quad \leftarrow \text{Feasibility constraints}$$

where $c_1 \in (0, 1)$ and $c_2 \in (0, 1/2)$. Here, the cost/practical constraint restricts the proportion of subjects receiving the treatment, and the feasibility constraint restricts the treatment assignment probability in each subgroup to be bounded away from zero and one.

Because the objective function is the minimum of $m-1$ rate function, the above optimization problem is nonlinear. We instead work with its equivalent epigraph representation:

Problem *B*:

$$\max_{e,z} z, \qquad \leftarrow \text{Linear objective function for simple optimization}$$

$$\text{s.t.} \sum_{j=1}^{m} p_j e_j \le c_1, \qquad \leftarrow \text{'Cost' / practical constraint}$$

$$c_2 \le e_j \le 1 - c_2,\ j = 1, \ldots, m, \qquad \leftarrow \text{Feasibility constraints}$$

$$\frac{(\tau_j - \tau_1)^2}{2\left(\mathbb{V}_1(e_1) + \mathbb{V}_j(e_j)\right)} - z \ge 0,\ j = 2, \ldots, m. \leftarrow \text{Equivalent to maximize correct selection probability}$$

The above epigraph representation has two key advantages. First, it formulates a concave optimization problem, enabling efficient solutions using open-source software such as `IPOPT` (Wächter & Biegler, 2006) and `GUROBI` (LLC Gurobi Optimization, 2018). Second, it facilitates exploration of the Lagrangian dual problem and allows us to obtain a simplified expression of the oracle treatment allocations in certain cases (Glynn & Juneja, 2004). For instance, suppose that the conditional variance of potential outcomes in the treatment and control arms is the same for each subgroup [i.e. $\sigma_j(1)^2 = \sigma_j(0)^2$], and assume that each subgroup has an equal enrolment proportion with $p_j = \frac{1}{m}$. In such cases, we can demonstrate that the oracle treatment allocation $\boldsymbol{e}^* = (e_1^*, \ldots, e_m^*)$ satisfies the following equation (see online supplementary material for the derivation):

$$\frac{(\tau_j - \tau_1)^2}{\frac{\sigma_1(1)^2}{e_1^*(1-e_1^*)} + \frac{\sigma_j(1)^2}{e_j^*(1-e_j^*)}} = \frac{(\tau_k - \tau_1)^2}{\frac{\sigma_1(1)^2}{e_1^*(1-e_1^*)} + \frac{\sigma_k(1)^2}{e_k^*(1-e_k^*)}}, \quad j \ne k, \text{ and } j, k \ne 1.$$

This equation suggests that the required number of participants in the treatment arm of subgroup j is reduced when it is relatively easier to distinguish subgroup j from the best subgroup. This occurs when there is a larger difference between τ_j and τ_1 or when the variance of subgroup j is higher. To provide a clearer understanding, we consider a simple scenario with $m = 3$. In Figure 1, we plot the relationship between e_2^*, τ_2, and $\sigma_2(1)^2$.

Having obtained the oracle treatment allocation, we aim to approximate it using accrued data in an adaptive experiment. Next, we will discuss our proposed adaptive treatment allocation strategy in both fully adaptive and multistage settings.

4.2 Fully adaptive case with large T and small n_t

In this section, we provide our proposed design strategy in the fully adaptive setting with large T and small n_t (Table 1). The derived fully adaptive response adjusted randomization (RAR) sequential policy $\boldsymbol{\pi}_{\text{RAR}} = (\pi_1, \ldots, \pi_{T-1})$ enables us to dynamically revise the treatment assignment probability so that the derived subgroup treatment effect estimator shares the same property as the one delivered by the oracle allocation strategies.

Stage 1 Randomly assign subjects in each subgroup to the treatment arm with a prespecified propensity score, such as $e_{1j} = \frac{1}{2}$, $j = 1, \ldots, m$.

As we have no prior information about enrolling participants, Stage 1 of our design serves as an exploration stage. Note that in theoretical investigations, we allow Stage 1 sample size n_1 to be a vanishing fraction of the total sample size N, that is $\frac{n_1}{N} \to 0$. In practical implementations, we

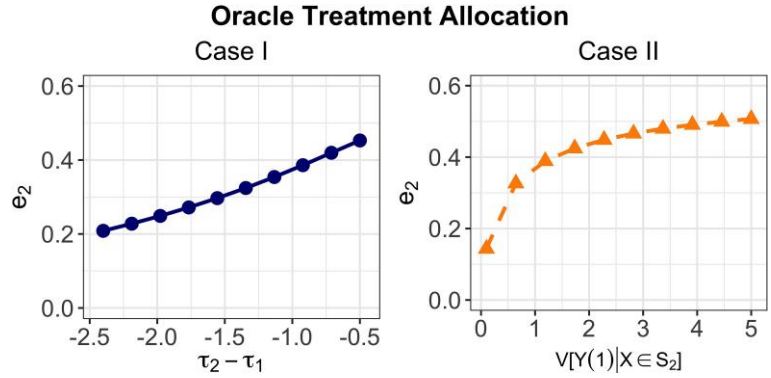

Figure 1. The change of oracle treatment allocation in the second subgroup in two different cases: (I) $\tau_1 = 3$, $\tau_3 = 0.5$, $\sigma_j(1)^2 = 2$, for $j = 1, \ldots, 3$ and (II) $\tau_1 = 3$, $\tau_2 = 2$, $\tau_3 = 1$, $\sigma_1(1)^2 = \sigma_3(1)^2 = 2$.

recommend enrolling at least 2 subjects in each subgroup under each treatment arm. Therefore, $n_1 \geq 4 \cdot m$.

Stage t, for $t = 2, \ldots, T-1$. Obtain $\hat{\boldsymbol{e}}_t^*$ by solving the sample analogue of Problem B: that is

$$\hat{\boldsymbol{e}}_t^* = \arg\max_{\boldsymbol{e}} \left\{ z : \sum_{l=1}^{m} \hat{p}_{tl} e_l \leq c_1, \;\; c_2 \leq e_l \leq 1 - c_2, \right. \\ \left. \min_{2 \leq j \leq m} \frac{(\hat{\tau}_{t-1,(j)} - \hat{\tau}_{t-1,(1)})^2}{2\left(\hat{\mathbb{V}}_{t-1,(1)}(e_1) + \hat{\mathbb{V}}_{t-1,(j)}(e_j)\right)} - z \geq 0 \right\}, \tag{4}$$

where the subscript (j) indexes the subgroup with the jth largest estimated treatment effect, and

$$\hat{\tau}_{t-1,(j)} = \frac{\sum_{s=1}^{t-1}\sum_{i=1}^{n_s} \mathbb{1}_{(X_{is} \in \mathcal{S}_{(j)})} D_{is} Y_{is}}{\sum_{s=1}^{t-1}\sum_{i=1}^{n_s} \mathbb{1}_{(X_{is} \in \mathcal{S}_{(j)})} D_{is}} - \frac{\sum_{s=1}^{t-1}\sum_{i=1}^{n_s} \mathbb{1}_{(X_{is} \in \mathcal{S}_{(j)})} (1 - D_{is}) Y_{is}}{\sum_{s=1}^{t-1}\sum_{i=1}^{n_s} \mathbb{1}_{(X_{is} \in \mathcal{S}_{(j)})} (1 - D_{is})},$$

$$\hat{\mathbb{V}}_{t-1,(j)}(e_j) = \frac{\sum_{s=1}^{t-1}\sum_{i=1}^{n_s} \mathbb{1}_{(X_{is} \in \mathcal{S}_{(j)})} D_{is} \left(Y_{is} - \dfrac{\sum_{s=1}^{t-1}\sum_{i=1}^{n_s} \mathbb{1}_{(X_{is} \in \mathcal{S}_{(j)})} D_{is} Y_{is}}{\sum_{s=1}^{t-1}\sum_{i=1}^{n_s} \mathbb{1}_{(X_{is} \in \mathcal{S}_{(j)})} D_{is}} \right)^2}{\sum_{s=1}^{t-1}\sum_{i=1}^{n_s} \mathbb{1}_{(X_{is} \in \mathcal{S}_{(j)})} D_{is}} \left(e_j \cdot \frac{\sum_{s=1}^{t-1}\sum_{i=1}^{n_s} \mathbb{1}_{(X_{is} \in \mathcal{S}_{(j)})}}{\sum_{s=1}^{t-1} n_s} \right)^{-1} + \frac{\sum_{s=1}^{t-1}\sum_{i=1}^{n_s} \mathbb{1}_{(X_{is} \in \mathcal{S}_{(j)})} (1 - D_{is}) \left(Y_{is} - \dfrac{\sum_{s=1}^{t-1}\sum_{i=1}^{n_s} \mathbb{1}_{(X_{is} \in \mathcal{S}_{(j)})} (1 - D_{is}) Y_{is}}{\sum_{s=1}^{t-1}\sum_{i=1}^{n_s} \mathbb{1}_{(X_{is} \in \mathcal{S}_{(j)})} (1 - D_{is})} \right)^2}{\sum_{s=1}^{t-1}\sum_{i=1}^{n_s} \mathbb{1}_{(X_{is} \in \mathcal{S}_{(j)})} (1 - D_{is})} \cdot \left((1 - e_j) \cdot \frac{\sum_{s=1}^{t-1}\sum_{i=1}^{n_s} \mathbb{1}_{(X_{is} \in \mathcal{S}_{(j)})}}{\sum_{s=1}^{t-1} n_s} \right)^{-1}.$$

Assign treatment with probability $\hat{\boldsymbol{e}}_t^*$.

In each Stage t, based on the newly collected data from the previous stage $\{\mathcal{H}_{t-1}\}$, we renew our understanding of the underlying data distribution and obtain a pair of updated estimates $(\hat{\tau}_{t-1,j}, \hat{\mathbb{V}}_{t-1,j})$ for each subgroup. These updated estimates thus enable us to better mimic the behaviour of the oracle treatment allocation strategy by solving a refined optimization problem defined in Eq. (6) and revise the treatment assignment accordingly. We then assign treatment according to $\hat{e}_t^*$. In Stage T, we calculate empirical treatment assignment probabilities $\hat{e}_T$ using the historical data collected up to Stage $T-1$.

If Eq. (4) suggests multiple possible solutions, one can either choose a treatment allocation proportion that minimizes costs or a treatment allocation that is higher for the benefitted subgroup (i.e. a subgroup with a larger treatment effect).

Statistical inference after Stage T Construct the final subgroup treatment effect estimator under the RAR design along with its standard error using

$$\hat{\tau}_j^{\mathrm{RAR}} = \frac{\sum_{s=1}^{T}\sum_{i=1}^{n_s} \mathbb{1}_{(X_{is}\in\mathcal{S}_j)} D_{is} Y_{is}}{\sum_{s=1}^{T}\sum_{i=1}^{n_s} \mathbb{1}_{(X_{is}\in\mathcal{S}_j)} D_{is}} - \frac{\sum_{s=1}^{T}\sum_{i=1}^{n_s} \mathbb{1}_{(X_{is}\in\mathcal{S}_j)} (1-D_{is}) Y_{is}}{\sum_{s=1}^{T}\sum_{i=1}^{n_s} \mathbb{1}_{(X_{is}\in\mathcal{S}_j)} (1-D_{is})}, \tag{5}$$

$$\begin{aligned}\hat{\mathbb{V}}_j^{\mathrm{RAR}} = {} & \frac{\sum_{s=1}^{T}\sum_{i=1}^{n_s} \mathbb{1}_{(X_{is}\in\mathcal{S}_j)} D_{is} \left(Y_{is} - \bar{Y}_{Tj}(1)\right)^2}{\sum_{s=1}^{T}\sum_{i=1}^{n_s} \mathbb{1}_{(X_{is}\in\mathcal{S}_j)} D_{is}} \left(\frac{\sum_{s=1}^{T}\sum_{i=1}^{n_s} \mathbb{1}_{(X_{is}\in\mathcal{S}_j)} D_{is}}{N}\right)^{-1} \\ & + \frac{\sum_{s=1}^{T}\sum_{i=1}^{n_s} \mathbb{1}_{(X_{is}\in\mathcal{S}_j)} (1-D_{is}) \left(Y_{is} - \bar{Y}_{Tj}(0)\right)^2}{\sum_{s=1}^{T}\sum_{i=1}^{n_s} \mathbb{1}_{(X_{is}\in\mathcal{S}_j)} (1-D_{is})} \left(\frac{\sum_{s=1}^{T}\sum_{i=1}^{n_s} \mathbb{1}_{(X_{is}\in\mathcal{S}_j)} (1-D_{is})}{N}\right)^{-1},\end{aligned} \tag{6}$$

where $\bar{Y}_{Tj}(d) = \frac{\sum_{s=1}^{T}\sum_{i=1}^{n_s} 1_{(D_{is}=d)} 1_{(X_{is}\in\mathcal{S}_j)} Y_{is}}{\sum_{s=1}^{T}\sum_{i=1}^{n_s} 1_{(D_{is}=d)} 1_{(X_{is}\in\mathcal{S}_j)}}$ for $d = 0, 1$.

Then, identify the best subgroup as the one exhibiting the maximal treatment effect size:

$$j^* = \underset{1\le j\le m}{\operatorname{argmax}}\ \hat{\tau}_j^{\mathrm{RAR}}. \tag{7}$$

Lastly, construct a two-sided level-α confidence interval for the selected best subgroup as

$$\left[\hat{\tau}_{j^*}^{\mathrm{RAR}} \pm \Phi^{-1}(1-\alpha/2)\cdot\sqrt{\hat{\mathbb{V}}_{j^*}^{\mathrm{RAR}}/N}\right]. \tag{8}$$

4.3 Multistage case with small T and large n_t

In this section, we provide an alternative multistage design strategy with small T and large n_t, when experimenters cannot revise the treatment assignment strategy too frequently. Stage 1 and the

statistical inference after Stage t are the same in fully adaptive and multistage settings. In Stage t, the multistage setting also requires an additional calibration step, as shown below:

Stage t, for t=2,…,T−1. (a) Solve for $\hat{e}_t^*$ as in the fully adaptive setting. (b) In each subgroup, assign subjects to the treatment arm with probability $\tilde{e}_{t,(j)}$, where

$$\tilde{e}_{t,(j)} = \frac{\left(\hat{e}^*_{t,(j)} \sum_{s=1}^{t} \sum_{i=1}^{n_s} \mathbb{1}_{(X_{is} \in \mathcal{S}_{(j)})}\right) - \sum_{s=1}^{t-1} \sum_{i=1}^{n_s} \mathbb{1}_{(X_{is} \in \mathcal{S}_{(j)})} D_{is}}{\sum_{i=1}^{n_t} \mathbb{1}_{(X_{it} \in \mathcal{S}_{(j)})}}, \quad j = 1, \ldots, m.$$

By incorporating this additional step, we can ensure that the treatment allocation closely approximates the oracle treatment allocation. This adjustment is crucial because it enables the subgroup treatment effect estimators to compete with those obtained under the oracle allocation strategies. As a result, we can obtain accurate estimations of the treatment effects for different subgroups, even in scenarios where the oracle allocation is not directly feasible.

Due to the page limit, we provide a detailed illustration of our proposed AE design in online supplementary material, Section C. We present the oracle subgroup enrichment proportions in online supplementary material, Section C.1 and introduce our proposed AE design in online supplementary material, Section C.2.

5 Theoretical investigation

In this section, we establish the theoretical properties of our proposed adaptive experiment design strategies. We start with introducing notations and assumptions in Section 5.1. We then provide a general result on the consistency of estimated moments of the potential outcomes in the adaptive setting (Lemma 1 in Section 5.2.1), which encompasses the proposed designs as special cases. It is also worth mentioning that the consistency result applies to both fully adaptive ($T \to \infty$) and multistage scenarios (T fixed), and hence it may be of independent interest. See the online supplementary material for additional results and discussions. Building on this lemma, we show in Section 5.2.2 that (i) the proposed treatment allocation (in the RAR design) and enrichment proportion (in the AE design) converge to their oracle counterparts; (ii) both the actual treatment and enrichment frequencies converge asymptotically to the oracle values; and (iii) in both RAR and AE settings, the estimated treatment effects are asymptotically normally distributed. Combined with a consistent variance estimator, results in Section 5.2 deliver a suite of estimation and statistical inference methods targeted at learning treatment effect heterogeneity with a rigorous statistical guarantee. In Section 5.3, we compare our proposed RAR design with the classical completely randomized design under simplified theoretical conditions. The theoretical comparison demonstrates three advantages of our design: (i) it corresponds to a smaller large deviation rate, suggesting a higher correct selection probability and a stronger estimation bias control (Theorem 1); (ii) our design improves the statistical efficiency of uncovering treatment effect heterogeneity (Proposition 2); and (iii) our design reduces the required sample size for best subgroup identification (Proposition 3).

5.1 Assumption and additional notation

We consider the asymptotic regime where the number of enrolled subjects, $N = \sum_{t=1}^{T} n_t$, grows, where we recall that n_t is the number of subjects in Stage t. In the fully adaptive setting, this is equivalent to letting the number of time stages, denoted by T, approach infinity. In the multistage setting, T is fixed, and n_t will increase. Additionally, we denote the total number of subjects enrolled in subgroup j as $N_j = \sum_{t=1}^{T} n_{tj}$, which is the sum of subjects in subgroup j across all stages.

We work under the following assumptions for our theoretical investigations:

Assumption 1 (i) The potential outcomes and the covariates, $(Y_{it}(0), Y_{it}(1), X_{it})$, are independently and identically distributed across $t = 1, \ldots, T$ and $i = 1, \ldots, n_t$.

(ii) The potential outcomes have bounded fourth moments: $\mathbb{E}[|Y_{it}(d)|^4] < \infty$ for $d = 0, 1$. (iii) The potential outcomes have nonvanishing conditional variances: there exists some $\delta > 0$, such that $\mathbb{V}[Y_{it}(d)|X_{it} \in \mathcal{S}_j] \geq \delta$ for $d = 0, 1$ and $j = 1, 2, \ldots, m$.

Assumption 2 There are $m \geq 2$ subgroups, and the subgroup treatment effects can be monotonically ordered with $\tau_1 > \tau_2 > \cdots > \tau_m$.

To simplify theoretical derivation, we assume that there are no exact ties among the population subgroup treatment effects. As an extension of our current framework, in the online supplementary material, Section I, we provide a tentative approach to handle potential ties based on our earlier work (W. Wei et al., 2023).

Assumption 3 (i) For the RAR design, the subgroup proportions $p_1, \ldots, p_m$ are bounded away from 0 by a positive constant, that is, there exists a constant $\delta \in (0, 1)$ such that $p_j \geq \delta$ for all j. (ii) For the AE design, the subgroup treatment assignment probabilities $e_1, \ldots, e_m$ are bounded away from 0 and 1; that is, there exists a constant $\delta \in (0, 1/2)$ such that $\delta \leq e_j \leq 1 - \delta$ for all j.

Assumption 3, together with the constraints in our optimization Problems A and C, ensures that each considered subgroup has a nonvanishing enrolment probability, and within each subgroup, participants will be assigned to both the treatment and control arms (Ma & Wang, 2020).

To facilitate theoretical discussions in the upcoming section, we will differentiate 'actual treatment allocations' and 'actual enrichment proportions' from those given by our algorithms. To be precise, the actual treatment allocations refer to the cumulative empirical treatment frequencies at each stage:

$$\hat{\boldsymbol{e}}_t = \left(\frac{\sum_{s=1}^{t} \sum_{i=1}^{n_s} \mathbb{1}_{(X_{is} \in \mathcal{S}_1)} D_{is}}{\sum_{s=1}^{t} \sum_{i=1}^{n_s} \mathbb{1}_{(X_{is} \in \mathcal{S}_1)}}, \ \ldots, \ \frac{\sum_{s=1}^{t} \sum_{i=1}^{n_s} \mathbb{1}_{(X_{is} \in \mathcal{S}_m)} D_{is}}{\sum_{s=1}^{t} \sum_{i=1}^{n_s} \mathbb{1}_{(X_{is} \in \mathcal{S}_m)}} \right).$$

We also adopt the convention $\hat{\boldsymbol{e}} = \hat{\boldsymbol{e}}_T$. Similarly, for AE designs, we define the actual enrichment proportions as

$$\hat{\boldsymbol{p}}_t = \left(\frac{\sum_{s=1}^{t} \sum_{i=1}^{n_s} \mathbb{1}_{(X_{is} \in \mathcal{S}_1)}}{\sum_{s=1}^{t} n_s}, \ \ldots, \ \frac{\sum_{s=1}^{t} \sum_{i=1}^{n_s} \mathbb{1}_{(X_{is} \in \mathcal{S}_m)}}{\sum_{s=1}^{t} n_s} \right),$$

and $\hat{\boldsymbol{p}} = \hat{\boldsymbol{p}}_T$. In contrast, we use the terms 'optimized treatment allocations' and 'optimized subgroup enrichment proportions' to refer to the proposed design, solved from Eq. (6) and online supplementary material, Eq. (3):

$$\hat{\boldsymbol{e}}_t^* = (\hat{e}_{t1}^*, \ \ldots, \ \hat{e}_{tm}^*), \quad \text{and} \quad \hat{\boldsymbol{p}}_t^* = (\hat{p}_{t1}^*, \ \ldots, \ \hat{p}_{tm}^*).$$

5.2 Theoretical properties of the proposed adaptive experiment strategy

We are now ready to introduce the theoretical properties of our proposed adaptive experiment strategies. Section 5.2.1 presents a general consistency result on the estimated moments of potential outcomes, which encompasses our proposed RAR and AE designs as special cases.

5.2.1 Consistency results in a general adaptive experiment setting

In the lemma below, we use the generic notation $\mathfrak{p}_{tj} = (X_{it} \in \mathcal{S}_j | \boldsymbol{\mathcal{H}}_{t-1})$ for the subgroup proportion, and $\mathfrak{e}_{tj} = \mathbb{P}(D_{it} = 1 | X_{it} \in \mathcal{S}_j, \boldsymbol{\mathcal{H}}_{t-1})$ for the treatment probability for subgroup j, where $\boldsymbol{\mathcal{H}}_{t-1}$ is the sigma-algebra formed by $(X_{is}, D_{is}, Y_{is})_{1 \leq i \leq n_s, 1 \leq s \leq t-1}$ for $t = 1, 2, \ldots, T$, and $\boldsymbol{\mathcal{H}}_0$ is the trivial

sigma-algebra. The notations suggest that $\mathfrak{p}_{tj}$ and $\mathfrak{e}_{tj}$ can depend on $\mathcal{H}_{t-1}$ hence allowing for adaptive designs. Notice that both the RAR and the AE designs are special cases: in the RAR design, we set $\mathfrak{p}_{tj} = p_j$ and $\mathfrak{e}_{tj} = \hat{e}^*_{tj}$; in the AE design, we have $\mathfrak{p}_{tj} = \hat{p}^*_{tj}$ and $\mathfrak{e}_{tj} = e_j$.

Lemma 1 Assume Assumption 1 holds, and that there exists some $\delta \in (0, 1/2)$ such that for all $j = 1, 2, \ldots, m$ and $t = 1, 2, \ldots,$

$$\mathfrak{p}_{tj} \geq \delta, \text{ and } \delta \leq \mathfrak{e}_{tj} \leq 1 - \delta.$$

Then for any $j = 1, 2, \ldots, m$, and any t satisfying $\sum_{s=1}^{t} n_s \to \infty$,

$$\frac{\sum_{s=1}^{t}\sum_{i=1}^{n_s} \mathbb{1}_{(X_{is}\in\mathcal{S}_j)} D_{is} Y^c_{is}}{\sum_{s=1}^{t}\sum_{i=1}^{n_s} \mathbb{1}_{(X_{is}\in\mathcal{S}_j)} D_{is}} = \mathbb{E}[Y_{it}(1)^c | X_{it} \in \mathcal{S}] + O_p\left(\frac{1}{\sqrt{\sum_{s=1}^{t} n_s}}\right),$$

$$\frac{\sum_{s=1}^{t}\sum_{i=1}^{n_s} \mathbb{1}_{(X_{is}\in\mathcal{S}_j)} (1 - D_{is}) Y^c_{is}}{\sum_{s=1}^{t}\sum_{i=1}^{n_s} \mathbb{1}_{(X_{is}\in\mathcal{S}_j)} (1 - D_{is})} = \mathbb{E}[Y_{it}(0)^c | X_{it} \in \mathcal{S}] + O_p\left(\frac{1}{\sqrt{\sum_{s=1}^{t} n_s}}\right).$$

In addition to encompassing the RAR and AE designs as special cases, Lemma 1 also applies to both fully adaptive and multistage settings. To be precise, fully adaptive corresponds to $T \to \infty$ and $n_t = 1$ (or a fixed constant), in which case the consistency holds as $t \to \infty$. On the other hand, a multistage setting involves a fixed T. Then, the consistency result holds for each fixed t as the cumulative sample size $\sum_{s=1}^{t} n_s$ tends to infinity.

A common challenge for proving consistency in adaptive experiments is that the treatment probability or the subgroup frequency can depend on historical data. Our proof strategy builds on explicit variance bounds, which is in contrast to the classical method that employs results on optional stopping (Doob, 1936; F. Hu & Zhang, 2004b).

5.2.2 Theoretical results under our proposed adaptive experiment strategies

In this section, we establish the theoretical properties of our proposed adaptive experimental design strategies, including (i) consistency of the optimized and actual treatment allocations for the RAR design, (ii) consistency of the optimized and actual enrichment proportions for the AE design; and (iii) consistency and asymptotic normality of the estimated subgroup treatment effects. We also provide a consistent estimator for the asymptotic variance of the estimated treatment effects, which allows for valid statistical inference. To save space, we focus on the full adaptive setting ($T \to \infty$ and $n_t = 1$). Similar results can be established for multistage designs (Regime 2 in Table 1), which we discuss in the online supplementary material.

To start, we show that the estimated variance is consistent as a function of the treatment probability in the RAR design or as a function of the subgroup proportion in the AE design.

Corollary 1 Assume Assumptions 1 and 3 hold. Let $\delta \in (0, 1/2)$ be some constant. Then

$$\text{RAR design:} \quad \sup_{\delta\leq e\leq 1-\delta} \left|\hat{\mathbb{V}}_{tj}(e) - \mathbb{V}_j(e)\right| = O_p\left(\frac{1}{\sqrt{\sum_{s=1}^{t} n_s}}\right),$$

$$\text{AE design:} \quad \sup_{\delta\leq p\leq 1-\delta} \left|\hat{\mathbb{V}}_{tj}(p) - \mathbb{V}_j(p)\right| = O_p\left(\frac{1}{\sqrt{\sum_{s=1}^{t} n_s}}\right),$$

for $j = 1, 2, \ldots, m$.

Another useful corollary of Lemma 1 is that the estimated subgroup treatment effects are consistent.

Corollary 2 Assume Assumptions 1 and 3 hold. Then for both RAR and AE designs,

$$\hat{\tau}_{tj} - \tau_j = O_p\left(\frac{1}{\sqrt{\sum_{s=1}^{t} n_s}}\right), \quad j = 1, \ldots, m.$$

If Assumption 2 also holds, then as $t \to \infty$,

$$\mathbb{P}\left(\hat{\tau}_{t,(j)} = \tau_j\right) \to 1, \quad \mathbb{P}\left(\tau_{(j)} = \tau_j\right) \to 1.$$

Building on Lemma 1 and Corollary 2, we now present the theoretical properties of our proposed adaptive experiment strategies. As our proposed adaptive experiment strategies are derived by sequentially solving the optimization problems in Section 4.2 and online supplementary material, Section C, we shall present Theorem 1 which includes two related but conceptually different consistency results: the convergence of the optimized treatment allocation or enrichment proportion to their oracle values, and the consistency of the actual treatment allocation or enrichment proportion.

Theorem 1 ((Asymptotic consistency of adaptive experiment strategies)). Assume Assumptions 1–3 hold. Assume that Problems A and Problem C (in online supplementary material, Section C) admit unique solutions. Then for any $\delta > 0$ and as $t \to \infty$, for the optimized treatment allocation and enrichment proportion:

$$\begin{aligned} \text{RAR design:} \quad & \mathbb{P}\left(\left\|\hat{\boldsymbol{e}}_t^* - \boldsymbol{e}^*\right\| \le \delta\right) \to 1, \\ \text{AE design:} \quad & \mathbb{P}\left(\left\|\hat{\boldsymbol{p}}_t^* - \boldsymbol{p}^*\right\| \le \delta\right) \to 1. \end{aligned}$$

In addition, for the actual treatment allocation and enrichment proportion:

$$\begin{aligned} \text{RAR design:} \quad & \mathbb{P}\left(\left\|\hat{\boldsymbol{e}}_t - \boldsymbol{e}^*\right\| \le \delta\right) \to 1, \\ \text{AE design:} \quad & \mathbb{P}\left(\left\|\hat{\boldsymbol{p}}_t - \boldsymbol{p}^*\right\| \le \delta\right) \to 1. \end{aligned}$$

The first part of Theorem 1 suggests that the empirically and sequentially optimized treatment allocations and enrichment proportions converge to their oracle counterparts. We assume that the optimization problems admit unique solutions in Theorem 1 because in practical implementations, nonuniqueness of the solution is not a serious concern in our setting. Whenever Eq. (6) produces multiple treatment allocations, the researcher can always choose one using some additional criteria (say, with the smallest cost). We provide more details of selecting the set of optimizers in Section 4.2 and online supplementary material, Section C. For this reason, we are able to assume that the solution to the optimization problems is unique throughout the rest of the paper.

The second part of Theorem 1 implies that the actual treatment allocations—the fraction of subjects assigned to receive treatment in each subgroup—converge to the oracle treatment allocation rule. The same consistency result holds for the enrichment design, as the actual subgroup proportions will converge to their oracle counterparts. In other words, although our proposed designs in Section 4.2 and online supplementary material, Section C have no prior knowledge about the underlying data distribution before the experiment starts, they can allocate experimental efforts in a similar fashion to the oracle strategies when the sample size is sufficiently large.

Theorem 2 ((Asymptotic normality and consistent variance estimation)). Assume Assumptions 1–3 hold. In addition, assume that Problem A and Problem C (in online supplementary material, Section C) admit unique solutions, which are denoted by $\boldsymbol{e}^*$ and $\boldsymbol{p}^*$. Then as $t \to \infty$,

$$\text{RAR design: } \sqrt{N}\left(\hat{\tau}_{j^*}^{\text{RAR}} - \tau_1\right) \underset{\to}{\mathcal{D}} \mathcal{N}\left(0, \mathbb{V}_1(e_1^*)\right), \quad \mathbb{V}_1(e_1^*) = \frac{\sigma_1(1)^2}{p_1 e_1^*} + \frac{\sigma_1(0)^2}{p_1(1-e_1^*)},$$

$$\text{AE design: } \sqrt{N}\left(\hat{\tau}_{j^*}^{\text{AE}} - \tau_1\right) \underset{\to}{\mathcal{D}} \mathcal{N}\left(0, \mathbb{V}_1(p_1^*)\right), \quad \mathbb{V}_1(p_1^*) = \frac{\sigma_1(1)^2}{p_1^* e_1} + \frac{\sigma_1(0)^2}{p_1^*(1-e_1)}.$$

In addition,

$$\hat{\mathbb{V}}_{j^*}^{\text{RAR}} - \mathbb{V}_1(e_1^*) = O_p\left(\frac{1}{\sqrt{N}}\right), \quad \hat{\mathbb{V}}_{j^*}^{\text{AE}} - \mathbb{V}_1(p_1^*) = O_p\left(\frac{1}{\sqrt{N}}\right).$$

The theoretical results established in Theorem 2 indicate that the selected best subgroup treatment effect is a $\sqrt{N}$-consistent estimate of the best subgroup treatment effect τ_1. In addition, the asymptotic variance can be consistently estimated by $\hat{\mathbb{V}}_{j^*}$. This further suggests that the constructed confidence interval for the best subgroup, as given by Eq. (10), has correct coverage asymptotically. The asymptotic normality result relies on the martingale central limit theorem (Hall & Heyde, 1980) and the consistency results of our proposed adaptive experiment strategies. For its formal proof, we refer readers to the online supplementary material.

5.3 Comparison with completely randomized experiments

In this section, we compare our proposed RAR design with completely randomized experiments, where the treatment is randomly assigned with a prefixed probability throughout the entire experiment. To simplify theoretical derivations, we work under the assumption that the outcome variables follow Gaussian distributions and the treatment assignments are independent, enabling us to conveniently compare the large deviation rates between our design and complete randomization. Concretely, the comparisons will be examined from three perspectives: (1) the large deviation rate and estimation bias (Proposition 1), (2) the asymptotic variance of the estimated best subgroup treatment effect (Proposition 2), and (3) the minimum sample size required to achieve a predetermined correct selection probability (Proposition 3).

In order to establish a fair comparison with completely randomized experiments, we employ the same IPW estimator with estimated propensity scores to estimate the treatment effect, denoted as

$$\hat{\tau}_j^{\text{CR}} = \frac{\sum_{i=1}^N \mathbb{1}_{(X_i \in \mathcal{S}_j)} D_i Y_i}{\sum_{i=1}^N \mathbb{1}_{(X_i \in \mathcal{S}_j)} D_i} - \frac{\sum_{i=1}^N \mathbb{1}_{(X_i \in \mathcal{S}_j)} (1 - D_i) Y_i}{\sum_{i=1}^N \mathbb{1}_{(X_i \in \mathcal{S}_j)} (1 - D_i)},$$

for $j = 1, \ldots, m$. In this section, we consider a setting where (1) subgroup proportions are equal: $p_1 = p_2 = \cdots = p_m = \frac{1}{m}$, and (2) there exists a cost constraint: $\sum_{j=1}^m p_j e_j \le c_1$, $c_1 \in (0, 1)$. In the completely randomized design, we set $\hat{e}_{ij}^* = c_1$ for every t and j. This ensures that this design is comparable to ours while also meeting the cost constraint. The variance of $\hat{\tau}_j^{\text{CR}}$ can be derived with a simple form:

$$\mathbb{V}_j(c_1) = \frac{\sigma_j(1)^2}{p_j \cdot c_1} + \frac{\sigma_j(0)^2}{p_j \cdot c_1}. \tag{9}$$

We start by comparing the large deviation rates under the proposed RAR design and the complete randomization design.

Proposition 1 (Large deviation rate comparison). Under Assumptions 1–3,

$$\lim_{N\to\infty}\frac{1}{N}\log\left(1-\mathbb{P}(\hat{\tau}_1^{\mathrm{RAR}}\geq\max_{2\leq j\leq m}\hat{\tau}_j^{\mathrm{RAR}})\right)\leq\lim_{N\to\infty}\frac{1}{N}\log\left(1-\mathbb{P}(\hat{\tau}_1^{\mathrm{CR}}\geq\max_{2\leq j\leq m}\hat{\tau}_j^{\mathrm{CR}})\right),$$

where $\hat{\tau}_1^{\mathrm{CR}}$ denotes the estimated treatment effect of the best subgroup under the complete randomization design.

Note that under Assumption 2, the correct selection probability $\mathbb{P}(\hat{\tau}_1^{\mathrm{RAR}}\geq\max_{2\leq j\leq m}\hat{\tau}_j^{\mathrm{RAR}})$ can be equivalently written as $\mathbb{P}(j^*=1)\to 1$, where $j^*=\underset{1\leq j\leq m}{\operatorname{argmax}}\,\hat{\tau}_j^{\mathrm{RAR}}$ is the index of the selected best subgroup, as defined in (7).

Proposition 1 suggests that our proposed RAR design has a faster large deviation rate than that obtained under completely randomized experiments (i.e. the rate function implied by our method is larger in magnitude). This result has two indications. First, it implies that the probability of correctly selecting the best subgroup in our RAR design converges to one exponentially faster than in complete randomization as the sample size increases; see Figure 2a for verification of Proposition 1. There, we provide a simulation study with a fixed sample size $N=500$ and set $\tau_1=1.6$, $\tau_3=0.5$, and $\tau_2=\tau_1-\delta$, where $\delta\in\{0.03, 0.04, \ldots, 0.4\}$. We compare the correct selection probability under the proposed design and the complete randomization design with respect to various distances between τ_1 and τ_2. Furthermore, a faster large deviation rate indicates that our design provides stronger bias control of the selected best subgroup compared to complete randomization. This is because the estimation bias of the best subgroup is proportional to the incorrect selection probability of the best subgroup, as shown in the following equation:

$$\mathbb{E}[\hat{\tau}_{j^*}]-\tau_1=-\tau_1\cdot\underbrace{\left(1-\mathbb{P}(\hat{\tau}_1\geq\max_{2\leq j\leq m}\hat{\tau}_j)\right)}_{\text{incorrect selection prob.}}.$$

Our design achieves stronger control over the incorrect selection probability, which in turn allows for better bias regulation compared to complete randomization. This conclusion can also be verified through our simulation results in Figure 5b.

Next, we compare the asymptotic efficiency gain of the proposed design for estimating the best subgroup treatment effect with the complete randomization design. Note that both variance lower bounds derived from our proposed design in Eq. (3) and the complete randomization design in Eq. (11) share a similar form, which allows us to compare the performance of our design with complete randomization. To provide some insights into the efficiency comparison, we consider a simple case formalized in Proposition 2 below. In online supplementary material, Section F.6, we consider more general settings and provide additional theoretical insights therein.

Proposition 2 ((Asymptotic variance comparison)). Assume (1) $\sigma_j(1)^2=\sigma_j(0)^2$, for $j=1,\ldots,m$, (2) $\sigma_1(1)^2=\cdots=\sigma_m(1)^2$, and (3) the cost constraint $c_1<0.5$. For all possible oracle treatment allocations $\boldsymbol{e}^*=(e_1^*,\ldots,e_m^*)\in\mathcal{E}^*$, we have for $j=1,\ldots,m$,

$$\begin{cases}\mathbb{V}_j(e_j^*)\leq\mathbb{V}_j(c_1), & \text{if } (\tau_j-\tau_1)^2-\frac{1}{e_1^*(1-e_1^*)}\leq\frac{1}{c_1^2},\\ \mathbb{V}_j(e_j^*)>\mathbb{V}_j(c_1), & \text{if } (\tau_j-\tau_1)^2-\frac{1}{e_1^*(1-e_1^*)}>\frac{1}{c_1^2}.\end{cases}$$

Proposition 2 shows the efficiency comparison between our proposed RAR design and complete randomization. When estimating the best subgroup treatment effect, the asymptotic variance under our proposed RAR design is smaller than the complete randomization design. However, when τ_j is far away from τ_1 or when the expected variance of the outcome in subgroup j is small, our proposed RAR design is less likely to have efficiency gain. Proposition 2 thus entails the efficiency trade-off between our proposed design and the complete randomization design.

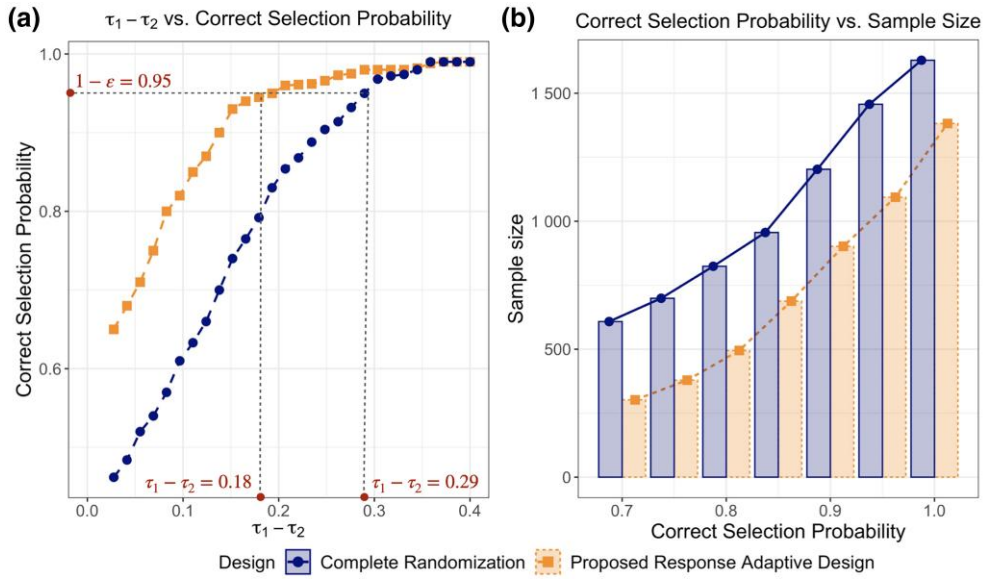

Figure 2. Verification of Propositions 1 and 3. (a) Correct selection probability comparison with respect to various distances between τ_1 and τ_2. (b) Sample size comparison between the proposed oracle RAR design and the complete randomization design when fixing $\tau_1 - \tau_2 = 0.1$.

The efficiency trade-off can also be seen in Figure 4. In online supplementary material, Section F.6, we provide another result without restricting $c_1 < 0.5$ and all variance terms to be equal.

Lastly, we compare the minimum sample size required to reach a fixed correct selection probability level.

Proposition 3 ((Sample size comparison)). Assume (1) $\sigma_j(1)^2 = \sigma_j(0)^2$ for $j = 1, \ldots, m$, (2) $\sigma_1(1)^2 = \cdots = \sigma_m(1)^2$, and (3) the cost constraint $c_1 < 0.5$. Suppose we aim to reach a correct selection probability of at least $1 - \varepsilon$. For some positive constants $C < \infty$ and $C' < \infty$, under the complete randomization design, the required sample size is characterized as

$$N \geq \frac{1}{(\tau_1 - \tau_2)^2} \cdot \left| C \cdot \log(\varepsilon) \cdot \sigma_1(1)^2 \right|.$$

Under our proposed RAR design, for all possible oracle treatment allocations $\boldsymbol{e}^* = (e_1^*, \ldots, e_m^*) \in \mathcal{E}^*$, the required sample size is characterized as

$$N \geq \left(\frac{1}{(\tau_1 - \tau_2)^2} \cdot \left| C' \cdot \log(\varepsilon) \cdot \sigma_1(1)^2 \right| \right)^{3/4}.$$

Proposition 3 says that to reach a correct selection probability level of at least $1 - \varepsilon$, our proposed adaptive design strategy often requires a smaller sample size compared to the complete randomization design. To verify Proposition 3, we provide a simple simulation in Figure 2b. In Figure 2b, we fix $\tau_1 - \tau_2 = 0.1$ and investigate the sample size needed to reach various correct selection probability levels. Figure 2b demonstrates that to reach a prespecified correct selection probability level, our proposed design requires smaller sample sizes. In other words, when τ_1 is close to τ_2, our proposed RAR design correctly distinguishes the best subgroup from the second-best subgroup with a higher probability.

6 A synthetic case study

In this section, we investigate the performance of our proposed RAR design and AE design in a synthetic case study using e-commerce data. We summarize four takeaways as follows: First, compared to several classical experimental design strategies, our proposed design requires the smallest sample size to reach a prefixed level of correct selection probability [panel (a) in Figures 5–8]. Benefiting from an improved correct selection probability, our design also yields the lowest estimation bias for the best subgroup [panel (b) in Figures 5–8]. Second, our proposed RAR design yields a smaller variance when estimating the best subgroup treatment effect (Figures 4–7). Third, the fully adaptive setting achieves an equivalent correct selection probability with less experimental data compared to the multistage setting, while the multistage setting can be more practical to implement as it requires fewer updates (Figure 5a versus Figure 7a).

6.1 Synthetic case study background

We design our synthetic case study using e-commerce data collected from ModCloth, a website specializing in women's apparel. A crucial marketing strategy for apparel-based websites is the use of human models to showcase their products. Various studies indicate a prevailing 'prothin' bias in fashion advertising, suggesting that such websites often tend to display idealized, size-small models wearing their clothes (Aagerup, 2011; Levine & Schweitzer, 2015). However, in light of the recent social campaigns advocating for inclusiveness in fashion marketing, some fashion companies have revised their advertising strategies to feature models of a wider range of body shapes (Cinelli & Yang, 2016). While it is hypothesized that the inclusive advertising strategy could improve customer satisfaction, it remains unknown which clothing category benefits the most from the inclusive advertising strategy (Joo & Wu, 2021). Through this case study, we aim to identify the clothing category that benefits most significantly from the display of a diverse range of body shapes and investigate the performance of various experimental strategies in identifying this best-performing clothing category.

The original ModCloth data are collected and processed as in Wan et al. (2020), and the dataset contains 99, 893 observations collected from 2010 to 2019. For each clothing product, the website displays one of the two types of human model images: (1) a model wearing a size small, or (2) two models, one wearing a size small and the other a size large (Figure 3). We define the treatment variable as $D = 1$ if both 'small' and 'large' images are displayed and $D = 0$ if only 'small' images are shown. We consider four clothing categories: (1) bottoms, (2) tops, (3) outwear, and (4) dresses. In the context of our manuscript, clothing categories are equivalent to 'subgroups.' To quantify customer satisfaction, we use customer ratings that range from 0 to 5. For this case study, we generate synthetic experimental data based on the original dataset, which shall be illustrated in the next section.

6.2 Synthetic data generation and simulation setup

Our data generation process mimics the ModCloth data, and we consider four nonoverlapping subgroups defined by clothing categories. Denote the subgroup membership for each subject i as $\mathcal{S} = (1_{(X_i \in \mathcal{S}_1)}, \ldots, 1_{(X_i \in \mathcal{S}_4)})$. We generate the potential outcome from

$$Y_i(d) | X_i \in \mathcal{S}_j \sim \mathcal{N}(\mu_{d,j}, \ \sigma_{d,j}), \quad j = 1, \ldots, m.$$

We obtain the subgroup mean and standard deviation parameters calibrated from the original dataset:

$$\boldsymbol{\mu}_1 = (4.14, 4.12, 4.43, 4.48)^{\mathrm{T}}, \quad \boldsymbol{\mu}_0 = (4.83, 3.74, 4.02, 4.31)^{\mathrm{T}},$$
$$\boldsymbol{\sigma}_1 = (1.17, 1.06, 0.80, 0.90)^{\mathrm{T}}, \quad \boldsymbol{\sigma}_0 = (0.39, 1.57, 1.23, 1.10)^{\mathrm{T}}.$$

The subgroup proportions are $\boldsymbol{p} = (0.20, 0.16, 0.56, 0.08)^{\mathrm{T}}$. We denote $\boldsymbol{\tau} = (-0.69, 0.38, 0.41, 0.18)^{\mathrm{T}}$ as the true subgroup treatment effects. The treatment assignment D_i is decided based on different experiment strategies, which shall be discussed later in the section. To generate synthetic data, we consider two design setups:

Figure 3. An example of two different advertising strategies taken from ModCloth website. The left panel shows an inclusive advertising strategy of displaying both small and plus-size human models. The right panel shows a conventional advertising strategy that only displays human models wearing size small.

Table 2. Comparison of designs in our synthetic case study

	Fully adaptive (Setup 1)	**Multistage (Setup 2)**
Response-adaptive design	Methods in comparison	Methods in comparison
	(a) Proposed design in Section 4.2	(a) Proposed design in Section 4.3
	(b) Complete randomization	(b) Complete randomization
	(c) Neyman allocation	
	(d) Proposed design combined with DBCD	
Enrichment design	(a) Proposed design in online supplementary material, Section C	(a) Proposed design in online supplementary material, Section C
		(b) Equal enrichment
	(b) Equal enrichment	(c) Adaptive enrichment with combination testing

Setup 1: We mimic the fully adaptive experiment and fix the total sample size as $N \in \{400, \ldots, 2,000\}$, and $n_1 \in \{80, \ldots, 400\}$. We assume subjects are enrolled sequentially across T experimental stages, where $T \in \{320, \ldots, 1,600\}$.

Setup 2: We mimic the multistage experiment and consider two settings: (a) We set $T = 2$, $n_1 \in \{300, \ldots, 1,900\}$ and $n_2 = 100$. (b) We set $T = 4$, $n_1 \in \{100, \ldots, 1,700\}$, $n_2 = n_3 = n_4 = 100$. Additional design setups and simulation results with a smaller first stage sample size are provided in the online supplementary material, Section G.

Under each design setup, we compare our proposed design strategy with other conventional designs as summarized in Table 2. In Table 2 , the 'complete randomization' design refers to setting the treatment assignment probability $e_{tj} = \frac{1}{2}$ in all experimental stages, for $t = 1, \ldots, T$, $j = 1, \ldots, m$. The 'Neyman allocation' refers to setting the treatment assignment probability $e_{tj} = \frac{\sigma_{1j}}{\sigma_{1j}+\sigma_{0j}}$. The 'Proposed design combined with DBCD' refers to enhancing our proposed design with the DBCD in L. Zhang and Rosenberger (2006). The DBCD is a RAR design that targets the current treatment allocation towards the optimal treatment allocation. As we consider assigning

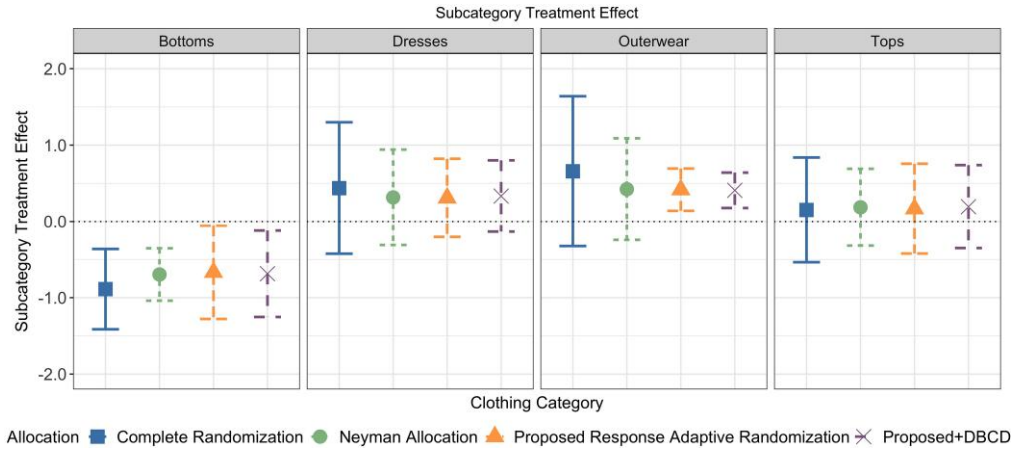

Figure 4. The estimated treatment effects and the associated standard errors in the four clothing categories under the complete randomization design, Neyman allocation, our proposed response-adaptive randomization design, and our proposed design in combination with the doubly adaptive biased coin design in the fully adaptive setting ($N = 400$).

treatments to multiple subgroups, we use the DBCD to target the optimal treatment allocation in each subgroup separately. Our implementation is summarized as follows:

1. At Stage t, obtain *optimal treatment allocations* $\hat{e}^*_{t,j}$, $j = 1, \ldots, m$ by solving the optimization problem as in Section 4.2. Calculate *current treatment allocation* up to Stage $t-1$, denoted as $\hat{e}_{t-1,j}$, $j = 1, \ldots, m$.
2. For each subgroup j, calculate treatment allocation under the DBCD proposed in F. Hu and Zhang (2004b):

$$\psi_{t,j}(\hat{e}^*_{t,j}, \hat{e}_{t-1,j}) = \frac{\hat{e}^*_{t,j}\left(\frac{\hat{e}^*_{t,j}}{\hat{e}_{t-1,j}}\right)^{\gamma}}{\hat{e}^*_{t,j}\left(\frac{\hat{e}^*_{t,j}}{\hat{e}_{t-1,j}}\right)^{\gamma} + (1-\hat{e}^*_{t,j})\left(\frac{1-\hat{e}^*_{t,j}}{1-\hat{e}_{t-1,j}}\right)^{\gamma}},$$

   where $\gamma \in [0, \infty)$ is a tuning parameter.
3. At Stage t, we assign treatments with probability $\psi_{t,j}(\hat{e}^*_{t,j}, \hat{e}_{t-1,j})$ in each subgroup j.

The 'equal enrichment design' refers to the design that sets the enrichment proportion as $p_{tj} = \frac{1}{m}$ across all the experimental stages. The 'AE with combination testing' approach is a method that distributes the Type I error rate across experimental stages. Based on the computed Type I error rate each stage aims to reach, the corresponding enrichment proportions can be estimated. We implement the combination testing approach using `R` package `rpact` (Lakens et al., 2021).

We evaluate the performance of each adaptive experiment strategy from two aspects. First, we compare the experimental efforts (i.e. sample size) needed to reach various correct selection probability levels: {0.75, 0.8, 0.85, 0.9, 0.95, 0.99}. Second, we compare the $\sqrt{N}$-scaled bias of the estimated best subgroup treatment effect. The synthetic case study results are summarized in the following section.

6.3 Synthetic case study results

In Figures 4–7, we compare our proposed RAR design with the other conventional designs in the fully adaptive setting and the multistage setting. We summarize our simulation results from three aspects, following the order outlined at the start of Section 6.

First, by comparing Figures 5a, 7a and c, and 8a and c, our proposed designs require smaller sample sizes to reach the same level of correct selection probability than other designs under

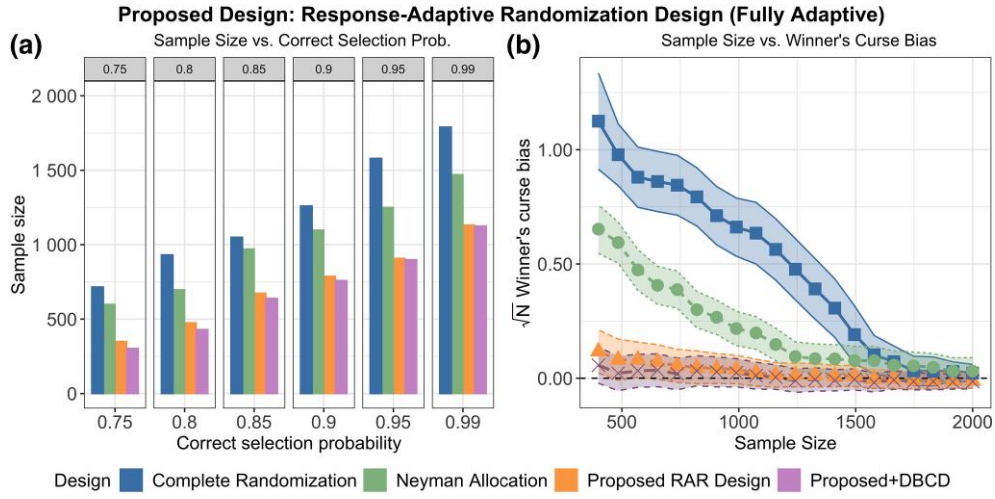

Figure 5. Comparison of the proposed response-adaptive randomization design, the complete randomization design, Neyman allocation, and our proposed design in combination with the doubly adaptive biased coin design under the fully adaptive setting. (a) The sample size comparison under various correct selection probability levels. (b) The $\sqrt{N}$-scaled winner's curse bias comparison with respect to different sample sizes.

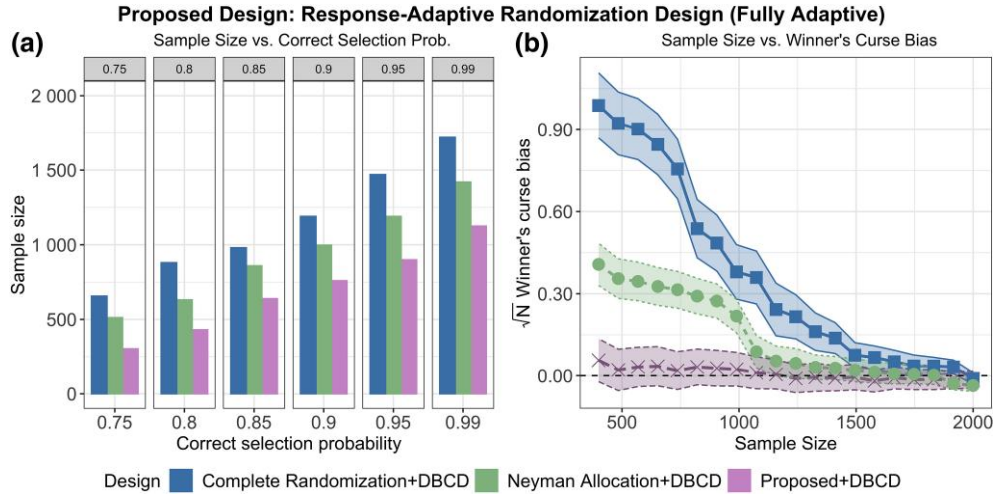

Figure 6. Comparison of the proposed response-adaptive randomization design, the complete randomization design, and the Neyman allocation in combination with the doubly adaptive biased coin design under the fully adaptive setting. (a) The sample size comparison under various correct selection probability levels. (b) The $\sqrt{N}$-scaled winner's curse bias comparison with respect to different sample sizes.

comparison. Benefiting from this design feature, our design yields the best subgroup treatment effect estimator with the lowest bias. This result supports our theoretical analysis in Proposition 1.

Second, in line with our theoretical analysis in Proposition 2, our proposed RAR design is efficient in estimating the best subgroup treatment effect and is less efficient for the worst subgroup, a trend we observe consistently in both fully adaptive and multistage settings. This can be seen from the results in Figure 4.

Third, the simulation results under both RAR design and AE design suggest that fully adaptive experiments can achieve equivalent levels of correct selection probability with smaller sample sizes compared to multistage experiments; see Figures 5a and 7a for example. Whenever the sample size is large, the difference between the fully adaptive and multistage is negligible. We conjecture that

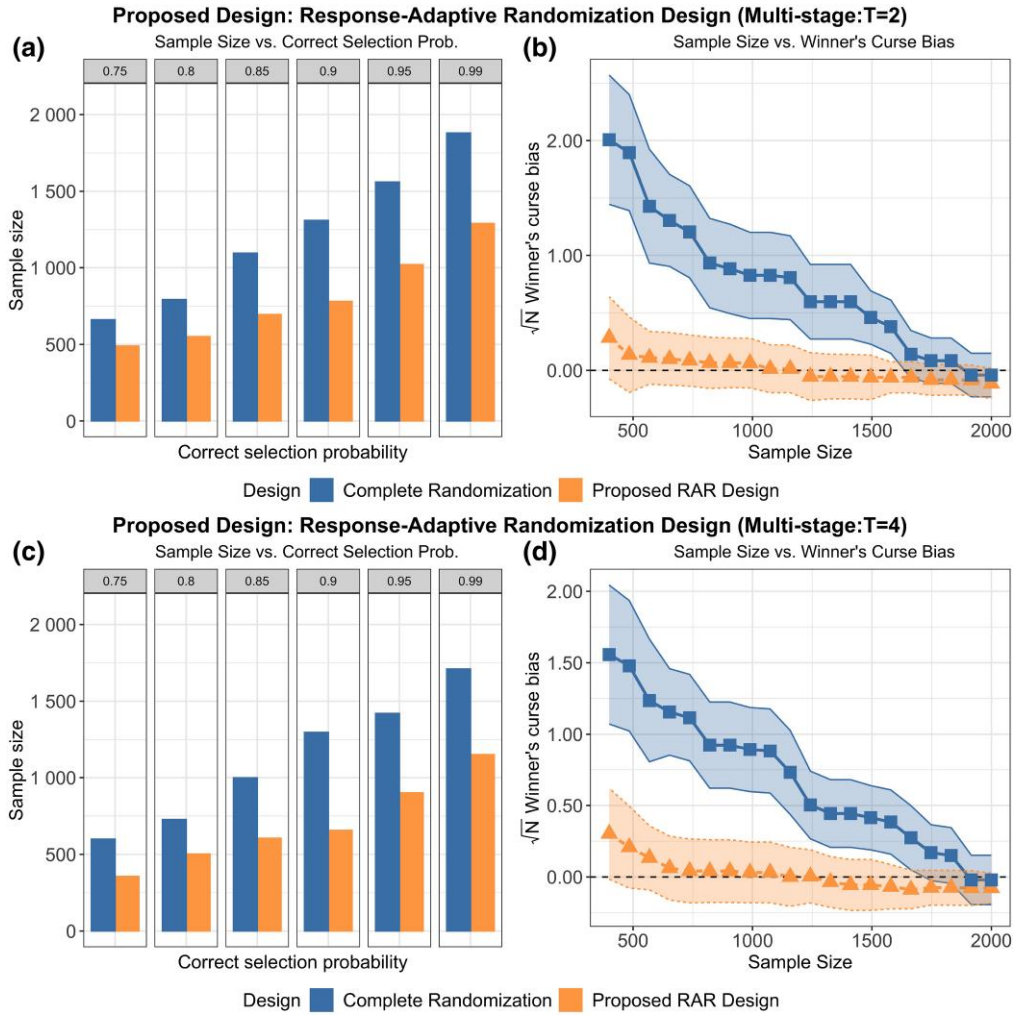

Figure 7. Comparison of the proposed response-adaptive randomization design and the complete randomization design under the multistage setting ($T = 2$ and $T = 4$). (a) and (c): The sample size comparison under various correct selection probability levels. (b) and (d): The $\sqrt{N}$-scaled winner's curse bias comparison with respect to different sample sizes.

this could be attributed to the fully adaptive design providing more opportunities for experimenters to adjust treatment assignment probabilities, potentially achieving the oracle at a faster asymptotic rate.

Lastly, in the RAR setting, combining our proposed design with the DBCD can further enhance the finite-sample performance. Figures 4 and 5 demonstrate that DBCD can enhance the performance of our method in finite samples. When using the DBCD to target our actual treatment allocation towards the optimal treatment allocation, the design strategy exhibits a smaller estimation bias for the best subgroup treatment effect and an increase in the correct selection probability. As the sample size increases, the performances of our proposed design and the DBCD-enhanced design tend to converge. In Figure 6, we provide an additional simulation study to highlight the broad benefits of DBCD in enhancing the finite-sample performance of various designs. We compare three designs: (i) complete randomization + DBCD, (ii) Neyman allocation + DBCD, and (iii) Proposed RAR design + DBCD. We use '+DBCD' to indicate that the DBCD is applied to each design to guide the treatment allocation closer to the optimal treatment allocation. Figure 6 shows that DBCD generally improves the finite-sample performance of these design

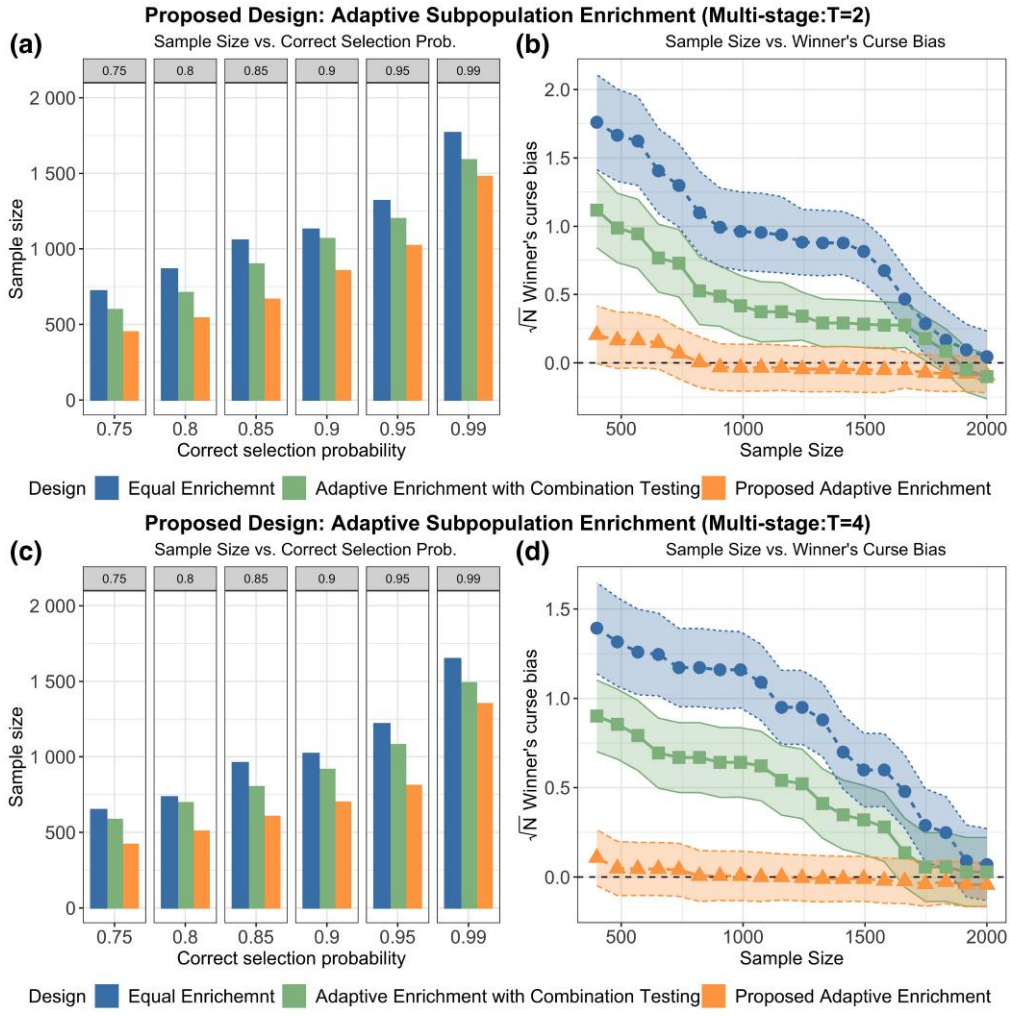

Figure 8. Comparison of the proposed adaptive subgroup enrichment design and the equal enrichment design under the multistage setting ($T = 2$ and $T = 4$). (a) and (c): The sample size comparison under various correct selection probability levels. (b) and (d): The $\sqrt{N}$-scaled winner's curse bias comparison with respect to different sample sizes.

strategies by effectively targeting the current treatment allocation towards the optimal treatment allocation.

In sum, from an application perspective, when an experimenter can only enrol a limited number of subjects in online experiments, our proposed adaptive experiment strategies demonstrate more efficient use of the samples to identify the best subgroup with a higher probability and can reduce the winner's curse bias on estimating the best subgroup treatment effect. Furthermore, the results of the synthetic case study suggest that the adoption of an inclusive advertising strategy could have practical marketing advantages and potentially positive social effects. As studied in the marketing literature (Cinelli & Yang, 2016; Joo & Wu, 2021), such an inclusive marketing strategy could improve customer satisfaction, elevate customer self-esteem, and reduce body-focused anxiety. As our designs may have applications beyond e-commons, we provide another synthetic case study in the context of health care in the online supplementary material, Section H.

7 Discussion

In this manuscript, we propose a unified adaptive experimental framework designed to study treatment effect heterogeneity. Three directions warrant future studies. First, it is possible to extend our

current framework to identify the top few subgroups instead of the best one. Take two subgroups for example; our goal can be formulated as finding the oracle treatment assignment that solves $\max_e \mathbb{P}(\min\{\hat{\tau}_1, \hat{\tau}_2\} \geq \max_{3\leq j\leq m} \hat{\tau}_j)$, which can be achieved by revising the objective function as $\max_e \min_{3\leq j\leq m} \max_{1\leq k\leq 2} G(\mathcal{S}_k, \mathcal{S}_j; e_k, e_j)$. Second, the theoretical analysis presented in Section 5.3 is based on the assumption that the collected sample is independent and identically distributed. We hypothesize that this condition can be relaxed by utilizing refined concentration inequalities that incorporate martingales (Bercu et al., 2015; Chung & Lu, 2006). Exploring these possibilities will be an avenue for future research. Third, our design considers the setting when outcomes are observed without delay. The significance of incorporating delayed responses in adaptive trials has been discussed and recognized in various adaptive design literature, including Rosenberger et al. (2012) and Robertson et al. (2023). Some existing adaptive experiment methods and theoretical results related delayed outcomes have been discussed under the urn models (Bai et al., 2002; F. Hu & Zhang, 2004a; L. J. Wei, 1988; L. J. Wei & Durham, 1978; L.-X. Zhang et al., 2007), the DBCD framework (F. Hu et al., 2008; Kim et al., 2014; L. Zhang & Rosenberger, 2006), and in the group sequential settings (Ghosh et al., 2022; Hampson & Jennison, 2013; Schüürhuis et al., 2024). We hope to extend our proposed design and adjust it for delayed outcomes in our future research.

Acknowledgments

The authors would like to thank the editor, the associate editor, and the referees for their valuable feedback that greatly improved the manuscript.

Conflicts of interest: None declared.

Funding

The research of J.W. is partially supported by the National Science Foundation (NSF) The Faculty Early Career Development (CAREER) award DMS-2239047 and the NSF award DMS-2220537.

Data availability

The data that support the findings of this study are openly available at https://dl.acm.org/doi/10.1145/3336191.3371855.

Supplementary material

Supplementary material is available online at *Journal of the Royal Statistical Society: Series B*.

References

Aagerup U. (2011). The influence of real women in advertising on mass market fashion brand perception. *Journal of Fashion Marketing and Management: An International Journal*, *15*(4), 486–502. https://doi.org/10.1108/13612021111169960

Andrews I., Kitagawa T., & McCloskey A. (2024). Inference on winners. *The Quarterly Journal of Economics*, *139*(1), 305–358. https://doi.org/10.1093/qje/qjad043

Assmann S. F., Pocock S. J., Enos L. E., & Kasten L. E. (2000). Subgroup analysis and other (mis) uses of baseline data in clinical trials. *The Lancet*, *355*(9209), 1064–1069. https://doi.org/10.1016/S0140-6736(00)02039-0

Athey S., Imbens G. W., & Wager S. (2018). Approximate residual balancing: Debiased inference of average treatment effects in high dimensions. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *80*(4), 597–623. https://doi.org/10.1111/rssb.12268

Atkinson A. C., & Biswas A. (2005). Bayesian adaptive biased-coin designs for clinical trials with normal responses. *Biometrics*, *61*(1), 118–125. https://doi.org/10.1111/biom.2005.61.issue-1

Audibert, J.Y., & Bubeck, S. (2010). Best arm identification in multi-armed bandits. In *COLT-23th Conference on learning theory-2010* (pp. 13-p).

Bai Z. D., Hu F., & Rosenberger W. F. (2002). Asymptotic properties of adaptive designs for clinical trials with delayed response. *The Annals of Statistics*, *30*(1), 122–139. https://doi.org/10.1214/aos/1015362187

Bandyopadhyay, U., & Biswas, A. (1999). Allocation by randomized play-the-winner rule in the presence of prognostic factors. *Sankhya: The Indian Journal of Statistics, Series B*, *61*(3), 397–412.

Bercu B., Delyon B., & Rio E. (2015). *Concentration inequalities for sums and martingales*. Springer.

Burnett T., Mozgunov P., Pallmann P., Villar S. S., Wheeler G. M., & Jaki T. (2020). Adding flexibility to clinical trial designs: An example-based guide to the practical use of adaptive designs. *BMC Medicine*, *18*(1), 1–21. https://doi.org/10.1186/s12916-020-01808-2

Butler N. A. (2001). Optimal and orthogonal Latin hypercube designs for computer experiments. *Biometrika*, *88*(3), 847–857. https://doi.org/10.1093/biomet/88.3.847

Cattaneo M. D., Jansson M., & Ma X. (2019). Two-step estimation and inference with possibly many included covariates. *The Review of Economic Studies*, *86*(3), 1095–1122. https://doi.org/10.1093/restud/rdy053

Cheng Y., & Shen Y. (2005). Bayesian adaptive designs for clinical trials. *Biometrika*, *92*(3), 633–646. https://doi.org/10.1093/biomet/92.3.633

Chung F., & Lu L. (2006). Concentration inequalities and martingale inequalities: A survey. *Internet Mathematics*, *3*(1), 79–127. https://doi.org/10.1080/15427951.2006.10129115

Cinelli M. D., & Yang L. (2016). The role of implicit theories in evaluations of "plus-size" advertising. *Journal of Advertising*, *45*(4), 472–481. https://doi.org/10.1080/00913367.2016.1230838

Dembo A. (2009). *Large deviations techniques and applications*. Springer.

Djebbari H., & Smith J. (2008). Heterogeneous impacts in PROGRESA. *Journal of Econometrics*, *145*(1-2), 64–80. https://doi.org/10.1016/j.jeconom.2008.05.012

Doob J. L. (1936). Note on probability. *The Annals of Mathematics*, *37*(2), 363–367. https://doi.org/10.2307/1968449

Dudik M., Hsu D., Kale S., Karampatziakis N., Langford J., Reyzin L., & Zhang T. (2011). 'Efficient optimal learning for contextual bandits', arXiv, arXiv:1106.2369, preprint: not peer reviewed.

Efron B. (1971). Forcing a sequential experiment to be balanced. *Biometrika*, *58*(3), 403–417. https://doi.org/10.1093/biomet/58.3.403

Eichelsbacher P., & Löwe M. (2003). Moderate deviations for IID random variables. *ESAIM: Probability and Statistics*, *7*, 209–218. https://doi.org/10.1051/ps:2003005

Eisele J. R. (1994). The doubly adaptive biased coin design for sequential clinical trials. *Journal of Statistical Planning and Inference*, *38*(2), 249–261. https://doi.org/10.1016/0378-3758(94)90038-8

Fisher R. A. (1936). Design of experiments. *British Medical Journal*, *1*(3923), 554. https://doi.org/10.1136/bmj.1.3923.554-a

Follmann, D. (1997). Adaptively changing subgroup proportions in clinical trials. *Statistica Sinica*, *7*(4), 1085–1102.

Garivier A., & Kaufmann E. (2016). Optimal best arm identification with fixed confidence. In *Conference on Learning Theory* (pp. 998–1027). PMLR.

Gertler P. J., Martinez S. W., & Rubio-Codina M. (2012). Investing cash transfers to raise long-term living standards. *American Economic Journal: Applied Economics*, *4*(1), 164–192.https://doi.org/10.1257/app.4.1.164

Ghosh P., Ristl R., König F., Posch M., Jennison C., Götte H., Schüler A., & Mehta C. (2022). Robust group sequential designs for trials with survival endpoints and delayed response. *Biometrical Journal*, *64*(2), 343–360. https://doi.org/10.1002/bimj.v64.2

Glynn P., & Juneja S. (2004). A large deviations perspective on ordinal optimization. In *Proceedings of the 2004 Winter Simulation Conference, 2004.* (Vol. 1). IEEE.

Gsponer T., Gerber F., Bornkamp B., Ohlssen D., Vandemeulebroecke M., & Schmidli H. (2014). A practical guide to Bayesian group sequential designs. *Pharmaceutical Statistics*, *13*(1), 71–80. https://doi.org/10.1002/pst.v13.1

Guo X., & He X. (2021). Inference on selected subgroups in clinical trials. *Journal of the American Statistical Association*, *116*(535), 1498–1506. https://doi.org/10.1080/01621459.2020.1740096

Hadad V., Hirshberg D. A., Zhan R., Wager S., & Athey S. (2021). Confidence intervals for policy evaluation in adaptive experiments. *Proceedings of the National Academy of Sciences USA*, *118*(15), e2014602118. https://doi.org/10.1073/pnas.2014602118

Hall P., & Heyde C. C. (1980). *Martingale limit theory and its application*. Academic Press.

Hampson L. V., & Jennison C. (2013). Group sequential tests for delayed responses (with discussion). *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *75*(1), 3–54. https://doi.org/10.1111/j.1467-9868.2012.01030.x

He, X., Madigan, D., Yu, B., Wellner, J. (2019). Statistics at a crossroad: Who is for the challenge? In *NSF Workshop report. National Science Foundation*. https://www.nsf.gov/mps/dms/documents/Statistics_at_a_Crossroads_Workshop_Report_2019.pdf

Hill J. L. (2011). Bayesian nonparametric modeling for causal inference. *Journal of Computational and Graphical Statistics*, *20*(1), 217–240. https://doi.org/10.1198/jcgs.2010.08162

Hirano K., Imbens G. W., & Ridder G. (2003). Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica*, *71*(4), 1161–1189. https://doi.org/10.1111/ecta.2003.71.issue-4

Hollander F. (2000). *Large deviations*. American Mathematical Society.

Hu F., & Rosenberger W. F. (2003). Optimality, variability, power: Evaluating response-adaptive randomization procedures for treatment comparisons. *Journal of the American Statistical Association*, *98*(463), 671–678. https://doi.org/10.1198/016214503000000576
Hu F., & Rosenberger W. F. (2006). *The theory of response-adaptive randomization in clinical trials*. John Wiley & Sons.
Hu F., & Zhang L.-X. (2004a). Asymptotic normality of urn models for clinical trials with delayed response. *Bernoulli*, *10*(3), 447–463. https://doi.org/10.3150/bj/1089206406
Hu F., & Zhang L.-X. (2004b). Asymptotic properties of doubly adaptive biased coin designs for multitreatment clinical trials. *The Annals of Statistics*, *32*(1), 268–301. https://doi.org/10.1214/aos/1079120137
Hu F., Zhang L.-X., Cheung S. H., & Chan W. S. (2008). Doubly adaptive biased coin designs with delayed responses. *Canadian Journal of Statistics*, *36*(4), 541–559. https://doi.org/10.1002/cjs.v36:4
Hu F., Zhang L.-X., & He X. (2009). Efficient randomized-adaptive designs. *The Annals of Statistics*, *37*, 2543–2560. https://doi.org/10.1214/08-AOS655
Hu J., Zhu H., & Hu F. (2015). A unified family of covariate-adjusted response-adaptive designs based on efficiency and ethics. *Journal of the American Statistical Association*, *110*(509), 357–367. https://doi.org/10.1080/01621459.2014.903846
Huang Y., Gilbert P. B., & Janes H. (2012). Assessing treatment-selection markers using a potential outcomes framework. *Biometrics*, *68*(3), 687–696. https://doi.org/10.1111/biom.2012.68.issue-3
Jeff Wu C. F., & Hamada M. S. (2011). *Experiments: Planning, analysis, and optimization*. John Wiley & Sons.
Joo B. R., & Wu J. (2021). The impact of inclusive fashion advertising with plus-size models on female consumers: The mediating role of brand warmth. *Journal of Global Fashion Marketing*, *12*(3), 260–273. https://doi.org/10.1080/20932685.2021.1905021
Karlan D. S., & Zinman J. (2008). Credit elasticities in less-developed economies: Implications for microfinance. *American Economic Review*, *98*(3), 1040–1068. https://doi.org/10.1257/aer.98.3.1040
Kasy M., & Sautmann A. (2021). Adaptive treatment assignment in experiments for policy choice. *Econometrica*, *89*(1), 113–132. https://doi.org/10.3982/ECTA17527
Kaufmann, E., Cappé, O., & Garivier, A. (2016). On the complexity of best arm identification in multi-armed bandit models. *Journal of Machine Learning Research*, *17*(1), 1–42.
Kharitonov, E., Vorobev, A., Macdonald, C., Serdyukov, P., & Ounis, I. (2015). Sequential testing for early stopping of online experiments. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*, *SIGIR'15* (pp. 473–482). Association for Computing Machinery.
Kim M.-O., Liu C., Hu F., & Lee J. J. (2014). Outcome-adaptive randomization for a delayed outcome with a short-term predictor: Imputation-based designs. *Statistics in Medicine*, *33*(23), 4029–4042. https://doi.org/10.1002/sim.6222
Kubota K., Ichinose Y., Scagliotti G., Spigel D., Kim J. H., Shinkai T., Takeda K., Kim S.-W., Hsia T.-C., Li R. K., & Tiangco B. J. (2014). Phase III study (MONET1) of motesanib plus carboplatin/paclitaxel in patients with advanced nonsquamous nonsmall-cell lung cancer (NSCLC): Asian subgroup analysis. *Annals of Oncology*, *25*(2), 529–536. https://doi.org/10.1093/annonc/mdt552
Lai T. L., Lavori P. W., & Tsang K. W. (2019). Adaptive enrichment designs for confirmatory trials. *Statistics in Medicine*, *38*(4), 613–624. https://doi.org/10.1002/sim.v38.4
Lakens, D., Pahlke, F., & Wassmer, G. (2021). Group sequential designs: A tutorial. *PsyArXiv*. https://psyarxiv.com/x4azm/
Levine E. E., & Schweitzer M. E. (2015). The affective and interpersonal consequences of obesity. *Organizational Behavior and Human Decision Processes*, *127*, 66–84. https://doi.org/10.1016/j.obhdp.2015.01.002
Lin, C. D., Bingham, D., Sitter, R. R., & Tang, B. (2010). A new and flexible method for constructing designs for computer experiments. *The Annals of Statistics*, *38*(3), 1460–1477. https://doi.org/10.1214/09-aos757
Lin Y., Zhu M., & Su Z. (2015). The pursuit of balance: An overview of covariate-adaptive randomization techniques in clinical trials. *Contemporary Clinical Trials*, *45*, 21–25. https://doi.org/10.1016/j.cct.2015.07.011
LLC Gurobi Optimization. (2018). Gurobi optimizer reference manual. https://www.gurobi.com/
Luedtke A. R., & Van Der Laan M. J. (2016). Statistical inference for the mean outcome under a possibly non-unique optimal treatment strategy. *The Annals of Statistics*, *44*(2), 713. https://doi.org/10.1214/15-AOS1384
Ma X., & Wang J. (2020). Robust inference using inverse probability weighting. *Journal of the American Statistical Association*, *115*(532), 1851–1860. https://doi.org/10.1080/01621459.2019.1660173
Ma X., Wang J., & Wu C. (2023). Breaking the winner's curse in Mendelian randomization: Rerandomized inverse variance weighted estimator. *The Annals of Statistics*, *51*(1), 211–232. https://doi.org/10.1214/22-AOS2247
Morgan K. L., & Rubin D. B. (2012). Rerandomization to improve covariate balance in experiments. *The Annals of Statistics*, *40*(2), 1263–1282. https://doi.org/10.1214/12-AOS1008
Morgan K. L., & Rubin D. B. (2015). Rerandomization to balance tiers of covariates. *Journal of the American Statistical Association*, *110*(512), 1412–1421. https://doi.org/10.1080/01621459.2015.1079528
Mukerjee R., & Wu C.-F. (2006). *A modern theory of factorial design*. Springer.

Park Y., Liu S., Thall P. F., & Yuan Y. (2022). Bayesian group sequential enrichment designs based on adaptive regression of response and survival time on baseline biomarkers. *Biometrics*, *78*(1), 60–71. https://doi.org/10.1111/biom.v78.1

Petrov V. V. (1975). *Sums of independent random variables*. Springer.

Pocock S. J. (1977). Group sequential methods in the design and analysis of clinical trials. *Biometrika*, *64*(2), 191–199. https://doi.org/10.1093/biomet/64.2.191

Raudenbush S. W., & Bloom H. S. (2015). Learning about and from a distribution of program impacts using multisite trials. *American Journal of Evaluation*, *36*(4), 475–499. https://doi.org/10.1177/1098214015600515

Robertson D. S., Lee K. M., López-Kolkovska B. C., & Villar S. S. (2023). Response-adaptive randomization in clinical trials: From myths to practical considerations. *Statistical Science: A Review Journal of the Institute of Mathematical Statistics*, *38*(2), 185. https://doi.org/10.1214/22-STS865

Robins J. M., Rotnitzky A., & Zhao L. P. (1994). Estimation of regression coefficients when some regressors are not always observed. *Journal of the American statistical Association*, *89*(427), 846–866. https://doi.org/10.1080/01621459.1994.10476818

Rosenberger, W. F. (2002). Randomized urn models and sequential design. *Sequential Analysis*, *21*(1-2), 1–41. https://doi.org/10.1081/sqa-120004166

Rosenberger W. F., & Hu F. (2004). Maximizing power and minimizing treatment failures in clinical trials. *Clinical Trials*, *1*(2), 141–147. https://doi.org/10.1191/1740774504cn016oa

Rosenberger W. F., & Lachin J. M. (1993). The use of response-adaptive designs in clinical trials. *Controlled Clinical Trials*, *14*(6), 471–484. https://doi.org/10.1016/0197-2456(93)90028-C

Rosenberger W. F., Sverdlov O., & Hu F. (2012). Adaptive randomization for clinical trials. *Journal of Biopharmaceutical Statistics*, *22*(4), 719–736. https://doi.org/10.1080/10543406.2012.676535

Rosenberger W. F., Vidyashankar A. N., & Agarwal D. K. (2001). Covariate-adjusted response-adaptive designs for binary response. *Journal of Biopharmaceutical Statistics*, *11*(4), 227–236. https://doi.org/10.1081/BIP-120008846

Rosenblum M., Fang E. X., & Liu H. (2020). Optimal, two-stage, adaptive enrichment designs for randomized trials, using sparse linear programming. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *82*(3), 749–772. https://doi.org/10.1111/rssb.12366

Rosenblum M., Liu H., & Yen E.-H. (2014). Optimal tests of treatment effects for the overall population and two subpopulations in randomized trials, using sparse linear programming. *Journal of the American Statistical Association*, *109*(507), 1216–1228. https://doi.org/10.1080/01621459.2013.879063

Rosenblum M., & van der Laan M. J. (2011). Optimizing randomized trial designs to distinguish which subpopulations benefit from treatment. *Biometrika*, *98*(4), 845–860. https://doi.org/10.1093/biomet/asr055

Russo D. (2020). Simple Bayesian algorithms for best-arm identification. *Operations Research*, *68*(6), 1625–1647. https://doi.org/10.1287/opre.2019.1911

Russo D. J., Van Roy B., Kazerouni A., Osband I., & Wen Z. (2018). A tutorial on Thompson sampling. *Foundations and Trends® in Machine Learning*, *11*(1), 1–96. https://doi.org/10.1561/2200000070

Schüürhuis, S., Konietschke, F., & Kunz, C. U. (2024). A two-stage group-sequential design for delayed treatment responses with the possibility of trial restart. *Statistics in Medicine*, *43*(12), 2368–2388. https://doi.org/10.1002/sim.10061

Simchi-Levi D., & Wang C. (2023a). Multi-armed bandit experimental design: Online decision-making and adaptive inference. In *International Conference on Artificial Intelligence and Statistics* (pp. 3086–3097). PMLR.

Simchi-Levi D., & Wang C. (2023b). Pricing experimental design: Causal effect, expected revenue and tail risk. In *International Conference on Machine Learning* (pp. 31788–31799). PMLR.

Simchi-Levi, D., Wang, C., & Zheng, Z. (2023). Non-stationary experimental design under structured trends. https://dx.doi.org/10.2139/ssrn.4514568

Simon N., & Simon R. (2013). Adaptive enrichment designs for clinical trials. *Biostatistics*, *14*(4), 613–625. https://doi.org/10.1093/biostatistics/kxt010

Stallard, N. (2022). Adaptive enrichment designs with a continuous biomarker. *Biometrics*, *79*(1), 9–19. https://doi.org/10.1111/biom.13644

Stallard N., Todd S., & Whitehead J. (2008). Estimation following selection of the largest of two normal means. *Journal of Statistical Planning and Inference*, *138*(6), 1629–1638. https://doi.org/10.1016/j.jspi.2007.05.045

Sun F., & Tang B. (2017). A method of constructing space-filling orthogonal designs. *Journal of the American Statistical Association*, *112*(518), 683–689. https://doi.org/10.1080/01621459.2016.1159211

Thall P. F., & Wathen J. K. (2007). Practical Bayesian adaptive randomisation in clinical trials. *European Journal of Cancer*, *43*(5), 859–866. https://doi.org/10.1016/j.ejca.2007.01.006

Tymofyeyev Y., Rosenberger W. F., & Hu F. (2007). Implementing optimal allocation in sequential binary response experiments. *Journal of the American Statistical Association*, *102*(477), 224–234. https://doi.org/10.1198/016214506000000906

van der Laan, M. J. (2008). The construction and analysis of adaptive group sequential designs. U.C. Berkeley Division of Biostatistics Working Paper Series 232, University of California Berkeley, Berkeley, CA. https://biostats.bepress.com/ucbbiostat/paper232/

Villar S. S., Bowden J., & Wason J. (2015). Multi-armed bandit models for the optimal design of clinical trials: Benefits and challenges. *Statistical Science: A Review Journal of the Institute of Mathematical Statistics*, *30*(2), 199. https://doi.org/10.1214/14-STS504

Villar S. S., & Rosenberger W. F. (2018). Covariate-adjusted response-adaptive randomization for multi-arm clinical trials using a modified forward looking Gittins index rule. *Biometrics*, *74*(1), 49–57. https://doi.org/10.1111/biom.12738

Wächter A., & Biegler L. T. (2006). On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical Programming*, *106*(1), 25–57. https://doi.org/10.1007/s10107-004-0559-y

Wan, M., Ni, J., Misra, R., & McAuley, J. (2020). Addressing marketing bias in product recommendations. In *Proceedings of the 13th international Conference on Web Search and Data Mining* (pp. 618–626). Association for Computing Machinery.

Wang S.-J., O'Neill R. T., & Hung H. M. J. (2007). Approaches to evaluation of treatment effect in randomized clinical trials with genomic subset. *Pharmaceutical Statistics: The Journal of Applied Statistics in the Pharmaceutical Industry*, *6*(3), 227–244. https://doi.org/10.1002/pst.v6:3

Wei L. J. (1988). Exact two-sample permutation tests based on the randomized play-the-winner rule. *Biometrika*, *75*(3), 603–606. https://doi.org/10.1093/biomet/75.3.603

Wei L. J., & Durham S. (1978). The randomized play-the-winner rule in medical trials. *Journal of the American Statistical Association*, *73*(364), 840–843. https://doi.org/10.1080/01621459.1978.10480109

Wei, W., Zhou, Y., Zheng, Z., & Wang, J. (2023). Inference on the best policies with many covariates. *Journal of Econometrics*, *239*(2), 105460. https://doi.org/10.1016/j.jeconom.2022.06.013

Wu C. F. J., & Xu H. (2001). Generalized minimum aberration for asymmetrical fractional factorial designs. *The Annals of Statistics*, *29*(2), 549–560. https://doi.org/10.1214/aos/1009210552

Xu Y., Trippa L., Müller P., & Ji Y. (2016). Subgroup-based adaptive (SUBA) designs for multi-arm biomarker trials. *Statistics in Biosciences*, *8*(1), 159–180. https://doi.org/10.1007/s12561-014-9117-1

Zelen, M. (1994). The randomization and stratification of patients to clinical trials. *Journal of Chronic Diseases*, *27*(7), 365–375. https://doi.org/10.1016/0021-9681(74)90015-0

Zhan, R., Hadad, V., Hirshberg, D. A., & Athey, S. (2021). Off-policy evaluation via adaptive weighting with data from contextual bandits. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining* (pp. 2125–2135). Association for Computing Machinery.

Zhang L., & Rosenberger W. F. (2006). Response-adaptive randomization for clinical trials with continuous outcomes. *Biometrics*, *62*(2), 562–569. https://doi.org/10.1111/biom.2006.62.issue-2

Zhang, L.-X., Chan, W. S., Cheung, S. H., & Hu, F. (2007). A generalized drop-the-loser urn for clinical trials with delayed responses. *Statistica Sinica*, *17*(1), 387–409. https://www.jstor.org/stable/26432528

Zhang, L.-X., Hu, F., Cheung, S. H., & Chan, W. S. (2007). Asymptotic properties of covariate-adjusted response-adaptive designs. *The Annals of Statistics*, *35*(3), 1166–1182. https://doi.org/10.1214/009053606000001424

Zhao, J. (2023). Adaptive neyman allocation. https://arxiv.org/abs/2309.08808

Zhao W., Ma W., Wang F., & Hu F. (2022). Incorporating covariates information in adaptive clinical trials for precision medicine. *Pharmaceutical Statistics*, *21*(1), 176–195. https://doi.org/10.1002/pst.v21.1

Zhou Q., Ernst P. A., Morgan K. L., Rubin D. B., & Zhang A. (2018). Sequential rerandomization. *Biometrika*, *105*(3), 745–752. https://doi.org/10.1093/biomet/asy031

Zhu, H., Yu, J., Lai, D., & Wang, L. (2023). Seamless clinical trials with doubly adaptive biased coin designs. *The New England Journal of Statistics in Data Science*, *1*(3), 1–9. https://doi.org/10.51387/23-NEJSDS25

Zhu, H., & Zhu, H. (2023). Covariate-adjusted response-adaptive designs based on semiparametric approaches. *Biometrics*, *79*(4), 2895–2906. https://doi.org/10.1111/biom.13849