Biometrics n, 1–??
Month Year
DOI: 10.1111/xxxxxxxxxxxxx

GO-SMART: Generalized Outcome-Adaptive Sequential Multiple Assignment Randomized Trial

Xue Yang, Yu Cheng, Peter F. Thall, Abdus S. Wahed

Summary: Dynamic treatment regimes (DTRs) are effective vehicles for individualizing treatments for chronic diseases such as cancers, AIDs, and mental illnesses. Sequential multiple assignment randomized trials (SMARTs) provide a systematic process of constructing and evaluating DTRs. However, standard SMARTs ignore the potential treatment effects observed in past patients from the trial and continue exposing patients to inferior treatments. This, in practice, could result in decreased retention and lower treatment adherence. We developed a generalized outcome-adaptive SMART (GO-SMART) design that allows between-subject adaption to alleviate the above concern by imbalancing randomization probabilities for future patients in favor of the treatments observed to be more effective in previous patients. We proposed estimators for making inferences about the embedded DTRs from GO-SMARTs. Extensive simulation studies were conducted to evaluate the performance of GO-SMART compared to the classical SMART design. The analytical and simulation results show that the proposed estimators are consistent and, compared with the SMART design, the GO-SMART design treats significantly more participants with the optimal treatment without sacrificing the statistical power and objective of the design.

Key words: Outcome-adaptive randomization; Dynamic treatment regimes; Sequential multiple assignment randomized trials; Inverse probability weighting; Causal inference

This paper has been submitted for consideration for publication in *Biometrics*

1 Introduction

Complex diseases such as cancers, AIDs, obesity, and mental illnesses commonly require several stages of treatment, necessitating sequential medical decisions accounting for individual patient characteristics and dynamic disease progression (Kidwell, 2014). Rules for such individualized sequential decisions across stages are collectively known as dynamic treatment regimes (DTRs) (Lavori and Dawson, 2000, 2004; Murphy, 2005). They are mappings of patients' intermediate disease outcome and treatment history to the current treatment assignment (Kidwell, 2014). Sequential multiple assignment randomized trials (SMARTs) provide a systematical process of constructing and evaluating DTRs from a causal inference point of view (Ogbagaber et al., 2016; Kidwell, 2014; Liu et al., 2017). In contrast with the standard RCTs, where individuals are assigned to all fixed pre-determined treatment strategies or combinations, SMARTs closely mimic the clinical treatment process where patients go through multiple stages of treatment and the treatment assignments at each stage could depend on patient's treatment history, disease status, and other characteristics (Lavori and Dawson, 2000; Murphy, 2005; Dawson and Lavori, 2012; Kidwell, 2014).

In a SMART, participants are randomized to the available first-stage treatments at enrollment. They are then followed until some pre-specified event (e.g., response) triggers the start of the next stage. Participants migrating to a new stage are re-randomized to available treatment options at that stage (Cheung et al., 2015). The treatment options across stages potentially include continuing the same regimen, switching to other alternatives, or augmenting with other options to intensify the current treatment (Wang et al., 2022). For example, Auyeung et al. (2009) proposed a SMART design of neurobehavioral treatment for patients with malignant melanoma where two potential agents, escitalopram (ESC) or methylphenidate (MPH), were assessed. At the first stage, participants were equally randomized to ESC and MPH. Patients were evaluated for their Hamilton depression (HAMD)

scores at week 6, and if the HAMD score was at most 11 (in clinical remission), the patient was to continue current treatment for another 8 weeks in the second stage; however, if the HAMD score was 12 or higher, they were equally randomized to the augmented treatment (ESC+MPH) or to switch to the other agent.

Similar to standard randomized control trials (RCTs), the ethical dilemma of randomizing patients equally to treatments, when data from earlier patients show clear benefit towards some treatments, persists in SMARTs that potentially impedes its widespread use in real-world (Bell et al., 2013; Liu et al., 2017). Traditional SMART design randomizes participants equally to available options at each stage regardless of the likely benefit of the treatment (Cheung et al., 2015; Wang et al., 2022). This, in theory, may potentially boost the statistical power for identifying the optimal DTR (Murphy, 2005); however, randomizing patients without considering the potential treatment effects estimated from previous therapeutic experience may potentially damage the trust between patients and physicians (Thall and Wathen, 2007). This may lead to reduced enthusiasm in participants about the trial and consequently result in increased attrition and poor compliance (Liu et al., 2017). For example, in the CATIE Schizophrenia Trial, 74% of patients (1061 out of 1432) discontinued their assigned treatment in phase 1 before 18 months (Lieberman et al., 2005), and in the ExTENd study of treatment for alcohol dependence, out of 302 participants, 52 and 41 dropped out in the first and second stages respectively (Lei et al., 2012). In Kasari et al. (2014), attrition rates increased over time with 10% by week 12, 14% by week 24, and 25% by week 36.

Response-adaptive randomization (RAR) has long been proposed as an alternative to the standard RCTs (Thompson, 1933; Eisele, 1994; Wei and Durham, 1978; Karrison et al., 2003; Rosenberger and Hu, 2004; Hu and Rosenberger, 2006; Zhang and Rosenberger, 2007; Thall and Wathen, 2007; Rosenberger et al., 2012; Thall et al., 2015) to alleviate these issues and to achieve collective benefit for more patients in the trial by assigning more patients

to the better treatment (Williamson and Villar, 2020; Cheung et al., 2015), yet obtaining the same goal of estimating treatment efficacy as would be in the case of RCTs. For newly enrolled patients, RAR allows the randomization probabilities of treatment assignments to be modified based on the response from previous patients (Chow and Chang, 2008).

Inspired by RAR, some attempts have been made to apply "between-patients" adaptive randomization to SMARTs. Lee et al. (2015) employed a model-based Q-learning Bayesian objective function and applied ϵ-greedy adaptive randomization in a two-cycle phase I/II does-finding trial, where the adaption is both "within-patients" and "between-patients" in the sense that the Cycle 2 actions depended on individual patient's cycle one dose and outcome, and the posterior predictive distribution of treatment effectiveness estimated from other patients' data. Cheung et al. (2015) proposed a SMART with adaptive randomization (SMART-AR) design in a Q-learning framework to adapt the randomization probabilities using the historical and empirical randomization probabilities. However, both of these approaches would potentially require a longer duration and larger sample size than SMART as they need sufficient data to accurately estimate the treatment effectiveness and the Q-function before adaptation. Besides, it may not be robust to misspecification of the models at each stage. Chao et al. (2020) proposed a Bayesian group sequential small n SMART (snSMART) design that allows for removing the worst-performing treatment arm through a pre-specified number of interim analyses. However, this design is only applicable to a three-arm trial and is proposed for comparing treatment effects, not DTRs. Recently, Wang et al. (2022) proposed a response-adaptive SMART (RA-SMART) design that used treatment response data from the first stage to skew randomization probabilities in the second stage in favor of more promising treatments. However, this design only used first-stage efficacy information to guide the adaptation in the second stage. The lack of utilizing the second stage efficacy information potentially results in reduced statistical power.

In this article, we propose a generalized outcome-adaptive SMART (GO-SMART) design that allows adaptation of the randomization probabilities in both stages of a two-stage SMART, based on the information from all previous patients in the trial. We aim to improve adherence and retention by assigning more patients to those promising DTRs based on previous patients data without sacrificing the objective of efficiently evaluating and comparing DTRs. Since there are many variations of SMARTs, we will primarily consider a two-stage SMART design described in Thall et al. (2007) as an example to simplify the description of our adaptive schemes. The remainder of this paper is organized as follows: In Section 2, we introduce the framework, notation, and assumptions relevant to this work. The new generalized outcome-adaptive randomization scheme is introduced in Section 3 for a two-stage SMART design. In Section 4, we define various estimators for valid inference from the proposed design and demonstrate their consistency. In Section 5, we evaluate the performance of the GO-SMART design compared to classical SMART and RA-SMART designs (Wang et al., 2022) through extensive simulation studies. We conclude in Section 6 with some discussions and future directions.

2 Framework, notation, and assumptions

We consider a two-stage SMART design described in Thall et al. (2007). Suppose that at enrollment (stage I), each patient is randomized to a set of treatments $\mathcal{A} = \{A_s, s = 1, \ldots, J\}$. The goal is to achieve a response, a binary indicator of cure, remission, or improvement. Following a fixed period of follow-up, responders to the initial treatment either discontinue treatment or continue the same treatment, depending on the type of disease and response. Patients who do not respond to the initial treatment at stage I are then re-randomized to a set of alternative treatments excluding the treatment they received initially. This study design allows one to estimate and make inferences about $J(J-1)$ dynamic treatment regimes (DTRs), namely, $d(A_j, A_l)$, $A_j, A_l \in \mathcal{A}, j \neq l$, where the regime $d(A_j, A_l)$ stands for "Treat

with treatment A_j, if no response, treat with treatment A_l." Figure 1 shows an example of such a two-stage SMART design with three treatment options $\mathcal{A} = \{A_1, A_2, A_3\}$. Overall, six DTRs, $d(A_1, A_2)$, $d(A_1, A_3)$, $d(A_2, A_1)$, $d(A_2, A_3)$, $d(A_3, A_1)$ and $d(A_3, A_2)$ can be assessed at the conclusion of the study.

[Figure 1 about here.]

We now formalize the quantities of interest in terms of patient-specific potential outcomes. Let Y_1^a denote the potential response indicator (1 for a response, 0, otherwise) at the end of stage I if the patient received treatment a, and let Y_2^{ab} be the potential response indicator at the end of stage II if the patient received treatment b ($b \neq a$) in that stage following no response at stage I, $a, b \in \mathcal{A} = \{A_1, \ldots, A_J\}, b \neq a$. Then, the overall potential outcome $Y\{d(a,b)\}$ under the treatment regime $d(a,b)$ can be written as

$$Y\{d(a,b)\} = Y_1^a + (1 - Y_1^a)Y_2^{ab},$$

where $a, b \in \mathcal{A}, a \neq b$. Our first goal is to estimate the overall response rates for the $J(J-1)$ regimes $d(a,b)$, $a, b \in \mathcal{A}, a \neq b$, and to find the corresponding optimal treatment regime. The overall response rate, $\pi(a,b)$ under regime $d(a,b)$ is defined as

$$\pi(a,b) = E[Y\{d(a,b)\}] = \pi_1^a + (1 - \pi_1^a)\pi_2^{ab}, \tag{1}$$

where $\pi_1^a = E(Y_1^a) = Pr\{Y_1^a = 1\}$ is the stage I response rate for patients receiving treatment a, and $\pi_2^{ab} = E(Y_2^{ab} \mid Y_1^a = 0) = Pr\{Y_2^{ab} = 1 \mid Y_1^a = 0\}$ is the response rate for patients who received a in stage I, did not respond, and subsequently received b ($b \neq a$). The optimal DTR among the $J(J-1)$ regimes, $d(a^*, b^*)$, is defined as the best treatment regime with the highest response rate, s.t. $(a^*, b^*) = \underset{a,b \in \mathcal{A}, a \neq b}{\operatorname{argmax}} \pi(a,b)$. Note that for patient i, $Y_i\{d(A_j, A_l)\}, A_j, A_l \in \mathcal{A}, j \neq l$ represent the set of all potential overall response outcomes. In practice, only one among $J(J-1)$ potential outcomes $Y_i\{d(a,b)\}$ can be observed if patient i follows strategy

$d(a,b), (a \neq b \in \mathcal{A})$. The other potential outcomes will be the unobserved counterfactual outcomes for patient i (Rubin, 1974; Splawa-Neyman et al., 1990).

Suppose there are n patients in the study, and the observed data can be described as

$$\{I_{1i}(A_j), Y_{1i}, (1-Y_{1i})[I_{2i}(A_l), Y_{2i}]\}, i = 1, \ldots, n; A_j, A_l \in \mathcal{A}, j \neq l,$$

where Y_{1i} is the observed response for patient i in stage I, Y_{2i} is the observed response for patient i if they proceed to stage II, and $I_{ki}(x)$ is the indicator for receiving treatment x in stage k, $k = 1, 2$, $x \in \mathcal{A}$. Note that if patient i has achieved response through the first stage of treatment, the second stage treatment/response will be non-existent, and hence, in our notation, these quantities are preceded by the first stage non-response indicator $(1 - Y_{1i})$. The observed overall response for patient i, Y_i, can then be written as $Y_i = Y_{1i} + (1 - Y_{1i})\, Y_{2i}$.

To estimate the overall response rates under DTRs and make valid inferences, as is customary in the literature of causal inference (Ko et al., 2003; Cole and Frangakis, 2009), we assume that consistency, sequential randomization, and positivity assumptions hold under the standard Neyman-Rubin causal framework (Splawa-Neyman et al., 1990; Rubin, 1974). In addition, we assume that, by design, the randomization probability of assigning patient i to treatment A_j could depend only on the observed outcome and treatment assignment of previous $\{i-1\}$ patients. In Web Appendix A, we provide details about the assumptions.

3 GO-SMART design

In this section, we propose a GO-SMART design that dynamically adapts the randomization probabilities in both stages based on accumulated information from previous patients in the study. GO-SMART design aims to assign more patients to more promising treatments/DTRs by skewing the randomization probabilities based on the accumulated information on response rate during the trial. This "between-patient" adaption is implemented for both stage I and stage II randomization probabilities. In stage I, the randomization probability for

assigning treatment A_j to subject i , namely $P^1_{j,i}$, is allowed to depend on all previous $(i-1)$ patients' stage I information:

$$P^1_{j,i} = Pr\{I_{1i}(A_j) = 1 \mid H_{1i}\},$$

where H_{1i} is the accumulated first-stage information in the trial until the randomization of the ith patient. More specifically,

$$H_{1i} = \{(Y_{1k}, I_{1k}(A_1), \ldots, I_{1k}(A_J))_{k=1,\ldots,i-1}\}.$$

Note that although for simplicity, it is assumed here that the first-stage response information is available for all previous $(i-1)$ patients, the algorithm described below will apply even if it is available only for a subset of the $(i-1)$ patients. Similarly, the randomization probability for assigning treatment A_l to subject i in Stage II following a non-response to A_j in stage I, denoted by $P^2_{jl,i}$, is allowed to depend on all previous $(i-1)$ patients' available stage I and stage II information as follows:

$$P^2_{jl,i} = Pr\{I_{2i}(A_l) = 1 \mid I_{1i}(A_j) = 1, Y_{1i} = 0, H_{2i}\}.$$

where H_{2i} is the accumulated information collected immediately prior to the second stage randomization of the ith patient. In particular, we proposed two types of adaptive schemes with different stage II adaptation rules. sFigure 1 describes the sketch of the adaptive randomization algorithm, which is detailed below.

GO-SMART ADAPTAION ALGORITHM

Stage I: Let p_0 be a pre-specified burn-in proportion, and $n_0 = \lfloor p_0 n \rfloor < n$ is referred to as the first-stage burn-in sample, where $\lfloor x \rfloor$ means the greatest integer less than or equal to x. Then, for $i = 1, \ldots, n_0$, we randomize patient i to treatment $A_j, A_j \in \mathcal{A}$ with equal probability, i.e. $P^1_{j,i} = 1/J$. For $i = n_0+1, \ldots, n$, patient i is randomized to treatment $A_j \in \mathcal{A}$ with probability

$$P^1_{j,i} \equiv P^1_{j,i}(c, \epsilon) = \min\left[\max\left\{P^1_{j,i}(c), \epsilon\right\}, 1-\epsilon\right],$$

where $P^1_{j,i}(c) = \left(\hat{\pi}^{A_j}_{\{1,\{i-1\}\}}\right)^c \Big/ \sum_{k=1}^{J}\left(\hat{\pi}^{A_k}_{\{1,\{i-1\}\}}\right)^c$, and $\hat{\pi}^{A_j}_{\{1,\{i-1\}\}} = \hat{Pr}\{Y_{1i}^{A_j} = 1 \mid H_{1i}\}$ is the first-stage response probability for treatment A_j from the previous $\{i-1\}$ patients in the trial. The constant $c \in [0,1]$ is a tuning parameter to control the degree of dependence of the randomization probability on the observed response rates, and $\epsilon \in (0,1)$ is a randomization constraint parameter that bounds the randomization probabilities away from the boundaries of 0 and 1.

Stage II: Let p_1 be a pre-specified stage II burn-in proportion and $n_1 = \lfloor p_1 n \rfloor$, $(n_0 < n_1 < n)$ be the second-stage burn-in sample. Responders in Stage I will stop the treatment (or continue the same treatment), whereas the non-responders will be re-randomized in this stage to alternative treatments with probability:

$$P^2_{jl,i} = \begin{cases} \frac{1}{J-1}, & i = 1,\ldots,n_0, \\ P^2_{jl,i}(c,\epsilon) \equiv \min\left[\max\left\{P^2_{jl,i}(c),\epsilon\right\}, 1-\epsilon\right], & i = n_0+1,\ldots,n_1, \\ P^{2*}_{jl,i}(c,\epsilon) \equiv \min\left[\max\left\{P^{2*}_{jl,i}(c),\epsilon\right\}, 1-\epsilon\right], & i = n_1+1,\ldots,n, \end{cases}$$

where $P^2_{jl,i}(c) = \left(\hat{\pi}^{A_l}_{\{1,\{i-1\}\}}\right)^c \Big/ \sum_{k\neq j}\left(\hat{\pi}^{A_k}_{\{1,\{i-1\}\}}\right)^c$, and $\hat{\pi}^{A_l}_{\{1,\{i-1\}\}} = \hat{Pr}\{Y_{1i}^{A_l} = 1 \mid H_{1i}\}$, as defined before. The adapted randomization probability $P^{2*}_{jl,i}(c)$ for patients randomized in the second stage after the first n_1 patients uses second-stage treatment information. We propose two ways of adaptation, leading to two forms of GO-SMART design.

GO-SMART AR-1 use the stage II cumulative conditional efficacy information among those who did not respond in stage I to guide the second stage randomization probabilities for non-responders. More specifically, define: $P^{2*}_{jl,i}(c) = (\hat{\pi}^{A_jA_l}_{\{2,\{i-1\}\}})^c \Big/ \sum_{k\neq j}(\hat{\pi}^{A_jA_k}_{\{2,\{i-1\}\}})^c$, and $\hat{\pi}^{A_jA_l}_{\{2,\{i-1\}\}} = \hat{Pr}\{Y_{2i}^{A_jA_l} = 1 \mid Y_{1i}^{A_j} = 0, I_{1i}(A_j) = 1, H_{2i}\}$ is the conditional probability of responding to A_l given the patient did not respond to A_j, estimated from the first $\{i-1\}$ patients.

GO-SMART AR-2 use the cumulative efficacy for the corresponding DTR to guide the

second stage randomization probabilities for non-responders to define

$$P^{2*}_{jl,i}(c) = \left\{\hat{\pi}^{\{i-1\}}(A_j, A_l)\right\}^c \Big/ \sum_{k\neq j} \left\{\hat{\pi}^{\{i-1\}}(A_j, A_k)\right\}^c,$$

with $\hat{\pi}^{\{i-1\}}(A_j, A_l) = \hat{Pr}\{Y_i\{d(A_j, A_l)\} = 1 \Big| I_{1i}(A_j) = 1, H_{2i}\}$ is the overall response rate of DTR $d(A_j, A_l)$ estimated from the previous $\{i-1\}$ patients.

4 Inference about Embedded DTR from GO-SMART design

In this section, we present four estimators for overall response rate $\pi(A_j, A_l)$ associated with DTR $d(A_j, A_l)$, $A_j, A_l \in \mathcal{A}, j \neq l$. We provide corresponding variance estimators and evaluate consistency. Details about the proof of the Theorems in this section can be found in Web Appendix B. To facilitate the construction of the estimators, let us define $n_{1j} = \sum_{i=1}^n I_{1i}(A_j)$, $r_{1j} = \sum_{i=1}^n I_{1i}(A_j)Y_{1i}$, $n_{2jl} = \sum_{i=1}^n I_{1i}(A_j)(1-Y_{1i})I_{2i}(A_l)$, and $r_{2jl} = \sum_{i=1}^n I_{1i}(A_j)(1-Y_{1i})I_{2i}(A_l)Y_{2i}$ to be the number of patients assigned to A_j in stage I, the number of patients responded to A_j in stage I, the number of patients assigned to A_l in stage II upon a no-response to A_j in stage I, and the number of patients responded to A_l in stage II after a no-response to A_j in stage I, respectively. We further define $\hat{\pi}_1^{A_j} = r_{1j}/n_{1j}$ and $\hat{\pi}_2^{A_jA_l} = r_{2jl}/n_{2jl}$ to be the proportion of first-stage responder among patients assigned to A_j in stage I and the proportion of second-stage responders among those who received A_l in the second stage following a non-response in the first stage to A_j.

4.1 Sample mean estimator

The simplest and perhaps the most naive estimator of $\pi(A_j, A_l)$ is the sample mean estimator, which in the case of our binary response is essentially the sample proportion of responders among all patients in the sample who are treated following the DTR $d(A_j, A_l)$. In terms of observed data from GO-SMART:

$$\hat{\pi}^{SM}(A_j, A_l) = \frac{r_{1j} + r_{2jl}}{r_{1j} + n_{2jl}}, \tag{2}$$

which is the ratio of the number of overall responders and the number of patients treated under $d(A_j, A_l)$. An approximate formula for the variance estimator is given by the corresponding formula for the variance of a sample proportion, namely,

$$\hat{Var}\{\hat{\pi}^{SM}(A_j, A_l)\} = \frac{\hat{\pi}^{SM}(A_j, A_l)(1 - \hat{\pi}^{SM}(A_j, A_l))}{r_{1j} + n_{2jl}}.$$

The sample mean estimator is an inconsistent estimator for the true DTR response rate (See Web Appendix B-2).

4.2 G-estimator

A G-estimator (Murphy, 2005; Bembom and van der Laan, 2007) for the overall response rate under DTR $d(A_j, A_l)$ is obtained by plugging in the estimated stage I and stage II sample response rates in the RHS of Equation (1) for $a = A_j$ and $b = A_l$. More specifically, the G-estimator of $\pi(A_j, A_l)$ can be written as:

$$\hat{\pi}^G(A_j, A_l) = \hat{\pi}_1^{A_j} + (1 - \hat{\pi}_1^{A_j})\hat{\pi}_2^{A_j A_l}. \tag{3}$$

Using the delta method, an approximate variance estimator for $\hat{\pi}^G(A_j, A_l)$ can be given by

$$\hat{Var}\{\hat{\pi}^G(A_j, A_l)\} \approx (1 - \hat{\pi}_2^{A_j A_l})^2 \hat{Var}(\hat{\pi}_1^{A_j}) + (1 - \hat{\pi}_1^{A_j})^2 \hat{Var}(\hat{\pi}_2^{A_j A_l}),$$

where $\hat{Var}(\hat{\pi}_1^{A_j}) = \hat{\pi}_1^{A_j}(1 - \hat{\pi}_1^{A_j})/n_{1j}$, $\hat{Var}(\hat{\pi}_2^{A_j A_l}) = \hat{\pi}_2^{A_j A_l}(1 - \hat{\pi}_2^{A_j A_l})/n_{2jl}$.

THEOREM 4.1 (Consistency for G-estimator): *Under the consistency, sequential randomization, and positivity assumptions, and assuming that the randomization probabilities in both stages are bounded away from zero for each patient, we have the following results:*

(1) $\hat{\pi}_1^{A_j} \xrightarrow{p} \pi_1^{A_j}$ *as* $n \to \infty$, *for any* $A_j \in \mathcal{A}$.

(2) $\hat{\pi}_2^{A_j A_l} \xrightarrow{p} \pi_2^{A_j A_l}$ *as* $n \to \infty$, *for any* $A_j, A_l \in \mathcal{A}, j \neq l$.

(3) $\hat{\pi}^G(A_j, A_l) \xrightarrow{p} \pi(A_j, A_l)$ *as* $n \to \infty$, *for any* $A_j, A_l \in \mathcal{A}, j \neq l$.

4.3 Inverse probability of randomization weighted estimators

Since patients in GO-SMART are potentially randomized with unequal probabilities, it is natural to consider inverse probability of randomization weighted (IPRW) estimators (Robins et al., 1994; Ko and Wahed, 2012). This estimator takes the form

$$\hat{\pi}^{IPRW}(A_j, A_l) = \frac{1}{n}\sum_{i=1}^{n}\left\{\frac{Y_{1i}I_{1i}(A_j)}{P^1_{j,i}} + \frac{(1-Y_{1i})I_{1i}(A_j)I_{2i}(A_l)}{P^1_{j,i}P^2_{jl,i}}\right\}Y_i. \tag{4}$$

The estimated variance of IPRW estimator can be approximated by

$$\hat{V}ar\{\hat{\pi}^{IPRW}(A_j, A_l)\} \approx \frac{1}{n^2}\sum_{i=1}^{n}\left[\left\{\frac{Y_{1i}I_{1i}(A_j)}{P^1_{j,i}} + \frac{(1-Y_{1i})I_{1i}(A_j)I_{2i}(A_l)}{P^1_{j,i}P^2_{jl,i}}\right\}Y_i - \hat{\pi}^{IPRW}(A_j, A_l)\right]^2.$$

THEOREM 4.2 (Consistency for IPRW estimator): *Under the assumptions of consistency, sequential randomization, and the randomization probabilities being bounded away from 0 and 1 for all patients, we have*

$$\hat{\pi}^{IPRW}(A_j, A_l) \xrightarrow{p} \pi(A_j, A_l) \quad as \quad n \to \infty, \forall A_j, A_l \in \mathcal{A}, j \neq l.$$

4.4 IPRW estimator with normalized weights

Another widely used estimator for DTR response rate is the normalized inverse probability of randomization weighted (NIPRW) estimator, which uses the sum of individual weights instead of the sample size in the denominator of IPRW estimator. The NIPRW estimator is defined as

$$\hat{\pi}^{NIPRW}(A_j, A_l) = \frac{\sum_{i=1}^{n} W_{jl,i}Y_i}{\sum_{i=1}^{n} W_{jl,i}}, \tag{5}$$

where

$$W_{jl,i} = \frac{Y_{1i}I_{1i}(A_j)}{P^1_{j,i}} + \frac{(1-Y_{1i})I_{1i}(A_j)I_{2i}(A_l)}{P^1_{j,i}P^2_{jl,i}}.$$

The variance of this estimator can be estimated approximately by

$$\hat{V}ar\{\hat{\pi}^{NIPRW}(A_j, A_l)\} \approx \frac{1}{n^2}\sum_{i=1}^{n}\left\{W_{jl,i}(Y_i - \hat{\pi}^{NIPRW}(A_j, A_l))\right\}^2.$$

THEOREM 4.3 (Consistency for NIPRW estimator): *Under Theorem 4.2, consistency, sequential randomization, and bounding the randomization probabilities away from 0 and 1 for all patients, we have*

$$\hat{\pi}^{NIPRW}(A_j, A_l) \xrightarrow{p} \pi(A_j, A_l) \quad as \quad n \to \infty, \forall A_j, A_l \in \mathcal{A}, j \neq l.$$

5 Simulation Study

In order to demonstrate the operating characteristics of the GO-SMART design in comparison to the classical SMART design, we simulated clinical trials under various conditions. The two designs are assessed along with the RA-SMART design (Wang et al., 2022) based on their ability to (1) produce unbiased estimates of overall response rates for DTRS, (2) identify optimal DTR, and (3) treat more patients with more effective treatment/DTR.

5.1 Population

We simulated trials with three potential treatment options, that is, $\mathcal{A} = (A_1, A_2, A_3)$ under 8 scenarios. Table 1 summarizes the true parameter values for stage I response rate, $\pi_1^{A_j}, A_j \in \mathcal{A}$, and conditional response rate in stage II given non-response in stage I, $\pi_2^{A_j A_l}, A_j, A_l \in \mathcal{A}, j \neq l$. The response rate under regime $d(A_j, A_l)$, $\pi(A_j, A_l)$, was calculated based on Equation (1). Under each scenario, we compared the GO-SMART AR-1 and GO-SMART AR-2 with the classical SMART design and the RA-SMART design. Each scenario is characterized by three aspects: (1) whether the optimal DTR involves the optimal first- and second- stage optimal treatment options, more specifically, whether $a_1^{opt} = \underset{a \in \mathcal{A}}{\text{argmax}}\, \pi_1^a = a^*$ such that $(a^*, b^*) = \underset{a,b \in \mathcal{A}, a \neq b}{\text{argmax}}\, \pi(a, b)$ and $(a_2^{opt}, b_2^{opt}) = \underset{a,b \in \mathcal{A}, a \neq b}{\text{argmax}}\, \pi_2^{ab} = (a^*, b^*)$, (2) the effect size (how different the optimal DTR is from the next optimal DTR in terms of overall response rate), and (3) how different the optimal DTR is from the worst DTR in terms of overall response rate.

Scenario S0 is the null scenario where all first-line and second-line options work the same,

leading to no (or all) optimal DTR. Scenarios S1 to S3 are designed so that the optimal DTR is $d(A_1, A_3)$ and it involves optimal stage I treatment ($A_1 = \underset{a \in \mathcal{A}}{\text{argmax}}\, \pi_1^a$) and conditional stage II optimal treatment options $(A_1, A_3) = \underset{(a,b) \in \mathcal{A}, a \neq b}{\text{argmax}}\, \pi_2^{ab}$. In S1, the optimal DTR response rate (0.70) is 0.42 higher than that for the worst DTR, while in S2, the response rate for optimal DTR is 0.30 higher than the worst DTR. Both in S1 and S2, the optimal DTR has only 0.05 higher response rate compared to the next best DTR, whereas in S3, the optimal DTR differs from the next best DTR by a larger margin of 0.21. In S4, the optimal DTR contains the optimal stage I treatment but does not contain the stage II optimal conditional treatment. From S5 to S7, the optimal DTR does not involve the optimal stage I treatment. Additionally, in S5, there are no differences in the effectiveness of the first-stage treatments. In scenarios S6 and S7 low DTR response rates (all less than 0.30) are investigated with varying effect sizes and the optimal DTR is initialized with the inferior treatment. In S7 though, the worst DTR does not involve the inferior treatment from stage I, whereas in all other scenarios, the worst DTR initializes with the inferior treatment in stage I.

[Table 1 about here.]

5.2 Data generation

In all scenarios, for each individual i in the population, the first-stage potential outcome $Y_{1i}^{A_j}$ was drawn from a Bernoulli distribution with probability $\pi_1^{A_j}$, for $j = 1, 2, 3$. The second-stage potential outcome given non-response in the first stage $Y_{2i}^{A_j A_l} \mid Y_{1i}^{A_j} = 0$ was drawn from a Bernoulli distribution with probability $\pi_2^{A_j A_l}, j = 1, 2, 3; j \neq l$. Overall, six DTRs, $d(A_j, A_l), j = 1, 2, 3; j \neq l$ were assessed at the conclusion of each trial.

Under the SMART design, as is customary, the first stage randomization probabilities were fixed to be $P_{j,i}^1 = \frac{1}{3}$ for all $i = 1, ..., n$ and $j = 1, 2, 3$, the second stage randomization probabilities were fixed at $\frac{1}{2}$ for nonresponders ($P_{jl,i}^2 = \frac{1}{2}$ for all $i = 1, ..., n$ and $l \neq j \in 1, 2, 3$). First-stage responders stopped treatment in stage II. Under the RA-SMART, data were

generated following Wang et al. (2022), where the first stage randomization probabilities were fixed to be $P^1_{j,i} = 1/3$ for all $i = 1, ..., n$ and $j = 1, 2, 3$, and the observed response rate for first $n_0 = \lfloor p_0 n \rfloor$ patients were used to determine the observed inferior treatment $A_{\hat{\tau}} = \underset{A_j \in T}{\operatorname{argmin}} \pi^{A_j}_{1(n_0)}$, $\hat{\pi}^{A_j}_{1(n_0)} = \sum_{i=1}^{n_0} I_{1i}(A_j) Y_{1i} \big/ \sum_{i=1}^{n_0} I_{1i}(A_j)$. In the second stage, the first $n_0 = \lfloor p_0 n \rfloor$ patients were randomized equally to the two treatments they did not receive in the first stage, that is, $P^2_{jl,i} = 0.5$ for all $i = 1, ..., n_0$ and $l \neq j \in 1, 2, 3$. For the remaining $n - n_0$ patients, the second-stage randomization probabilities were modified to be smaller for the identified inferior treatment. For details, please see Wang et al. (2022). The burn-in proportion p_0 was varied between 0.25 and 0.5, and the adjusted randomization probability for inferior treatment in the second stage was set to be 0.2.

For the GO-SMART design, patients were assigned to treatments following the algorithm proposed in Section 3. As suggested in Wathen and Thall (2017), we set $\epsilon = 0.1$ to bound the randomization probabilities between $[0.1, 0.9]$. For the tuning parameter c, common values considered in the literature include $c = 0.5, 1$, and $i/(2n)$, where i is the current sample size when a new patient is to be randomized, and n is the trial's total sample size. sFigure 2 in the Web Appendix D shows how the tuning parameter c skews the randomization probability away from 0.5 in the second stage of a three-treatment scenario for various values of the previously observed response rates. The lower the c, the closer to 0.5 the randomization probability is. As c increases the randomization probability skews more towards the better-performing treatment. The more separation there is between the response rates, the further away the randomization probabilities are from 0.5 regardless of the value of c. A value of $c = i/(2n)$ was recommended by (Thall and Wathen, 2007), where c ranges from $(n_0+1)/n$ to 0.5. We tested multiple values of $c \in \{0.25, 0.5, 0.75, 1, i/(2n), i/n\}$ for GO-SMART designs. However, in our simulations with GO-SMART, we found that $c = 1$ and $c = i/n$ performed better than other choices of c. In the following section, we presented the results for $c = \frac{i}{n}$.

The simulation results for $c = 1$ and $c = \frac{i}{2n}$ are presented in the Web Appendix E. The response rate for each DTR was estimated using the sample mean estimator in Equation (2), the G-estimator in Equation (3), the IPRW estimator in Equation (4), and the NIPRW estimator in Equation (5). The burn-in proportions (p_0, p_1) were varied between $(0.25, 0.5)$ and $(0.5, 0.75)$. For each scenario, we simulated 10,000 Monte Carlo trials. The operating characteristics were evaluated under two different total sample sizes, $n = 300$ and $n = 600$.

5.3 Simulation results

5.3.1 Estimation of DTR response rate

We evaluate estimators based on the coverage probability, calculated as the proportion of times the corresponding 95% Wald confidence interval contained the true response rate for the DTR. As illustrated in Figure 2, under both the null and the alternative scenarios, the sample mean estimators resulted in a very low coverage ranging from 0.10 to 0.60. This is not surprising because the sample mean estimator is an inconsistent estimator of the DTR response rate, as demonstrated theoretically in Web Appendix B-2, and shown in the Web Appendix E sFigure 3 that the sample mean estimators had large biases ranging from 0.06 to 0.16. This is consistent with the findings reported in Thall et al. (2015). Therefore, we excluded this estimator from further evaluations regarding the power and the expected number of responses. Overall, the remaining three estimators (G-, IPRW, NIPRW) performed well under SMART, RA-SMART, and GO-SMART designs with coverage probabilities close to the nominal level of 0.95.

[Figure 2 about here.]

5.3.2 Type I error and power of identifying the optimal DTR

[Figure 3 about here.]

The Type I error and power are calculated as the proportion of $10,000$ Monte-Carlo simulations for which the designed optimal DTR is estimated to be optimal under the null and alternative scenarios, respectively. Figure 3 presents the Type I error for sample sizes $n = 300$ and $n = 600$. SMART, RA-SMART, and GO-SMART - all three designs lead to Type I errors close to the nominal level of 1/6. Note that the coverage probabilities are sometimes slightly lower than 1/6 for IPRW estimators under the RA-SMART design, however, they are all within the Monte Carlo margin of error $\pm 1.96 * \sqrt{(1/6) * (1 - 1/6)/10000} = \pm 0.007$.

Figure 4 compares the power of SMART, RA-SMART, and GO-SMART designs for $n = 600$ under different scenarios. First we note that the powers achieved for scenarios S1 and S3 are similar to those for scenarios S2 and S4, respectively, even though S3 and S4 result in larger powers than S1 and S2 regardless of the designs and estimation methods. This shows that the difference between the optimal DTR and the worst DTR in terms of response rates does not affect the statistical power of the design; however, the difference between the optimal DTR and the next best DTR, as one would expect, affects the power. The larger this difference is, the higher the power is. The similarity of estimated powers in scenarios S3 and S4 shows that when the optimal DTR involves the optimal stage I treatment, the powers do not get affected by whether it involves the optimal stage II treatment or not. A comparison of scenarios S1, S2, S3, and S4 to scenarios S5 and S6 shows that whether the optimal DTR involves the optimal stage I treatment or not affects the power depending on the design. In scenarios S5 and S6, where the optimal DTR does not involve the optimal stage I treatment, both the RA-SMART and GO-SMART adaptive designs are less powerful compared to the traditional SMART design. However, GO-SMART is still more powerful than RA-SMART design in these scenarios. On the other hand, in scenarios S1, S2, S3, and S4, where the optimal DTR contains the optimal stage I treatment, the GO-SMART design is more powerful than the SMART and RA-SMART designs. The power estimates in scenario

S6, where low true DTR response rates are assumed (ranging in $[0.107, 0.297]$), show that the power is not affected by the overall response rate; rather, it depends on the difference in the response rate between the optimal and the next best DTRs.

Across almost all scenarios, the IPRW estimators have lower power compared to the G-estimator and the NIPRW-estimator. This is due to the fact that the IPRW estimator has a larger variance compared to the G- and NIPRW estimators. Similar results have been observed in Ko and Wahed (2012) and Valente et al. (2020).

[Figure 4 about here.]

When the sample size is small and/or the response rates for DTRs are close to each other, it is challenging to identify the optimal DTR. In such cases, it might be of interest to identify a set of best DTRs rather than a single optimal DTR. The set consists of the best DTR and the adjacent ones with response rates within a pre-specified neighborhood of the optimal DTR. In our simulations, we examined the power of identifying the best set of DTRs within a 0.05 neighborhood of the optimal DTR response rates. In such cases, the power is defined as the proportion of simulated trials for which the optimal DTR is contained in the estimated best DTR set. For $n = 300$ and scenario S1, sFigure 4 in the Web Appendix E shows that the powers are effectively increased by identifying the proportion of sets containing best DTRs.

5.3.3 Total number of responders in the trial, and number of patients treated with the true optimal and worst DTRs

As is evident from the results above, GO-SMART design maintains Type I error at the nominal level and achieves similar or better power in identifying the optimal DTR compared to traditional SMART. The biggest advantage of the GO-SMART design over the SMART and the RA-SMART designs is its ability to benefit more patients in the trial than its counterparts. This can be examined by computing the average of the total number of responses in the trial as well as the number of patients treated with the true optimal and

worst DTRs. For $n = 600$, sFigure 5 in Web Appendix E shows the average of the total number of responders in the trial. For scenario S1, where the DTR response rate varied between 0.28 and 0.70, when $(p_0, p_1) = (0.25, 0.5)$, GO-SMART AR-1 and AR-2 designs resulted in 326 and 324 total number of responders out of the 600 randomized compared to 309 and 311 under SMART and RA-SMART designs respectively. Similarly, in scenario S6 where DTR response rate varied between 0.107 and 0.297, when $(p_0, p_1) = (0.25, 0.5)$, GO-SMART AR-1 and AR-2 designs had on average 111 and 108 total responders compared to 106 and 104 in SMART and RA-SMART designs, respectively.

Figure 5 shows the average number of patients treated with the optimal and the worst DTRs. Regardless of the value of (p_0, p_1) and scenario, the GO-SMART designs assign more patients to the optimal DTR and fewer patients to the worst than the SMART and RA-SMART designs; and between two forms of adaptations within GO-SMART, AR-1 assigns more patients to optimal DTR and fewer patients to worst DTR than AR-2 does. For example, in S2, when $(p_0, p_1) = (0.5, 0.75)$ GO-SMART AR-1 and AR-2 designs on average treat 174 and 172 patients with the optimal DTR compared to 150 for SMART and 135 for RA-SMART. Whereas in the same scenario, GO-SMART AR-1 and AR-2 designs, on average, treat 99 and 100 patients with the worst DTR compared to 120 for SMART and 120 for RA-SMART. Noticeably, in S2 and S3, where the optimal DTR involves inferior treatment in stage II, the RA-SMART design assigns less patients to the optimal DTR, and in S7, where the worst DTR does not involve the inferior treatment, the RA-SMART design assigns more patients to the worst DTR compared to the SMART design. However, GO-SMART is still favorable in this scenario because the adaptive randomization scheme depends on the rank of the response rates in both stages.

[Figure 5 about here.]

6 Discussion

In this article, we propose a general outcome-adaptive SMART or GO-SMART design that allows the randomization probabilities in all stages to be modified based on the accumulated information from all previous patients. As shown through simulation studies, the GO-SMART design achieves similar power as classical SMART in identifying the optimal DTR but results in more patients achieving response in the trial compared to SMART. Thus, GO-SMART is expected to improve the acceptability of treatment assignments and, consequently, the adherence as it increases the likelihood of response by assigning significantly more patients to the optimal DTR and fewer patients to the worst DTR than the classical SMART.

Compared to the existing RA-SMART design, our design has the following advantages: First, GO-SMART allows "between-patients" adaption in both the first and second stages. Instead of reducing the randomization probability of the observed inferior arm in adapting second-stage randomization probabilities, we propose two types of adaptive schemes that allow the second-stage randomization probabilities to be modified based on the conditional treatment efficacy or the dynamic treatment regime. Both schemes utilize first and second-stage treatment efficacy information from previous patients and hence result in increased efficiency. Second, unlike RA-SMART, GO-SMART successfully treats more patients using optimal DTR than SMART when the optimal DTR involves the inferior treatment in stage II. Third, the proposed GO-SMART design can be flexibly generalized to other SMART designs with different treatment options across stages. This is beyond the scope of this paper and will be elaborated in a future paper.

Despite excellent operating characteristics achieved by GO-SMART, some limitations of the design deserve further consideration. First, in the GO-SMART design, we aim to find the optimal DTR, which is defined as the DTR with the maximum estimated response

rate. This is implemented in the proposed design based on a simple numeric comparison of the point estimates and fails to incorporate uncertainty in the estimation. One alternative we considered in the simulation studies is to identify a set of best DTRs that are within some pre-specified threshold. There are also other ways to define the set of best DTRs, e.g., those with estimated response rates above a fixed threshold. Second, our formulation ignores multiple comparisons of DTR response rates. The multiple comparisons with the best (MCB) methodology might potentially be employed to formalize this idea (Hsu, 1981, 1984). Third, to simplify the conceptual framework, we did not incorporate covariates in the design. However, in practice, it is also very common to observe auxiliary covariate information. It is worthy to generalize the GO-SMART by including the observed covariates in the design process (e.g., in determining the updated randomization probabilities). Fourth, we put the same weight on the response in the first stage and that in the second stage. However, in reality, there are different viewpoints on what is desirable in a given response trajectory. Perhaps the quicker the response, the better the health outcomes. Finally, the proposed GO-SMART design will not be easily applied to clinical trials where clinical endpoints are delayed or need to wait a very long time to observe. In this case, instead of the "true" clinical endpoint, a surrogate endpoint that is quicker and easier to measure may be considered as the outcome to apply the proposed adaptive scheme.

References

Auyeung, S. F., Long, Q., Royster, E. B., Murthy, S., McNutt, M. D., Lawson, D., Miller, A., Manatunga, A., and Musselman, D. L. (2009). Sequential multiple-assignment randomized trial design of neurobehavioral treatment for patients with metastatic malignant melanoma undergoing high-dose interferon-alpha therapy. *Clinical trials (London, England)* **6,** 480–490.

Bell, M. L., Kenward, M. G., Fairclough, D. L., and Horton, N. J. (2013). Differential dropout

and bias in randomised controlled trials: when it matters and when it may not. *The BMJ* **346,** e8668.

Bembom, O. and van der Laan, M. J. (2007). Statistical methods for analyzing sequentially randomized trials. *Journal of the National Cancer Institute* **99,** 1577–1582.

Chao, Y.-C., Braun, T. M., Tamura, R. N., and Kidwell, K. M. (2020). A Bayesian Group Sequential Small n Sequential Multiple-Assignment Randomized Trial. *Journal of the Royal Statistical Society Series C: Applied Statistics* **69,** 663–680.

Cheung, Y. K., Chakraborty, B., and Davidson, K. W. (2015). Sequential multiple assignment randomized trial (SMART) with adaptive randomization for quality improvement in depression treatment program. *Biometrics* **71,** 450–459.

Chow, S.-C. and Chang, M. (2008). Adaptive design methods in clinical trials a review. *Orphanet Journal of Rare Diseases* **3,** 1–13.

Cole, S. R. and Frangakis, C. E. (2009). The Consistency Statement in Causal Inference: A Definition or an Assumption? *Epidemiology* **20,** 3–5.

Dawson, R. and Lavori, P. W. (2012). Efficient design and inference for multistage randomized trials of individualized treatment policies. *Biostatistics (Oxford, England)* **13,** 142–152.

Eisele, J. R. (1994). The doubly adaptive biased coin design for sequential clinical trials. *Journal of Statistical Planning and Inference* **38,** 249–261.

Hsu, J. C. (1981). Simultaneous Confidence Intervals for all Distances from the "Best". *The Annals of Statistics* **9,** 1026–1034. Publisher: Institute of Mathematical Statistics.

Hsu, J. C. (1984). Constrained Simultaneous Confidence Intervals for Multiple Comparisons with the Best. *The Annals of Statistics* **12,** 1136–1144.

Hu, F. and Rosenberger, W. F. (2006). *The Theory of Response-Adaptive Randomization in Clinical Trials*. John Wiley & Sons.

Karrison, T. G., Huo, D., and Chappell, R. (2003). A group sequential, response-adaptive design for randomized clinical trials. *Controlled Clinical Trials* **24,** 506–522.

Kasari, C., Kaiser, A., Goods, K., Nietfeld, J., Mathy, P., Landa, R., Murphy, S., and Almirall, D. (2014). Communication Interventions for Minimally Verbal Children With Autism: Sequential Multiple Assignment Randomized Trial. *Journal of the American Academy of Child and Adolescent Psychiatry* **53,** 635–646.

Kidwell, K. M. (2014). SMART Designs in Cancer Research: Past, Present and Future. *Clinical trials (London, England)* **11,** 445–456.

Ko, H., Hogan, J. W., and Mayer, K. H. (2003). Estimating Causal Treatment Effects from Longitudinal HIV Natural History Studies Using Marginal Structural Models. *Biometrics* **59,** 152–162.

Ko, J. H. and Wahed, A. S. (2012). Up-front versus sequential randomizations for inference on adaptive treatment strategies. *Statistics in Medicine* **31,** 812–830.

Lavori, P. W. and Dawson, R. (2000). A design for testing clinical strategies: biased adaptive within-subject randomization. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* **163,** 29–38.

Lavori, P. W. and Dawson, R. (2004). Dynamic treatment regimes: practical design considerations. *Clinical Trials* **1,** 9–20.

Lee, J., Thall, P. F., Ji, Y., and Mller, P. (2015). Bayesian Dose-Finding in Two Treatment Cycles Based on the Joint Utility of Efficacy and Toxicity. *Journal of the American Statistical Association* **110,** 711–722.

Lei, H., Nahum-Shani, I., Lynch, K., Oslin, D., and Murphy, S. (2012). A SMART Design for Building Individualized Treatment Sequences. *Annual review of clinical psychology* **8,** 21–48.

Lieberman, J. A., Stroup, T. S., McEvoy, J. P., Swartz, M. S., Rosenheck, R. A., Perkins,

D. O., Keefe, R. S., Davis, S. M., Davis, C. E., Lebowitz, B. D., Severe, J., and Hsiao, J. K. (2005). Effectiveness of Antipsychotic Drugs in Patients with Chronic Schizophrenia. *New England Journal of Medicine* **353,** 1209–1223.

Liu, Y., Wang, Y., and Zeng, D. (2017). Sequential Multiple Assignment Randomization Trials with Enrichment Design. *Biometrics* **73,** 378–390.

Murphy, S. A. (2005). An experimental design for the development of adaptive treatment strategies: DEVELOPMENT OF ADAPTIVE TREATMENT STRATEGIES. *Statistics in Medicine* **24,** 1455–1481.

Ogbagaber, S. B., Karp, J., and Wahed, A. S. (2016). Design of sequentially randomized trials for testing adaptive treatment strategies: S. B. OGBAGABER, J. KARP AND A. S. WAHED. *Statistics in Medicine* **35,** 840–858.

Robins, J. M., Rotnitzky, A., and Zhao, L. P. (1994). Estimation of Regression Coefficients When Some Regressors Are Not Always Observed. *Journal of the American Statistical Association* **89,** 846–866.

Rosenberger, W. F. and Hu, F. (2004). Maximizing power and minimizing treatment failures in clinical trials. *Clinical Trials* **1,** 141–147.

Rosenberger, W. F., Sverdlov, O., and Hu, F. (2012). Adaptive Randomization for Clinical Trials. *Journal of Biopharmaceutical Statistics* **22,** 719–736.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology* **66,** 688–701.

Splawa-Neyman, J., Dabrowska, D. M., and Speed, T. P. (1990). On the Application of Probability Theory to Agricultural Experiments. Essay on Principles. Section 9. *Statistical Science* **5,** 465–472.

Thall, P., Fox, P., and Wathen, J. (2015). Statistical controversies in clinical research: scientific and ethical problems with adaptive randomization in comparative clinical trials.

Annals of Oncology **26,** 1621–1628.

Thall, P. F. and Wathen, J. K. (2007). Practical Bayesian adaptive randomisation in clinical trials. *European Journal of Cancer* **43,** 859–866.

Thall, P. F., Wooten, L. H., Logothetis, C. J., Millikan, R. E., and Tannir, N. M. (2007). Bayesian and frequentist two-stage treatment strategies based on sequential failure times subject to interval censoring. *Statistics in Medicine* **26,** 4687–4702.

Thompson, W. R. (1933). On the Likelihood that One Unknown Probability Exceeds Another in View of the Evidence of Two Samples. *Biometrika* **25,** 285–294.

Valente, M. J., MacKinnon, D. P., and Mazza, G. L. (2020). A Viable Alternative when Propensity Scores Fail: Evaluation of Inverse Propensity Weighting and Sequential G-estimation in a Two-Wave Mediation Model. *Multivariate behavioral research* **55,** 165–187.

Wang, J., Wu, L., and Wahed, A. S. (2022). Adaptive randomization in a two-stage sequential multiple assignment randomized trial. *Biostatistics* **23,** 1182–1199.

Wang, X., Deliu, N., Narita, Y., and Chakraborty, B. (2022). SMART-EXAM: Incorporating Participants' Welfare into Sequential Multiple Assignment Randomized Trials. arXiv:2210.16255 [stat.ME].

Wathen, J. K. and Thall, P. F. (2017). A simulation study of outcome adaptive randomization in multi-arm clinical trials. *Clinical Trials* **14,** 432–440.

Wei, L. J. and Durham, S. (1978). The Randomized Play-the-Winner Rule in Medical Trials. *Journal of the American Statistical Association* **73,** 840–843.

Williamson, S. F. and Villar, S. S. (2020). A response-adaptive randomization procedure for multi-armed clinical trials with normally distributed outcomes. *Biometrics* **76,** 197–209.

Zhang, L. and Rosenberger, W. F. (2007). Response-adaptive randomization for survival trials: the parametric approach. *Journal of the Royal Statistical Society: Series C*

(Applied Statistics) **56,** 153–165.

Received Month 2023. *Revised Month* Year. *Accepted Month* Year.

Figure 1. A two-stage SMART design with three potential treatment choices A_1, A_2, A_3.

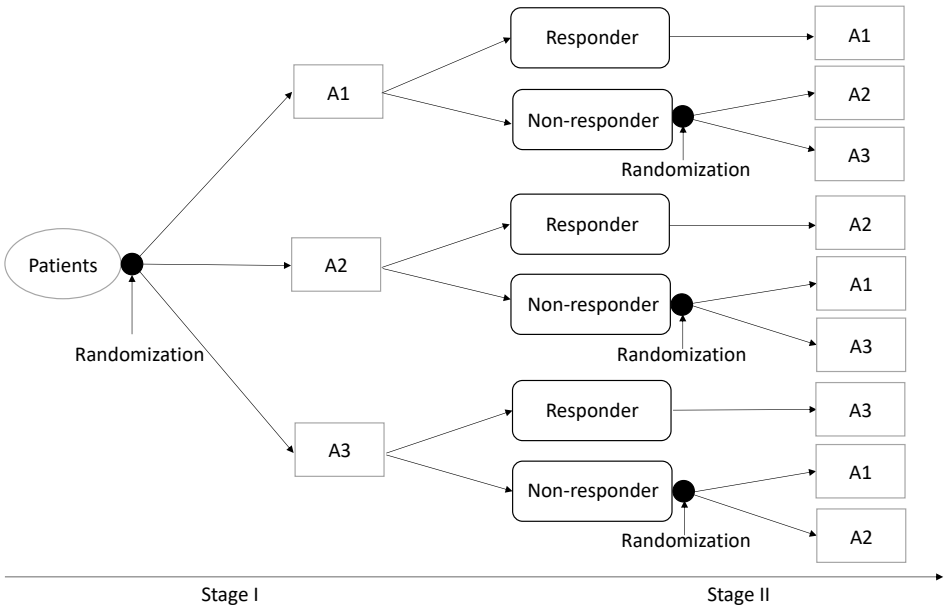

Figure 2. Coverage probability for estimating response rates under various DTRS for SMART, RA-SMART, and GO-SMART designs when $n = 600$ under scenarios S0 and S1. The black dashed lines represent 0.95 expected coverage probability.

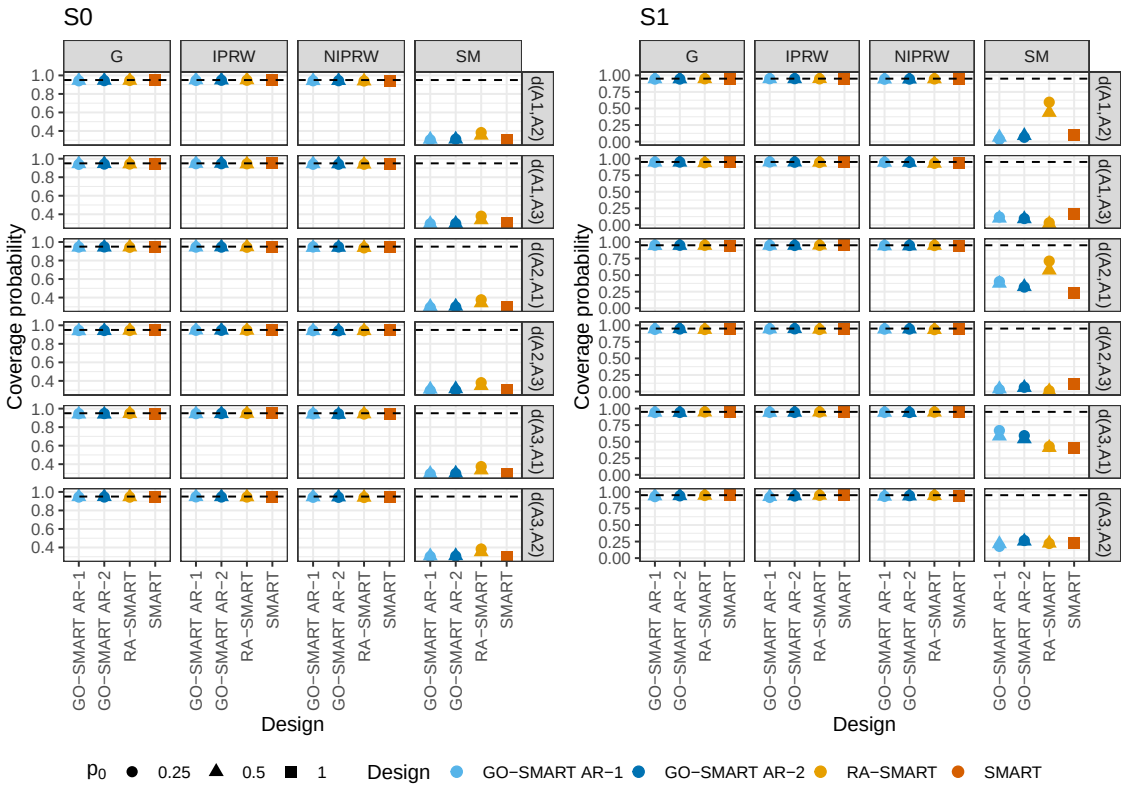

Figure 3. Type I error of identifying the designated optimal DTR under SMART, RA-SMART, and GO-SMART designs for sample sizes $n = 300$ and $n = 600$. Black horizontal line equals 1/6, the expected type I error under the null, where all DTRs have the same response rate.

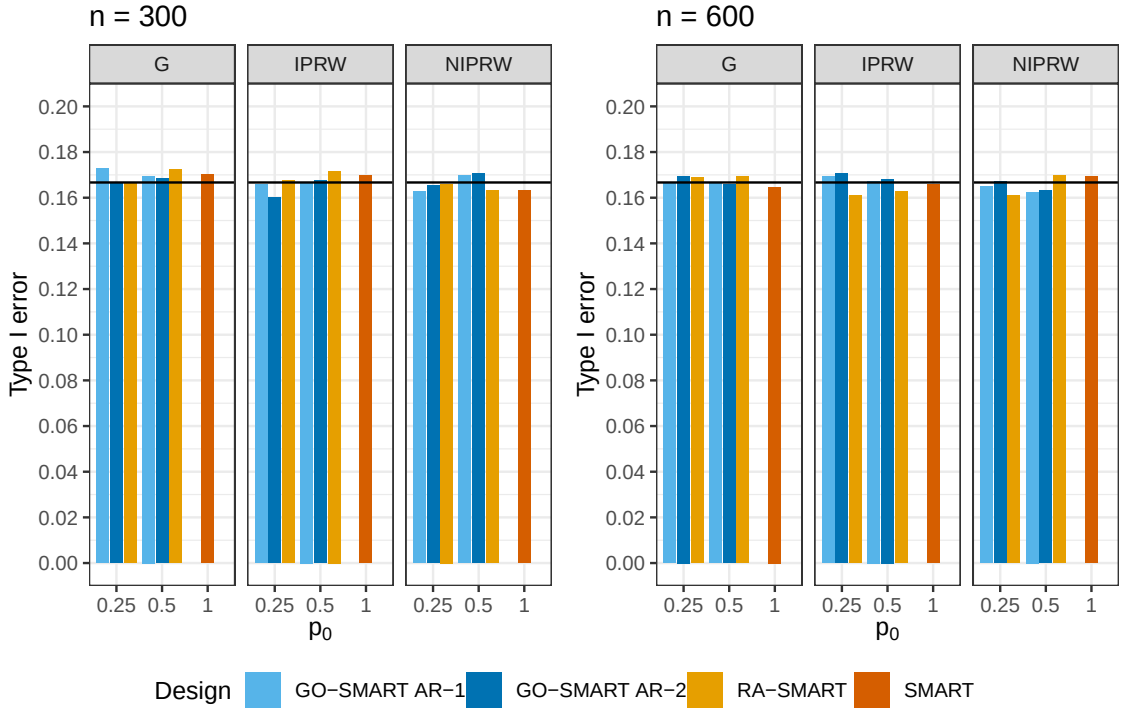

Figure 4. Power of identifying the designated optimal DTR through various designs: SMART, RA-SMART, and GO-SMART (AR-1 and AR-2) for $n = 600$ under scenarios S1-S6 described in Table 1.

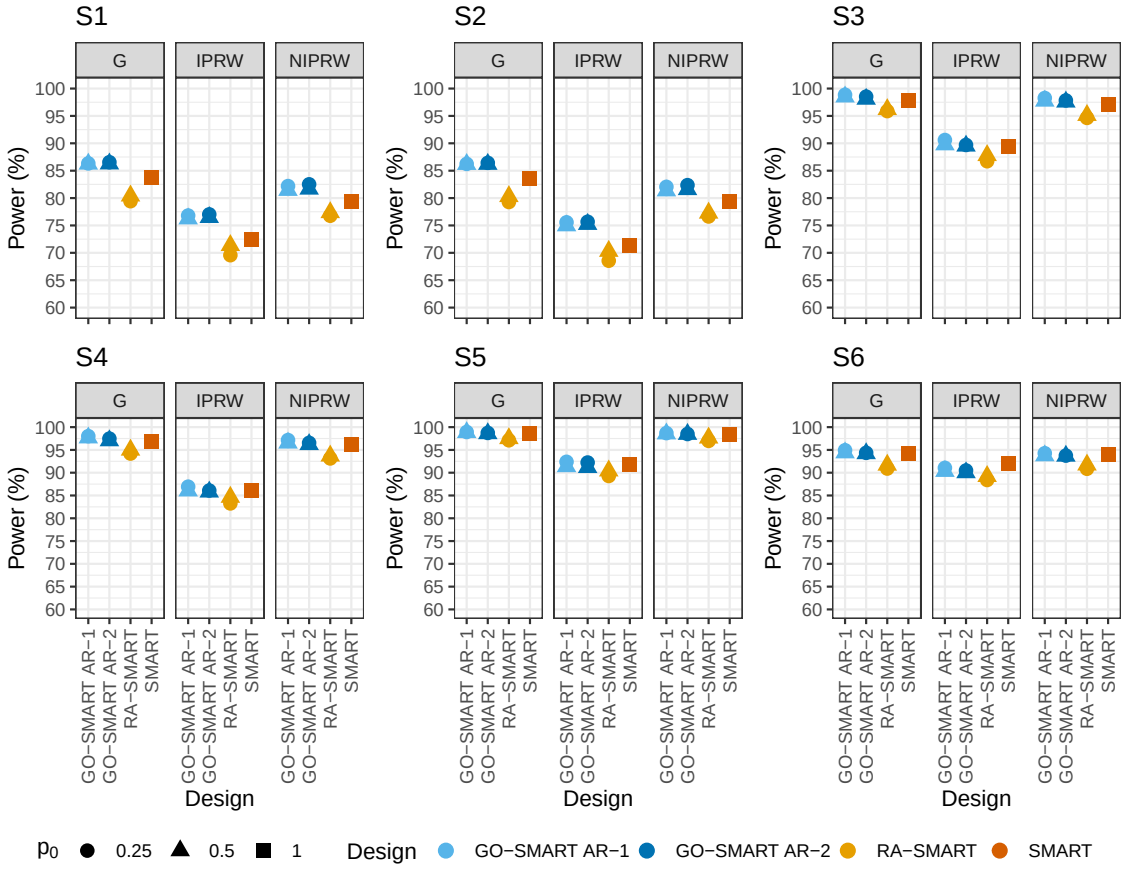

Figure 5. Number of patients treated with the true optimal DTR and the worst DTR in SMART, RA-SMART, and GO-SMART designs with $n = 600$ for scenario S2, S3, S5, and S7.

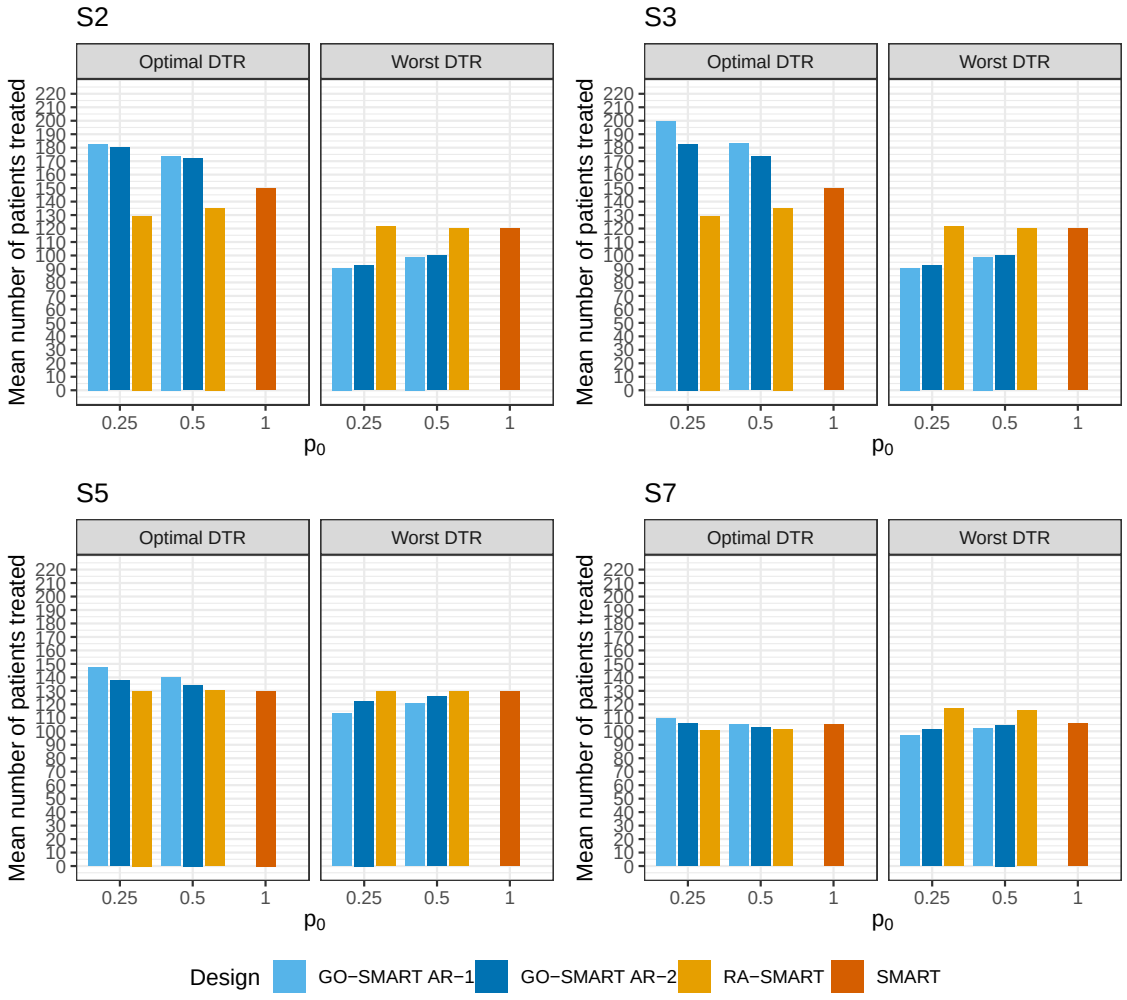

Table 1

Eight scenarios are studied in the simulations reflecting various features of the two-stage three-treatment DTR framework. First three rows provides the first stage response rates, next six rows in the first block of the table provides the assumed second-stage response rates under six potential treatment sequences, and the last six rows provide the overall response rates under the six DTRs.

Parameter	S0 (H_0)	S1	S2	S3	S4	S5	S6	S7
$\pi_1^{A_1}$	0.3	**0.5***	**0.5**	**0.5**	**0.5**	0.3	0.05	0.05
$\pi_1^{A_2}$	0.3	0.35	0.35	0.35	0.35	0.3	**0.07**	**0.07**
$\pi_1^{A_3}$	0.3	0.2	0.2	0.2	0.2	0.3	0.06	0.06
$\pi_2^{A_1A_2}$	0.35	0.3	0.3	0.06	0.06	**0.6**	0.06	0.1
$\pi_2^{A_1A_3}$	0.35	**0.4**	**0.4**	**0.4**	0.4	0.15	**0.26**	**0.26**
$\pi_2^{A_2A_1}$	0.35	0.35	0.35	0.35	0.35	0.2	0.1	0.1
$\pi_2^{A_2A_3}$	0.35	0.2	0.2	0.2	0.2	0.25	0.09	0.09
$\pi_2^{A_3A_1}$	0.35	0.25	0.25	0.25	0.25	0.3	0.15	0.15
$\pi_2^{A_3A_2}$	0.35	0.1	0.4	0.4	**0.45**	0.4	0.08	0.08
$\pi(A_1, A_2)$	0.545	*0.65***	*0.65*	0.53	0.53	**0.72**	0.107	0.145
$\pi(A_1, A_3)$	0.545	**0.7**	**0.7**	**0.7**	**0.7**	0.405	**0.297**	**0.297**
$\pi(A_2, A_1)$	0.545	0.578	0.578	*0.578*	*0.578*	0.44	0.163	0.163
$\pi(A_2, A_3)$	0.545	0.48	0.48	0.48	0.48	0.475	0.154	0.154
$\pi(A_3, A_1)$	0.545	0.4	0.4	0.4	0.4	0.51	*0.201*	*0.201*
$\pi(A_3, A_2)$	0.545	0.28***	0.52	0.52	0.56	*0.58*	0.135	0.135

* **Bold** indicates the response rate for the optimal DTR and optimal stage I/II treatment;
** *Italicized* indicates the response rate for the second best DTR;
*** Underlined indicates the response rate for the worst DTR.