Journal of Statistical Research *https://doi.org/10.3329/jsr.v57i12.72967*
2023, Vol. 57, No. 1-2, pp. 21-33 *ISSN 0256 - 422 X*

INCORPORATING TREATMENT-COVARIATE INTERACTION IN DESIGNING A COVARIATE ADJUSTED RESPONSE ADAPTIVE ALLOCATION FOR BINARY RESPONSE TRIALS

RAHUL BHATTACHARYA*
Department of Statistics, University of Calcutta, Kolkata, India
Email: rahul_bhattya@yahoo.com

ATANU BISWAS
Applied Statistics Unit, Indian Statistical Institute, Kolkata, India
Email: atanu@isical.ac.in

SUMMARY

Assuming binary treatment outcomes, a covariate adjusted response adaptive (CARA) design is developed for multi-arm clinical trials assuming the presence of treatment-covariate interaction. The proposed allocation skews the allocation dynamically according to the performance of the treatments and the covariate profile of any incoming subject. Relevant design and precision based measures of the performance are further investigated empirically, which established the usefulness of the proposed allocation design over its competitors.

Keywords and phrases: Response adaptive design, Covariate adjusted allocation, Treatment-covariate interaction.

AMS Classification: 62P10

1 Introduction

With the passage of the pandemic situation, the phrases *Clinical Trials*, *Efficacy*, *Effectiveness*, among others, are attracting public attention. Everyone's concern lies in the successful development of an effective vaccine (or a *treatment*) to make life normal, as earlier. Clinical trials are, therefore, the new form of saviour for their ability to identify an effective treatment in a scientific way. Broadly, a clinical trial identifies the best treatment through the assignment of eligible subjects to a set of prospective treatments.

Usually *Fixed Randomization* is adopted for treatment assignment, where subjects are assigned either of the treatments randomly with fixed probability. But such a randomization procedure suffers from the subjectivity as the allocation ratio (e.g. 1:1 or 2:1) is fixed on an ad hoc basis with a view to ensure higher allocation to the prospective treatment. But if the prospective treatment performs poorly, more subjects are at a risk of developing adverse effects. For an example, in a recent trial on safety and efficacy evaluation of *Sputnik V* (Logunov et al., 2021), the allocation ratio is kept

* Corresponding author

fixed at 3:1 throughout the trial. Naturally adaptive allocation designs are always the better option for their ability to update the allocation dynamically in favor of the treatment doing better for every incoming patient.

A number of adaptive allocation designs have been developed in the last few decades assuming homogeneous subjects. But response of a subject is always influenced by his/her covariate information. Information on covariate is, in general, available at the time of recruitment and hence a reasonable allocation design must be *personalized*, that is, *covariate adjusted* to decide the current patient assignment by utilizing his/her covariate profile. Precisely, if the available response, allocation and covariate data is used altogether for the allocation of incoming subjects, the resulting allocation is termed covariate adjusted response adaptive or CARA (Hu and Rosenberger, 2006). A brief account of different opinions on the need for incorporating the covariate information in the design phase of clinical trials together with recommendations can be found in Rosenberger and Sverdlov (2008). However, for the development of the current work, we consider multiple treatments, binary treatment outcome and the presence of covariate information. Instances of CARA designs with logistic regression based binary response models can be found in Basak et al. (2009), Bandyopadhyay and Bhattacharya (2012), among others. But these designs are actually developed to favour the better performing treatment considering some treatment specific effective measure like failure and odds. For example, a treatment is declared better performing, if it possesses the lowest failure rate. That is, the better treatment is decided through a comparison of treatment specific performance measures. However, instead of comparing treatment specific performance measures, a better option could be to use a relative performance based effective measure, where simultaneous comparison of all the treatments under consideration is performed. In a recent work, Biswas and Bhattacharya (2018) developed a CARA allocation design under the same framework by combining relevant ethical and precision metrics and investigated from the viewpoint of ethics and precision. Another CARA design can be found in the work of Villar and Rosenberger (2018), which was developed from a Bayesian viewpoint using the Gittins index solution to the classic multi-armed bandit problem.

Existence of treatment-covariate interaction (Zhu, 2015), that is, when effectiveness of treatments vary sensibly for different types of subjects, is another important concern in clinical trials. For example, in the real clinical trial on Stroke Prevention in Atrial Fibrillation study (Hart et al., 2003), a significant difference between aspirin and placebo in reducing the number of strokes among patients receiving anticoagulation (a binary covariate) is observed. However, such a difference becomes insignificant among the patients without receiving anticoagulation. Interestingly, if we ignore the interaction between the treatments and the covariate (i.e. anticoagulation status), aspirin would be prescribed for every patient to reduce stroke, and naturally, leave a certain fraction of patients missing the appropriate treatment. Thus a reasonable allocation design must be sensitive to the presence of such an interaction. However, none of the available CARA designs are assessed theoretically from the viewpoint of treatment-covariate interaction. A full analysis from the viewpoint of making inference for few existing CARA designs under the presence of treatment-covariate interaction is found in the work of Zhu (2015). But, the main objective of such work was to derive large sample results related to CARA designs rather than developing new CARA allocation designs.

In contrast to the above, assuming binary treatment outcomes and multiple treatments, we focus on developing a CARA allocation design, which will be sensitive to the presence of different kinds of treatment-covariate interaction. Specifically, we start from defining an effectiveness measure for treatments with binary outcome in the presence of covariates and subsequently derive an allocation function. The development of an intuitive measure of relative treatment effect with the follow up CARA allocation design and related asymptotic results are described and studied thoroughly in Section 2. Sensitivity of few existing CARA allocation designs are further investigated in the light of treatment-covariate interaction. Empirical investigation of the design and precision measures of the proposed and competing allocation designs are provided in Section 3. Redesign results of a real life clinical trial are further added in Section 4. Finally, Section 5 ends with a discussion of the relevant and upcoming issues.

2 Developing the Covariate Adjusted Allocation

2.1 A relative treatment effectiveness measure

Consider a trial involving $t(\geq 2)$ treatments indicated by $1, 2, \ldots, t$. Incoming subjects are assigned to one of these treatments and their responses are observed. Denoting the response of the ith subject assigned to treatment k by Y_{ki} and the associated d component vector of covariate information by $\mathbf{Z}_i$, we assume the binary response model

$$P(Y_{ki} = 1|\mathbf{Z}_i) = p_k(\mathbf{Z}_i),$$

where $p_k(\mathbf{Z})$ is the unknown success probability for the k th treatment and $p_k(\mathbf{Z})$ is higher for higher $\mathbf{Z}$. As a pre-phase to the development of an allocation function, we start with developing a reasonable treatment effectiveness measure. For fixed $\mathbf{Z}$, the most common measure is $q_k(\mathbf{Z}) = 1 - p_k(\mathbf{Z})$, the failure probability of treatment k. However, such a choice is treatment specific. That is, a lower $q_k(\mathbf{Z})$ does not imply that treatment k is the best unless it is the lowest. Thus a treatment effectiveness measure should consider individual as well as the relative performance among the available treatments. Simultaneous consideration of the individual and relative performances within the same framework for binary response clinical trials can be found in recent considerations of Yin et al. (2012), Trippa et al. (2012) and Wason and Trippa (2014), but from a Bayesian perspective. Specifically, for each experimental treatment, they derived the posterior probability that the experimental treatment is better than a control treatment and normalized suitably to derive the randomization probabilities.

We, therefore, consider both treatment specific and relative performance of a treatment among the available treatments to define a meaningful treatment effectiveness measure. For the development, we consider a hypothetical situation with *t* patients, each having the same covariate profile $\mathbf{Z}$ and t patients are assigned to one of the t treatments exclusively. If $Y_k(\mathbf{Z})$ denotes the potential outcome of the patient assigned to treatment $k, k = 1, 2, \ldots, t$, treatment k is worst performing if its response is lower than the lowest among the available responses. That is, treatment k is worst performing, if $Y_k(\mathbf{Z}) \leq min_{s \neq k} Y_s(\mathbf{Z})$. Naturally, treatment k can be regarded as better performing

if the converse $Y_k(\mathbf{Z}) > min_{s\neq k} Y_s(\mathbf{Z})$ holds. Since, the response is either 0 or 1, some responses can be identical and hence, the event of tie $Y_k(\mathbf{Z}) = min_{s\neq k} Y_s(\mathbf{Z})$ occurs with positive probability. Consequently, we use a weight $\frac{1}{2}$ for tie and define the tie adjusted quantity

$$\pi_k(\mathbf{Z}) = P(Y_k(\mathbf{Z}) > min_{s(\neq k)} Y_s(\mathbf{Z})|\mathbf{Z}) + \frac{1}{2} P(Y_k(\mathbf{Z}) = min_{s(\neq k)} Y_s(\mathbf{Z})|\mathbf{Z})$$

to measure the effectiveness of treatment k. Further, for given $\mathbf{Z}$, $min_{s(\neq k)} Y_s(\mathbf{Z})$ has a Bernoulli distribution with success probability $P(Y_s(\mathbf{Z}) = 1 \forall s \neq k|\mathbf{Z}) = \prod_{s\neq k} p_s(\mathbf{Z})$, and hence we get the equivalent expression

$$\pi_k(\mathbf{Z}) = \frac{1}{2} + \frac{1}{2}\Big\{p_k(\mathbf{Z}) - \prod_{s\neq k} p_s(\mathbf{Z})\Big\}.$$

Now to justify the suitability of the $\pi_k(\mathbf{Z})$ as an effectiveness measure, we evaluate it from the viewpoints of treatment superiority and treatment-covariate interaction. First assume that treatment k is best for fixed covariate value Z, that is, $p_k(\mathbf{Z}) > p_{k'}(\mathbf{Z})$ for all $k \neq k'$. Then a simple algebra shows

$$\pi_k(\mathbf{Z}) - \pi_{k'}(\mathbf{Z}) = \frac{1}{2}\Big\{p_k(\mathbf{Z}) - p_{k'}(\mathbf{Z})\Big\}\Big\{1 + \prod_{s\neq(k,k')} p_s(\mathbf{Z})\Big\},$$

and hence $\pi_k(\mathbf{Z}) > \pi_{k'}(\mathbf{Z})$ whenever $p_k(\mathbf{Z}) > p_{k'}(\mathbf{Z})$ is satisfied. Thus $\pi_k(\mathbf{Z})$ is the highest for the best treatment and $\pi_k(\mathbf{Z})$'s are ordered according to treatment effectiveness. Next to investigate the effect of presence of treatment-covariate interaction, we assume $p_k(\mathbf{Z}) > p_{k'}(\mathbf{Z})$ but $p_k(\mathbf{Z}') < p_{k'}(\mathbf{Z}')$ for some covariate $\mathbf{Z} \neq \mathbf{Z}'$. Then it is found that $\pi_k(\mathbf{Z}) > \pi_{k'}(\mathbf{Z})$ but $\pi_k(\mathbf{Z}') < \pi_{k'}(\mathbf{Z}')$ and hence the defined measure $\pi_k(\mathbf{Z})$ is suitably sensitive to the presence of treatment-covariate interaction. However, the defined measure is not normalised and, therefore, we suggest the normalised version $\rho_k(\mathbf{Z}) = \frac{\pi_k(\mathbf{Z})}{\sum_{s=1}^t \pi_s(\mathbf{Z})}$ to define the allocation probability to treatment k in ideal situation.

Now we shall judge few existing allocation designs in the light of treatment-covariate interaction. First of all, we consider the heuristic multi-treatment extension of the odds based allocation design of Rosenberger et al. (2001) by Villar and Rosenberger (2018), which assigns to treatment k using the allocation function

$$\rho_{1k}(\mathbf{Z}) = \frac{[q_k(\mathbf{Z})/p_k(\mathbf{Z})]}{\sum_{s=1}^t q_s(\mathbf{Z})/p_s(\mathbf{Z})}.$$

Then it is easy to observe that $p_k(\mathbf{Z}) > p_{k'}(\mathbf{Z})$ implies $\rho_{1k}(\mathbf{Z}) > \rho_{1k'}(\mathbf{Z})$ whereas $p_k(\mathbf{Z}') < p_{k'}(\mathbf{Z}')$ gives $\rho_{1k}(\mathbf{Z}') > \rho_{1k'}(\mathbf{Z}')$. Thus the above allocation function lacks sensitivity when treatment-covariate interaction is present.

Next, we consider the allocation function

$$\rho_{2k}(\mathbf{Z}) = \frac{\sqrt{q_k(\mathbf{Z})p_k(\mathbf{Z})}}{\sum_{s=1}^t \sqrt{q_s(\mathbf{Z})p_s(\mathbf{Z})}},$$

which corresponds to Neyman optimum allocation of survey sampling. Then $\rho_{2k}(\mathbf{Z}) > \rho_{2k'}(\mathbf{Z})$ whenever $p_k(\mathbf{Z}) > p_{k'}(\mathbf{Z})$ and $p_k(\mathbf{Z}) + p_{k'}(\mathbf{Z}) < 1$ holds. On the other hand $\rho_{2k}(\mathbf{Z}') < \rho_{2k'}(\mathbf{Z}')$

under $p_k(\mathbf{Z}') < p_{k'}(\mathbf{Z}')$ and $p_k(\mathbf{Z}')+p_{k'}(\mathbf{Z}') > 1$. Naturally, the above allocation design is sensitive to treatment-covariate interaction only when some additional conditions are satisfied.

Finally, we consider the optimal allocation function of Biswas and Bhattacharya (2018)

$$\rho_{3k}(\mathbf{Z}) = \frac{\sqrt{w_k}\left\{q_k(\mathbf{Z})p_k(\mathbf{Z})\gamma_k(\mathbf{Z})\right\}^{-1/2}}{\sum_{s=1}^{t}\sqrt{w_s}\left\{q_s(\mathbf{Z})p_s(\mathbf{Z})\gamma_s(\mathbf{Z})\right\}^{-1/2}},$$

where w_k are positive weights based on some inferential objective and $\gamma_k(> 0)$ are clinically relevant criteria. Specifically, for the choice $\gamma_k(\mathbf{Z}) = \frac{q_k(\mathbf{Z})}{p_k(\mathbf{Z})}$, the above allocation function reduces to

$$\rho_{3k}(\mathbf{Z}) = \frac{\sqrt{w_k}/q_k(\mathbf{Z})}{\sum_{s=1}^{t}\sqrt{w_s}/q_s(\mathbf{Z})}.$$

Since, for any $\mathbf{Z}$,

$$\frac{\rho_{3k}(\mathbf{Z})}{\rho_{3k'}(\mathbf{Z})} = \frac{\sqrt{w_{k'}}q_{k'}(\mathbf{Z})}{\sqrt{w_k}q_k(\mathbf{Z})},$$

it follows that $p_k(\mathbf{Z}) > p_{k'}(\mathbf{Z})$ implies $\rho_{3k}(\mathbf{Z}) > \rho_{3k'}(\mathbf{Z})$ provided $w_k < w_{k'}$ is satisfied. Further, $p_k(\mathbf{Z}') < p_{k'}(\mathbf{Z}')$ together with $w_k > w_{k'}$ implies $\rho_{3k}(\mathbf{Z}') < \rho_{3k'}(\mathbf{Z}')$. Therefore, the above allocation design can not act sensibly in the presence of treatment-covariate interaction for arbitrary set of weights. In fact, the same conclusion continues for most of the clinically meaningful choices of γ_k.

2.2 Implementation

It is easy to observe that the defined allocation probability $\rho_k(\mathbf{Z})$ depends on $p_k(\mathbf{Z}) = p_k(\boldsymbol{\eta}_k, \mathbf{Z})$, that is, on the unknown parameters of the response distributions (i.e. $\boldsymbol{\eta}_k$) and the vector of covariate information $\mathbf{Z}$. Therefore, to assign patients according to their covariate profile using all available response, allocation and covariate data, we suggest to replace $\mathbf{Z}$ by the current patient's covariate and substitute the estimates of $\boldsymbol{\eta}_k$ based on the available data. Therefore, if the $(i+1)$ th subject is the current subject, we suggest to replace $\mathbf{Z}$ by the current patient's covariate $\mathbf{Z}_{i+1}$ and plug in $\boldsymbol{\eta}_k$ by an appropriate estimate $\hat{\boldsymbol{\eta}}_{ki}$, based on the available data for the i previously assigned subjects. For the purpose of estimation of the unknown parameters, we suggest to use sequentially updated maximum likelihood estimates. If $\widehat{\rho}_{ki}(\mathbf{Z}_{i+1})$ denotes the estimated allocation probability for treatment k based on the relevant data on i assigned subjects at $\mathbf{Z} = \mathbf{Z}_{i+1}$, the allocation design can be described by the collection of conditional allocation probabilities

$$P(\delta_{ki+1} = 1|Y_{sj}, \delta_{sj}, \mathbf{Z}_j, j \leq i, s = 1(1)t, \mathbf{Z}_{i+1}) = \widehat{\rho}_{ki}(\mathbf{Z}_{i+1}), k = 1, 2, \ldots, t,$$

where δ_{ki} is the treatment indicator (=1 if treatment k is assigned and =0 otherwise) of the ith subject. However, to get initial estimates of the parameters, we assign n_0(fixed in advance) subjects to each treatment arm using some restricted randomization procedure and calculate the relevant estimates to start adaptive allocation from the $tn_0 + 1$ th subject onwards.

2.3 Limiting proportions

To facilitate the theoretical assessment of the proposed allocation procedure, we provide below the limiting values of the observed allocation and failure proportions. The observed allocation proportion to treatment k can be expressed as $n^{-1}N_{kn}$ with $N_{kn} = \sum_{i=1}^{n} \delta_{ki} \;\; k = 1, 2, \ldots, t$ and the observed failure proportion by $F_n = n^{-1} \sum_{k=1}^{t} \sum_{i=1}^{n} (1 - Y_{ki})$. Then it follows from the results of Zhang et al. (2007) that the limiting allocation proportion to treatment k is $E_{\mathbf{Z}} \{\rho_k(\mathbf{Z})\}$ and that for the observed failure proportion is $E_{\mathbf{Z}} \left\{ \sum_{j=1}^{t} \rho_j(\mathbf{Z}) q_j(\mathbf{Z}) \right\}$, where $\mathbf{Z}$ has some d variate distribution with positive definite dispersion matrix. However, for better explanation of the role of covariate information, we consider a categorical covariate $\mathbf{Z}$ and based on n assignments, define $N_{kn}(\mathbf{z})$, the number of subjects having covariate category $\mathbf{z}$ assigned to treatment k and the total number of subjects with covariate $\mathbf{z}$ by $N_n(\mathbf{z})$. Then for a given covariate category $\mathbf{z}$, the conditional proportion of subjects assigned to treatment k can be expressed as $\frac{N_{kn}(\mathbf{z})}{N_n(\mathbf{z})}$, which converges (Zhang et al., 2007) to $\rho_k(\mathbf{z})$, provided $P(\mathbf{Z} = \mathbf{z}) > 0$.

3 Evaluating the Performance

3.1 Performance measures

For the evaluation of the proposed CARA design, we consider $t = 3$ treatments, a single binary covariate Z with $P(Z = z) = p_z = \frac{1}{2}, z = -1, 1$. Although the CARA design is developed for a general class of response models, performance evaluation requires specification of a response model. We, therefore, consider the popular logistic model of responses, that is,

$$\log \left\{ \frac{P(Y_{ki} = 1|Z_i)}{1 - P(Y_{ki} = 1|Z_i)} \right\} = \alpha_k + \beta_k Z_i, k = 1, 2, 3, i \geq 1,$$

where Z_i is the covariate associated with the ith patient and $\boldsymbol{\eta}_k = (\alpha_k, \beta_k)^T, k = 1, 2, 3$ are the unknown parameters. For performance evaluation of the proposed CARA design, we consider the following measures:

- The expected overall and conditional allocation proportions (denoted by EAP) to different treatments together with the standard deviations;
- The expected overall and conditional proportions of failure (denoted by EFP) together with the relevant standard deviations;
- The powers of the LR test at the 5% significance level for testing the homogeneity hypothesis H_0: $\alpha_1 = \alpha_2 = \alpha_3, \beta_1 = \beta_2 = \beta_3$ against H_a: Atleast one pair of $\alpha_k' s$ and/or $\beta_k' s$ are different; and
- The type I error rates.

For the calculation of power, we consider the statistic λ_n, which is the logarithm of the LR criterion, multiplied by -2. Naturally, we reject the null hypothesis for large values of λ_n. Moreover, under

the null hypothesis, it follows from Zhang et al. (2007) that the asymptotic distribution of λ_n is chi square with 4 degrees of freedom. Therefore, for the calculation of type I error rate, we simulate the quantity $P_{H_0}(\lambda_n > \chi^2_{4,.05})$, where $\chi^2_{4,.05}$ is the upper 5% points of a chi square distribution with 4 degrees of freedom.

In order to judge the utility of incorporating treatment-covariate interaction, we consider the already mentioned odds ratio based CARA allocation design of Villar and Rosenberger (2018). It is already established that such allocation design fails to assign subjects sensibly in the presence of treatment-covariate interaction. As another competitor, we consider the popularly used complete randomization (CR), where each treatment is assigned with equal probability $\frac{1}{3}$.

3.2 Simulation study

Assessing the proposed CARA design in small samples is difficult as there is no dedicated software for the purpose. However, we have written programs in `R` using in-built `glm` function conveniently to compute the performance measures. Specifically, we carried out a detailed simulation study with $n = 120$ to extract the performance measures of the proposed CARA design empirically. We start with assigning twelve subjects, two for each category, to each treatment arm and obtain the initial parameter estimates through numerical methods. For the allocation, we update the estimates after each response using numerical procedures and substitute these in the allocation function to obtain the allocation probabilities for incoming subjects. For simulation, we have considered three groups (i.e. groups I, II and III) of parameter configurations ensuring different types of treatment-covariate interaction and report them in Table 1 together with the expected success rates (ESR) $E_Z\{p_k(Z)\}, k = 1, 2, 3$. Configurations under I assume that $\beta_1 = \beta_2 = \beta_3$ and ensure that treatment 1 is most successful for each of the covariate values. On the other hand, set of parameters under II assumes that $\alpha_1 = \alpha_2 = \alpha_3$ and ensures that the ESR for treatment 1 is the highest under the alternative. However, the success probability for $Z = -1$ decreases and that for $Z = +1$ increases as we move from the null to an alternative configuration. Thus the parameter values under II ensures qualitative interaction. Finally, we vary both sets of $(\alpha_1, \alpha_2, \alpha_3)$ and $(\beta_1, \beta_2, \beta_3)$ to get configurations under III. As earlier, the parameters are set to have highest expected success rate for treatment 1 producing the least and the highest success rates for $Z = -1$ and $Z = +1$, respectively, under the alternatives.

For any group of parameter configurations, the null hypothesis H_0 together with three alternatives $H_{ai}, i = 1, 2, 3$ are considered. For the computation of power, we consider H_0 as the null hypothesis and calculate power considering these alternatives. But the design properties like expected allocation and failure proportions are explored only under the alternatives. The numerical features for the proposed allocation design are explored in Table 2, whereas those for the Odds based design are provided in Table 3. However, the observations arising out of these tables are discussed and explained in details in the Appendix.

Table 1: Parameter configuration with ESR for logistic response model

	Hypothesis	Parameters	Expected Success Rates (ESR)
Configuration		$(\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3)$	$(E_Z p_1(Z), E_Z p_2(Z), E_Z p_3(Z))$
	H_0	(-.50,-.50,-.50,-1.0,-1.0,-1.0)	(.40,.40,.40)
I	H_{a1}	(.52,.52,-.05,-1.0,-1.0,-1.0)	(.60,.60,.40)
	H_{a2}	(1.68,.52,-.50,-1.0,-1.0,-1.0)	(.80,.60,.40)
	H_{a3}	(1.68,1.68,-.50,-1.0,-1.0,-1.0)	(.80,.80,.40)
	H_0	(1.0,1.0,1.0,.85,.85,.85)	(.55,.55,.55)
II	H_{a1}	(1.0,1.0,1.0,3.0,.85,.85)	(.70,.55,.55)
	H_{a2}	(1.0,1.0,1.0,3.0,3.0,.85)	(.70,.70,.55)
	H_{a3}	(1.0,1.0,1.0,3.0,1.51,.85)	(.70,.65,.55)
	H_0	(-1.1,-1.1,-1.1,1.1,1.1,1.1)	(.30,.30,.30)
III	H_{a1}	(-.01,-1.1,-1.1,2.0,1.1,1.1)	(.50,.30,.30)
	H_{a2}	(.85,-.01,-1.1,.05,2.0,1.1)	(.70,.50,.30)
	H_{a3}	(.85,.85,-.01,.05,.05,2.0)	(.70,.70,.50)

4 Redesigning a Real Life Clinical Trial

With a view to justify the suitability of the proposed procedure, we redesign a real clinical trial adopting the proposed procedure. For our purpose, we use the data reported in Tamura et al. (1994), which is a description of an Eli Lilly sponsored adaptive stratified trial of the anti-depression drug with two treatment arms, Control and *Fluoxetine*. The patients in the actual trial were classified either Shortened or Normal according to their Rapid Eye Movement Latency (REML), which is presumed to be a marker for endogenous depression. The actual trial reported a significant difference between the treatments for the shortened REML stratum.

The shortened REML status (i.e. whether shortened or normal) is taken as covariate in our study and the final patient outcome (i.e. whether success or failure) is taken as the binary response. However, excluding misclassified, missing and unavailable patient responses, we have considered the data comprising 80 patients (i.e. excluding the data from patient numbers 4,5,56,57,63,68,73,79 and 88). Designating Fluoxetine as Treatment 1 and Control as Treatment 2, we estimate the parameters of the logistic regression based response model for each treatment as $\hat{\alpha}_1 = .486, \hat{\beta}_1 = -.034$ and $\hat{\alpha}_2 = -.201, \hat{\beta}_2 = -.492$. Moreover, for the considered data we obtain the proportion of patients with shortened REML as $\frac{39}{80} = .4875$. Treating these estimates as the true parameter values and generating the binary covariate values from the `Bernoulli(.4875)` distribution, we simulate the trial with 10000 repetitions, where the sequence of allocation probabilities to different treatments are computed according to the proposed allocation rule. Based on the data thus generated, we compute estimates of EAP and EFP to different treatments and the allocation probability to *Fluoxetine* for the next entering (i.e. the 81st) patient for each covariate value.

The computation reveals the expected number of allocations to Fluoxetine as 49 (actual trial

Table 2: Performance measures for the proposed CARA design

	Expected Allocation and Failure Proportions (SD)					
Configuration	(z, p_z)	Treatment 1	Treatment 2	Treatment 3	**EFP (SD)**	**Power (ER)**
I(H_{a1})	(-1, 0.5)	0.363 (0.08)	0.361 (0.08)	0.277 (0.08)	0.234 (0.06)	0.358 (.073)
	(+1 ,0.5)	0.353 (0.07)	0.353 (0.07)	0.294 (0.06)	0.677 (0.06)	**0.386 (.057)**
	Overall	0.358 (0.05)	0.357 (0.05)	0.285 (0.05)	0.455 (0.05)	
I(H_{a2})	(-1, 0.5)	0.394 (0.07)	0.348 (0.08)	0.258 (0.08)	0.185 (0.05)	0.908 (.070)
	(+1, 0.5)	0.407 (0.07)	0.331 (0.07)	0.263 (0.06)	0.556 (0.07)	**0.921 (.054)**
	Overall	0.401 (0.05)	0.339 (0.05)	0.260 (0.05)	0.370 (0.04)	
I(H_{a3})	(-1, 0.5)	0.384 (0.07)	0.383 (0.07)	0.233 (0.08)	0.137 (0.04)	0.949 (.072)
	(+1, 0.5)	0.389 (0.07)	0.389 (0.07)	0.223 (0.06)	0.443 (0.07)	**0.968 (.062)**
	Overall	0.387 (0.05)	0.386 (0.05)	0.228 (0.05)	0.291 (0.04)	
II(H_{a1})	(-1, 0.5)	0.404 (0.07)	0.299 (0.06)	0.297 (0.06)	0.713 (0.06)	0.827 (.071)
	(+1, 0.5)	0.289 (0.07)	0.355 (0.07)	0.356 (0.07)	0.052 (0.03)	**0.861 (.054)**
	Overall	0.347 (0.05)	0.327 (0.05)	0.326 (0.05)	0.383 (0.04)	
II(H_{a2})	(-1 ,0.5)	0.376 (0.07)	0.378 (0.07)	0.245 (0.06)	0.565 (0.07)	0.878 (.075)
	(+1, 0.5)	0.314 (0.08)	0.314 (0.08)	0.372 (0.07)	0.092 (0.04)	**0.868 (.055)**
	Overall	0.345 (0.05)	0.346 (0.05)	0.309 (0.05)	0.328 (0.04)	
II(H_{a3})	(-1 ,0.5)	0.366 (0.07)	0.373 (0.07)	0.208 (0.06)	0.551 (0.07)	0.701 (.077)
	(+1, 0.5)	0.310 (0.08)	0.308 (0.08)	0.367 (0.07)	0.088 (0.04)	**0.868 (.055)**
	Overall	0.338 (0.05)	0.340 (0.05)	0.288 (0.05)	0.319 (0.04)	
III(H_{a1})	(-1, 0.5)	0.336 (0.06)	0.332 (0.06)	0.332 (0.06)	0.894 (0.04)	0.334 (.101)
	(+1, 0.5)	0.419 (0.07)	0.291 (0.08)	0.290 (0.07)	0.341 (0.06)	**0.685 (.053)**
	Overall	0.378 (0.05)	0.311 (0.05)	0.311 (0.05)	0.617 (0.04)	
III(H_{a2})	(-1, 0.5)	0.429 (0.07)	0.289 (0.06)	0.282 (0.06)	0.641 (0.07)	0.696 (.103)
	(+1, 0.5)	0.343 (0.08)	0.404 (0.07)	0.254 (0.08)	0.274 (0.06)	**0.991 (.061)**
	Overall	0.386 (0.05)	0.346 (0.05)	0.268 (0.05)	0.458 (0.05)	
III(H_{a3})	(-1, 0.5)	0.398 (0.07)	0.399 (0.07)	0.203 (0.06)	0.426 (0.07)	0.789 (.104)
	(+1, 0.5)	0.309 (0.08)	0.311 (0.08)	0.381 (0.07)	0.224 (0.05)	**0.978 (.059)**
	Overall	0.354 (0.05)	0.355 (0.05)	0.291 (0.05)	0.324 (0.05)	

Boldface figures indicate Power and Error rates (ER) for CR procedure

Table 3: Performance measures for the Odds based CARA design

	Expected Allocation and Failure Proportions (SD)					
Configuration	(z, p_z)	Treatment 1	Treatment 2	Treatment 3	**EFP (SD)**	**Power (ER)**
I(H_{a1})	(-1, 0.5)	0.243 (0.12)	0.247 (0.12)	0.51 (0.15)	0.28 (0.07)	0.381 (.108)
	(+1, 0.5)	0.231 (0.14)	0.23 (0.14)	0.539 (0.19)	0.726 (0.07)	**0.386 (.057)**
	Overall	0.237 (0.09)	0.239 (0.09)	0.525 (0.12)	0.503 (0.05)	
I(H_{a2})	(-1, 0.5)	0.17 (0.08)	0.265 (0.13)	0.566 (0.14)	0.272 (0.07)	0.778 (.106)
	(+1, 0.5)	0.116 (0.08)	0.265 (0.16)	0.619 (0.17)	0.709 (0.07)	**0.921 (.054)**
	Overall	0.142 (0.05)	0.265 (0.1)	0.593 (0.11)	0.491 (0.05)	
I(H_{a3})	(-1, 0.5)	0.182 (0.08)	0.181 (0.08)	0.637 (0.12)	0.263 (0.07)	0.964 (.109)
	(+1, 0.5)	0.13 (0.09)	0.132 (0.09)	0.738 (0.13)	0.691 (0.08)	**0.968 (.062)**
	Overall	0.156 (0.06)	0.156 (0.06)	0.688 (0.09)	0.478 (0.06)	
II(H_{a1})	(-1, 0.5)	0.734 (0.14)	0.133 (0.1)	0.134 (0.1)	0.769 (0.08)	0.806 (.095)
	(+1, 0.5)	0.211 (0.07)	0.394 (0.14)	0.394 (0.14)	0.111 (0.04)	**0.861 (.054)**
	Overall	0.473 (0.08)	0.263 (0.09)	0.264 (0.09)	0.441 (0.05)	
II(H_{a2})	(-1 ,0.5)	0.453 (0.22)	0.453 (0.22)	0.094 (0.07)	0.843 (0.05)	0.514 (.094)
	(+1, 0.5)	0.254 (0.08)	0.25 (0.08)	0.496 (0.12)	0.076 (0.04)	**0.868 (.055)**
	Overall	0.354 (0.12)	0.351 (0.12)	0.295 (0.07)	0.46 (0.05)	
II(H_{a3})	(-1 ,0.5)	0.67 (0.17)	0.203 (0.14)	0.127 (0.09)	0.776 (0.07)	0.634 (.097)
	(+1, 0.5)	0.231 (0.08)	0.321 (0.12)	0.448 (0.13)	0.089 (0.04)	**0.868 (.055)**
	Overall	0.451 (0.10)	0.262 (0.09)	0.287 (0.08)	0.433 (0.05)	
III(H_{a1})	(-1, 0.5)	0.306 (0.19)	0.348 (0.21)	0.346 (0.21)	0.894 (0.04)	0.234 (.109)
	(+1, 0.5)	0.127 (0.07)	0.437 (0.17)	0.436 (0.17)	0.452 (0.07)	**0.685 (.053)**
	Overall	0.216 (0.11)	0.392 (0.13)	0.391 (0.13)	0.674 (0.04)	
III(H_{a2})	(-1, 0.5)	0.064 (0.05)	0.437 (0.22)	0.499 (0.22)	0.854 (0.05)	0.782 (.114)
	(+1, 0.5)	0.274 (0.14)	0.154 (0.08)	0.572 (0.15)	0.384 (0.08)	**0.991 (.061)**
	Overall	0.169 (0.07)	0.296 (0.12)	0.536 (0.14)	0.619 (0.05)	
III(H_{a3})	(-1, 0.5)	0.096 (0.07)	0.096 (0.07)	0.808 (0.12)	0.772 (0.08)	0.931 (.116)
	(+1, 0.5)	0.395 (0.15)	0.397 (0.15)	0.208 (0.11)	0.254 (0.06)	**0.978 (.059)**
	Overall	0.246 (0.09)	0.246 (0.09)	0.508 (0.08)	0.513 (0.05)	

Boldface figures indicate Power and Error rates (ER) for CR procedure.

figure 40) and expected number of treatment failures as 37 (actual trial figure 40), respectively. Naturally an increase in the expected number of assignments to Fluoxetine (i.e. better treatment) and a decrease in the expected number of failures are observed. Further, we project the allocation probability for the hypothetical 81st subject to Fluoxetine as 0.64 and 0.58 for the shortened and normal REML stratum, respectively. As the actual trial reported superiority of Fluoxetine for the shortened REML stratum patients, presence of treatment-covariate interaction is pertinent and therefore, higher allocation probabilities to Fluoxetine for such patients are expected. Thus the proposed procedure not only assigns sensibly to different treatments but also takes into account the presence of treatment-covariate interaction. Consequently, the proposed allocation design has the potential to produce reasonable outcomes considering available data and hence can be a suitable candidate in real life trials.

5 Discussion

A CARA design is developed for binary response multi-treatment clinical trials considering the presence of treatment-covariate interaction. We have investigated different features of the proposed design empirically considering a single binary covariate and compared with relevant competitors. A real clinical trial is further redesigned using the proposed design to assess the design from practitioner's point of view.

Although the assessment of the proposed design used a single binary categorical covariate, the allocation design can be used for continuous covariates or mixed covariates. Moreover, the proposed design is best suited for instantaneous responses. But delay in getting the response is natural in any clinical trial and in such a situation we suggest to update the allocation probabilities based on the available data. Although such a modification makes the adaptive mechanism weak, we presume that moderate delay does not affect the limiting properties. Details in these regards are intended for a future work.

Acknowledgements

The authors sincerely thank the anonymous Reviewer and Associate Editor for their insightful suggestions, which led to an improvement over the submitted version.

A Appendix

Discussion on the findings in Table 2 and Table 3 The numerical figures of Table 2, depicting the design as well as precision measures, came from a simulation study with 20000 iterations of the allocation procedures under consideration. However, in the absence of any package for CARA designs, programs were written in R to compute the performance measures empirically. The expected allocation and failure proportions given for each covariate category (i.e. $Z = -1$ or $Z = +1$) represent the expected conditional proportions or covariate specific values. After a close examination of Table 2, we observe that the performance measuring figures of Table 2 are in well agreement with

the assumed parameter configuration. The expected overall allocation proportions vary according to the treatment effectiveness, that is, highest allocation to the best treatment (i.e. treatment 1 in this computation) and less allocation to the less effective treatments (i.e. treatments 2 and 3). The expected overall failure proportion also decreases with increasing treatment effectiveness.

Now, we examine the performance of the proposed CARA design for different covariate categories. As indicated earlier, the figures in Table 2 under the configuration I are computed in the presence of quantitative treatment-covariate interaction, that is, for the computation, treatment 1 is assumed best for every covariate category. The corresponding figures (i.e. under configuration I) of Table 2 reveal the same, that is, the highest allocation proportion to treatment 1, on an average. We further observe that the expected conditional proportion of failures is always the minimum for $Z = -1$, which is expected as the treatments are set to be most effective for $Z = -1$. The same, that is, variation according to treatment effectiveness is observed for other sets of parameters. In addition, the type I error rate is consistently maintained in the range 7-10%, which is slightly higher than the nominal 5%. However, such inflation is common in response adaptive procedures. Moreover, apart from minor exceptions, higher powers are observed, which shows the ability of the proposed allocation design to capture a little departure from equality of treatment effects.

Regarding the odds based competitor, the consideration of treatment-covariate interaction is not observed in the figures of Table 3 (e.g. figures corresponding to H_{a3} for configuration II). In addition, significantly higher expected failures are observed. However, the precision level for both the designs are more or less the same with the complete randomization (CR) procedure, where each treatment is assigned with equal probability without considering the performance of the treatments and incoming patients' covariate. Thus, the derived class of CARA allocation designs uses all available information, even the covariate information of the incoming subject and assigns subjects according to the prevailing effectiveness of treatments without a significant compromise in precision level. A high level of precision translates into a relatively small trial and consequently produces lower number of allocations to the inferior treatment to make the design ethically appealing from the viewpoint of a real practitioner of clinical trial.

References

Bandyopadhyay, U. and Bhattacharya, R. (2012), "An urn based covariate adjusted response adaptive allocation design," *Statistical Methods in Medical Research*, 21, 135–148.

Basak, G. K., Biswas, A., and Volkov, S. (2009), "An urn model for odds-ratio-based response-adaptive phase III clinical trials for two or more treatments," *Journal of Biopharmaceutical Statistics*, 19, 838–856.

Biswas, A. and Bhattacharya, R. (2018), "A class of covariate-adjusted response-adaptive allocation designs for multitreatment binary response trials," *Journal of biopharmaceutical statistics*, 28, 809–823.

Hart, R. G., Halperin, J. L., Pearce, L. A., Anderson, D. C., Kronmal, R. A., McBride, R., Nasco, E.,

Sherman, D. G., Talbert, R. L., Marler, J. R., et al. (2003), "Lessons from the stroke prevention in atrial fibrillation trials," *Annals of Internal Medicine*, 138, 831–838.

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

Logunov, D. Y., Dolzhikova, I. V., Shcheblyakov, D. V., Tukhvatulin, A. I., Zubkova, O. V., Dzharullaeva, A. S., Kovyrshina, A. V., Lubenets, N. L., Grousova, D. M., Erokhova, A. S., et al. (2021), "Safety and efficacy of an rAd26 and rAd5 vector-based heterologous prime-boost COVID-19 vaccine: an interim analysis of a randomised controlled phase 3 trial in Russia," *The Lancet*, 397, 671–681.

Rosenberger, W. F. and Sverdlov, O. (2008), "Handling Covariates in the Design of Clinical Trials," *Statistical Science*, 23, 404–419.

Rosenberger, W. F., Vidyashankar, A., and Agarwal, D. K. (2001), "Covariate-adjusted response-adaptive designs for binary response," *Journal of biopharmaceutical statistics*, 11, 227–236.

Tamura, R., Faries, D., Andersen, J., and Heiligenstein, J. (1994), "A case study of an adaptive clinical trial in the treatment of out-patients with depressive disorder," *Journal of the American Statistical Association*, 89, 768–776.

Trippa, L., Lee, E. Q., Wen, P. Y., Batchelor, T. T., Cloughesy, T., Parmigiani, G., and Alexander, B. M. (2012), "Bayesian Adaptive Randomized Trial Design for Patients With Recurrent Glioblastoma," *Journal of Clinical Oncology*, 30, 3258–3263.

Villar, S. S. and Rosenberger, W. F. (2018), "Covariate-adjusted response-adaptive randomization for multi-arm clinical trials using a modified forward looking Gittins index rule," *Biometrics*, 74, 49–57.

Wason, J. M. and Trippa, L. (2014), "A comparison of Bayesian adaptive randomization and multistage designs for multi-arm clinical trials," *Statistics in medicine*, 33, 2206–2221.

Yin, G., Chen, N., and Jack Lee, J. (2012), "Phase II trial design with Bayesian adaptive randomization and predictive probability," *Journal of the Royal Statistical Society Series C: Applied Statistics*, 61, 219–235.

Zhang, L.-X., Hu, F., Cheung, S. H., and Chan, W. S. (2007), "Asymptotic properties of covariate-adjusted response-adaptive designs," *Annals of Statistics*, 35, 1166–1182.

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

Received: July 22, 2023

Accepted: April 3, 2024