Check for updates

Received: 25 April 2022 | Accepted: 16 February 2023

DOI: 10.1111/biom.13849

BIOMETRIC METHODOLOGY

Biometrics A JOURNAL OF THE INTERNATIONAL BIOMETRIC SOCIETY WILEY

Covariate-adjusted response-adaptive designs based on semiparametric approaches

Hai Zhu[1] | **Hongjian Zhu**[2]

[1]SystImmune Inc., Redmond, Washington, USA

[2]Statistical Innovation Group, AbbVie Inc, Virtual Office, Sugar Land, Texas, USA

Correspondence
Hongjian Zhu, Statistical Innovation Group, AbbVie Inc, Virtual Office, Sugar Land, Texas, USA.
Email: Hongjian.Zhu@abbvie.com

Funding information
National Science Foundation, Grant/Award Number: 2014951; PRIME award from the University of Texas Health Science Center at Houston

Abstract

We consider theoretical and practical issues for innovatively using a large number of covariates in clinical trials to achieve various design objectives without model misspecification. Specifically, we propose a new family of semiparametric covariate-adjusted response-adaptive randomization (CARA) designs and we use the target maximum likelihood estimation (TMLE) to analyze the correlated data from CARA designs. Our approach can flexibly achieve multiple objectives and correctly incorporate the effect of a large number of covariates on the responses without model misspecification. We also obtain the consistency and asymptotic normality of the target parameters, allocation probabilities, and allocation proportions. Numerical studies demonstrate that our approach has advantages over existing approaches, even when the data-generating distribution is complicated.

KEYWORDS
CARA, clinical trials, efficiency, ethics, semiparametric

1 | INTRODUCTION

With the rapidly increasing ability to collect and store large quantities of data, such as the Electronic Health Record (EHR) dataset, it is of great importance to study how to incorporate a large number of covariates into both the design and the analysis of clinical trials without compromising the integrity and validity of the process. Both the U.S. Food and Drug Administration (FDA) (FDA, 2021) and the European Medicines Agency (EMA, 2015) have paid lots of attention to this problem and issued guidelines on using covariates in clinical trials. The covariate-adjusted response-adaptive randomization (CARA) design (Hu & Rosenberger, 2006; Rosenberger et al., 2001; Villar & Rosenberger, 2018) provides an innovative way to use covariates in clinical trials, but it suffers from several theoretical and practical problems that have hindered its development and application. We propose a new family of CARA designs with innovative semiparametric data analytical approaches to provide a unified solution to these issues and satisfy the new demands of the data-rich world.

Hu and Rosenberger (2006) classified adaptive randomization procedures into four categories: restricted randomization (RR), covariate adaptive randomization (CAR), response adaptive randomization (RAR), and CARA. CARA can combine the advantages of the other three types of adaptive randomization designs by adjusting the allocation probability for each new patient based on his or her baseline covariates and the full history of the previous patients' treatment assignments, responses, and baseline covariates. Conceptually, CARA designs have at least the following four advantages, but each comes with theoretical or practical difficulties. First, CARA designs can use all the observed data for randomization, but deriving theoretical results is challenging because the observed treatment assignments, covariates, and responses are related to each other in a complicated manner. Second, CARA designs can seize opportunities in the big-data era by making use of a large number of available covariates, but it is hard to avoid

 © 2023 The International Biometric Society.

model misspecification. Third, compared to traditional designs such as complete randomization (CR), RR, and CAR, CARA designs can pursue different objectives, but developing a family of CARA designs that can achieve multiple and often competing objectives in a single clinical trial is a delicate task that often further complicates the theoretical investigation. Finally, CARA designs can adjust the allocation probability for each patient because more data are available, but it is challenging to make the design easy to implement. These problems have prevented the rapid development of CARA despite its numerous advantages.

Existing work on CARA has tried to solve some of the above problems, but there is no unified solution addressing all the issues. Atkinson and Biswas (2005a, 2005b) considered different objectives, but they did not derive the asymptotic properties, which makes it difficult to assess the validity of statistical inference based on their designs. Hu et al. (2015) also considered multiple objectives with theoretical results, but their asymptotic results are valid only when two covariates are involved in a highly restrictive parameter model. Zhang et al. (2007) and Zhu (2015) focused on asymptotic results, but they established the validity of statistical inference following CARA designs only under correctly specified models. Chambaz and van der Laan (2014) solved the problem of model misspecification, but their designs only considered one allocation proportion. Moreover, they used TMLE (van der Laan & Rubin, 2006; van der Laan & Rose, 2011) to update the allocation probability; this is not easy for clinical trialists to implement and may have an excessive computational burden. In addition, Zheng (2014) and Chambaz et al. (2014) greatly advanced this field by using Least Absolute Shrinkage and Selection Operator (LASSO) regression to estimate the conditional outcome given treatment assignments and baseline covariates, and updating the allocation probability after each block of patients, but again, they also focused only on Neyman allocation. In short, no approach in the literature can solve all the above problems and take full advantage of CARA.

In this article, we propose a new semiparametric family of CARA designs flexibly addressing both efficiency and ethics to meet various research needs. In particular, we develop sequential parameter estimators based on martingale estimating equations to update the allocation probability for the next patient when implementing the design. Then we use the TMLE to estimate the target parameters of the clinical trial to incorporate a large number of covariates and avoid model misspecification. TMLE is a semiparametric approach that obtains an unbiased double-robust substitution estimator for a target parameter of interest instead of the whole distribution. We overcome the theoretical difficulties caused by the mechanism of CARA designs and the advanced TMLE technique to provide rigorous proofs of the asymptotic properties of the proposed design and analysis approaches in the semiparametric setting. We also present numerical studies to demonstrate the advantages of our CARA design over other commonly used randomizations. In summary, the main contributions of this article are the following:

(1) Our design offers flexibility in combining different (optimal) design objectives and meeting different research needs.
(2) We update the allocation probability via sequential parameter estimators based on martingale estimating equations instead of other semiparametric estimators such as TMLE. This significantly reduces the computational intensity, simplifies the theoretical investigation, ensures theoretical validity, and makes the design relatively easier to be implemented.
(3) We use TMLE to analyze the correlated data from CARA designs in order to correctly incorporate the effect of a large number of covariates on the responses without model misspecification and to tackle the issue of restrictive modeling assumptions.
(4) We carefully select the working model and fluctuation model in TMLE to achieve good interpretability.
(5) Our design combines all the above advantages and addresses the four problems mentioned above in a single clinical trial. More importantly, we overcome the theoretical difficulties caused by each of these elements and their combinations and provide a sound theoretical foundation for the proposed design.

The rest of the article is organized as follows. In Section 2, we present our new family of CARA designs with the semiparametric sequential parameter estimators based on martingale estimating equations. We then discuss the asymptotic properties of the design in terms of parameter estimators, allocation probabilities, and allocation proportions. In Section 3, we define the target parameter in clinical trials and provide the asymptotic properties of the TMLE for statistical inference. In Section 4, we present numerical studies for three scenarios to demonstrate that our methodology is preferable to existing methods. In Section 5, we study the performance of our design by redesigning a clinical trial using real data. Section 6 provides concluding remarks. We provide detailed steps to calculate the TMLE based on our design and its proof, as well as proofs of all the other theorems in the Appendix.

2 | CARA DESIGNS BASED ON SEMIPARAMETRIC ESTIMATORS

2.1 | Data structure

Consider a clinical trial with K experimental arms and one control arm. Assume that n patients sequentially

15410420, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/biom.13849 by Readcube (Labtiva Inc.), Wiley Online Library on [30/08/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

enter the trial. Let $A_i \in \mathbb{A} = \{0, \dots, K\}$, $i = 1, \dots, n$, denote the treatment assignment of the ith patient. Let Y_i be the binary ($Y_i \in \{0, 1\}$) or continuous ($Y_i \in \mathbb{R}$) one-dimensional primary endpoint of the ith patient. Let $\boldsymbol{W}_i = (W_{i,1}, \dots, W_{i,n_W})$ represent the ith patient's baseline characteristics. Assume we are interested in a biomarker/subgroup indicator V_i that is a function of the baseline characteristics: $V_i = f_V(\boldsymbol{W}_i) \in \mathbb{V} = \{v_1, \dots, v_q\}$. The biomarker V might be chosen based on the previous translational research and represent a comprehensive understanding of the impact of the baseline characteristics on the treatment effects. Let $X = (Y(a), a \in \mathbb{A}, \boldsymbol{W}) \sim P_0$ be the full data structure, where $Y(a)$ denotes the realization of Y under $A = a$ and P_0 represents the true data-generating probability distribution. According to the notation of counterfactuals (van der Laan & Rose, 2011), the full data structure X contains all possible realizations of Y under different treatments $a \in \mathbb{A}$. The observed data for the ith patient are a censored version of X_i, denoted as $O_i = (Y_i(A_i), A_i, \boldsymbol{W}_i)$.

We use $G_i(\cdot)$ to denote the censoring mechanism for the ith patient. For CARA designs, $G_i(\cdot)$ is the conditional probability of treatment assignment A_i given $(X_1, \dots, X_i)$. Based on the coarsening-at-random (Heitjan & Rubin, 1991) assumption that the censoring mechanism depends only on the observed data, we assume that G_i is conditioned on the historical observed data $\boldsymbol{O}_{i-1} = (O_1, \dots, O_{i-1})$ and the baseline characteristics $\boldsymbol{W}_i$ through the subgroup indicator V_i. Mathematically, $G_i(a, v) = \text{pr}(A_i = a \mid V_i = f_V(\boldsymbol{W}_i) = v, \boldsymbol{O}_{i-1})$. For convenience, we write G_i for the conditional probability and omit the $\boldsymbol{O}_{i-1}$. The likelihood of the ith observed data O_i is factorized as $P_0(O_i) = Q_0(Y_i, A_i, \boldsymbol{W}_i)G_i(A_i, V_i)$, where $Q_0(Y_i, A_i, \boldsymbol{W}_i) = P_0(Y_i \mid A_i, \boldsymbol{W}_i)P_0(\boldsymbol{W}_i)$ is a parameter of the full-data distribution P_0. We use the notation Q_0G_i as a subscript to denote the data-generating mechanism for the ith observed data. Let $N_{a,v}(n)/n$ be the proportion of n patients assigned to treatment a in subgroup v.

2.2 | A family of CARA designs

Clinical trials may have a variety of design objectives, such as assigning more patients to the superior treatment group with a higher efficiency of detecting the treatment effects. In addition, patients with different baseline characteristics may respond to the treatments differently. We propose a new family of CARA designs that takes into account efficiency and ethics simultaneously, acknowledges the heterogeneity of the patients, and avoids unnecessary model assumptions by using semiparametric estimators. We will now introduce two important elements of the proposed design: sequential parameter estimators and measurements of efficiency and ethics. We will then present the new family of designs.

Sequential parameter estimators: We first define our design parameter vector $\boldsymbol{\theta}_0 = \{\boldsymbol{\theta}_0^{a,v}, a \in \mathbb{A}, v \in \mathbb{V}\}$, where $\boldsymbol{\theta}_0^{a,v} = (\theta_{0,1}^{a,v}, \theta_{0,2}^{a,v})$ for all pairs of (a, v), with

$$\theta_{0,1}^{a,v} = E_{P_0}(Y \mid A = a, V = v),\ \theta_{0,2}^{a,v} = E_{P_0}(Y^2 \mid A = a, V = v). \quad (1)$$

These parameters are the conditional first and second moments of the observed Y. We define an extra parameter $\sigma_0^{a,v} = \theta_{0,2}^{a,v} - (\theta_{0,1}^{a,v})^2$, the conditional variance of Y given $(A, V) = (a, v)$ under the true probability distribution P_0. Note that when Y is binary, $\theta_{0,2}^{a,v}$ is redundant and $\sigma_0^{a,v} = \theta_{0,1}^{a,v}(1 - \theta_{0,1}^{a,v})$. $\theta_{0,1}^{a,v}$ can be treated as the ethical parameter as it usually presents the efficacy or toxicity of the treatment. $\theta_{0,2}^{a,v}$ and $\sigma_0^{a,v}$ are used to optimize the efficiency of the design such as in Neyman allocation.

Next, we discuss how to obtain an appropriate semiparametric estimator $\widehat{\boldsymbol{\theta}}_n = \{\widehat{\boldsymbol{\theta}}_n^{a,v}, a \in \mathbb{A}, v \in \mathbb{V}\}$ based on the accumulated data $\boldsymbol{O}_n$, where $\widehat{\boldsymbol{\theta}}_n^{a,v} = (\widehat{\theta}_{n1}^{a,v}, \widehat{\theta}_{n2}^{a,v})$ and the subscript n refers to the sample size. For an arbitrary parameter $\theta \in \Theta$, we first define two estimating functions $M_1^{a,v}(\theta)(\boldsymbol{O}_i)$ and $M_2^{a,v}(\theta)(\boldsymbol{O}_i)$ for all pairs of (a, v) as follows:

$$M_1^{a,v}(\theta)(\boldsymbol{O}_i) = \frac{I_i(a, v)(Y_i - \theta)}{G_i(a, v)(\boldsymbol{O}_{i-1})},$$

$$M_2^{a,v}(\theta)(\boldsymbol{O}_i) = \frac{I_i(a, v)\left(Y_i^2 - \theta\right)}{G_i(a, v)(\boldsymbol{O}_{i-1})}, \quad (2)$$

where $I_i(a, v)$ is shorthand for the indicator function $I(A_i = a, V_i = v)$. We can redefine $\theta_{0,1}^{a,v}$ and $\theta_{0,2}^{a,v}$ as the true parameters of the martingale estimating functions (2) such that $E_{Q_0G_i}\{M_1^{a,v}(\theta_{0,1}^{a,v})(\boldsymbol{O}_i)\} = 0$ and $E_{Q_0G_i}\{M_2^{a,v}(\theta_{0,2}^{a,v})(\boldsymbol{O}_i)\} = 0$ for all (a, v). The estimators $\widehat{\theta}_n$ are the solutions of $\sum_{i=1}^n\{M_1^{a,v}(\theta_{n,l}^{a,v})(\boldsymbol{O}_i)\} = 0$, $l = 1, 2$, with the closed form

$$\widehat{\theta}_{n,1}^{a,v} = \frac{\sum_{i=1}^n \frac{I_i(a,v)}{G_i(a,v)} Y_i}{\sum_{i=1}^n \frac{I_i(a,v)}{G_i(a,v)}}, \quad \widehat{\theta}_{n,2}^{a,v} = \frac{\sum_{i=1}^n \frac{I_i(a,v)}{G_i(a,v)} Y_i^2}{\sum_{i=1}^n \frac{I_i(a,v)}{G_i(a,v)}}. \quad (3)$$

The solutions are the weighted sum of corresponding responses. The responses are weighted by the reciprocal of the probability of receiving the treatment that they actually received based on their covariates, which is the same concept of the inverse probability of treatment weighting (IPTW). As a result, the practical meaning can also be understood similarly to IPTW.

Compared to other methods such as directly using TMLE (van der Laan & Rose, 2011; van der Laan &

15410420, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/biom.13849 by Readcube (Labtiva Inc.), Wiley Online Library on [30/08/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Rubin, 2006) for the sequential parameter estimators, the proposed estimators (3) have theoretical and practical advantages. First, our estimators are easy for practitioners to understand and calculate, which will help to promote the proposed design. Second, the martingale estimating equations defined by the estimating functions $M_1^{a,v}(\theta)(\boldsymbol{O}_i)$ and $M_2^{a,v}(\theta)(\boldsymbol{O}_i)$ become independent of the CARA allocation probability G_i. Finally, these estimators obtained from martingale estimating equations will simplify the theoretical investigation.

Measurements of efficiency and ethics: We define $d(a, v, \theta_0)$ and $e(a, v, \theta_0), a \in \mathbb{A}, v \in \mathbb{V}$, as finite one-dimensional measurements of efficiency and ethics for treatment a in subgroup v, respectively, where $d(\cdot,\cdot,\cdot)$ and $e(\cdot,\cdot,\cdot)$ are given positive functions such that for $a \in \mathbb{A}, v \in \mathbb{V}$, $d(a, v, \cdot) : \mathbb{R}^{2(K+1)q+2} \to \mathbb{R}^+$ and $e(a, v, \cdot) : \mathbb{R}^{2(K+1)q+2} \to \mathbb{R}^+$. Efficiency means the power of detecting treatment differences. Ethics means fewer patients assigned to inferior treatments. In this paper, we focus on the individual patient's clinical response performance in a trial, not the so-called population ethics (Rosenberger & Lachin, 2015). For example, we can use the reciprocal of the failure rate as an ethics measurement and the variance of the response as an efficiency measurement. The choice of these measurements is determined by the design objectives and will lead to different target allocation proportions. It is natural that these two measurements are functions of the unknown parameters θ_0 and depend on the treatment assignments. We also allow these measurements to vary with the subgroups V, which is consistent with the idea of precision medicine. Specific examples of the measurements of efficiency and ethics are given in Section 4.

New family of CARA designs: We propose a new general family of CARA designs that assign the ith subject in subgroup $V_i = v$ to treatment $A = a, a = 0, \ldots, K$, with probability $G_i(a, v) = \text{pr}(A_i = a \mid V_i = v) = h_{av}(\widehat{\theta}_{i-1}, e(l, v, \widehat{\theta}_{i-1}), d(l, v, \widehat{\theta}_{i-1}), l = 0, \ldots, K)$, where the $h_{av}(\cdot)$ are $\mathbb{R}^{2(K+1)(q+1)} \to \mathbb{R}(0, 1)$ functions. For each new patient, we sequentially update the parameter estimators $\widehat{\theta}_{i-1}$, the values of the measurements of efficiency and ethics, and the allocation probability functions to achieve different aims. Logistically, there is no difficulty in sequentially updating these values after each group of patients. We emphasize that the sequential parameter estimators $\widehat{\theta}_{i-1}$ must be obtained without parameter model assumptions, and the final data analysis is performed in the semiparametric setting to correctly incorporate the effect of a large number of covariates on the responses without model misspecification. There are some points to emphasize about the proposed design.

(1) Compared to Hu et al. (2015), our design sequentially estimates the parameters and updates the allocation probability based on semiparametric approaches, avoiding unnecessary model assumptions. Our analysis approaches to be discussed below are also semiparametric, addressing the difficulty of dealing with many covariates and complex parametric models in Hu et al. (2015).
(2) It is well known that the objectives relating to efficiency and ethics often conflict with each other in a clinical trial. Therefore, it is of importance to balance these two objectives, which necessitate the proposed design. Such design may not always target certain mathematically optimal objective, but offer clinical trialists the flexibility to adjust the weight of the two objectives. For example, one can ignore the ethics and consider only the measurement of efficiency, similar to the D-optimality and D_A-optimality proposed by Atkinson (1982).
(3) Under certain conditions, our design can also target the optimal allocation proportion such as Neyman allocation (see Chambaz & van der Laan, 2014 as a special case of our general framework). Some other A-optimality and D-optimality designs such as Fackle-Fornius and Nyquist (2015) can also be directly implemented in our framework.
(4) As a general framework, in addition to the estimated measurements of efficiency and ethics, we added $\widehat{\theta}_{i-1}$ in the allocation probability function to increase the flexibility of future specific designs.

2.3 | Asymptotic results for the CARA designs

We have proposed a general family of CARA designs based on semiparametric approaches. We now derive the asymptotic results for a specific family of designs in the above general family. Our design assigns the ith subject in subgroup $V_i = v$ to treatment a with probability

$$G_i(a, v) = pr(A_i = a \mid V_i = v, \widehat{\theta}_{i-1}) = \frac{e(a, v, \widehat{\theta}_{i-1})^{\gamma_1} d(a, v, \widehat{\theta}_{i-1})^{\gamma_2}}{\sum_{k \in \mathbb{A}} e(k, v, \widehat{\theta}_{i-1})^{\gamma_1} d(k, v, \widehat{\theta}_{i-1})^{\gamma_2}}, \quad (4)$$

where $(\gamma_1, \gamma_2) \in [0, +\infty)^2$ are tuning parameters that determine the balance between ethics and efficiency. The ratio form makes the allocation function a legitimate probability and guarantees the scale-invariant property of the efficiency and ethics measurements. This family of CARA designs has very few restrictions on the efficiency and

15410420, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/biom.13849 by Readcube (Labtiva Inc.), Wiley Online Library on [30/08/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

ethics component, so it can satisfy diverse practical needs in clinical trials.

We introduce the following conditions for the asymptotic results:

Condition 1. $\sup_{a\in\mathbb{A},v\in\mathbb{V}} E_0(Y^2 \mid A=a, V=v) < \infty$.

Condition 2. G_i is bounded in $[g_L, g_U]$, where $0 < g_L < g_U < 1$.

Condition 3. For any fixed pair $(a,v) \in \mathbb{A}\times\mathbb{V}$, $d(a,v,\theta)$ and $e(a,v,\theta)$ are both continuous in terms of θ.

Condition 2 indicates that the CARA designs should avoid assigning a probability of zero or one to any treatment when allocating patients.

Theorem 1. *Under Conditions (1)–(3)*

$$\widehat{\theta}_n \xrightarrow{a.s.} \theta_0\,, \quad G_n(a,v) \xrightarrow{a.s.} G_0(a,v)$$
$$N_{a,v}(n)/n \xrightarrow{a.s.} p_0(v)G_0(a,v), \tag{5}$$

for all (a,v) as $n\to\infty$, where $p_0(v)=P_0(V=v)$ is the marginal probability of $V=v$, and

$$G_0(a,v) = pr(A=a \mid V=v) = \frac{e(a,v,\theta_0)^{\gamma_1} d(a,v,\theta_0)^{\gamma_2}}{\sum_{k\in\mathbb{A}} e(k,v,\theta_0)^{\gamma_1} d(k,v,\theta_0)^{\gamma_2}}.$$

Theorem 1 shows the consistency of $\widehat{\theta}_n$, G_n, and $N_{a,v}(n)/n$. To study the asymptotic normality of the semiparametric estimator and the allocation proportions of the CARA design, we introduce the following conditions:

Condition 4. $\sup_{a\in\mathbb{A},v\in\mathbb{V}} E_0(Y^4 \mid A=a, V=v) < \infty$.

Condition 5. For any fixed pair $(a,v) \in \mathbb{A}\times\mathbb{V}$, $d(a,v,\theta)$ and $e(a,v,\theta)$ are both differentiable in terms of θ.

Theorem 2. *Under Conditions (1)–(5)*

$$\sqrt{n}(\widehat{\theta}_n - \theta_0) \xrightarrow{D} N(0, \Sigma_0^{CARA}), \tag{6}$$

$$\sqrt{n}\left(\frac{N_{a,v}(n)}{n} - p_0(v)G_0(a,v)\right) \xrightarrow{D} N$$
$$\left(0, p_0(v)G_0(a,v) - p_0(v)^2 G_0(a,v)^2\right), \tag{7}$$

where $\Sigma_0^{CARA} = \mathrm{diag}\{\Sigma_0^{a,v}, (a,v)\in\mathbb{A}\times\mathbb{V}\}$ is a block diagonal matrix. Each element matrix of Σ_0^{CARA} has the form

$$\Sigma_0^{a,v} = \frac{1}{p_0(v)G_0(a,v)} \begin{pmatrix} \theta_{0,2}^{a,v} - (\theta_{0,1}^{a,v})^2 & \theta_{0,3}^{a,v} - (\theta_{0,1}^{a,v})\theta_{0,2}^{a,v} \\ \theta_{0,3}^{a,v} - (\theta_{0,1}^{a,v})\theta_{0,2}^{a,v} & \theta_{0,3}^{a,v} - (\theta_{0,2}^{a,v})^2 \end{pmatrix},$$

where $\theta_{0,3}^{a,v}$ and $\theta_{0,4}^{a,v}$ are defined as the third and fourth conditional moments of Y given $(A,V)=(a,v)$ under P_0.

Theorems 1 and 2 reveal characteristics of the randomization procedure. Next, we discuss the data analysis at the end of the clinical trial with well-defined target parameters.

3 | ANALYSIS OF CARA DESIGNS BASED ON SEMIPARAMETRIC APPROACHES

3.1 | Target parameters of TMLE

Suppose that the response Y is rescaled into [0,1], and it can be either binary or continuous. The rescaling can be implemented through any continuous one-to-one mapping. This rescaled response can be easily transformed back and has no impact on our calculations and inference. By doing this, one can adopt the quasi-log-likelihood loss function that is valid not only for binary Y but also for $Y\in[0,1]$ (Gruber & van der Laan, 2010). This widely applied TMLE technique is robust and retains global constraints (van der Laan & Rubin, 2006). The target parameter can be defined as a $(K+1)$-dimensional parameter $\boldsymbol{\psi}_0 = \Psi(P_0) = (\psi_{0,0},\ldots,\psi_{0,K})$, where Ψ is the target mapping $\Psi : \mathcal{M} \to \mathbb{R}^{k+1}$, and

$$\psi_{0,j} = E_{P_0}(Y \mid A=j) = E_{P_0}\left\{E_{P_0}(Y \mid A=j, \boldsymbol{W})\right\}, j=0,\ldots,K. \tag{8}$$

We emphasize here that our approaches consider the effect of the covariates on the responses without unnecessary model assumptions. At the end of the trial, we perform the following hypothesis test:

$$H_0 : C\boldsymbol{\psi}_0 = \mathbf{0} \text{ versus } H_1 : C\boldsymbol{\psi}_0 \neq \mathbf{0}, \tag{9}$$

where C is a $K\times(K+1)$ contrast matrix representing the additive treatment differences between the K experimental arms and the control arm:

$$C_{K\times(K+1)} = \begin{bmatrix} -1 & 1 & & \ldots \\ -1 & & 1 & \\ \vdots & & & \ddots \\ -1 & & & 1 \end{bmatrix}.$$

Other target parameters such as the relative risk and the odds ratio in scenarios with binary outcomes can also be studied. To save space and keep a coherent main text, we leave the detailed calculation of TMLE ($\widehat{\boldsymbol{\psi}}_n^{TMLE}$), a detailed and lengthy theoretical derivation and explanation for the calculation procedure with additional important theorems in the Appendix.

15410420, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/biom.13849 by Readcube (Labtiva Inc.), Wiley Online Library on [30/08/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

3.2 | Statistical inference for the TMLE

In the following theorem, we construct the normality of $\widehat{\boldsymbol{\psi}}_n^{TMLE} = (\widehat{\psi}_{n,0}^{TMLE}, \ldots, \widehat{\psi}_{n,K}^{TMLE})$ directly through the martingale estimating equation (S2.6) in the Appendix. We need another condition.

Condition 6. $E_0(W_j^2) < \infty$ for all $j \in \{1, \ldots, n_W\}$.

Theorem 3. *Under Conditions (1)–(6), we have*

$$\sqrt{n}\left(\widehat{\boldsymbol{\psi}}_n^{TMLE} - \boldsymbol{\psi}_0\right) \xrightarrow{D} N(0, \Sigma_0^{TMLE}) \text{ as } n \to \infty, \quad (10)$$

where $\Sigma_0^{TMLE} = [\sigma_0^{TMLE}(j,k)]$, $j,k = 0, \ldots, K$, *is a* $(K+1) \times (K+1)$ *covariance matrix with*

$$\sigma_0^{TMLE}(j,k) = E_{Q_0 G_0}\left\{\mathrm{IC}_j(Q_{G_0}(\theta_0, \boldsymbol{\beta}_0, \epsilon_0), G_0) \mathrm{IC}_k(Q_{G_0}(\theta_0, \boldsymbol{\beta}_0, \epsilon_0), G_0)\right\},$$

and the influence curve $\mathrm{IC}_j(Q_{G_0}(\theta_0, \boldsymbol{\beta}_0, \epsilon_0), G_0)$, $j = 0, \ldots, K$, *is defined in (S2.4) in the Supporting Information.*

Then $\sigma_0^{TMLE}(j,k)$ can be consistently estimated by

$$\widehat{\sigma}_n^{TMLE}(j,k) = \frac{1}{n}\sum_{i=1}^{n}\left\{\frac{G_n^\star(A_i, V_i)}{G_i(A_i, V_i)}\right\}^2 \left\{\mathrm{IC}_j(Q_n^\star, G_n^\star)(O_i)\, \mathrm{IC}_k(Q_n^\star, G_n^\star)(O_i)\right\}, \quad (11)$$

and the hypothesis (9) can be tested using the statistic

$$T^\star = \left(C\widehat{\boldsymbol{\psi}}_n^{TMLE}\right)^T \left(\frac{1}{n} C \widehat{\boldsymbol{\Sigma}}_n^{TMLE} C^T\right)^{-1} \left(C\widehat{\boldsymbol{\psi}}_n^{TMLE}\right). \quad (12)$$

The null hypothesis is rejected at level α if $T^\star > \chi_K^2(1-\alpha)$. We have shown that, without additional model assumptions, our approach leads to consistent estimators with asymptotic normality, so the hypothesis testing can easily be performed. In addition, the double robust nature of TMLE ensures its asymptotic efficiency in the light of semiparametric statistical model efficiency theory (van der Laan & Rose, 2011).

4 | NUMERICAL STUDIES

Having obtained the asymptotic properties of the proposed family of CARA designs, in this section, we numerically evaluate its finite-sample operating characteristics regarding the type I error, power, and ethics properties. We study four designs representing different ethics measurements:

- CARA_1: $e(a,v,\widehat{\theta}_{i-1}) = \widehat{\theta}_{i-1,1}^{a,v}$, $d(a,v,\widehat{\theta}_{i-1}) = \sqrt{\widehat{\theta}_{i-1,2}^{a,v} - (\widehat{\theta}_{i-1,1}^{a,v})^2}$,
- CARA_2: $e(a,v,\widehat{\theta}_{i-1}) = (1 - \widehat{\theta}_{i-1,1}^{a,v})^{-1}$, $d(a,v,\widehat{\theta}_{i-1}) = \sqrt{\widehat{\theta}_{i-1,2}^{a,v} - (\widehat{\theta}_{i-1,1}^{a,v})^2}$,
- CARA_3: $e(a,v,\widehat{\theta}_{i-1}) = \widehat{\theta}_{i-1,1}^{a,v} * (1 - \widehat{\theta}_{i-1,1}^{a,v})^{-1}$, $d(a,v,\widehat{\theta}_{i-1}) = \sqrt{\widehat{\theta}_{i-1,2}^{a,v} - (\widehat{\theta}_{i-1,1}^{a,v})^2}$,
- CARA_4: $e(a,v,\widehat{\theta}_{i-1}) = \Phi(\widehat{\theta}_{i-1,1}^{a,v} - \sum_{k=1}^{n_A} \widehat{\theta}_{i-1,1}^{k,v}/n_A)$, $d(a,v,\widehat{\theta}_{i-1}) = \sqrt{\widehat{\theta}_{i-1,2}^{a,v} - (\widehat{\theta}_{i-1,1}^{a,v})^2}$,

where n_A denotes the number of treatment arms, and $\Phi(\cdot)$ denotes the cumulative distribution function (CDF) of a standard normal distribution. All four ethics measurements return a larger value for the superior treatment arm in terms of additive treatment effect. The efficiency measurement was chosen based on the idea of Neyman allocation. These ethical and efficient measures are either intuitive or have attracted much attention in the field of RAR. In the tables, we use $\mathrm{CARA}_k(\gamma_1, \gamma_2)$ to represent the kth design above with tuning parameters γ_1 and γ_2. We report results for three scenarios: (1) two treatment arms with continuous endpoints, (2) three treatment arms with continuous endpoints, and (3) three treatment arms with binary endpoints. The advantages of our methods can also be seen in other scenarios, but we omit these results to save space.

Our simulations compare the operating characteristics of the proposed CARA design with TMLE and the traditional CR with the t-test (the chi-square test for the three-treatment scenario). When implementing CARA designs, we assign the first 10–20% of the patients to the treatments with stratified permuted block (SPB) randomization. We prespecify the nominal level at $\alpha = 0.05$, and all the results are based on 10,000 simulation replications.

Scenario 1: Two treatments with a continuous endpoint. Consider a clinical trial with two arms and a bounded continuous endpoint $Y \in \mathbb{R}$, and the sample size is $n = 600$. Suppose we have a covariate vector $\boldsymbol{W} = (W_1, \ldots, W_{10})$ and a binary subgroup indicator $V(\boldsymbol{W}) = I(\sum_{p=1}^{10} W_p > 3.38)$, where $I(\cdot)$ is the indicator function, W_1, W_2, W_3 all follow a uniform distribution in [0,1], W_4, W_5, W_6 all follow a Bernoulli distribution with $p = 0.5$, W_7, W_8, W_9, W_{10} all follow a standard normal distribution, and all these covariates are independent. $\boldsymbol{W}$ and $V(\boldsymbol{W})$ will be generated in the same manner in Scenarios 2 and 3. To study the robustness of CARA, we propose the following two distribution models to generate the endpoint Y:

15410420, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/biom.13849 by Readcube (Labtiva Inc.), Wiley Online Library on [30/08/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 Type I error under CR and CARA procedures for Scenario 1: two arms and continuous endpoint.

Allocation ($n = 600$)	Model 1 ($\beta_A = 0$)		Model 2 ($\beta_A = 0$)	
	TMLE	*t*-test	TMLE	*t*-test
CR	5.58	5.18	5.57	5.01
$CARA_1(0,1)$	5.28	5.15	5.73	5.47
$CARA_1(0.5,1)$	5.27	4.87	5.83	5.65
$CARA_1(1,1)$	5.19	4.78	5.70	5.89
$CARA_2(0.5,1)$	5.20	4.99	5.80	5.53
$CARA_2(1,1)$	5.24	4.86	5.82	5.63
$CARA_3(0.5,1)$	5.17	4.77	5.80	5.74
$CARA_3(1,1)$	5.24	4.95	5.85	6.09
$CARA_4(0.5,1)$	5.15	5.06	5.76	5.43
$CARA_4(1,1)$	5.19	5.04	5.74	5.57

M1: $Y \sim N(\mu, \sigma^2)$, Y is truncated if $Y < 0$ or $Y > 12$,

$$\mu = \mu_0 + (1 + \beta_A A)(1 + \beta_V V + \sum_{p=1}^{10} \beta_{W,p} * W_p),$$

$$\sigma = \frac{1 + \beta_A A}{1 + \beta_V V}.$$

M2: $Y \sim \text{Gamma}(a, b)$, Y is truncated if

$$Y > 10, a = 1 + (1 + \beta_A A)(1 + \beta_V V) + \sum_{p=1}^{10}$$

$$\beta_{W,p} * W_p, \quad b = \frac{1 + \beta_A A}{1 + \beta_V V}.$$

For a given A and V, M1 generates a symmetric distribution of Y, whereas M2 generates a skewed distribution of Y. The complexity in the models acknowledges not only the treatment effect and the difference between subgroups but also their interaction effect and within-group heterogeneity due to unmeasured factors. Such models represent the possible complexity of the real world, and our simulation will show that the proposed methods could capture the complexity without model misspecification. We vary the values of β_A to study the properties of the proposed CARA design while fixing $\beta_V = -0.3$, $\beta_W = (-0.52, 0.48, -1.51, -0.39, -0.35, -0.81, 0.16, -0.20, 0.25, -0.32)$ in model M1 and $\beta_V = -0.2$, $\beta_W = (0.015, 0.63, -0.6, 0.012, 0.37, -0.34, 0.084, -0.1, 0.13, -0.17)$ in model M2.

We show results under H_0 in Table 1, and we can see that our CARA design with derived TMLE can control the type I error very well. In addition, in Table 2, we study a revised version of our design: we sequentially update the allocation probability using our proposed design after each group of 10% of the sample size. We can see from Table 2 that we can still control the type I error rate when group-sequentially updating the allocation probability using our design. We report the operating characteristics of our design without and with group-sequential updating of the allocation probability under H_1 in Tables 3 and 4. Both versions of our design lead to a higher power and more patients in the better arm. In particular, when using $CARA_3(1,1)$, we assign around 17% more patients to the superior arm, and we have a higher power than traditional design.

TABLE 2 Type I error under CR and CARA procedures with group-sequential updating of allocation probability for Scenario 1: two arms and continuous endpoint.

Allocation ($n = 600$)	Model 1 ($\beta_A = 0$)		Model 2 ($\beta_A = 0$)	
	TMLE	*t*-test	TMLE	*t*-test
CR	5.58	5.18	5.57	5.01
$CARA_1(0,1)$	5.35	4.93	5.55	5.10
$CARA_1(0.5,1)$	5.34	4.93	5.69	5.14
$CARA_1(1,1)$	5.32	4.95	5.73	5.05
$CARA_2(0.5,1)$	5.22	4.93	5.56	4.98
$CARA_2(1,1)$	5.24	4.83	5.51	5.04
$CARA_3(0.5,1)$	5.23	4.91	5.70	5.24
$CARA_3(1,1)$	5.43	4.87	5.64	5.67
$CARA_4(0.5,1)$	5.42	4.84	5.54	5.06
$CARA_4(1,1)$	5.20	4.88	5.47	4.93

In Table 5, we studied the sequential monitoring of our proposed design with one interim look at information time 0.75. The O'Brien–Fleming-like spending function will be used to control the type I error rate. We compare our design with sequential monitoring of CR with TMLE and *t*-test as data analysis approaches. The type I error rate, power, average sample size (Avg N), and treatment proportions are reported in Table 5, and the comparisons are visualized using a radar plot in Figure 1. We can see that group sequential monitoring can reduce the sample size, TMLE can help incorporate multiple covariates without model specification and increase the power compared to *t*-test, and CARA can achieve ethical objectives by assigning more patients to the better treatment. The ethical effect of CARA can be clearly seen in Table 5. Both TMLE and CARA in our proposed design and group sequential monitoring help achieve different aims.

Scenario 2: Three treatments with continuous endpoint. Consider a clinical trial with three arms ($A = \{0, 1, 2\}$) and a bounded continuous endpoint $Y \in \mathbb{R}$. We propose the following two models to generate Y:

M3: $Y \sim N(\mu, \sigma^2)$, Y is truncated if $Y < 0$ or $Y > 8$,

$$\mu = \mu_0 + \{1 + \beta_{A1} I(A = 1) + \beta_{A2} I(A = 2)\}(1 + \beta_V V)$$

15410420, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/biom.13849 by Readcube (Labtiva Inc.), Wiley Online Library on [30/08/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 3 Power and allocation proportion under CR and CARA procedures for Scenario 1: two arms and continuous endpoint.

Allocation	Model 1 ($\beta_A = 0.42$)		Model 2 ($\beta_A = 0.125$)	
($n = 600$)	Power (%)	Trt Prop (%)	Power (%)	Trt Prop (%)
CR	74.2	50.0, 50.0	82.2	50.0, 50.0
$CARA_1(0.5,1)$	85.1	43.5, 56.5	85.5	44.9, 55.1
$CARA_1(1,1)$	85.0	42.8, 57.2	84.9	43.0, 57.0
$CARA_2(0.5,1)$	85.1	43.6, 56.4	85.5	46.3, 53.7
$CARA_2(1,1)$	85.0	43.0, 57.0	85.4	45.7, 54.3
$CARA_3(0.5,1)$	85.1	42.9, 57.1	85.4	44.3, 55.7
$CARA_3(1,1)$	85.0	41.6, 58.4	84.7	41.9, 58.1
$CARA_4(0.5,1)$	84.8	44.0, 56.0	85.4	46.5, 53.5
$CARA_4(1,1)$	85.1	43.7, 56.3	85.6	46.1, 53.9

TABLE 4 Power and allocation proportion under CR and CARA procedures with group-sequential updating of allocation probability for Scenario 1: two arms and continuous endpoint.

Allocation	Model 1 ($\beta_A = 0.42$)		Model 2 ($\beta_A = 0.125$)	
($n = 600$)	Power (%)	Trt Prop (%)	Power (%)	Trt Prop (%)
CR	74.2	50.0, 50.0	82.2	50.0, 50.0
$CARA_1(0.5,1)$	84.8	43.6, 56.4	85.5	44.9, 55.1
$CARA_1(1,1)$	84.8	42.8, 57.2	85.4	43.1, 56.9
$CARA_2(0.5,1)$	84.8	43.6, 56.4	85.4	46.3, 53.7
$CARA_2(1,1)$	84.8	43.0, 57.0	85.4	45.7, 54.3
$CARA_3(0.5,1)$	84.7	42.9, 57.1	85.4	44.4, 55.6
$CARA_3(1,1)$	84.7	41.6, 58.4	84.7	42.0, 58.0
$CARA_4(0.5,1)$	85.0	44.0, 56.0	85.5	46.5, 53.5
$CARA_4(1,1)$	85.0	43.7, 56.3	85.5	46.2, 53.8

TABLE 5 Type I error, power, average sample size (Avg N), and allocation proportion for group-sequential monitoring of CR and CARA for Model M1 in Scenario 1: two arms and continuous endpoint.

Allocation	Type I error (%) ($\beta_A = 0$)		Power (%) ($\beta_A = 0.42$)		Avg N		
($n = 600$)	TMLE	*t*-test	TMLE	*t*-test	TMLE	*t*-test	Trt Prop (%)
CR	5.23	4.90	84.09	73.25	506.0	526.5	50.0, 50.0
$CARA_1(0.5,1)$	5.48	5.19	84.43	72.15	505.3	526.9	43.7, 56.3
$CARA_1(1,1)$	5.46	4.92	84.52	72.82	504.6	526.5	53.0, 57.0
$CARA_2(0.5,1)$	5.37	5.12	84.47	72.26	505.1	526.7	43.8, 56.2
$CARA_2(1,1)$	5.51	5.14	84.49	73.11	504.9	525.9	43.2, 56.8
$CARA_3(0.5,1)$	5.43	4.96	84.64	72.87	504.7	526.2	43.1, 56.9
$CARA_3(1,1)$	5.39	5.00	84.42	73.80	505.2	525.4	41.8, 58.2
$CARA_4(0.5,1)$	5.35	5.10	84.36	71.86	505.2	527.4	44.2, 55.8
$CARA_4(1,1)$	5.41	5.10	84.43	72.19	505.3	526.8	43.9, 56.1

$$+\sum_{p=1}^{10}\beta_{W,p} * W_p, \sigma = \frac{1+\beta_{A1}I(A=1)+\beta_{A2}I(A=2)}{1+\beta_V V}.$$

M4: $Y \sim \text{Gamma}(a,b)$, Y is truncated if $Y > 12$,

$$a = 1 + \{1+\beta_{A1}I(A=1)+\beta_{A2}I(A=2)\}(1+\beta_V V)$$

$$+\sum_{p=1}^{10}\beta_{W,p} * W_p, b = \frac{1+\beta_{A1}I(A=1)+\beta_{A2}I(A=2)}{1+\beta_V V}.$$

We vary the values of β_{A1} and β_{A2} in the two models to study the properties of the proposed CARA design

15410420, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/biom.13849 by Readcube (Labtiva Inc.), Wiley Online Library on [30/08/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

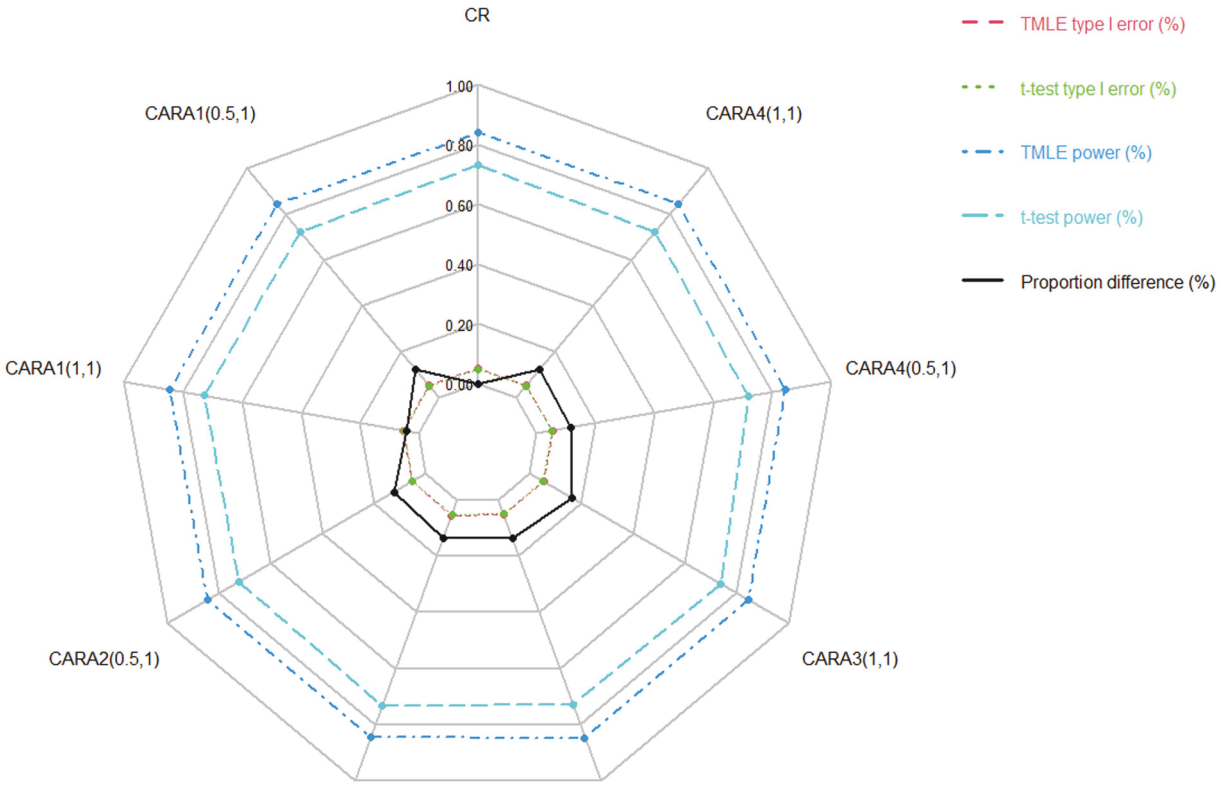

FIGURE 1 Radar plot for the comparison of type I error, power, and allocation proportion difference among group-sequential monitoring designs using CR and CARA for Model M1 in Scenario 1: two arms and continuous endpoint.

TABLE 6 Type I error under CR and CARA procedures for Scenario 2: three arms and continuous endpoint.

Allocation ($n = 1200$)	Model 3 ($\beta_{A1} = \beta_{A2} = 0$)		Model 4 ($\beta_{A1} = \beta_{A2} = 0$)	
	TMLE	Chi-sq	TMLE	Chi-sq
CR	5.42	4.89	5.28	4.50
$CARA_1(0,1)$	5.55	5.47	5.66	7.04
$CARA_1(1,0)$	5.13	4.46	5.56	6.69
$CARA_1(1,1)$	5.34	4.86	5.79	10.07
$CARA_2(1,0)$	5.17	4.57	5.72	5.52
$CARA_2(1,1)$	5.41	5.16	5.59	7.89
$CARA_3(1,0)$	5.01	4.36	5.58	7.49
$CARA_3(1,1)$	5.27	4.66	5.86	11.58
$CARA_4(1,0)$	5.16	4.86	5.76	5.32
$CARA_4(1,1)$	5.28	5.36	5.15	7.42

while fixing $\beta_V = 0.5$, $\beta_W = (-0.15, 0.05, -0.39, -0.12, -0.17, -0.22, 0.02, -0.04, 0.05, -0.06)$ in model M3, and $\beta_V = -0.4$, $\beta_W = (0.055, 0.3, -0.19, 0.04, 0.19, -0.07, 0.035, -0.04, 0.055, -0.08)$ in model M4.

In Table 6, we report the type I error. For model M3, both TMLE and the chi-square test demonstrate well-controlled type I error under CR and the CARA designs. However, for the complicated M4, TMLE outperforms the chi-square test in terms of type I error. In Table 7, the power and proportion of treatment are compared under CR and CARA procedures for the sample size $n = 1200$. Our method can assign more patients to the better arm and lead to a higher power.

Scenario 3: Three treatments with binary endpoints. Consider a clinical trial with three arms ($A = \{0, 1, 2\}$) with binary endpoints, and the sample

15410420, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/biom.13849 by Readcube (Labtiva Inc.), Wiley Online Library on [30/08/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 7 Power and allocation proportion under CR and CARA procedures for Scenario 2: three arms and continuous endpoint.

Allocation ($n = 1200$)	Model 3 ($\beta_{A1} = 0, \beta_{A2} = 0.16$)		Model 4 ($\beta_{A1} = 0.03, \beta_{A2} = 0.12$)	
	Power (%)	Trt Prop (%)	Power (%)	Trt Prop (%)
CR	81.2	33.3, 33.3, 33.3	81.4	33.3, 33.3, 33.3
$CARA_1$(0,1)	86.3	31.9, 31.9, 36.2	84.3	31.7, 32.7, 35.6
$CARA_1$(1,0)	84.9	32.8, 32.8, 34.4	84.4	31.3, 32.5, 36.2
$CARA_1$(1,1)	86.2	31.3, 31.4, 37.3	84.4	29.7, 31.8, 38.5
$CARA_2$(1,0)	84.9	32.9, 32.9, 34.3	84.5	32.8, 33.1, 34.2
$CARA_2$(1,1)	86.3	31.4, 31.4, 37.2	84.4	31.1, 32.4, 36.5
$CARA_3$(1,0)	84.9	32.3, 32.3, 35.3	85.1	30.8, 32.2, 37.0
$CARA_3$(1,1)	86.4	30.8, 30.9, 38.3	84.2	29.1, 31.5, 39.4
$CARA_4$(1,0)	84.4	33.1, 33.1, 33.7	84.6	33.0, 33.2, 33.8
$CARA_4$(1,1)	86.3	31.7, 31.7, 36.6	84.4	31.3, 32.5, 36.1

TABLE 8 Type I error, power, and allocation proportion under CR and CARA procedures for Scenario 3: three arms and binary endpoint.

Allocation ($n = 1200$)	Type I error (%) ($\beta_{A1} = \beta_{A2} = \beta_{AV} = 0$)		($\beta_{A1} = 0, \beta_{A2} = 0.25, \beta_{AV} = 0.20$)	
	TMLE	Chi-sq	Power (%)	Trt Prop (%)
CR	5.66	5.36	81.9	33.3, 33.3, 33.3
$CARA_1$(0,1)	5.61	5.01	83.8	34.0, 33.6, 32.4
$CARA_1$(1,0)	5.57	5.34	83.8	31.3, 32.6, 36.0
$CARA_1$(1,1)	5.54	5.22	83.9	32.0, 33.0, 35.0
$CARA_2$(1,0)	5.72	5.54	84.0	30.1, 31.9, 38.1
$CARA_2$(1,1)	5.58	5.41	83.8	30.7, 32.3, 37.0
$CARA_3$(1,0)	5.42	5.57	83.9	28.4, 30.9, 40.7
$CARA_3$(1,1)	5.57	5.44	83.1	28.9, 31.4, 39.8
$CARA_4$(1,0)	5.61	5.30	84.0	32.4, 33.0, 34.6
$CARA_4$(1,1)	5.52	5.05	84.1	33.0, 33.3, 33.6

size is $n = 1200$. Assume that the success rate of the binary endpoint Y is M5: $p = \Phi(\beta_{A1} I(A = 1) + \beta_{A2} I(A = 2) + \beta_V V + \beta_{AV} AV + \sum_{p=1}^{10} \beta_{W,p} * W_p)$, where $(\beta_{A1}, \beta_{A2}, \beta_V, \beta_{AV}, \beta_{W,p})$ are unknown parameters. We vary the values of $(\beta_{A1}, \beta_{A2}, \beta_{AV})$ to study the type I error, power, and other properties of our design while fixing $\beta_V = 0.2$, $\beta_W = (-0.01, 0.11, -0.125, -0.006, 0.06, -0.08, 0.016, -0.02, 0.026, -0.033)$.

In Table 8, we report the type I error with $(\beta_{A1}, \beta_{A2}, \beta_{AV}) = (0, 0, 0)$ and the operating characteristics under H_1. Both TMLE and the chi-square test lead to a well-controlled type I error rate for both CARA and CR, whereas CARA is better than CR in terms of power. In terms of treatment allocation proportion, CARA can assign more patients to the superior arm, especially when the choice of γ_1 and γ_2 emphasizes the ethics properties.

5 | REDESIGN OF A CLINICAL TRIAL

In this section, we study the performance of our design by redesigning a clinical trial (ACTG-320 study) where a three-drug treatment (indinavir, zidovudine, and lamivudine) was compared to a two-drug treatment (zidovudine and lamivudine) in HIV patients (Hammer et al. (1997)). The endpoint Y is CD4 count and the covariates are the stratified CD4 count at screening (W_1), the Karnofsky performance score (W_2), and the months of prior Zidovudine (ZDV) use (W_3). We denote $A = 1$ as the three-drug treatment and $A = 0$ as the two-drug treatment. All covariates are generated based on the empirical distribution from the trial samples with nonzero CD4 count in the study:

$$W_1 \sim \text{Bern}(0.63), W_2 \sim \text{Bern}(0.185), W_3 \sim 0.8 * \text{Unif}(1.5, 4.5) + 0.2 * \text{Unif}(1, 5).$$

15410420, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/biom.13849 by Readcube (Labtiva Inc.), Wiley Online Library on [30/08/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 9 Power and allocation proportion under CR and CARA procedures for redesiging the ACTG-320 study.

Allocation ($n = 600$)	Power (%)		Trt prop (%)
	TMLE	*t*-test	
CR	85.6	69.1	50.0, 50.0
$CARA_1(0,1)$	85.6	68.9	47.9, 52.1
$CARA_1(0.5,1)$	85.3	68.5	46.2, 53.8
$CARA_1(1,1)$	85.0	68.7	44.5, 55.5
$CARA_2(0.5,1)$	85.5	68.7	47.6, 52.4
$CARA_2(1,1)$	85.5	68.8	47.3, 52.7
$CARA_3(0.5,1)$	85.3	68.8	45.9, 54.1
$CARA_3(1,1)$	84.7	68.8	43.9, 56.1
$CARA_4(0.5,1)$	85.7	69.0	47.7, 52.3
$CARA_4(1,1)$	85.3	68.8	47.5, 52.5

The V used in CARA randomization approach is the dichotomized W_3, where $V = 1$ if $W_3 \geq 24$ and $V = 2$ if $W_3 < 24$. We fitted the model using original clinical trial data and adjusted the treatment effects to make the power around 0.85 with sample size 600. Then the following model will be used to generate the response Y: $Y = (3.65 + 0.838A + 6.443W_1 - 0.199W_2 - 0.388AW_2 + 0.006W_3 + \epsilon)^2$, where $\epsilon \sim N(0, 2.419^2)$. In Table 9, we can see that TMLE returns considerably higher power than the *t*-test when the data-generating function is nonlinear and complex, and CARA help assigns more subjects to the better treatment arm without sacrificing power. In particular, the $CARA_3(1, 1)$ assigns 12% more patients to the better treatment.

6 | DISCUSSION

CARA designs have been proposed to make use of all the observed information in a clinical trial to update the allocation probabilities and achieve different objectives. Different types of CARA designs have been studied in the literature, including target-based approaches (Zhang et al., 2007; Zhu, 2015), utility-based approaches (Atkinson & Biswas, 2005a, 2005b; Hu et al., 2015), Bayesian approach (Thall & Wathen, 2005), and treatment-effect mappings (Bandyopadhyay & Biswas, 2001; Rosenberger et al., 2001). It is worth mentioning that Villar and Rosenberger (2018) innovatively made use of future sequences of allocations and covariate values under the Gittins rule to introduce the covariates and randomization to the Gittins index and achieved a nonmyopic near optimal mean total rewards criterion in a computationally tractable way, significantly advancing CARA's development. Note that the current paper differs from Villar and Rosenberger (2018) and will promote the development and application of CARA designs by focusing on using many covariates without model misspecification and incorporating both ethics and efficiency measures.

In this article, we have proposed a novel unified solution to the challenges of CARA designs. We have achieved several theoretical and practical objectives in a single clinical trial. Our approach can help clinical trialists to pursue various design objectives and make use of a large number of covariates in the design and analysis without unnecessary model assumptions. Moreover, we have offered a theoretical foundation for the proposed methods, which will help researchers to further investigate the properties of this approach and help regulation agencies to understand the fundamental properties of the new design to promote its application.

Finally, we discuss some limitations and future research about our proposed method. First, we update the allocation probability patient by patient while implementing the CARA design in the numerical studies. In practice, we can group sequentially update the information and assign the next group of patients. Second, there is no logical difficulty in incorporating delayed responses in the procedure. For RAR, the large sample results are not affected by delayed responses under widely applicable conditions (Bai et al., 2002; Hu et al., 2008; Zhang et al., 2007). Explicit theoretical results about the effect of delayed responses on CARA designs are worth investigating. Third, an ideal case for adaptive randomization is that it can target certain optimal allocations for corresponding objectives. This topic is demanding, especially when the covariates are continuous. We leave these for future research.

ACKNOWLEDGMENTS

This work was partly supported by grant DMS-2014951 from the National Science Foundation (USA) and the PRIME award from the University of Texas Health Science Center at Houston. This publication was neither originated nor managed by AbbVie, and it does not communicate results of AbbVie-sponsored Scientific Research. Thus, it is not in scope of the AbbVie Publication Procedure (PUB-100).

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this paper as no new data were created or analyzed in this paper.

ORCID

Hongjian Zhu https://orcid.org/0000-0002-8584-3650

REFERENCES

Atkinson, A.C. (1982) Optimum biased coin designs for sequential clinical trials with prognostic factors. *Biometrika*, 69(1), 61–67.

15410420, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/biom.13849 by Readcube (Labtiva Inc.), Wiley Online Library on [30/08/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Atkinson, A.C. & Biswas, A. (2005a) Adaptive biased-coin designs for skewing the allocation proportion in clinical trials with normal responses. *Statistics in Medicine*, 24(16), 2477–2492.

Atkinson, A.C. & Biswas, A. (2005b) Bayesian adaptive biased-coin designs for clinical trials with normal responses. *Biometrics*, 61(1), 118–125.

Bai, Z., 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.

Bandyopadhyay, U. & Biswas, A. (2001) Adaptive designs for normal responses with prognostic factors. *Biometrika*, 88(2), 409–419.

Chambaz, A. & van der Laan, M.J. (2014) Inference in targeted group-sequential covariate-adjusted randomized clinical trials. *Scandinavian Journal of Statistics*, 41(1), 104–140.

Chambaz, A., van der Laan, M.J. & Zheng, W. (2014) Targeted covariate-adjusted response-adaptive lasso-based randomized controlled trials. In: Sverdlov, O. (Ed.) *Modern adaptive randomized clinical trials: statistical, operational, and regulatory aspects*. Boca Raton: Taylor & Francis, pp. 345–368.

EMA (2015) Guideline on adjustment for baseline covariates in clinical trials.

Fackle-Fornius, E. & Nyquist, H. (2015) Optimal allocation to treatment groups under variance heterogeneity. *Statistica Sinica*, 25, 537–549.

FDA (2021) Adjusting for Covariates in Randomized Clinical Trials for Drugs and Biological Products.

Gruber, S. & van der Laan, M.J. (2010) A targeted maximum likelihood estimator of a causal effect on a bounded continuous outcome. *The International Journal of Biostatistics*, 6(1).

Hammer, S.M., Squires, K.E., Hughes, M.D., Grimes, J.M., Demeter, L.M., Currier, J.S., Eron Jr, J.J., Feinberg, J.E., Balfour Jr, H.H., Deyton, L.R. et al. (1997) A controlled trial of two nucleoside analogues plus indinavir in persons with human immunodeficiency virus infection and cd4 cell counts of 200 per cubic millimeter or less. *New England Journal of Medicine*, 337(11), 725–733.

Heitjan, D.F. & Rubin, D.B. (1991) Ignorability and coarse data. *The Annals of Statistics*, 19(4), 2244–2253.

Hu, F. & Rosenberger, W.F. (2006) *The theory of response-adaptive randomization in clinical trials*, volume 525. Hoboken, NJ: John Wiley & Sons.

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.

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.

Rosenberger, W.F. & Lachin, J.M. (2015) *Randomization in clinical trials: theory and practice*. Hoboken, NJ: John Wiley & Sons.

Rosenberger, W.F., Vidyashankar, A. & Agarwal, D.K. (2001) Covariate-adjusted response-adaptive designs for binary response. *Journal of Biopharmaceutical Statistics*, 11(4), 227–236.

Thall, P.F. & Wathen, J.K. (2005) Covariate-adjusted adaptive randomization in a sarcoma trial with multi-stage treatments. *Statistics in Medicine*, 24(13), 1947–1964.

van der Laan, M.J. & Rose, S. (2011) *Targeted learning: causal inference for observational and experimental data*. New York: Springer Science & Business Media.

van der Laan, M.J. & Rubin, D. (2006) Targeted maximum likelihood learning. *The International Journal of Biostatistics*, 2(1).

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.

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.

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.

Zheng, W. (2014) *Semiparametric and robust methods for complex parameters in causal inference*. PhD thesis, UC Berkeley.

Zhu, H. (2015) Covariate-adjusted response adaptive designs incorporating covariates with and without treatment interactions. *Canadian Journal of Statistics*, 43(4), 534–553.

SUPPORTING INFORMATION

The proofs, derivations, and related formulas referenced in Sections 2 and 3 are available with this paper at the Biometrics website on Wiley Online Library.

How to cite this article: Zhu, H. & Zhu, H. (2023) Covariate-adjusted response-adaptive designs based on semiparametric approaches. *Biometrics*, 1–12. https://doi.org/10.1111/biom.13849

15410420, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/biom.13849 by Readcube (Labtiva Inc.), Wiley Online Library on [30/08/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License