Statistical Methods in Medical Research
http://smm.sagepub.com/

Multiple-stage sampling procedure for covariate-adjusted response-adaptive designs

Eunsik Park and Yuan-chin Ivan Chang

Stat Methods Med Res published online 30 May 2013
DOI: 10.1177/0962280213490091

The online version of this article can be found at:
http://smm.sagepub.com/content/early/2013/05/29/0962280213490091

Published by:
SAGE
http://www.sagepublications.com

Original Article

SMMR
STATISTICAL METHODS IN MEDICAL RESEARCH

Statistical Methods in Medical Research
0(0) 1–22

DOI: 10.1177/0962280213490091
smm.sagepub.com
SAGE

Multiple-stage sampling procedure for covariate-adjusted response-adaptive designs

Eunsik Park[1] and Yuan-chin Ivan Chang[2]

Abstract

Covariate-adjusted response-adaptive (CARA) design becomes an important statistical tool for evaluating and comparing the performance of treatments when targeted medicine and adaptive therapy become important medical innovations. Due to the nature of the adaptive therapies of interest and how subjects accrue to a sampling procedure, it is of interest how to control the sample size sequentially such that the estimates of treatment effects have satisfactory precision in addition to its asymptotic properties. In this paper, we apply a multiple-stage sequential sampling method to CARA design in such a way that the control of the sample size is more feasible. The theoretical properties of the proposed method, including the estimates of regression parameters and the allocation probabilities under this randomly stopped sampling procedure, are discussed. The numerical results based on synthesized data and a real example are presented.

Keywords

covaraite adjustment, response-adaptive design, multiple-stage, stopping rule, confidence set

1 Introduction

The sequential method has been widely applied to many clinical trials to match how the patients accrue in practice to a sampling procedure. The covariate-adjusted method becomes a good additional property to the response-adaptive clinical trials, as tailor-made or targeted medicine draws much attention in biomedical research. In addition, the response-adaptive concept allows us to put ethical considerations into designing clinical trials. The design of clinical trials with covariate adjustment and response adaptiveness based on previously collected information therefore causes the trials to inherit their sequential sampling nature. Moreover, the use of sequential sampling can accommodate a balance between the benefit of personalization for the better treatment of participants as well as the statistical optimality in these types of clinical trials, simultaneously. That makes sequential methods a suitable statistical tool for covariate-adjusted

[1]Department of Statistics, Chonnam National University, Gwangju, Korea
[2]Institute of Statistical Science, Academia Sinica, Taipei, Taiwan

Both authors contributed equally to this work.
Corresponding author:
Eunsik Park, Department of Statistics, Chonnam National University, Gwangju 500757, Korea.
Email: espark02@gmail.com

response-adaptive (CARA) designs.[1–6] Although the response-adaptive design alone has been studied by many authors,[7] and the properties of the response-adaptive designs with adjustment according to individual covariate information have been studied by some authors,[8–11] there is still a lack of discussions about sample size calculation and about stopping criterion for this type of design, which motivates this study.

Sampling one subject at a time, as in a standard fully sequential method, introduces some operational inconvenience in practice. Thus, multiple-stage methods, a compromise between the theoretical beauty of fully sequential methods and practical usefulness, become good alternatives. Application of a two-stage CARA design is, of course, theoretically justifiable. Bandyopadhyay et al.[12] worked on two-stage design for binary responses under CARA designs with the prefixed total sample sizes. However, allocating subjects based on the information obtained only in the initial stage may be risky, as the initial stage is relatively unstable due to many factors, such as sample size. Hence, a three-stage CARA design is a better choice, as it considers the operational convenience in addition to taking full advantage of a CARA design.

In this paper, from a more practical prospective, we study a three-stage method to CARA-designed clinical trials with K treatments under a generalized linear model assumption. The multiple-stage methods may not be new in clinical trials and sequential analysis. However, in many clinical trials, the initial stage is usually not included in the final analysis, which may not be affordable for many trials. In this study, the subjects in all stages are included in the final analysis and, thus, our study is more efficient. Our goal is to estimate the treatment effects such that the estimates satisfy a prescribed precision with the minimum sample size, and also so that subjects can be allocated to the superior treatment while maintaining the quality and efficiency of the estimation of treatment effects. The asymptotic properties of the proposed three-stage method are obtained under a rather general assumption about CARA designs. We show that under the proposed method, the allocation rule maintains the same asymptotic properties as those obtained in its non-sequential counterpart. In our numerical study, for illustration purposes, we adopt the idea of using a utility function to balance the ethical consideration and the efficiency of the estimate for treatment allocation.[12] We, then, allow the utility function to vary the tuning parameters depending on the precision of the estimate at different allocation stages such that subjects are allocated to a ''more adequate'' treatment.

The rest of this paper is organized as follows. The details and asymptotic properties of the proposed method are presented in Section 2. Empirical results from simulation studies, for illustration purposes, where we focus on a CARA design under logistic models, are presented in Section 3 together with results based on a real example. Some discussion is given in Section 4. In addition, the proof of the theorem is given in Appendix A of the supplementary material.

2 Method

In this section, we study a multiple-stage confidence set estimation under a general CARA design after introducing some necessary notations for this design. Our method allows the sample size to depend on the information on observations such that a prescribed estimation accuracy is fulfilled. Specially, a three-stage method is presented and recommended.

Let $Y_{n,k}$, $n = 1, 2, \ldots$, $k = 1, \ldots, K$, denote the response of the n-th subject to the k-th treatment, and let $\boldsymbol{\xi}_n$ be the p dimensional vector of covariates of the n-th subject. Suppose that for each $n \geq 1$ and $k \in \{1, \ldots, K\}$, the responses and covariates satisfy

$$E[Y_{n,k}|\boldsymbol{\xi}_n] = \mu_k(\boldsymbol{\theta}_k, \boldsymbol{\xi}_n), \tag{2.1}$$

where $\boldsymbol{\theta}_k \in \Theta_k \subseteq R^p$ is an unknown vector of parameters and $\mu_k(\cdot,\cdot)$ is a known function. Assume that $\mu_k(\boldsymbol{\theta}_k, \boldsymbol{\xi}_n) = \mu_k(\boldsymbol{\xi}_n'\boldsymbol{\theta}_k)$ for each n and k. Let $\Theta \equiv \prod_{k=1}^K \Theta_k$ and nonsingular $\boldsymbol{V} = \text{diag}\{V_1, \ldots, V_K\}$, where V_k denotes the covariance matrix of the estimator of $\boldsymbol{\theta}_k$ corresponding to equation (2.1). Then, from equation (2.1) and the definition of $\boldsymbol{V}$, the method of the quasi-likelihood of generalized linear models[13] can be applied to estimate $\boldsymbol{\theta}_k$ for each k.

Let $\boldsymbol{X}_n = (X_{n,1}, \ldots, X_{n,K})$, where $X_{n,k} \in \{0,1\}$ denotes assignment of treatment k to the n-th subject. Thus, $\boldsymbol{X}_i$ is a vector that denotes the random treatment assignments for subject i, for $i = 1, 2, \ldots$ with components equal to 0 or 1. Note that each subject is allocated to one treatment only, hence $X_{n,k} = 1$ for only one $k \in \{1, \ldots, K\}$, which implies that $\sum_{k=1}^K X_{n,k} = 1$. That is, $Y_{n,k}$ is observed only if $X_{n,k} = 1$. Hence, for each n, the vector $\boldsymbol{Y}_n \equiv (Y_{n,1}, \ldots, Y_{n,K})$ has only one component actually observed. Now, suppose $N_{n,k}$ is the number of subjects assigned to treatment k during the first n assignments, and let vector $\boldsymbol{N}_n \equiv (N_{n,1}, \ldots, N_{n,K})$. Hence, it follows from the definitions above that $\boldsymbol{N}_n = \sum_{i=1}^n \boldsymbol{X}_i$. Let $\mathcal{X}_n = \sigma(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n)$, $\mathcal{Y}_n = \sigma(\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_n)$, and $\mathcal{Z}_n = \sigma(\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_n)$, $\boldsymbol{\xi}_i \in \mathbb{R}^p$ denote the corresponding σ-fields, and let $\mathcal{F}_n = \sigma(\mathcal{X}_n, \mathcal{Y}_n, \mathcal{Z}_n)$. Then, a general CARA design is defined as the conditional probabilities of assigning treatments $1, \ldots, K$ to the n-th patient, conditioning on all observed information (individual responses and variables) up to previous $n-1$ assignments and the covariate information on the current subject, which is denoted as

$$\boldsymbol{\psi}_n = E[\boldsymbol{X}_n | \mathcal{F}_{n-1}, \boldsymbol{\xi}_n] = E[\boldsymbol{X}_n | \mathcal{X}_{n-1}, \mathcal{Y}_{n-1}, \mathcal{Z}_n] = \pi(\hat{\boldsymbol{\theta}}_{n-1}, \boldsymbol{\xi}_n),$$

where $\hat{\boldsymbol{\theta}}_n$ denotes the estimator of $\boldsymbol{\theta}$ at the current stage. Assume that the target allocation function $\boldsymbol{\pi}(\cdot,\cdot) = (\pi_1(\cdot,\cdot), \ldots, \pi_K(\cdot,\cdot))$ with $\sum_{k=1}^K \pi_k = 1$ and $0 < \nu_k = E_\xi[\pi_k(\theta, \boldsymbol{\xi})] < 1$, for $k = 1, \ldots, K$, and let $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_K)$. Suppose that Θ_k is bounded for all k and that for each fixed $\boldsymbol{\xi}$, $\pi_k(\boldsymbol{\theta}, \boldsymbol{\xi}) > 0$ is continuous and differentiable with respect to $\boldsymbol{\theta}$ such that $\nu_k(\tilde{\boldsymbol{\theta}}) = \nu_k(\theta) + (\tilde{\boldsymbol{\theta}} - \theta)(\partial \nu_k / \partial \tilde{\boldsymbol{\theta}})' + o(||\tilde{\boldsymbol{\theta}} - \theta||^{1+\zeta})$ for some $\zeta > 0$. Then, it has been proved that $\hat{\boldsymbol{\theta}}_n$ is strongly consistent with $\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) \to_L N(0, \boldsymbol{V})$ as $\min\{N_{n,k}, k = 1, \ldots, K\}$ goes to infinity (Zhang et al.,[10] Theorem 2.1). (The notation "$\to_L$" denotes the convergence in distribution.)

Now, we introduce a confidence ellipsoid and extend it to multiple-stage design under the CARA design. Let C_α^2 be the constant satisfying $P(\chi^2(p) \geq C_\alpha^2) = \alpha$, and

$$R_n = \{\boldsymbol{\theta} \in \Theta : n(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta})' \boldsymbol{V}^{-1} (\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) \leq C_\alpha^2\}. \tag{2.2}$$

Then, from the asymptotic normality of $\hat{\boldsymbol{\theta}}_n$, equation (2.2) defines a confidence ellipsoid for $\boldsymbol{\theta}$ such that $P(\boldsymbol{\theta} \in R_n) \approx 1 - \alpha$ as $\min\{N_{n,k}, k = 1, \ldots, K\}$ becomes large. However, no matter how high the coverage probability of a confidence ellipsoid is, it becomes less useful/informative if the size of a confidence set is too big. Therefore, we would require more in terms of the size of a confidence set in addition to the coverage probability. Suppose that we require the maximum axis of R_n to be no larger than 2δ for some $\delta > 0$ to control its size; then, to guarantee that R_n has the prescribed coverage probability $1 - \alpha$, we must have a sample size that satisfies the following inequality:

$$n\Lambda_{\min}(\boldsymbol{V}^{-1}) \geq \frac{C_\alpha^2}{\delta^2} \Leftrightarrow n \geq \frac{C_\alpha^2 \Lambda_{\max}(\boldsymbol{V})}{\delta^2}, \tag{2.3}$$

where $\Lambda_{\min}(\boldsymbol{V})$ and $\Lambda_{\max}(\boldsymbol{V})$ are minimum and maximum eigenvalues of V, respectively. It is clear that δ also means the precision of the estimate. As we can see from equation (2.3), the smaller the δ, the larger the sample size. Hence, the choice of δ will depend on the under-studied problem and the actual needs. From a practical prospective, besides the asymptotic properties of CARA designs,

we would also want to know how large a sample size must be to guarantee a satisfactory performance of the CARA-designed clinical trial. However, if $\boldsymbol{V}$ is unknown, equation (2.3) cannot provide such information.

Let R_δ denote the confidence ellipsoid of θ with the length of its maximum axis no larger than 2δ and its coverage probability no less than $1-\alpha$. Then, following equation (2.3), if $\boldsymbol{V}$ is known, the optimal sample size required to construct a confidence ellipsoid R_δ with the required coverage and specific precision is

$$n_{opt} = \text{ first } n \geq n_0 \quad \text{such that } n \geq \frac{C_\alpha^2 \Lambda_{\max}(\boldsymbol{V})}{\delta^2}, \tag{2.4}$$

where n_0 is the initial sample size. The variance matrix $\boldsymbol{V}$ is usually unknown due to the adaptive feature of the CARA-designed clinical trials and the model used in a trial. Hence, determining the sample size that can guarantee the desired properties beforehand is unlikely. Although this optimal sample size is not available, equation (2.4) still suggests an estimate of the optimal sample size through the estimate of $\boldsymbol{V}$, and the sequential method is a classical approach under such a situation. Here, in particular, a sequential three-stage sampling procedure is proposed to construct a confidence ellipsoid R_δ for $\boldsymbol{\theta}$ with a prefixed accuracy and a given coverage probability, simultaneously.

When a fully sequential procedure is adopted, both parameters in the allocation rule and the formula of sample size are adjusted and re-checked whenever a new subject is included in the trial. In theory, the fully sequential method is usually more precise and efficient in terms of sample size than its multiple-stage alternatives. However, such a fully sequential sampling scheme also introduces some operational difficulties in practice, as we need to re-estimate the parameters and check the stopping criterion whenever a new subject enters the trial. Hence, using multiple-stage methods as some practical alternatives[4] is usually preferred by practitioners. The two- and three-stage methods are the two most popular multiple-stage methods discussed in the literature.

For a two-stage method,[14] we first estimate the total sample size based on the initial samples, and then we collect an additional batch of subjects if the estimated sample size exceeds the size of the initial samples. It is obvious that its performance highly depends on the estimate obtained from the initial samples; otherwise, stop sampling and the estimation of the treatment effect will be based on the initial samples only. Thus, the determination of an initial sample size is crucial: If the initial sample size is too small, then the estimated total sample size becomes unstable; if it is too large, then there are no second-stage samples to be used.

Generally speaking, when a multiple-stage sequential sampling procedure is employed, we usually decide the number of stages first. Starting with an initial batch of subjects, we have an initial estimate of the required sample size. Instead of using this initial estimated sample size directly, we then only use its fraction as the sample size for the next batch of subjects. Once this new batch of subjects is collected, the sample size will be re-estimated using all of the samples up to the current stage. This type of operation will be repeated until the predetermined number of stages is reached or until a stopping criterion is satisfied. That is, we only re-estimate or re-check the sample size inequality when a batch of subjects is included. Note that the initial samples are also included in the final analysis in a multiple-stage sequential method, which is different from the samples used in a pilot study of some clinical trials.

Moreover, as described above, the allocation rule in a CARA-based clinical trial will also rely on the information obtained from its previous samples. Hence, when the initial sample size is too small, the allocation may be unreliable; conversely, when the initial sample size is too large, then only a

small number (or even zero) subjects will be included as the second-stage samples, who are the only subjects that can benefit from the CARA design. Thus, this CARA-designed clinical trial becomes less useful. Hence, as a compromise between the two-stage and the fully sequential method, we propose using a three-stage method in the CARA-designed clinical trial that has both the advantages of sequential sampling and the benefit of the CARA design. Of course, this idea can be easily generated for procedures with more than three stages. For further details regarding general sequential confidence set estimations, please refer to Siegmund.[15]

2.1 Three-stage method

Assume a randomized trial is used at the initial stage, and let $n_{0,1}, \ldots, n_{0,K}$ be the initial sample sizes for treatments 1 to K, respectively. Denote the total sample size, $n_{0,1} + \cdots + n_{0,K}$, at the initial stage by n_0. Replacing $\boldsymbol{V}$ in equation (2.4) with its estimate based on the initial samples $\hat{\boldsymbol{V}}_{n_0}$, we have the estimated sample size to construct a confidence ellipsoid with the required coverage probability $1-\alpha$ and precision ($\delta > 0$):

$$\tau_1 \equiv \tau_{1,\delta} = \left[\frac{C_\alpha^2 \Lambda_{\max}(\hat{\boldsymbol{V}}_{n_0})}{\delta^2} \right]. \tag{2.5}$$

(The notation $[a]$, for $a \in R$, denotes the smallest positive integer greater than a.) If a two-stage approach is adopted, then an extra batch of $\tau_1 - n_0$ subjects will be included in the study if $\tau_1 > n_0$; otherwise, no new subjects will be included, and the statistical inference will be based on the initial samples. However, if a three-stage approach is used, then we only collect a fraction of $\tau_1 - n_0$ subjects at the second stage, say $n_1 = r \cdot (\tau_1 - n_0)$ for some $0 < r < 1$, and these subjects are then allocated to the most suitable treatment for them based on the estimate $\hat{\boldsymbol{\theta}}_{n_0}$, a prescribed allocation rule, and the updated information up to the current stage. The required optimal sample size is then re-estimated based on all available information up to the current stage; that is, the information is updated using the initial samples and the newly included second batch of subjects together. Denote the second-stage estimate of the total sample size with

$$\tau_2 \equiv \left[\frac{C_\alpha^2 \Lambda_{\max}(\hat{\boldsymbol{V}}_{n_0+n_1})}{\delta^2} \right]. \tag{2.6}$$

When $\tau_2 > n_0 + n_1$, then we collect additional $n_2 = \tau_2 - (n_0 + n_1)$ new subjects at the third stage, and we allocate them based on the allocation rule with the updated parameters at the second stage. No extra subjects are included when $\tau_2 \leq (n_0 + n_1)$; that is, $n_2 = 0$ in this case. The final inference is then based on the total $n_{_T} \equiv n_{total} = n_0 + n_1 + n_2$ subjects recruited in the trial; the final sample size is equal to τ_2 if $\tau_1 > n_0$ and $\tau_2 > n_1$. As in the fully sequential procedure, τ_2 is random and depends on the estimate of the variance and the criteria of the confidence set of the parameters to be estimated. Let $\hat{\boldsymbol{\theta}}_{n_{_T}}$ be the estimate of θ with random sample size $n_{_T}$. Then, for the three-stage method described above, we have the following theorem:

Theorem 2.1: *If* $n_0(\delta) = o(\delta^2)$, *then*

(i) $\hat{\boldsymbol{\theta}}_{n_{_T}}$ *is a strongly consistent estimate of* $\boldsymbol{\theta}$ *as* $\delta \to 0$,
(ii) $\lim_{\delta \to 0} P(\boldsymbol{\theta} \in R_{\delta, n_{_T}}) = 1 - \alpha$, *and*

(iii) $\lim_{\delta\to 0} E[n_T/n_{opt}] = 1$,
where R_{δ,n_T} denotes the confidence ellipsoid R_δ of sample size n_T and precision δ. Moreover, for a given allocation function,

(iv) $\lim_{\delta\to 0}(\boldsymbol{N}_{n_T}/n_T) = \boldsymbol{v}$, *and* $(N_{\delta,k|\xi}/N_{n_T|\xi}) \to \pi_k(\boldsymbol{\theta}, \boldsymbol{\xi})$, *for* $k = 1, \ldots, K$, *almost surely as* $\delta \to 0$,
where $N_{\delta,k|\xi}$ is number of subjects assigned to treatment k with covariate $\boldsymbol{\xi}$ up to the n_T-th subject and $\boldsymbol{N}_{n_T|\xi}$ is the total number of subjects with covariate $\boldsymbol{\xi}$ up to the n_T-th subject.

Theorem 2.1 (i) and (ii) simply say that the estimate of $\boldsymbol{\theta}$ and the confidence set R_{δ,n_T} behave as well as their fixed sample size counterparts do as the sample size becomes large. Theorem 2.1 (iii) means the random sample size based on the proposed three-stage method is as efficient as the optimal sample size as $\delta \to 0$, asymptotically. Theorem 2.1 (iv) states that all allocation proportions converge to their expectations π for each treatment, asymptotically, under the proposed three-stage sampling scheme; moreover, for subjects with covariate $\boldsymbol{\xi}$ (i.e. given $\boldsymbol{\xi}$), the allocation probability will also converge to that of the non-random sample size as well.

Remark 2.2: *As mentioned, it is possible that the initial sample size is already too large such that the inequality in equation (2.3) is satisfied. In this case, the confidence set for $\boldsymbol{\theta}$ still has a coverage probability approximately equal to that required. However, in this case, the allocation rule is based on the initial design only, and the designs in the second and third stages have no effect at all. Thus, there is no benefit from the CARA design. The assumption $n_0(\delta) = o(\delta^2)$ is to prevent such a situation. This simply means that the initial sample size should not "increase too fast" as δ becomes small, which is required in the proof of (iii) of the above theorem.*

Remark 2.3: *It is known that a clinical trial based on a CARA design is naturally a sequential procedure. Although sequential methods are very popular in clinical trials, the discussions on the stopping criterion of the trials under CARA designs are rare. Zhang et al.[10] proved some asymptotic properties without providing any stopping criterion. However, it is also known that, from the literature of sequential analysis,[15] when the sample size is random and depends on the observed information, the asymptotic properties are no longer guaranteed[16] with the arguments of Zhang et al.[10] That is the reason why this theorem is important and necessary.*

The same idea can be extended to the estimation of the contrasts of parameters, which allows us to compare parameters/treatment effects. Suppose that H is a contrast matrix with $rank(H) = q > 0$. Let $\gamma_k = H'\boldsymbol{\theta}_k$ for $k = 1, \ldots, K$, $\boldsymbol{\gamma} = (H'\boldsymbol{\theta}_1, \ldots, H'\boldsymbol{\theta}_K)'$ with $\boldsymbol{\theta}_k \in \Theta_k$, for $k = 1, \ldots, K$, and $\hat{\boldsymbol{\gamma}} = (\hat{\gamma}_1, \ldots, \hat{\gamma}_K)'$. In Theorem 2.1 (iii), we have proved that $\lim_{\delta\to 0} E[n_T/n_{opt}] = 1$. Hence, following the similar arguments to that of Chang and Park,[17] the asymptotic normality of $\hat{\boldsymbol{\theta}}_{n_T}$ follows easily. Similarly, we have that $\hat{\gamma}$ is a strongly consistent estimate of γ, and $\sqrt{n}(\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}) \to_L N(0, \boldsymbol{V}^H)$ as $\min(N_{n,k}, k = 1, \ldots, K) \to \infty$, where $\boldsymbol{V}^H = diag\{H'V_1H, \ldots, H'V_KH\}$. Therefore, for $\boldsymbol{\theta}_k \in \Theta_k$, $k = 1, \ldots, K$, and a given $\delta > 0$ and $\alpha \in (0, 1)$, let us define

$$R^H_{\delta,n} = \left\{\boldsymbol{\gamma} : n(\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma})'(\boldsymbol{V}^{\boldsymbol{H}})^{-1}(\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}) \le C^2_{\alpha,qK}\right\},$$

where $C^2_{\alpha,qK}$ denotes the constant such that $P(\chi^2(q) \ge C^2_{\alpha,qK}) = \alpha$. Hence the required sample size must satisfy the following inequality:

$$n \ge \frac{C^2_{\alpha,qK}\Lambda_{\max}(\boldsymbol{V}^H)}{\delta^2}.$$

Similarly, suppose that n_0^H is the initial sample size. Parallel to equations (2.5) and (2.6), let

$$\tau_1^H = \left[\frac{C_{\alpha,qK}^2 \Lambda_{\max}(\hat{\boldsymbol{V}}_{n_0^H}^H)}{\delta^2}\right] \quad \text{and} \quad \tau_2^H = \left[\frac{C_{\alpha,qK}^2 \Lambda_{\max}(\hat{\boldsymbol{V}}_{n_0^H+n_1^H}^H)}{\delta^2}\right],$$

where similarly, $n_1^H = [\tau_1^H - n_0^H]^+$ and $n_2^H = [\tau_2^H - n_1^H - n_0^H]^+$ are sample sizes of the first and second stages, respectively.

Hence, using these two random sample sizes and letting $n_T^H = n_0^H + n_1^H + n_2^H$, we have a three-stage procedure to estimate $\boldsymbol{\gamma}$. Let $\hat{\boldsymbol{\gamma}} \equiv \hat{\boldsymbol{\gamma}}_{n_T^H}$ and $R_{\delta,n_T^H}^H$ be the estimate and confidence ellipsoid of $\boldsymbol{\gamma}$ under this three-stage sampling procedure. Then, the corollary below follows directly from Theorem 2.1.

Corollary 2.4: *If* $n_0(\delta) = o(\delta^2)$, *then*

(i) $\hat{\boldsymbol{\gamma}}_{n_T^H}$ *is a strong consistent estimate of* $\boldsymbol{\gamma}$ *as* $\delta \to 0$,
(ii) $\lim_{\delta\to 0} P(\boldsymbol{\gamma} \in R_{\delta,n_T^H}^H) = 1 - \alpha$, *and*
(iii) $\lim_{\delta\to 0} E[n_T^H / n_{opt}^H] = 1$,
where

$$n_{opt}^H = \textit{ first } n \geq n_0^H \quad \textit{such that } n \geq C_{\alpha,qK}^2 \Lambda_{\max}(\boldsymbol{V}^H)/\delta^2. \tag{2.7}$$

In addition, for a given allocation function,
(iv) $\lim_{\delta\to 0}(\boldsymbol{N}_{n_T^H}/n_T^H) = \boldsymbol{v}$, *and* $(N_{\delta,k|\xi}/N_{n_T^H|\xi}) \to \pi_k(\boldsymbol{\theta}, \boldsymbol{\xi})$, *for* $k = 1, \ldots, K$, *almost surely as* $\delta \to 0$.

The proof of Corollary 2.4 is similar to that of Theorem 2.1, so it is omitted.

Remark 2.5: *The sample size estimate based on the initial sample is usually unstable due to a lack of the information needed to decide the satisfactory initial sample size for a CARA design. On the other hand, it is well known that in the multiple-stage procedure, the estimates of the sample size are refined at each stage and become more stable as a new stage is added. Thus, a three-stage is a compromise that can provide satisfactory results under a CARA design, and that is the reason we focus on a three-stage procedure here. Of course, the proposed results can be easily extended to other multiple-stage procedures.*

3 Numerical study

Our method is built on some common assumptions of the quasi-likelihood method and mild regularity conditions on the CARA design. To see the performance of three-stage methods in terms of stopping time, coverage probability, and correct allocation probability (CAP), we apply our method to a CARA design based on a logistic model with a utility function, as follows. The numerical results based on both synthesized data and a real example are presented.

3.1 Treatment allocation rule

Atkinson and Biswas[18] and Bandyopadhyay et al.[12] suggest using a utility function to skew the treatment allocation proportion so that the best treatment is allocated more often. For K treatments,

their utility function is defined as

$$U(\boldsymbol{p}) = \log |\hat{I}_{\sum_{i=0}^{s+1} n_i}| - \eta \left\{ \sum_{k=1}^{K} p_k \log \left(\frac{p_k}{\pi_k(\hat{\boldsymbol{\theta}}, \boldsymbol{\xi})} \right) \right\}, \tag{3.1}$$

where $\hat{I}_n$ is the estimated Fisher information with cumulative samples of size n, and $\pi_k(\hat{\boldsymbol{\theta}}, \boldsymbol{\xi})$ is the estimate of $\pi_k(\hat{\boldsymbol{\theta}}, \boldsymbol{\xi})$, denoting the estimate of the target allocation proportion for treatment k up to the current, s-th, stage. The first term of the right-hand side of equation (3.1) depends on the information of the parameters to be estimated, and the second term is a relative entropy (or the Kullback-Leibler distance) of allocation proportions (p_k's) to optimize with respect to the estimated target allocation probabilities ($\pi_k(\hat{\boldsymbol{\theta}}, \boldsymbol{\xi})$). We can allow π_k, let us say as a function of T_n, and η to vary in different sampling stages. For example, T_n and η can vary according to the estimation precision or be defined as some function of the standard deviation of the estimated treatment effect at the current stage. Thus, for the given covariate $\boldsymbol{\xi}$ and the estimate of $\boldsymbol{\theta}$, this utility function balances the information on $\boldsymbol{\theta}$ and the ethical treatment allocation by tuning parameters T_n and/or η.

For a given $\boldsymbol{\xi}$ and the current estimate of $\boldsymbol{\theta}$, the optimal allocation rule is to find the vector of probabilities $\boldsymbol{p} = (p_1, \ldots, p_K)'$ that maximize this utility function above. That is, the design at the $(s+1)$th stage is to allocate n_{s+1} subjects to the treatment that maximizes the utility function given the observed information up to the s-th stage. The first term of the utility function is a log determinant of the information matrix, and the second term involves $\pi_k(\hat{\boldsymbol{\theta}}, \boldsymbol{\xi})$. Thus, if $\eta = 0$, the new subject is selected to maximize the Fisher information matrix, which is referred to as the piecewise D-optimal design, as mentioned in Bandyopadhyay et al.[12]. Conversely, if η goes to ∞, then the optimal value of $\boldsymbol{p}$ is to maximize the relative entropy function – the second term of equation (3.1).[19] Hence, the parameter η can be used to adjust the ethical and efficiency balance. Here, we adopt the idea of using a utility function to balance the needs for estimation precision of treatment effects and the ethical consideration.

3.2 Simulation study

We conduct the simulation study using a logistic model. The model and parameter setting are described below.

3.2.1 Logistic model

Suppose $Y_k = 1(0)$ denotes a variable with a positive (negative) response of a subject assigned to treatment k, for $k = 1, \ldots, K$. Let $\mu_k(\boldsymbol{\theta}_k, \boldsymbol{\xi}) = E[Y_k = 1|\boldsymbol{\xi}]$, and $\boldsymbol{\theta}_k = (\alpha_k, \theta_k^*)'$. Assume that

$$\text{logit}\,(\mu_k(\boldsymbol{\theta}_k, \boldsymbol{\xi})) = \alpha_k + \theta_k^* \boldsymbol{\xi}, \quad k = 1, \ldots, K. \tag{3.2}$$

Assume that we are at the s-th stage with n subjects recruited. Then, the maximum quasi-likelihood estimator $\hat{\boldsymbol{\theta}}_{n,k}$ of $\boldsymbol{\theta}_k$, for $k = 1, \ldots, K$, is the one that maximizes

$$L_k = \prod_{i=1}^{n} \mu_{i,k}^{X_{ik} Y_{i,k}} (1 - \mu_{i,k})^{X_{ik}(1 - Y_{i,k})}, \tag{3.3}$$

where $\mu_{i,k} = \mu_k(\boldsymbol{\theta}_k, \boldsymbol{\xi}_i)$. It follows that the conditional Fisher information matrix, for given $\boldsymbol{\xi}$, is

$$I_k(\boldsymbol{\theta}_k|\boldsymbol{\xi}) = \mu_k(\boldsymbol{\theta}_k, \boldsymbol{\xi})(1 - \mu_k(\boldsymbol{\theta}_k, \boldsymbol{\xi}))\boldsymbol{\xi}\boldsymbol{\xi}'.$$

Let $\hat{I}_{n,k} = n^{-1}\sum_{i=1}^{n} X_{i,k} I_k(\hat{\boldsymbol{\theta}}_{n,k}|\boldsymbol{\xi}_i)$ be the estimate of I_k for all k. If we assume $\eta = 0$, then for a K treatment problem, when the current stage is s, the new n_{s+1} designs (subjects) are chosen such that the estimated Fisher information matrix $\hat{I}_{n+n_{s+1}}$ is maximized, where $\hat{I}_{n+n_{s+1}} = \hat{I}_n + \hat{I}_{n_{s+1}}$ and $\hat{I}_{n_{s+1}} = \sum_{j=n+1}^{n+n_{s+1}} \hat{I}_j$,

$$\hat{I}_n = \begin{pmatrix} \frac{1}{n}\sum_{i=1}^{n} X_{i,1}\hat{\lambda}_{i,1}\boldsymbol{\xi}_i\boldsymbol{\xi}_i' & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \frac{1}{n}\sum_{i=1}^{n} X_{i,K}\hat{\lambda}_{i,K}\boldsymbol{\xi}_i\boldsymbol{\xi}_i' \end{pmatrix}, \qquad (3.4)$$

$$\hat{I}_j \equiv \begin{pmatrix} p_1\hat{\lambda}_{j,1}\boldsymbol{\xi}_j\boldsymbol{\xi}_j' & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & p_K\hat{\lambda}_{j,K}\boldsymbol{\xi}_j\boldsymbol{\xi}_j' \end{pmatrix},$$

and $\hat{\lambda}_{i,k} = \hat{\mu}_{i,k}(1 - \hat{\mu}_{i,k})$ for $i = 1, \ldots, n,\ j = n+1, \ldots, n+n_{s+1}$, and $k = 1, \ldots, K$.

Through numerical studies[12] provide tables with estimates of allocation proportions for several ηs and given T_n for two-stage CARA designs. For illustration purposes, we also apply the proposed method to logistic regression models with $K = 2$, and then we modify the utility function in equation (3.1) by defining the estimated target allocation proportion $\pi_k(\hat{\boldsymbol{\theta}}, \boldsymbol{\xi})$ with some function $J(t)$ symmetric at zero:

$$\pi_1(\hat{\boldsymbol{\theta}}, \boldsymbol{\xi}) = J\left(\frac{\boldsymbol{\xi}'\hat{\boldsymbol{\theta}}_1 - \boldsymbol{\xi}'\hat{\boldsymbol{\theta}}_2}{T_n}\right) \quad \text{and} \quad \pi_2(\hat{\boldsymbol{\theta}}, \boldsymbol{\xi}) = 1 - \pi_1(\hat{\boldsymbol{\theta}}, \boldsymbol{\xi}),$$

where T_n and/or η may vary in different sampling stages, serving as tuning parameters to balance efficiency and ethical consideration. Hence, $\pi_k(\hat{\boldsymbol{\theta}}, \boldsymbol{\xi})$ can vary sequentially at each stage through $\hat{\boldsymbol{\theta}}$ and T_n. In Section 3.2, we present numerical results with some suggestions for parameters T_n and η. The proposed three-stage procedure is then evaluated by the accuracy of the estimation of the treatment effect and the correct treatment allocation probability (CAP).

3.2.2 *Parameters setup and simulation results*

Consider two treatments A and B (i.e. $K = 2$) with one continuous covariate $\boldsymbol{\xi}$, and assume that the binary response and covariate $\boldsymbol{\xi}$ satisfy a logistic model, as described in equation (3.2). Assume further that intercepts of logistic models for both treatments are equal with $(\alpha_1, \alpha_2) = (0.1, 0.1)$ and regression coefficients $(\theta_1^*, \theta_2^*) = (-1, 1)$. The covariate is generated from a mixed normal distribution with means 2 and -2, and equal variance 1, and with the mixing probability 0.5. Because these two treatment groups have opposite slopes and the covariate is distributed symmetrically around the intersection point of the slopes, the correct allocation is to assign subjects with the covariate coming from a normal distribution with a positive (negative) mean to the treatment group with a positive (negative) slope.

Because the difference between treatment effects in logistic models is a function of the differences between intercepts and regression coefficients of treatments, we apply the stopping rule for the contrasts of parameters, $\boldsymbol{\gamma} = H'\boldsymbol{\theta}$, given in Section 2.1. Thus, the transpose of the contrast H is

defined as a matrix with its first row $(1, 0, -1, 0)$ and its second row $(0, 1, 0, -1)$, and the vector of parameters θ is $(\alpha_1, \theta_1^*, \alpha_2, \theta_2^*)'$.

Let precision δ be 0.3 for all simulation studies, and several cases of initial sample sizes for each treatment, m_0, are 5, 10, or 15. Several combinations of tuning parameters T_n and η are assumed: 0.5, 1, and 2 for T_n and 0, 0.1, and 1 for η. Both fixed and varying tuning parameters, T_n and η, are considered; that is, for comparison purposes, we consider cases in which both T_n and η are fixed throughout the study as well as cases in which these two parameters vary as the stage changes. For the cases with varying parameters, we let T_n and η be proportional and inversely proportional to the standard deviation of the treatment effect for a given covariate of a new observation, respectively. Findings from the simulation studies are as follows:

It is found that the average of the stopping time is very unstable when the initial sample size m_0 is small, such as 5. This is due to the unstable regression coefficient estimates using too small a sample size at the initial stage. As the initial sample size gets larger, the average of the stopping time and its variation become much smaller. The coverage probabilities of treatment differences are mostly more than the nominal level 0.95, and, for some cases, they are close to 1. This observation implies the possibility of over-sampling. That is, our procedure is conservative. Based on these findings, it is recommended that the initial sample size should not be too small in order to stabilize the stopping time in the early stage.

When $\eta = 0$, correct allocation probabilities are about 0.5 for both treatments. This case is equivalent to a randomized allocation, as there is no ethical consideration in the utility function, and our simulation setup is symmetric for two groups, so the results are as expected. As η gets larger, the correct treatment allocation gets better with similar performance for positive η. This confirms that η plays a role as a tuning parameter for ethical consideration and a small nonzero η is sufficient for correct allocation in our studies. Large correct allocation probabilities for positive η, in Table 1, illustrate that the CARA design with our three-stage sequential procedure successfully implements the idea of CARA designs. That is, our three-stage procedure achieves higher allocation proportions to the more suitable treatment for individual subjects than its non-sequential counterpart does.

For positive η, correct allocation is high (low) when $T_n = 1$, initial $T_n = 0.5$ with varying (fixed) T_n, or initial $T_n = 2$ with fixed (varying) T_n. This implies that if T_n varies depending on treatment effect variation, T_n becomes larger than the initial T_n. Thus, varying small T_n gives better allocation due to the reasonably tuned size of T_n. That is, varying large T_n gives worse allocation due to a too-liberal tuning of T_n. This emphasizes the importance of selecting a reasonably sized T_n.

Table 2 states results when the covariate effect is ignored. When the sampling is stopped, the average numbers of subjects are around 31–51, 36–38, and 30, respectively, for $m_0 = 5, 10, 15$; i.e. the sampling tends to stop very early with similar sizes for positive η, except when initial sample sizes are very small. Varying η and/or T_n does not make much difference in terms of the stopping time. The coverage probabilities of treatment differences cannot be computed because the covariate is ignored. Instead, coverage probabilities of intercept differences are given in Table 2, which are mostly larger than the nominal level 0.95 and vary more when initial sample size gets smaller.

Correct allocation probabilities in Table 2 are mostly close to 0.5, as the covariate is distributed symmetrically around the intersection in which varying η and/or T_n do not make much difference. Variations observed in cases with $m_0 = 15$ are due to small additional samples collected after the initial stage. This indicates that the response-adaptive design that ignores a significantly interacting covariate with treatment groups, fails to skew the allocation rule. Consequently, the response-adaptive design could not play an ethical role regardless of how large the initial samples are given or how big η is assumed.

Table 1. Mean (M) and standard deviation (SD) of stopping time (τ_{δ_γ}), coverage probability (CP), and correct allocation probability (CAP) of sequential 95% confidence interval estimation with $\delta = 0.3$, for three-stage design in simulation studies, when the covariate interacting with treatment groups is considered in the treatment allocation process.

			Variation		τ_{δ_γ}							Variation		τ_{δ_γ}			
m_0	T_n	η	T_{nV}	η_V	M	SD	CP	CAP	m_0	T_n	η	T_{nV}	η_V	M	SD	CP	CAP
5	0.5	0.0	N	N	238	406	0.96	0.45	10	1.0	0.1	Y	Y	79	25	0.96	0.78
5	0.5	0.0	Y	N	203	425	1.00	0.49	10	1.0	1.0	N	N	84	29	0.98	0.93
5	0.5	0.1	N	N	268	705	0.96	0.80	10	1.0	1.0	N	Y	87	34	0.99	0.91
5	0.5	0.1	N	Y	214	255	0.98	0.75	10	1.0	1.0	Y	N	84	31	0.96	0.87
5	0.5	0.1	Y	N	188	202	0.98	0.80	10	1.0	1.0	Y	Y	85	42	0.98	0.88
5	0.5	0.1	Y	Y	174	166	0.99	0.64									
5	0.5	1.0	N	N	161	117	0.98	0.80	10	2.0	0.0	N	N	91	73	0.93	0.50
5	0.5	1.0	N	Y	173	152	0.95	0.76	10	2.0	0.0	Y	N	86	44	0.99	0.43
5	0.5	1.0	Y	N	166	117	0.95	0.85	10	2.0	0.1	N	N	103	186	0.98	0.81
5	0.5	1.0	Y	Y	149	85	0.98	0.83	10	2.0	0.1	N	Y	84	44	0.95	0.77
									10	2.0	0.1	Y	N	84	32	1.00	0.68
5	1.0	0.0	N	N	230	370	0.99	0.50	10	2.0	0.1	Y	Y	81	26	0.99	0.68
5	1.0	0.0	Y	N	230	331	0.98	0.51	10	2.0	1.0	N	N	88	35	0.97	0.87
5	1.0	0.1	N	N	195	265	1.00	0.79	10	2.0	1.0	N	Y	82	24	0.97	0.87
5	1.0	0.1	N	Y	222	364	0.99	0.71	10	2.0	1.0	Y	N	84	42	0.97	0.77
5	1.0	0.1	Y	N	192	225	0.99	0.69	10	2.0	1.0	Y	Y	90	51	0.99	0.75
5	1.0	0.1	Y	Y	266	496	0.97	0.57									
5	1.0	1.0	N	N	196	248	0.97	0.85	15	0.5	0.0	N	N	70	25	0.98	0.42
5	1.0	1.0	N	Y	165	175	0.96	0.85	15	0.5	0.0	Y	N	67	26	0.98	0.47
5	1.0	1.0	Y	N	313	1220	1.00	0.78	15	0.5	0.1	N	N	64	17	0.99	0.75
5	1.0	1.0	Y	Y	213	363	1.00	0.74	15	0.5	0.1	N	Y	63	20	0.99	0.75
									15	0.5	0.1	Y	N	67	22	0.97	0.93
5	2.0	0.0	N	N	244	457	0.97	0.47	15	0.5	0.1	Y	Y	67	19	0.96	0.89
5	2.0	0.0	Y	N	331	1015	0.99	0.50	15	0.5	1.0	N	N	66	22	0.96	0.74
5	2.0	0.1	N	N	211	375	1.00	0.70	15	0.5	1.0	N	Y	66	22	0.99	0.73
5	2.0	0.1	N	Y	205	276	1.00	0.62	15	0.5	1.0	Y	N	68	27	0.97	0.95
5	2.0	0.1	Y	N	192	204	1.00	0.57	15	0.5	1.0	Y	Y	72	30	0.97	0.92
5	2.0	0.1	Y	Y	242	369	0.97	0.58									
5	2.0	1.0	N	N	219	402	0.98	0.78	15	1.0	0.0	N	N	64	18	0.97	0.44
5	2.0	1.0	N	Y	157	132	1.00	0.76	15	1.0	0.0	Y	N	65	18	0.98	0.44
5	2.0	1.0	Y	N	241	365	1.00	0.67	15	1.0	0.1	N	N	66	17	0.99	0.92
5	2.0	1.0	Y	Y	206	346	1.00	0.62	15	1.0	0.1	N	Y	75	43	0.98	0.89
									15	1.0	0.1	Y	N	64	17	0.95	0.91
10	0.5	0.0	N	N	81	31	0.97	0.41	15	1.0	0.1	Y	Y	63	17	1.00	0.88
10	0.5	0.0	Y	N	77	25	0.98	0.49	15	1.0	1.0	N	N	66	20	0.98	0.91
10	0.5	0.1	N	N	81	38	0.99	0.77	15	1.0	1.0	N	Y	67	21	1.00	0.91
10	0.5	0.1	N	Y	82	33	0.99	0.79	15	1.0	1.0	Y	N	65	19	1.00	0.93
10	0.5	0.1	Y	N	87	48	0.99	0.91	15	1.0	1.0	Y	Y	69	29	0.96	0.92
10	0.5	0.1	Y	Y	90	51	0.99	0.87									
10	0.5	1.0	N	N	83	38	1.00	0.80	15	2.0	0.0	N	N	64	16	0.98	0.43
10	0.5	1.0	N	Y	91	63	0.97	0.78	15	2.0	0.0	Y	N	57	14	0.99	0.46

(continued)

Table 1. Continued

			Variation		τ_{δ_γ}							Variation		τ_{δ_γ}			
m_0	T_n	η	T_{nV}	η_V	M	SD	CP	CAP	m_0	T_n	η	T_{nV}	η_V	M	SD	CP	CAP
10	0.5	1.0	Y	N	82	27	0.99	0.93	15	2.0	0.1	N	N	70	34	0.97	0.87
10	0.5	1.0	Y	Y	87	40	0.96	0.95	15	2.0	0.1	N	Y	68	23	1.00	0.85
									15	2.0	0.1	Y	N	67	18	0.97	0.78
10	1.0	0.0	N	N	81	27	0.97	0.48	15	2.0	0.1	Y	Y	65	23	0.97	0.75
10	1.0	0.0	Y	N	87	32	0.98	0.45	15	2.0	1.0	N	N	67	22	0.98	0.90
10	1.0	0.1	N	N	85	37	0.99	0.89	15	2.0	1.0	N	Y	72	37	1.00	0.90
10	1.0	0.1	N	Y	87	42	0.97	0.84	15	2.0	1.0	Y	N	64	20	0.97	0.84
10	1.0	0.1	Y	N	83	30	0.94	0.86	15	2.0	1.0	Y	Y	59	13	0.97	0.82

T_{nV} and η_V indicate if T_n and η vary.

Table 2. Mean (M) and standard deviation (SD) of stopping time (τ_{δ_γ}), coverage probability (CP), and correct allocation probability (CAP) of sequential 95% confidence interval estimation with $\delta = 0.3$, for three-stage design in simulation studies, when the covariate interacting with treatment groups is ignored in the treatment allocation process.

				Variation		τ_{δ_γ}							Variation		τ_{δ_γ}			
m_0	T_n	η	T_{nV}	η_V	M	SD	CP	CAP	m_0	T_n	η	T_{nV}	η_V	M	SD	CP	CAP
5	0.5	0.0	N	N	32	5	1.00	0.5	10	1.0	0.1	Y	Y	37	3	1.00	0.47
5	0.5	0.0	Y	N	32	5	1.00	0.49	10	1.0	1.0	N	N	36	2	1.00	0.51
5	0.5	0.1	N	N	42	17	0.96	0.49	10	1.0	1.0	N	Y	36	3	0.98	0.48
5	0.5	0.1	N	Y	40	9	0.98	0.5	10	1.0	1.0	Y	N	36	2	1.00	0.51
5	0.5	0.1	Y	N	44	20	0.96	0.5	10	1.0	1.0	Y	Y	36	2	1.00	0.50
5	0.5	0.1	Y	Y	36	7	1.00	0.5									
5	0.5	1.0	N	N	43	13	0.98	0.49	10	2.0	0.0	N	N	36	2	1.00	0.50
5	0.5	1.0	N	Y	44	17	0.96	0.5	10	2.0	0.0	Y	N	37	2	1.00	0.49
5	0.5	1.0	Y	N	51	28	0.85	0.49	10	2.0	0.1	N	N	36	3	1.00	0.50
5	0.5	1.0	Y	Y	39	9	1.00	0.52	10	2.0	0.1	N	Y	37	3	1.00	0.49
									10	2.0	0.1	Y	N	37	4	1.00	0.48
5	1.0	0.0	N	N	32	4	1.00	0.49	10	2.0	0.1	Y	Y	37	3	1.00	0.50
5	1.0	0.0	Y	N	33	5	1.00	0.47	10	2.0	1.0	N	N	36	3	1.00	0.49
5	1.0	0.1	N	N	37	13	0.98	0.52	10	2.0	1.0	N	Y	36	2	1.00	0.50
5	1.0	0.1	N	Y	39	20	1.00	0.51	10	2.0	1.0	Y	N	36	3	1.00	0.50
5	1.0	0.1	Y	N	37	12	0.96	0.51	10	2.0	1.0	Y	Y	36	2	1.00	0.47
5	1.0	0.1	Y	Y	36	12	1.00	0.52									
5	1.0	1.0	N	N	40	17	1.00	0.51	15	0.5	0.0	N	N	30	0	1.00	.
5	1.0	1.0	N	Y	42	23	0.96	0.49	15	0.5	0.0	Y	N	30	0	1.00	.
5	1.0	1.0	Y	N	36	6	1.00	0.49	15	0.5	0.1	N	N	30	0	1.00	0.95
5	1.0	1.0	Y	Y	38	16	1.00	0.51	15	0.5	0.1	N	Y	30	0	1.00	0.67
									15	0.5	0.1	Y	N	30	0	1.00	.
5	2.0	0.0	N	N	31	4	1.00	0.47	15	0.5	0.1	Y	Y	30	1	1.00	0.66
5	2.0	0.0	Y	N	33	5	1.00	0.52	15	0.5	1.0	N	N	30	2	1.00	0.21

(continued)

Table 2. Continued

				Variation		τ_{δ_γ}						Variation		τ_{δ_γ}			
m_0	T_n	η	T_{nV}	η_V	M	SD	CP	CAP	m_0	T_n	η	T_{nV}	η_V	M	SD	CP	CAP
5	2.0	0.1	N	N	35	12	1.00	0.51	15	0.5	1.0	N	Y	30	0	1.00	0.46
5	2.0	0.1	N	Y	33	6	0.98	0.49	15	0.5	1.0	Y	N	30	0	1.00	0.33
5	2.0	0.1	Y	N	34	5	1.00	0.53	15	0.5	1.0	Y	Y	30	0	1.00	0.60
5	2.0	0.1	Y	Y	32	5	1.00	0.51									
5	2.0	1.0	N	N	33	5	1.00	0.50	15	1.0	0.0	N	N	30	0	1.00	0.67
5	2.0	1.0	N	Y	35	12	0.98	0.50	15	1.0	0.0	Y	N	30	0	1.00	0.35
5	2.0	1.0	Y	N	33	5	1.00	0.50	15	1.0	0.1	N	N	30	0	1.00	0.33
5	2.0	1.0	Y	Y	32	5	1.00	0.49	15	1.0	0.1	N	Y	30	1	1.00	0.71
									15	1.0	0.1	Y	N	30	0	1.00	0.67
10	0.5	0.0	N	N	36	2	1.00	0.47	15	1.0	0.1	Y	Y	30	1	1.00	0.33
10	0.5	0.0	Y	N	36	2	1.00	0.47	15	1.0	1.0	N	N	30	0	1.00	0.58
10	0.5	0.1	N	N	36	3	1.00	0.51	15	1.0	1.0	N	Y	30	0	1.00	0.75
10	0.5	0.1	N	Y	36	2	1.00	0.52	15	1.0	1.0	Y	N	30	0	1.00	0.33
10	0.5	0.1	Y	N	38	10	0.98	0.52	15	1.0	1.0	Y	Y	30	0	1.00	0.63
10	0.5	0.1	Y	Y	36	3	1.00	0.49								1.00	
10	0.5	1.0	N	N	36	2	1.00	0.49	15	2.0	0.0	N	N	30	1	1.00	0.27
10	0.5	1.0	N	Y	36	3	1.00	0.52	15	2.0	0.0	Y	N	30	0	1.00	.
10	0.5	1.0	Y	N	36	2	1.00	0.49	15	2.0	0.1	N	N	30	0	1.00	0.33
10	0.5	1.0	Y	Y	36	3	0.98	0.53	15	2.0	0.1	N	Y	30	0	1.00	0.43
									15	2.0	0.1	Y	N	30	2	1.00	0.64
10	1.0	0.0	N	N	36	2	1.00	0.50	15	2.0	0.1	Y	Y	30	1	1.00	0.46
10	1.0	0.0	Y	N	37	3	1.00	0.49	15	2.0	1.0	N	N	30	1	1.00	0.32
10	1.0	0.1	N	N	36	2	1.00	0.50	15	2.0	1.0	N	Y	30	0	1.00	0.33
10	1.0	0.1	N	Y	36	3	1.00	0.50	15	2.0	1.0	Y	N	30	0	1.00	.
10	1.0	0.1	Y	N	36	3	0.98	0.51	15	2.0	1.0	Y	Y	30	0	1.00	0.35

T_{nV} and η_V indicate if T_n and η vary.

In Table 3, simulation results for two-stage designs are provided. We found that three-stage designs produced short and stable stopping times compared with two-stage designs as the former produced more stable estimates than the latter while giving similar correct allocation probabilities to the latter.

3.3 Illustrative example

Because the real data of the CARA-designed clinical trials are rarely available to the public, we illustrate a real example by modifying the study that applied the non-response-adaptive design of Cutsem et al.[20] in the CARA design while still employing the same relationship among the covariate, the treatment group, and the response variable as the one in Cutsem et al.[20] In this study, 599 patients, with either epidermal growth factor receptor-positive colorectal cancer or unresectable metastases, were randomly assigned to either receive FOLFIRI alone or in combination with cetuximab, respectively. A significant interaction was reported between the treatment group and KRAS mutation status for tumor response when analyzed using a logistic regression model. Tumor response is defined as the proportion of patients with a confirmed complete response or partial

Table 3. Mean (M) and standard deviation (SD) of stopping time (τ_{δ_γ}), coverage probability (CP), and correct allocation probability (CAP) of sequential 95% confidence interval estimation with $\delta = 0.3$, for two-stage design in simulation studies, when the covariate interacting with treatment groups is considered in the treatment allocation process.

			Variation		τ_{δ_γ}							Variation		τ_{δ_γ}			
m_0	T_n	η	T_{nV}	η_V	M	SD	CP	CAP	m_0	T_n	η	T_{nV}	η_V	M	SD	CP	CAP
5	0.5	0.0	N	N	741	3069	0.96	0.53	10	1.0	0.1	Y	Y	148	71	1.00	0.81
5	0.5	0.0	Y	N	336	235	0.96	0.46	10	1.0	1.0	N	N	141	63	1.00	0.92
5	0.5	0.1	N	N	350	408	1.00	0.81	10	1.0	1.0	N	Y	134	52	0.97	0.92
5	0.5	0.1	N	Y	354	477	1.00	0.75	10	1.0	1.0	Y	N	144	90	1.00	0.89
5	0.5	0.1	Y	N	522	1315	1.00	0.77	10	1.0	1.0	Y	Y	136	50	0.98	0.89
5	0.5	0.1	Y	Y	313	259	0.96	0.68									
5	0.5	1.0	N	N	452	1221	0.94	0.80	10	2.0	0.0	N	N	126	42	0.98	0.45
5	0.5	1.0	N	Y	408	1123	1.00	0.81	10	2.0	0.0	Y	N	145	124	0.98	0.50
5	0.5	1.0	Y	N	305	335	0.96	0.85	10	2.0	0.1	N	N	145	70	1.00	0.84
5	0.5	1.0	Y	Y	358	525	0.98	0.82	10	2.0	0.1	N	Y	134	40	0.98	0.77
									10	2.0	0.1	Y	N	136	63	1.00	0.74
5	1.0	0.0	N	N	404	598	0.98	0.50	10	2.0	0.1	Y	Y	137	51	1.00	0.64
5	1.0	0.0	Y	N	261	195	0.98	0.47	10	2.0	1.0	N	N	168	254	1.00	0.87
5	1.0	0.1	N	N	425	726	1.00	0.80	10	2.0	1.0	N	Y	166	282	1.00	0.86
5	1.0	0.1	N	Y	512	699	1.00	0.68	10	2.0	1.0	Y	N	134	58	1.00	0.78
5	1.0	0.1	Y	N	347	397	1.00	0.62	10	2.0	1.0	Y	Y	137	63	0.99	0.78
5	1.0	0.1	Y	Y	310	238	1.00	0.63									
5	1.0	1.0	N	N	345	485	0.99	0.86	15	0.5	0.0	N	N	95	32	0.96	0.44
5	1.0	1.0	N	Y	390	622	0.98	0.82	15	0.5	0.0	Y	N	94	42	1.00	0.52
5	1.0	1.0	Y	N	558	1064	1.00	0.75	15	0.5	0.1	N	N	93	36	0.99	0.76
5	1.0	1.0	Y	Y	415	461	1.00	0.72	15	0.5	0.1	N	Y	96	34	0.99	0.77
									15	0.5	0.1	Y	N	87	32	0.98	0.93
5	2.0	0.0	N	N	1321	5907	1.00	0.52	15	0.5	0.1	Y	Y	92	41	0.98	0.92
5	2.0	0.0	Y	N	522	1267	1.00	0.51	15	0.5	1.0	N	N	100	50	1.00	0.76
5	2.0	0.1	N	N	361	430	1.00	0.71	15	0.5	1.0	N	Y	96	39	0.97	0.76
5	2.0	0.1	N	Y	354	358	1.00	0.65	15	0.5	1.0	Y	N	92	36	0.99	0.93
5	2.0	0.1	Y	N	287	248	1.00	0.54	15	0.5	1.0	Y	Y	89	36	0.98	0.94
5	2.0	0.1	Y	Y	373	376	1.00	0.54									
5	2.0	1.0	N	N	457	777	1.00	0.79	15	1.0	0.0	N	N	92	46	0.98	0.54
5	2.0	1.0	N	Y	319	480	0.99	0.79	15	1.0	0.0	Y	N	98	46	1.00	0.45
5	2.0	1.0	Y	N	470	968	1.00	0.66	15	1.0	0.1	N	N	96	39	0.99	0.91
5	2.0	1.0	Y	Y	351	713	1.00	0.63	15	1.0	0.1	N	Y	97	43	0.99	0.89
									15	1.0	0.1	Y	N	90	28	0.99	0.91
10	0.5	0.0	N	N	142	48	0.98	0.54	15	1.0	0.1	Y	Y	93	41	0.99	0.89
10	0.5	0.0	Y	N	137	60	1.00	0.42	15	1.0	1.0	N	N	87	33	1.00	0.92
10	0.5	0.1	N	N	188	224	0.94	0.75	15	1.0	1.0	N	Y	89	32	1.00	0.93
10	0.5	0.1	N	Y	147	82	1.00	0.74	15	1.0	1.0	Y	N	92	53	0.99	0.91
10	0.5	0.1	Y	N	140	57	0.98	0.90	15	1.0	1.0	Y	Y	99	48	0.98	0.93
10	0.5	0.1	Y	Y	136	62	0.99	0.87									
10	0.5	1.0	N	N	149	187	1.00	0.79	15	2.0	0.0	N	N	93	36	1.00	0.45
10	0.5	1.0	N	Y	143	54	0.99	0.76	15	2.0	0.0	Y	N	88	30	1.00	0.47

(continued)

Table 3. Continued

			Variation		τ_{δ_γ}							Variation		τ_{δ_γ}			
m_0	T_n	η	T_{nV}	η_V	M	SD	CP	CAP	m_0	T_n	η	T_{nV}	η_V	M	SD	CP	CAP
10	0.5	1.0	Y	N	142	77	1.00	0.95	15	2.0	0.1	N	N	91	30	0.99	0.87
10	0.5	1.0	Y	Y	128	47	0.97	0.94	15	2.0	0.1	N	Y	89	39	0.97	0.84
									15	2.0	0.1	Y	N	88	32	0.98	0.79
10	1.0	0.0	N	N	126	45	0.98	0.44	15	2.0	0.1	Y	Y	92	32	0.98	0.77
10	1.0	0.0	Y	N	140	68	1.00	0.47	15	2.0	1.0	N	N	89	49	0.98	0.89
10	1.0	0.1	N	N	146	100	1.00	0.88	15	2.0	1.0	N	Y	93	32	1.00	0.90
10	1.0	0.1	N	Y	149	81	1.00	0.88	15	2.0	1.0	Y	N	93	39	1.00	0.83
10	1.0	0.1	Y	N	154	79	1.00	0.82	15	2.0	1.0	Y	Y	99	51	1.00	0.83

T_{nV} and η_V indicate if T_n and η vary.

Table 4. Tumor response summary to cetuximab plus FOLFIRI vs. FOLFIRI-alone treatment by KRAS status.

	Cetuximab plus FOLFIRI		FOLFIRI alone		
Tumor response	Yes	No	Yes	No	Odds ratio
KRAS population	140	137	111	152	1.38 (0.98–1.95)
Mutant KRAS	38	67	35	52	0.80 (0.44–1.45)
Wild-type KRAS	102	70	76	100	1.91 (1.24–2.93)

response, defined as a response persisting for at least 28 days. This was observed in 281 patients (46.9%) who were receiving cetuximab plus FOLFIRI as well as in 232 patients (38.7%) who received FOLFIRI alone. The adjusted significant odds ratio for a tumor response with cetuximab plus FOLFIRI treatment, when compared with FOLFIRI alone, was 1.40 (see Panel B of Figure 2 in Cutsem et al.[20]). Tumors of 348 patients (64.4%) had wild-type KRAS, and those of 192 patients (35.6%) had mutated KRAS. Table 4[20] gives the number of tumor responses and the odds ratios, along with their confidence intervals, for each KRAS mutation status.

Even though this study was performed under the random treatment allocation with possible response delays, we simulate a situation as if it were done under the three-stage CARA design. We assume logistic models for both groups with binary tumor responses, two treatment groups, and one binary covariate, KRAS status. The intercepts and regression coefficients of logistic models are chosen such that the same odds ratios as in Cutsem et al.[20] are achieved (Table 4).

Because the treatment effect is defined as a function of differences of intercepts and regression coefficients between the two treatments, we apply the stopping rule for the contrasts of parameters, $\boldsymbol{\gamma} = H'\boldsymbol{\theta}$, given in Section 2.1. Thus, the transpose of the contrast H is defined as a matrix with its first row $(1, 0, -1, 0)$ and its second row $(0, 1, 0, -1)$, and the vector of parameters θ is $(\alpha_A, \theta_A^*, \alpha_B, \theta_B^*)'$.

Precision δ is assumed to be 0.2, and the initial sample size for each treatment, m_0, is assumed to be 15, 20, and 25. Several combinations of tuning parameters T_n and η are assumed: 0.1, 0.3, and 0.5 for T_n and 0, 0.01, and 0.1 for η. As in the previous simulation study, both fixed and varying tuning parameters, T_n and η, are considered. Table 5 summarizes the results when the KRAS status is considered for the three-stage design.

Table 5. Mean (M) and standard deviation (SD) of stopping time (τ_{δ_γ}), coverage probability (CP), and correct allocation probabilities (CAP) of sequential 95% confidence interval estimation with $\delta = 0.2$, for three-stage design in an illustrative example, when the covariate interacting with treatment groups is considered in the treatment allocation process.

			Variation		τ_{δ_γ}								Variation		τ_{δ_γ}				
m_0	T_n	η	T_{nV}	η_V	M	SD	CP	CAP_w	CAP_m	m_0	T_n	η	T_{nV}	η_V	M	SD	CP	CAP_w	CAP_m
15	0.1	0.00	N	N	317	49	0.88	0.76	0.81	20	0.3	0.01	Y	Y	247	17	0.89	0.55	0.94
15	0.1	0.00	Y	N	327	58	0.70	0.71	0.75	20	0.3	0.10	N	N	245	17	0.97	0.78	0.81
15	0.1	0.01	N	N	313	51	0.98	0.34	0.70	20	0.3	0.10	N	Y	242	20	1.00	0.68	0.70
15	0.1	0.01	N	Y	317	32	0.97	0.52	0.83	20	0.3	0.10	Y	N	254	23	0.73	0.80	0.70
15	0.1	0.01	Y	N	331	45	0.73	0.44	0.85	20	0.3	0.10	Y	Y	248	20	0.92	0.87	0.65
15	0.1	0.01	Y	Y	318	46	0.88	0.58	0.70										
15	0.1	0.10	N	N	319	45	0.79	0.40	0.90	20	0.5	0.00	N	N	251	28	0.90	0.44	0.85
15	0.1	0.10	N	Y	313	29	0.95	0.50	0.73	20	0.5	0.00	Y	N	244	16	0.99	0.44	0.97
15	0.1	0.10	Y	N	322	45	0.89	0.35	0.64	20	0.5	0.01	N	N	251	20	0.75	0.71	0.83
15	0.1	0.10	Y	Y	314	49	0.97	0.37	0.81	20	0.5	0.01	N	Y	248	17	0.98	0.83	0.79
										20	0.5	0.01	Y	N	251	23	0.88	0.68	0.84
15	0.3	0.00	N	N	321	45	0.81	0.88	0.91	20	0.5	0.01	Y	Y	246	16	0.83	0.46	0.61
15	0.3	0.00	Y	N	320	35	0.64	0.70	0.79	20	0.5	0.10	N	N	251	26	0.96	0.85	0.52
15	0.3	0.01	N	N	307	29	1.00	0.77	0.76	20	0.5	0.10	N	Y	259	44	0.99	0.74	0.57
15	0.3	0.01	N	Y	326	27	0.81	0.67	0.78	20	0.5	0.10	Y	N	255	25	0.99	0.91	0.49
15	0.3	0.01	Y	N	310	38	0.96	0.73	0.70	20	0.5	0.10	Y	Y	253	34	0.98	0.82	0.44
15	0.3	0.01	Y	Y	344	50	0.86	0.80	0.74										
15	0.3	0.10	N	N	315	31	0.96	0.47	0.51	25	0.1	0.00	N	N	199	11	1.00	0.34	0.73
15	0.3	0.10	N	Y	314	47	0.89	0.55	0.73	25	0.1	0.00	Y	N	196	14	0.91	0.45	0.95
15	0.3	0.10	Y	N	328	51	0.90	0.69	0.61	25	0.1	0.01	N	N	203	14	1.00	0.18	0.73
15	0.3	0.10	Y	Y	321	30	0.72	0.69	0.52	25	0.1	0.01	N	Y	199	12	0.99	0.51	0.64
										25	0.1	0.01	Y	N	202	18	1.00	0.23	0.71
15	0.5	0.00	N	N	319	40	0.91	0.81	0.77	25	0.1	0.01	Y	Y	202	14	0.91	0.31	0.68
15	0.5	0.00	Y	N	323	36	0.77	0.68	0.72	25	0.1	0.10	N	N	202	13	0.92	0.44	0.80
15	0.5	0.01	N	N	307	27	0.98	0.75	0.65	25	0.1	0.10	N	Y	209	19	0.92	0.24	0.55
15	0.5	0.01	N	Y	320	37	0.91	0.69	0.63	25	0.1	0.10	Y	N	196	12	0.99	0.28	0.84
15	0.5	0.01	Y	N	310	36	0.91	0.68	0.66	25	0.1	0.10	Y	Y	202	14	0.99	0.29	0.58
15	0.5	0.01	Y	Y	325	38	0.89	0.80	0.75										
15	0.5	0.10	N	N	322	43	0.84	0.66	0.71	25	0.3	0.00	N	N	198	13	0.89	0.37	0.79
15	0.5	0.10	N	Y	302	42	0.98	0.70	0.54	25	0.3	0.00	Y	N	198	10	0.89	0.42	0.89
15	0.5	0.10	Y	N	345	87	0.91	0.70	0.55	25	0.3	0.01	N	N	208	19	1.00	0.79	0.77
15	0.5	0.10	Y	Y	311	32	0.91	0.76	0.67	25	0.3	0.01	N	Y	195	7	1.00	0.72	0.62
										25	0.3	0.01	Y	N	212	42	0.84	0.85	0.57
20	0.1	0.00	N	N	260	27	0.83	0.41	0.74	25	0.3	0.01	Y	Y	198	11	0.99	0.69	0.62
20	0.1	0.00	Y	N	257	28	0.88	0.37	0.72	25	0.3	0.10	N	N	203	17	0.98	0.81	0.51
20	0.1	0.01	N	N	250	15	0.97	0.31	0.76	25	0.3	0.10	N	Y	205	25	0.83	0.83	0.44
20	0.1	0.01	N	Y	249	15	0.92	0.28	0.76	25	0.3	0.10	Y	N	202	17	1.00	0.81	0.70
20	0.1	0.01	Y	N	240	12	0.91	0.20	0.95	25	0.3	0.10	Y	Y	197	13	0.98	0.79	0.67
20	0.1	0.01	Y	Y	251	22	0.81	0.37	0.76										
20	0.1	0.10	N	N	248	20	0.83	0.26	0.74	25	0.5	0.00	N	N	197	12	0.82	0.43	0.78
20	0.1	0.10	N	Y	250	24	1.00	0.32	0.73	25	0.5	0.00	Y	N	203	11	0.83	0.31	0.69

(continued)

Table 5. Continued

			Variation		τ_{δ_γ}								Variation		τ_{δ_γ}				
m_0	T_n	η	T_{nV}	η_V	M	SD	CP	CAP_w	CAP_m	m_0	T_n	η	T_{nV}	η_V	M	SD	CP	CAP_w	CAP_m
20	0.1	0.10	Y	N	249	15	0.89	0.28	0.69	25	0.5	0.01	N	N	208	16	0.75	0.86	0.80
20	0.1	0.10	Y	Y	249	21	0.98	0.36	0.71	25	0.5	0.01	N	Y	199	13	0.98	0.77	0.65
										25	0.5	0.01	Y	N	202	12	0.99	0.71	0.78
20	0.3	0.00	N	N	244	17	0.81	0.51	0.87	25	0.5	0.01	Y	Y	203	13	0.91	0.70	0.78
20	0.3	0.00	Y	N	244	19	0.91	0.65	0.76	25	0.5	0.10	N	N	207	24	0.82	0.72	0.67
20	0.3	0.01	N	N	256	30	0.99	0.62	0.51	25	0.5	0.10	N	Y	206	19	0.98	0.85	0.55
20	0.3	0.01	N	Y	245	27	0.84	0.70	0.84	25	0.5	0.10	Y	N	201	12	0.99	0.78	0.47
20	0.3	0.01	Y	N	253	22	0.83	0.72	0.63	25	0.5	0.10	Y	Y	197	10	1.00	0.88	0.60

T_{nV} and η_V indicate if T_n and η vary.

The average sample sizes when the sampling is stopped are much smaller than those in the original studies, and they become larger and less stable if the initial sample size m_0 gets smaller due to unstable regression coefficient estimates at the initial stage. This is consistent with previous simulation studies. The coverage probabilities of treatment differences vary rather largely depending on T_n or η values and become closer to 0.95 as the initial sample size m_0 gets larger and as η gets smaller.

For the wild-type KRAS population, as η gets larger, if $T_n = 0.3$ and 0.5, the correct treatment allocation gets better with similar performance for positive η. This confirms that η plays a role as a tuning parameter for ethical consideration and that a small nonzero η is sufficient for making the correct allocation. However, if $T_n = 0.1$, the correct treatment allocation worsens for positive η when T_n varies. $T_n = 0.1$ is too small to secure an accurate treatment estimate and thus inflates the small treatment effect in the utility function to cause an incorrect treatment allocation. For small T_n, non-increasing sensitivities are due to giving too much weight to the ethical consideration by increasing η before getting accurate estimates of the treatment effect. CAP also greatly decreases as η gets larger contrary to expectations when the initial sample size is small for the same reason. This implies that we do not need to sacrifice accuracy by increasing η. Small positive η is sufficient to obtain the ethical emphasis.

Most correct allocation probabilities are not increased when η is increased for the mutant KRAS population. This is also due to too much weight assigned to the ethical consideration with large η before accurately estimating treatment effects. For positive η, treatment allocation is the best for mutant KRAS when $T_n = 0.1$ due to easy detection of small treatment differences. (The odds ratio is small, 0.80, for mutant KRAS while it is large, 1.91, for wild-type KRAS.) If we have to use the same T_n for both KRAS statuses, then $T_n = 0.3$ or 0.5 are recommended because differences of correct allocation are larger when $T_n = 0.3$ or 0.5 compared with $T_n = 0.1$ for each KRAS status. In summary, small positive η is sufficiently large for ethical consideration, and reasonably sized T_n is recommended for correct allocation because too small of a T_n may yield a decreasing correct allocation from 0.5 as η is increased. Varying η does not seriously affect the correct allocation, but varying too small of a T_n may decrease the correct allocation greatly to less than 0.5.

In Table 6, results for two-stage CARA designs are provided as well. The two-stage designs produce averages of stopping times larger by 212, standard deviations of the stopping times larger by 22, differences from nominal coverage smaller by 0.01, and correct allocation probabilities larger by 0.00 than those of the three-stage designs. This occurs when we take an

Table 6. Mean (M) and standard deviation (SD) of stopping time (τ_{δ_γ}), coverage probability (CP), and correct allocation probabilities (CAP) of sequential 95% confidence interval estimation with $\delta = 0.2$, for two-stage design in an illustrative example, when the covariate interacting with treatment groups is considered in the treatment allocation process.

			Variation		τ_{δ_γ}								Variation		τ_{δ_γ}				
m_0	T_n	η	T_{nV}	η_V	M	SD	CP	CAP_w	CAP_m	m_0	T_n	η	T_{nV}	η_V	M	SD	CP	CAP_w	CAP_m
15	0.1	0.00	N	N	588	59	0.96	0.75	0.82	20	0.3	0.01	Y	Y	450	41	0.88	0.70	0.73
15	0.1	0.00	Y	N	613	100	0.87	0.73	0.74	20	0.3	0.10	N	N	448	36	0.90	0.91	0.63
15	0.1	0.01	N	N	628	81	0.80	0.24	0.81	20	0.3	0.10	N	Y	441	27	0.98	0.81	0.67
15	0.1	0.01	N	Y	602	48	0.96	0.36	0.81	20	0.3	0.10	Y	N	456	35	0.99	0.87	0.56
15	0.1	0.01	Y	N	610	93	0.89	0.32	0.74	20	0.3	0.10	Y	Y	458	41	0.83	0.66	0.46
15	0.1	0.01	Y	Y	597	86	0.83	0.47	0.68										
15	0.1	0.10	N	N	570	66	0.95	0.36	0.82	20	0.5	0.00	N	N	455	37	0.90	0.36	0.78
15	0.1	0.10	N	Y	597	61	0.88	0.63	0.81	20	0.5	0.00	Y	N	447	28	0.99	0.44	0.91
15	0.1	0.10	Y	N	601	76	0.89	0.51	0.85	20	0.5	0.01	N	N	441	28	0.88	0.66	0.80
15	0.1	0.10	Y	Y	620	81	0.74	0.12	0.85	20	0.5	0.01	N	Y	474	60	0.90	0.85	0.73
										20	0.5	0.01	Y	N	463	48	0.98	0.75	0.67
15	0.3	0.00	N	N	635	89	0.72	0.59	0.63	20	0.5	0.01	Y	Y	464	43	0.91	0.78	0.86
15	0.3	0.00	Y	N	573	36	0.91	0.82	0.84	20	0.5	0.10	N	N	453	32	0.91	0.89	0.64
15	0.3	0.01	N	N	595	65	0.89	0.66	0.69	20	0.5	0.10	N	Y	443	34	0.82	0.63	0.63
15	0.3	0.01	N	Y	586	62	0.97	0.73	0.78	20	0.5	0.10	Y	N	443	21	0.99	0.78	0.54
15	0.3	0.01	Y	N	641	114	0.73	0.75	0.63	20	0.5	0.10	Y	Y	466	70	0.82	0.75	0.51
15	0.3	0.01	Y	Y	573	51	0.98	0.80	0.72										
15	0.3	0.10	N	N	653	106	0.83	0.86	0.47	25	0.1	0.00	N	N	342	24	0.97	0.25	0.93
15	0.3	0.10	N	Y	614	63	0.87	0.83	0.38	25	0.1	0.00	Y	N	329	13	1.00	0.40	0.94
15	0.3	0.10	Y	N	628	90	0.89	0.63	0.48	25	0.1	0.01	N	N	354	24	0.91	0.43	0.71
15	0.3	0.10	Y	Y	612	82	0.72	0.60	0.71	25	0.1	0.01	N	Y	350	25	0.91	0.43	0.87
										25	0.1	0.01	Y	N	362	29	0.92	0.18	0.60
15	0.5	0.00	N	N	563	46	0.94	0.96	0.97	25	0.1	0.01	Y	Y	350	32	1.00	0.32	0.71
15	0.5	0.00	Y	N	602	85	0.80	0.83	0.83	25	0.1	0.10	N	N	353	29	1.00	0.47	0.74
15	0.5	0.01	N	N	595	68	0.81	0.61	0.67	25	0.1	0.10	N	Y	346	27	1.00	0.16	0.62
15	0.5	0.01	N	Y	629	108	0.73	0.65	0.85	25	0.1	0.10	Y	N	339	19	0.91	0.30	0.78
15	0.5	0.01	Y	N	586	51	0.90	0.64	0.92	25	0.1	0.10	Y	Y	337	23	1.00	0.61	0.66
15	0.5	0.01	Y	Y	614	65	0.83	0.75	0.74										
15	0.5	0.10	N	N	608	61	0.89	0.65	0.70	25	0.3	0.00	N	N	351	30	0.99	0.49	0.81
15	0.5	0.10	N	Y	602	84	0.91	0.73	0.44	25	0.3	0.00	Y	N	345	16	0.98	0.29	0.64
15	0.5	0.10	Y	N	602	100	0.92	0.78	0.48	25	0.3	0.01	N	N	353	20	0.97	0.87	0.45
15	0.5	0.10	Y	Y	589	64	0.97	0.83	0.51	25	0.3	0.01	N	Y	338	21	0.99	0.64	0.80
										25	0.3	0.01	Y	N	345	22	0.99	0.79	0.65
20	0.1	0.00	N	N	476	68	0.84	0.35	0.86	25	0.3	0.01	Y	Y	352	52	0.92	0.57	0.72
20	0.1	0.00	Y	N	457	37	0.97	0.44	0.88	25	0.3	0.10	N	N	347	23	0.91	0.76	0.56
20	0.1	0.01	N	N	455	36	0.90	0.33	0.86	25	0.3	0.10	N	Y	352	30	0.98	0.68	0.59
20	0.1	0.01	N	Y	485	84	1.00	0.34	0.68	25	0.3	0.10	Y	N	344	22	0.97	0.75	0.61
20	0.1	0.01	Y	N	469	73	0.89	0.33	0.77	25	0.3	0.10	Y	Y	340	27	0.98	0.76	0.59
20	0.1	0.01	Y	Y	458	32	0.99	0.14	0.81										
20	0.1	0.10	N	N	445	21	0.84	0.26	0.81	25	0.5	0.00	N	N	344	31	0.97	0.39	0.94
20	0.1	0.10	N	Y	445	33	1.00	0.47	0.60	25	0.5	0.00	Y	N	358	26	0.99	0.53	0.83

(continued)

Table 6. Continued

			Variation		τ_{δ_γ}								Variation		τ_{δ_γ}				
m_0	T_n	η	T_{nV}	η_V	M	SD	CP	CAP_w	CAP_m	m_0	T_n	η	T_{nV}	η_V	M	SD	CP	CAP_w	CAP_m
20	0.1	0.10	Y	N	453	41	0.98	0.29	0.66	25	0.5	0.01	N	N	356	30	0.74	0.69	0.53
20	0.1	0.10	Y	Y	455	26	0.83	0.33	0.80	25	0.5	0.01	N	Y	342	23	1.00	0.62	0.79
										25	0.5	0.01	Y	N	366	53	0.98	0.66	0.66
20	0.3	0.00	N	N	456	44	0.90	0.39	0.83	25	0.5	0.01	Y	Y	337	15	0.99	0.79	0.72
20	0.3	0.00	Y	N	450	22	0.82	0.52	0.88	25	0.5	0.10	N	N	351	28	0.99	0.83	0.41
20	0.3	0.01	N	N	438	23	0.96	0.77	0.75	25	0.5	0.10	N	Y	393	137	0.99	0.84	0.61
20	0.3	0.01	N	Y	448	30	1.00	0.57	0.83	25	0.5	0.10	Y	N	354	24	1.00	0.90	0.70
20	0.3	0.01	Y	N	437	24	0.98	0.77	0.66	25	0.5	0.10	Y	Y	346	19	1.00	0.90	0.84

T_{nV} and η_V indicate if T_n and η vary.

average of all of the combinations of parameters in Tables 5 and 6, respectively. These findings are the same as the one from simulation studies, which, again, provides evidence for the support of three-stage design.

Table 7 summarizes numerical results when KRAS status is ignored. When sampling is stopped, the average sample sizes are around $54-57$, $41-42$, and 50 for the total initial sample sizes equal to 30, 40, and 50, respectively; i.e. they tend to stop very early and are homogeneous. Varying η or T_n does not make much difference in terms of stopping time, as before. As mentioned, the coverage probabilities of treatment differences cannot be computed because the covariate is ignored. Instead, coverage probabilities of the intercept differences are given and are mostly larger than the nominal level 0.95, as there are few additional samples collected for larger m_0.

The CAPs for $m_0 = 25$ are not computed because no additional samples are collected due to the large initial samples. The CAPs for $m_0 = 20$ are unreliable, as only one or two additional samples were collected after the initial stage. The CAPs for $m_0 = 15$ are computed with the $14-17$ additional samples and similar observations with the ones under the response-adaptive design using fully sequential estimation.[17] This is due to the asymmetric distribution of the KRAS population on both sides of the intersection of two logistic curves for treatment effects, not because of covariate consideration. This, again, shows that response-adaptive allocation, ignoring the significantly interacting covariate with treatment groups, fails to assign subjects to their corresponding better treatment groups, regardless of initial samples or ethical consideration in the allocation process.

Remark 3.1: *We use the design of Bandyopadhyay et al.*[12] *only as an example, and our approach is not their extension. They assume that the total sample size is fixed and independent of the treatment effects. In addition, their asymptotic properties are built on the assumptions (CI and CII) about the ratio of the initial sample size to the total (non-random) sample size. Moreover, they place more emphasis on the allocation properties, and the estimate of the treatment effects in their approach becomes less important. On the contrary, in our approach, both the second and third sample sizes are random and depend on the estimate of parameters. As far as we know, the estimation and allocation properties under a randomly stopped CARA-designed clinical trial are new in this area.*[7] *Moreover, we emphasize both the estimates of the treatment effects and allocation properties. After all, the estimate of the treatment effects should be the main goal of a clinical trial. Our results are quite general and can be applied to generalized linear models with other types of CARA designs, as mentioned in Zhang et al.*[10]

Table 7. Mean (M) and standard deviation (SD) of stopping time (τ_{δ_γ}), coverage probability (CP), and correct allocation probabilities (CAP) of sequential 95% confidence interval estimation with $\delta = 0.2$, for three-stage design in an illustrative example, when the covariate interacting with treatment groups is ignored in the treatment allocation process.

			Variation		τ_{δ_γ}							Variation		τ_{δ_γ}			
m_0	T_n	η	T_{nV}	η_V	M	SD	CAP_w	CAP_m	m_0	T_n	η	T_{nV}	η_V	M	SD	CAP_w	CAP_m
15	0.1	0.00	N	N	55	3	0.50	0.50	20	0.3	0.01	Y	Y	41	2	0.84	0.10
15	0.1	0.00	Y	N	54	3	0.51	0.50	20	0.3	0.10	N	N	41	3	0.79	0.14
15	0.1	0.01	N	N	54	3	0.46	0.55	20	0.3	0.10	N	Y	41	2	0.88	0.09
15	0.1	0.01	N	Y	54	3	0.46	0.55	20	0.3	0.10	Y	N	41	2	0.87	0.11
15	0.1	0.01	Y	N	55	4	0.32	0.68	20	0.3	0.10	Y	Y	42	3	0.72	0.33
15	0.1	0.01	Y	Y	55	4	0.25	0.74									
15	0.1	0.10	N	N	55	3	0.56	0.45	20	0.5	0.00	N	N	41	1	0.00	1.00
15	0.1	0.10	N	Y	55	3	0.48	0.54	20	0.5	0.00	Y	N	41	3	0.00	1.00
15	0.1	0.10	Y	N	54	3	0.34	0.67	20	0.5	0.01	N	N	41	2	0.71	0.25
15	0.1	0.10	Y	Y	55	4	0.38	0.63	20	0.5	0.01	N	Y	41	3	0.81	0.12
									20	0.5	0.01	Y	N	41	2	0.79	0.11
15	0.3	0.00	N	N	55	3	0.48	0.51	20	0.5	0.01	Y	Y	41	2	0.88	0.20
15	0.3	0.00	Y	N	55	3	0.53	0.52	20	0.5	0.10	N	N	41	4	0.79	0.16
15	0.3	0.01	N	N	55	5	0.69	0.30	20	0.5	0.10	N	Y	42	4	0.81	0.15
15	0.3	0.01	N	Y	55	4	0.68	0.32	20	0.5	0.10	Y	N	42	3	0.81	0.10
15	0.3	0.01	Y	N	55	5	0.67	0.32	20	0.5	0.10	Y	Y	41	2	0.86	0.16
15	0.3	0.01	Y	Y	55	4	0.72	0.27									
15	0.3	0.10	N	N	54	5	0.58	0.43	25	0.1	0.00	N	N	50	0	–	–
15	0.3	0.10	N	Y	55	5	0.69	0.33	25	0.1	0.00	Y	N	50	0	–	–
15	0.3	0.10	Y	N	54	3	0.62	0.36	25	0.1	0.01	N	N	50	1	0.00	1.00
15	0.3	0.10	Y	Y	55	3	0.66	0.34	25	0.1	0.01	N	Y	50	0	–	–
									25	0.1	0.01	Y	N	50	0	–	–
15	0.5	0.00	N	N	54	3	0.48	0.51	25	0.1	0.01	Y	Y	50	0	–	–
15	0.5	0.00	Y	N	55	4	0.49	0.51	25	0.1	0.10	N	N	50	0	–	–
15	0.5	0.01	N	N	55	4	0.67	0.34	25	0.1	0.10	N	Y	50	0	–	–
15	0.5	0.01	N	Y	57	17	0.66	0.32	25	0.1	0.10	Y	N	50	0	–	–
15	0.5	0.01	Y	N	55	5	0.66	0.35	25	0.1	0.10	Y	Y	50	0	–	–
15	0.5	0.01	Y	Y	55	4	0.58	0.42									
15	0.5	0.10	N	N	57	17	0.67	0.31	25	0.3	0.00	N	N	50	0	–	–
15	0.5	0.10	N	Y	55	5	0.68	0.33	25	0.3	0.00	Y	N	50	0	–	–
15	0.5	0.10	Y	N	55	4	0.65	0.31	25	0.3	0.01	N	N	50	0	–	–
15	0.5	0.10	Y	Y	55	5	0.68	0.34	25	0.3	0.01	N	Y	50	0	–	–
									25	0.3	0.01	Y	N	50	0	–	–
20	0.1	0.00	N	N	41	3	0.00	1.00	25	0.3	0.01	Y	Y	50	0	–	–
20	0.1	0.00	Y	N	41	2	0.00	1.00	25	0.3	0.10	N	N	50	0	–	–
20	0.1	0.01	N	N	42	4	0.19	0.86	25	0.3	0.10	N	Y	50	0	–	–
20	0.1	0.01	N	Y	41	3	0.28	0.81	25	0.3	0.10	Y	N	50	0	–	–
20	0.1	0.01	Y	N	41	3	0.13	0.83	25	0.3	0.10	Y	Y	50	0	–	–
20	0.1	0.01	Y	Y	41	2	0.09	0.89									
20	0.1	0.10	N	N	42	4	0.24	0.67	25	0.5	0.00	N	N	50	0	–	–
20	0.1	0.10	N	Y	41	3	0.21	0.77	25	0.5	0.00	Y	N	50	0	–	–

(continued)

Table 7. Continued

			Variation		τ_{δ_γ}							Variation		τ_{δ_γ}			
m_0	T_n	η	T_{nV}	η_V	M	SD	CAP_w	CAP_m	m_0	T_n	η	T_{nV}	η_V	M	SD	CAP_w	CAP_m
20	0.1	0.10	Y	N	41	3	0.04	0.79	25	0.5	0.01	N	N	50	0	–	–
20	0.1	0.10	Y	Y	41	3	0.22	0.80	25	0.5	0.01	N	Y	50	0	–	–
									25	0.5	0.01	Y	N	50	0	–	–
20	0.3	0.00	N	N	41	2	0.00	1.00	25	0.5	0.01	Y	Y	50	0	–	–
20	0.3	0.00	Y	N	42	3	0.00	1.00	25	0.5	0.10	N	N	50	0	.	.
20	0.3	0.01	N	N	41	2	0.68	0.14	25	0.5	0.10	N	Y	50	0	.	.
20	0.3	0.01	N	Y	41	3	0.78	0.03	25	0.5	0.10	Y	N	50	0	.	.
20	0.3	0.01	Y	N	41	2	0.80	0.16	25	0.5	0.10	Y	Y	50	0	.	.

$T_{nV}$and η_V indicate if T_n and η vary.

4 Discussion

We have studied the application of the three-stage sequential sampling procedure to the CARA-designed clinical trial. We show that under such a multiple-stage scheme, the estimate of the vector of parameters is strongly consistent with the prescribed accuracy, and the properties of the allocation rule maintains its nice properties, just like those of its non-sequential counterparts discussed in the literature. Thus, the proposed method retains not only the nice characteristics of CARA design but also the significance of statistical properties.

At the beginning of a study, the proposed method allocates more samples, according to the needs of estimation, to improve the precision of the estimation of treatment effects. That is, it places more emphasis on the information part of a utility function when the estimation of treatment effects is unreliable. When the estimate of treatment effects becomes stable at the later stage of the study, as the sample size becomes large, the proposed method moves the weight gradually toward the ethical part of a utility function. Thus, based on the estimated information on treatment effects, we tend to allocate more patients to the better treatments depending on their covariate statuses. The advantage of this method is that we have more flexibility to sequentially alter the parameters of the utility function as sampling goes on, unlike the two-stage design in Bandyopadhyay et al.,[12] such that the needs for estimating treatment effects and the ethical consideration can be uniformly fulfilled. In addition, our numerical studies provide useful information on choices of parameters in the proposed method; fixed small positive η is sufficiently large for ethical consideration, and reasonably sized m_0 and T_n is recommended for correct allocation. There is no rule of thumb on how to choose these two parameters because it depends on a problem. We would recommend performing some simulation studies in advance based on possible parameter combinations. Numerically, we also found that when significant interaction between treatment and covariate is ignored, the study stops early with only a few additional samples collected. This makes it difficult to try skewed allocation to the better treatment and results in incorrect treatment allocation regardless of covariate values.

From a practical prospect, the multi-stage methods are usually more convenient than fully sequential methods due to operational convenience, such as the small number of evaluation times. However, these kinds of methods may require more samples because the information of treatment effects are updated less often. As far as we know, a theoretical guideline regarding how many stages should be adopted for a clinical trial is still lacking in the literature. The number of stages may be chosen depending on other practical issues, which is beyond the discussion here.

However, it is clear that the method proposed here can be extended to more than three sampling stages. It is also easy to see that when more stages are used, the procedure gets closer to the fully sequential one and becomes less convenient in practice. For example, we may need to re-estimate the sample size more often.

5 Supplementary material

The proof of Theorem 2.1 relies on the methods in Woodroofe[21] and Chang.[22] The reader is referred to the online supplementary material for the technical appendix of the proof of Theorem 2.1.

Funding

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2010-0006158), and National Science Council (NSC), Taiwan (99-2118-M-001-001-MY2).

References

1. Pocock SJ and Simon R. Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. *Biometrics* 1975; **31**: 103–115.
2. Wei LJ. A class of designs for sequential clinical trials. *J Am Stat Assoc* 1977; **72**: 382–386.
3. Wei LJ. The adaptive biased coin design for sequential experiments. *Ann Stat* 1978; **6**: 92–100.
4. Whitehead J. *The design and analysis of sequential clinical trials*, 2nd rev. ed. New York: John Wiley & Sons, 1997.
5. Jennison C and Turnbull B. *Group sequential methods with applications to clinical trials*. New York: Chapman and Hall, 2000.
6. Hu F, Rosenberger WF and Zhang L. Asymptotically best response-adaptive randomization procedures. *J Stat Plan Inference* 2006; **136**: 1911–1922.
7. Hu F and Rosenberger WF. *The theory of the response adaptive randomization in clinical trials*. London: Wiley Interscience, 2006.
8. Rosenberger WF, Vidyashankar AN and Agarwal D. Covariate-adjusted response-adaptive designs for binary response. *J Biopharm Stat* 2001; **1**: 227–236.
9. Antognini AB and Giovagnoli A. Compound optimal allocation for individual and collective ethics in binary clinical trials. *Biometrika* 2010; **97**: 935–946.
10. Zhang L, Hu F, Cheung S, et al. Asymptotic properties of covariate-adjusted response-adaptive designs. *Ann Stat* 2007; **35**: 1166–1182.
11. Biswas A, Park E and Bhattacharya R. Covariate-adjusted response-adaptive designs for longitudinal treatment responses: PEMF trial revisited. *Stat Meth Med Res* 2012; **21**: 379–392.
12. Bandyopadhyay U, Biswas A and Bhattacharya R. A covariate adjusted two-stage allocation design for binary responses in randomized clinical trials. *Stat Med* 2007; **26**: 4386–4399.
13. McCullagh P and Nelder JA. *Generalized linear models*, 2nd ed. London, UK: Chapman and Hall, 1989.
14. Stein C. A two-sample test for a linear hypothesis whose power is independent of the variance. *Ann Math Stat* 1945; **16**: 243–258.
15. Siegmund D. *Sequential analysis*. New York: Springer-Verlag, 1985.
16. Chow YS and Teicher H. *Probability theory*, 2nd ed. New York, USA: Springer, 1988.
17. Chang Y-cI and Park E. Sequential estimation for covariate-adjusted response-adaptive designs. *J Korean Stat Soc* 2013; **42**: 105–116.
18. Atkinson AC and Biswas A. Bayesian adaptive biased-coin designs for clinical trials with normal responses. *Biometrics* 2005; **61**: 118–125.
19. Bandyopadhyay U and Biswas A. Adaptive designs for normal responses with prognostic factors. *Biometrika* 2001; **88**: 409–419.
20. Cutsem E, Kohne C-H, Hitre E, et al. Cetuximab and chemotherapy as initial treatment for metastatic colorectal cancer. *N Engl J Med* 2009; **360**: 1408–1417.
21. Woodroofe M. *Nonlinear renewal theory in sequential analysis*. Philadelphia, PA: Society for Industrial and Applied Mathematics, 1982.
22. Chang Y-cI. Strong consistency of maximum quasi-likelihood estimate in generalized linear models via a last time. *Stat Prob Lett* 1999; **45**: 237–246.