

KERNEL-BASED ADAPTIVE RANDOMIZATION TOWARD BALANCE IN CONTINUOUS AND DISCRETE COVARIATES

Fei Jiang¹, Yanyuan Ma² and Guosheng Yin¹

¹*University of Hong Kong* and ²*Pennsylvania State University*

Abstract: Covariate balance among different treatment arms is critical in clinical trials, as confounding effects can be effectively eliminated when patients in different arms are alike. To balance the prognostic factors across different arms, we propose a new dynamic scheme for patient allocation. Our approach does not require discretizing continuous covariates to multiple categories, and can handle both continuous and discrete covariates naturally. This is achieved through devising a statistical measure to characterize the similarity between a new patient and all the existing patients in the trial. Under the similarity weighting scheme, we develop a covariate-adaptive biased coin design and establish its theoretical properties, thus improving the original Pocock–Simon design. We conduct extensive simulation studies to examine the design operating characteristics and we illustrate our method with a data example. The new approach is thereby demonstrated to be superior to existing methods in terms of performance.

Key words and phrases: Biased coin design, clinical trial, covariate-adaptive randomization, covariate balance, pocock and simon design, similarity measure, stratification.

1. Introduction

Peter Hall was one of the most influential and prolific researchers in modern statistics. His contributions are broad and cover many important areas. From interactions with him, the authors have been greatly influenced by his statistical thinking, especially in how to use “smoothing” methods to increase modeling flexibility and reduce estimation error. One of the nonparametric devices, called kernel smoothing, is widely used in density estimation and nonparametric regression. In density estimation, Hall (1981) derived the law of the iterated logarithm for the kernel estimator, discussed the choice of the order of kernels (Hall and Marron (1988)), and addressed the issues on constructing confidence intervals (Hall (1992)). In nonparametric regression, Hall (1984) investigated the asymptotic properties of the kernel regression estimator. A series of his follow-up works

focused on the confidence intervals and confidence bands for kernel estimators, which include Hall and Marron (1988), Hall (1992), Hall (1993), and Hall and Horowitz (2013). Motivated by kernel estimation, we propose a kernel-based covariate-adaptive randomization design. We apply the martingale convergence theorem in Hall and Heyde (1980) extensively in deriving the asymptotic properties of the proposed design, which reinforces Peter Hall's impact, especially in the area of sequential analysis.

The primary goals of randomized clinical trials are to differentiate the treatment effects efficiently as well as to treat patients effectively. If the treatment effects of different drugs can be quickly discriminated, then patients outside of the trial would benefit from the more effective therapy sooner. To achieve this goal, allocation of patients is random to balance out both known and unknown prognostic factors that may affect the response of interest, and the numbers of patients should also be balanced across different treatment arms to achieve high statistical power. For discrete covariates, various approaches have been developed for patient allocation to achieve covariate balancing (Hu and Hu (2012)). These include the biased coin covariate-adaptive randomization design (Wei (1978); Antognini and Giovagnoli (2004)), which is an extension of the biased coin design (Efron (1971)) for balancing the sample size, and the Pocock–Simon design which is based on a minimization method for sequential treatment assignment (Taves (1974); Pocock and Simon (1975)). Despite their popularity, the main drawbacks of these designs are that continuous covariates must be categorized into several groups, while clinical trials often collect a large number of continuous covariates and different ways of categorization may lead to different imbalanced structures. In addition, breaking down continuous covariates into sub-categories often changes the nature of the covariates and makes distributional balance unattainable (Ma and Hu (2013)). If the sub-categories are not appropriately defined, it can even lead to error and loss of efficiency in the randomization procedure (Stigsby and Taves (2010)).

Such a problem has arisen in many clinical trials, which is illustrated with an AIDS Clinical Trials Group study (Campbell et al. (2012)). To evaluate several antiretroviral regimens in diverse populations, patients in the A5175 trial were randomly assigned to the antiretroviral therapies with efavirenz plus lamivudine-zidovudine (arm 1) and atazanavir, didanosine-EC plus emtricitabine (arm 2). The study endpoint was the CD4 count at week 96. The baseline covariate CD4 cell count at screening was found to be strongly associated with the endpoint with a p -value less than 2×10^{-16} in a simple linear regression analysis. To

balance the CD4 cell count at screening, there was a controversy over the choice of the cutoffs, either the clinically meaningful low CD4 count 200 or the sample average 169. In a simulated clinical trial study, we compared the performances of using these two cutoffs under the same covariate-adaptive procedure. The resulting absolute mean difference between the two groups was 71.76 for the cutoff 200, and 43.73 for the cutoff 169 with corresponding p -values of 0.002 and 0.06 for the two sample t -test of the mean differences. This suggested that a slight variation in the cutoff can lead to substantially different allocation results. To handle continuous covariates, Frane (1998) proposed to calculate the p -value for the mean difference of each covariate, presuming that a new patient is assigned to each treatment group. Using the minimal p -value as a representation of the imbalance of assigning a new patient to a specific treatment, the new patient is then assigned to the treatment with the largest minimal p -value. Stigsby and Taves (2010) considered the rank-sum based covariate adaptive procedure, and Su (2011) discussed a method using quantiles of the covariate differences. Ma and Hu (2013) proposed a randomization procedure by defining the imbalance of the covariates through kernel density estimators, which summarize all the information in the covariate distributions.

To improve the overall balance among both continuous and discrete covariates, we develop a kernel-based adaptive randomization framework that can simultaneously handle a large number of continuous covariates in a single step. In particular, we define a similarity measure between each incoming patient and all the existing patients, and then allocate the new patient with the largest probability to the arm that has the least overall similarity to the new patient. Through weighing each observation by taking into account his/her similarity with the new patient, the proposed method handles both discrete and continuous covariates in a natural way and further broadens the traditional counting from integer values to all nonnegative values.

The rest of the paper is organized as follows. Section 2 describes our covariate-adaptive randomization procedure via introducing the similarity measure and modifying the biased coin design. In Section 3, we cast the Pocock-Simon design in our new framework so as to accommodate continuous covariates. We carried out simulation studies and a data example to illustrate the performance of the new designs in Section 4. Section 5 concludes with some remarks. Theoretical results are delineated in the Appendix and the corresponding proofs are presented in the Supplementary Materials.

2. Similarity Weighted Biased Coin Design

In a randomized clinical trial with m treatments, suppose that we have already assigned n patients to different arms, and a new patient arrives and is ready for treatment assignment. Let \mathbf{X}_i be the p -dimensional covariate vector for the i th patient, and I_{iu} be the indicator of assigning the i th patient to treatment arm u , $u = 1, \dots, m$.

We define a similarity measure w_i between the i th existing patient and the incoming $(n+1)$ th patient, whose covariate vector is \mathbf{X}_{n+1} with $X_{(n+1)k}$ denoting their k th component. For ease of exposition, we standardize all the covariate values to be within the range of $[-1, 1]$. The similarity measure between the new patient and the i th patient in the trial is defined as

$$w_i = \prod_{k=1}^p w_{ik}, \quad i = 1, \dots, n, \quad (2.1)$$

where

$$w_{ik} = K_{h_n}(X_{ik} - X_{(n+1)k}), \quad (2.2)$$

$K_{h_n}(x) = K(x/h_n)/h_n$, and $K(\cdot)$ is a kernel function satisfying $K(\cdot) \geq 0$, and $K(0) = 1$, and $h_n > 0$ is a bandwidth. There are standard kernel functions, and, although the selection of kernels does not affect the large sample properties of our allocation procedure, we recommend use of the Epanechnikov kernel for the bounded covariates, which is the most efficient one in minimizing the averaged mean squared error (Epanechnikov (1969)). The similarity measure w_{ik} indicates a higher level of similarity for patients whose k th covariate values are closer to $X_{(n+1)k}$, and the similarity decreases to zero as the difference between X_{ik} and $X_{(n+1)k}$ reaches the bandwidth h_n , for $k = 1, \dots, p$. By raising the value of h_n , the kernel takes into account more X_{ik} 's with larger distances from $X_{(n+1)k}$.

We propose the similarity weighted biased coin design for balancing covariates, described as follows.

1. Calculate the similarity measure of the new patient with each of the existing n patients in the trial to obtain w_1, \dots, w_n using (2.1).
2. For $u = 1, \dots, m$, calculate the weighted total number of patients in treatment arm u ,

$$n_u = \sum_{i=1}^n w_i I_{iu}, \quad (2.3)$$

and obtain the imbalance measure of arm u as $g_{nu} = n_u / (\sum_{u=1}^m n_u)$.

3. Define the allocation probability π_u to be a function of $\mathbf{g}_n = (g_{n1}, \dots, g_{n(m-1)})^\top$ that is decreasing with respect to each component g_{nu} . We assign the new patient to treatment arm u with probability $\pi_u(\mathbf{g}_n)$, $u = 1, \dots, m$.

In the construction of the similarity weighted biased coin design, a patient who is more similar to the new patient receives a larger weight, and is counted more towards the total number of patients in a specific arm. Compared with the biased coin design where $n_u = \sum_{i=1}^n I_{iu}$, our definition of n_u in (2.3) is a weighted sum of the treatment indicators I_{iu} . If the covariate vector \mathbf{X}_i contains only discrete variables, our method reduces to the existing discrete covariate-adaptive randomization by choosing h_n to be smaller than the smallest difference in different categories, $h_n < \min_{X_{ik} \neq X_{i'k}} |X_{ik} - X_{i'k}|$. Such a construction leads to $w_i = 0$ whenever $X_{ik} \neq X_{(n+1)k}$ for at least one k , and w_i reaches its maximum if $\mathbf{X}_i = \mathbf{X}_{n+1}$. As a result, our method reduces to the biased coin randomization procedure within each stratum defined by the discrete covariates.

We highlight several advantages of the proposed similarity weighting scheme. First, it overcomes the difficulties caused by the high dimensionality of the covariates. To accommodate high-dimensional covariates, Yuan, Huang and Liu (2011) resorted to a linear model structure. But it is subject to model misspecification. Second, even when all the covariates are discrete, we can choose the bandwidth h_n sufficiently large to avoid the situation of too many strata and too few or even zero observations within some strata. Finally, the procedure is automatic and flexible as reflected in the various ways of constructing the similarity measure.

To study the asymptotic properties of the imbalance measure $\mathbf{D}_{nu} = \sum_{i=1}^n (I_{iu} - \kappa_u)\mathbf{X}_i$, we explore the properties of $\mathbf{D}_{nu}^\top \mathbf{z} = \sum_{i=1}^n (I_{iu} - \kappa_u)\mathbf{X}_i^\top \mathbf{z}$, $u = 1, \dots, m-1$, where \mathbf{z} is an arbitrary p dimensional vector. We show that the allocation achieves the target ratio in the long run, as follows.

Theorem 1. *Assume that Conditions (C1) – (C6) hold and let $Q_z = E(\mathbf{z}^\top \mathbf{X}_i \mathbf{X}_i^\top \mathbf{z})$. Then the similarity weighted biased coin design with an allocation probability $\pi_u\{\mathbf{U}_n(\mathbf{X}_{n+1})\}$, $n^{-1/2}(\mathbf{D}_{n1}^\top \mathbf{z}, \dots, \mathbf{D}_{n(m-1)}^\top \mathbf{z})^\top$ converges to a zero-mean multivariate Gaussian distribution, with the variance-covariance matrix having the u th diagonal element $(1+2\rho)^{-1}(1-\kappa_u)\kappa_u Q_z$ and the (u, v) entry $-(1+2\rho)^{-1}\kappa_u\kappa_v Q_z$, for $u, v = 1, \dots, m-1, u \neq v$.*

Due to the arbitrariness of \mathbf{z} , we readily obtain the asymptotic normality of \mathbf{D}_n . The proof relies heavily on the martingale convergence theorem (Hall and Heyde (1980)), which is provided in the Supplementary Material.

3. Similarity Weighted Pocock–Simon Design

The proposed similarity measure can also be incorporated into the Pocock–Simon design (Pocock and Simon (1975)), namely the similarity weighted Pocock–Simon design, so that continuous covariates no longer need to be discretized. For a situation in which we assign the new patient to treatment 1, the similarity weighted Pocock–Simon design can be implemented as follows.

1. For the k th covariate, let $n_{ku} = \sum_{i=1}^n w_{ik} I_{iu}$ be the weighted total number of subjects assigned to treatment arm u , where w_{ik} is defined in (2.2).
2. Calculate the aggregated variation in the form of

$$d_k = \frac{1}{2 \sum_{u,v \in 1, \dots, m} (n_{ku} - n_{kv})^2}$$

for the k th covariate.

3. Sum the d_k 's across all the covariates, leading to the imbalance measure $g_{n1} = \sum_{k=1}^p d_k$.
4. Calculate g_{n2}, \dots, g_{nm} by presuming that the new patient is assigned to treatment arms 2, \dots , m , respectively.
5. Order the g_{nu} 's as $g_{n1} \leq \dots \leq g_{nm}$, create the randomization probabilities satisfying $\pi_{(n+1)1} \geq \dots \geq \pi_{(n+1)m}$, and assign the new patient to the m treatment arms with probabilities $\pi_{(n+1)1}, \dots, \pi_{(n+1)m}$.

The selection of d_k is not unique. For example, the sum of absolute differences, $d_k = 1/2 \sum_{u,v \in 1, \dots, m} |n_{ku} - n_{kv}|$, can also be used to measure the total imbalance among the treatment arms for the k th covariate. Our modified procedure can be viewed as a generalized version of the original Pocock–Simon design procedure: the former calculates the n_{ku} 's using a similarity weight w_{ik} , while the latter sets the weight $w_{ik} = 1$ if the i th patient has the same k th covariate value as the new patient, and $w_{ik} = 0$ otherwise. Our approach handles continuous covariates through a similarity-based weighting scheme and does not require discretization.

4. Numerical Studies

4.1. Simulation study

To evaluate the finite sample properties of the proposed similarity weighted biased coin design and the similarity weighted Pocock–Simon design, we simulated 1,000 two-arm clinical trials, each containing $n = 50$ subjects. We generated covariates in the form of

Table 1. Comparison of the sample size imbalance, $|n_1 - n_2|$, among the similarity weighted biased coin design, biased coin design, similarity weighted Pocock–Simon design and Pocock–Simon design for different dimensions (p) of covariates.

Design	Dimension of covariates p							
	1	2	3	4	5	6	7	8
Weighted biased coin	1.277	1.289	1.265	1.219	1.325	1.343	1.286	1.411
Biased coin	1.279	1.276	1.334	1.687	2.047	2.247	2.456	2.605
Weighted Pocock–Simon	0.122	0.159	0.159	0.183	0.199	0.211	0.233	0.246
Pocock–Simon	0.387	0.241	0.221	0.262	0.288	0.294	0.354	0.351

Table 2. Comparison of the covariate imbalance using the F statistics, among the similarity weighted biased coin design, biased coin design, similarity weighted Pocock–Simon design and Pocock–Simon design for different dimensions (p) of covariates.

Design	Dimension of covariates p							
	1	2	3	4	5	6	7	8
Weighted biased coin	0.278	0.297	0.299	0.311	0.326	0.345	0.373	0.389
Biased coin	0.280	0.299	0.312	0.402	0.525	0.663	0.809	0.872
Weighted Pocock–Simon	0.028	0.053	0.085	0.128	0.166	0.207	0.256	0.308
Pocock–Simon	0.149	0.145	0.162	0.193	0.225	0.271	0.322	0.358

$$X_{ik} = \frac{2 \exp(\xi_{ik})}{1 + \exp(\xi_{ik})} - 1, \quad k = 1, \dots, p, \quad (4.1)$$

where the dimension p of covariates ranged from 1 to 8, and ξ_{ik} was a normal random variable with mean $k/2$ and standard deviation 5. To implement the biased coin design and the Pocock–Simon design, we discretized the X_{ik} 's to be 0 or 1 according to the negative or positive signs of the covariates. We used the allocation probability function $\phi_u(\mathbf{y}) = (y_u^{-1} - 1) / \sum_{u=1}^m (y_u^{-1} - 1)$ in Atkinson (1982), which satisfies Conditions (C1) and (C2) as shown in Smith (1984). In all the numerical studies, we took the bandwidth to be 2.1, so that the support of the kernel would completely cover all the covariates. We experimented with other bandwidths between 2 and 2.5, and the results turned out to be similar as long as the bandwidth was chosen to be slightly larger than the covariate range.

We first make comparisons from two aspects: the imbalance of the sample sizes and the imbalance of the covariates between the two arms. To quantify the former, with n_u the sample size in arm u , $u = 1, 2$, we obtained $|n_1 - n_2|$ averaged over the 1,000 simulated trials for all four methods: similarity weighted biased coin design, biased coin design, similarity weighted Pocock–Simon design and Pocock–Simon design. Table 1 summarizes the imbalance measure on sample sizes, that demonstrates their similarity weighted designs tend to induce more

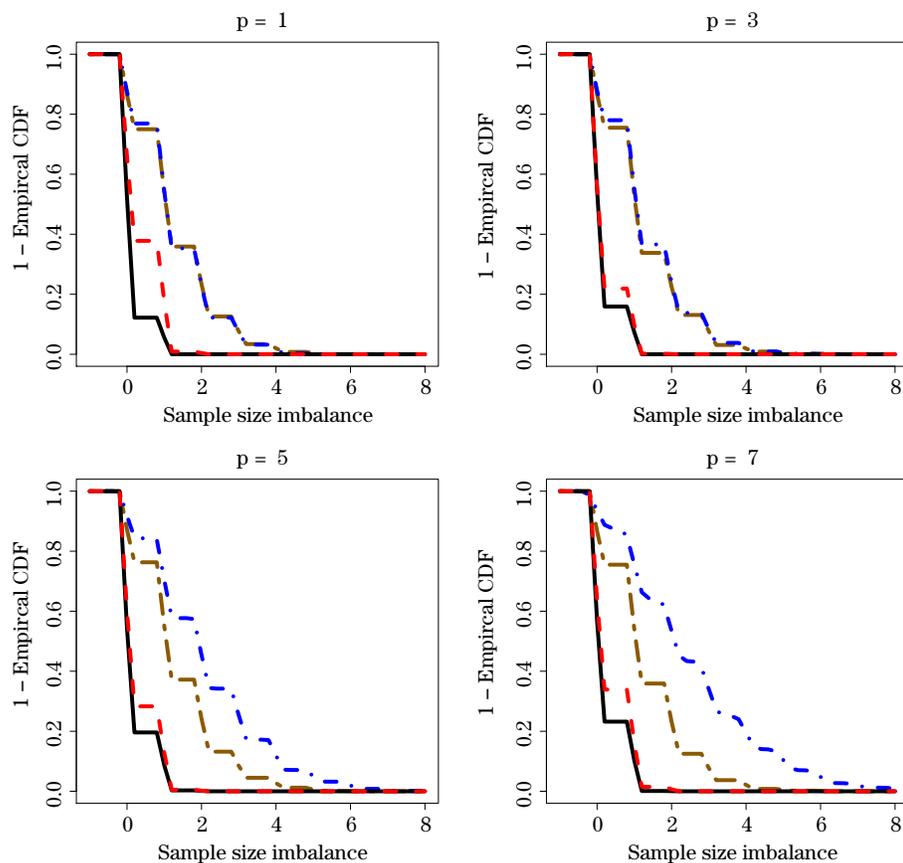


Figure 1. Survival functions of sample size imbalance for the similarity weighted biased coin design (dotted line), biased coin design (dot-dashed line), similarity weighted Pocock–Simon design (solid line), and Pocock–Simon design (dashed line).

balanced numbers of subjects between the two arms. Figure 1 shows the survival function, one minus the empirical cumulative distribution function (CDF), of $|n_1 - n_2|$ over the 1,000 replicated data sets. Clearly the similarity weighted designs outperform their counterparts. To compare the covariate imbalance, we borrow the idea from the analysis of variance to construct an F test statistic for each covariate,

$$F_k = \frac{\text{SSB}_k / (m - 1)}{(\text{SST}_k - \text{SSB}_k) / (n - m)}, \quad k = 1, \dots, p, \quad (4.2)$$

where the between-arm sum of squared errors (SSB_k) is

$$\text{SSB}_k = n_1 \left(n_1^{-1} \sum_{i=1}^n I_{i1} X_{ik} - n^{-1} \sum_{i=1}^n X_{ik} \right)^2 + n_2 \left(n_2^{-1} \sum_{i=1}^n I_{i2} X_{ik} - n^{-1} \sum_{i=1}^n X_{ik} \right)^2,$$

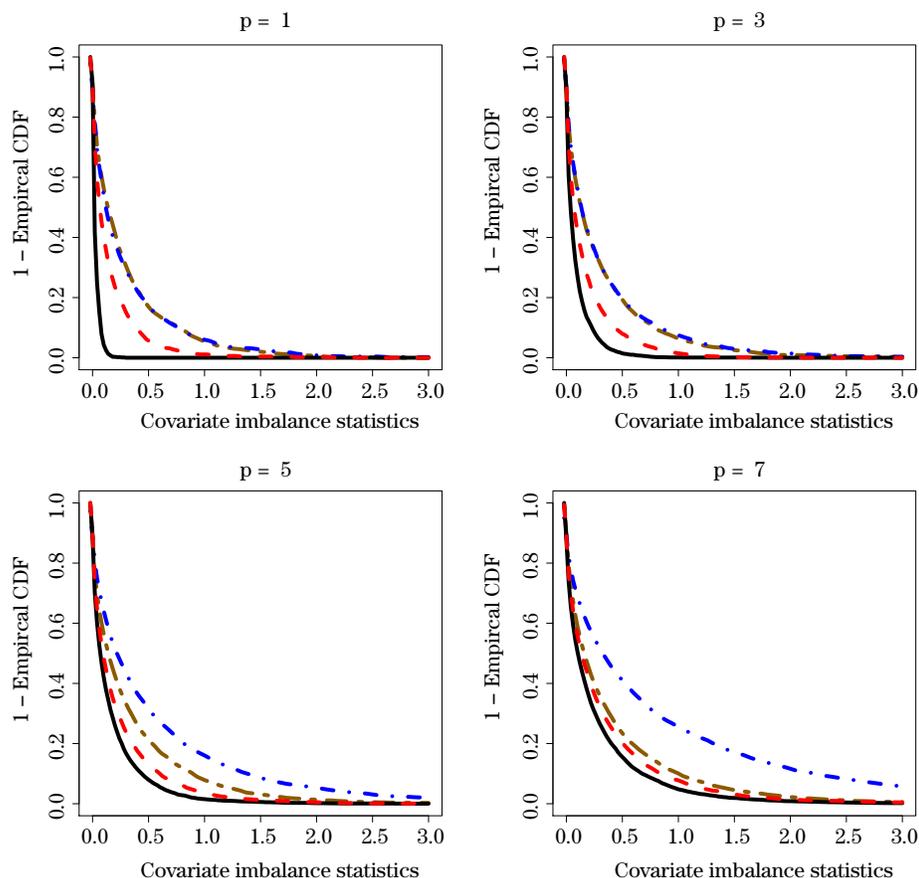


Figure 2. Survival functions of sample size imbalance for the similarity weighted biased coin design (dotted line), biased coin design (dot-dashed line), similarity weighted Pocock–Simon design (solid line), and Pocock–Simon design (dashed line).

and the total sum of squared errors (SST_k) is given by

$$SST_k \equiv \sum_{i=1}^n \left(X_{ik} - n^{-1} \sum_{i=1}^n X_{ik} \right)^2.$$

As the F statistic has the same distribution across all the covariates, we summarize the overall mean of the F statistics for all the covariates in Table 2, and plot the survival functions of the F statistics in Figure 2. Both the similarity weighted biased coin design and similarity weighted Pocock–Simon design outperform their counterparts in terms of balancing the covariates. For the biased coin designs, the improvement by using similarity weights enhances as the dimension p increases, while the opposite is true for the Pocock–Simon designs. Figure

Table 3. Comparison of the estimated treatment effect under the similarity weighted biased coin design, biased coin design, similarity weighted Pocock–Simon design and Pocock–Simon design, where $\hat{\mu}$ is the estimate of $\mu = -5$, and SD and MSE are the corresponding empirical standard deviation and mean squared error, respectively.

p	Weighted biased coin			Weighted Pocock–Simon		
	$\hat{\mu}$	SD	MSE	$\hat{\mu}$	SD	MSE
1	-4.983	0.314	0.332	-4.992	0.134	0.141
2	-5.030	0.830	0.860	-4.999	0.686	0.687
3	-5.022	1.723	1.745	-4.949	1.698	1.749
4	-4.864	3.050	3.186	-5.058	3.164	3.223
5	-4.924	4.776	4.852	-4.726	5.169	5.442
6	-5.130	6.778	6.909	-4.573	7.328	7.755
7	-4.997	8.224	8.227	-4.959	9.401	9.442
8	-5.153	9.482	9.636	-5.016	10.582	10.598
	Biased coin			Pocock–Simon		
1	-4.996	0.332	0.335	-4.996	0.268	0.272
2	-5.028	0.822	0.850	-4.994	0.865	0.871
3	-4.989	1.783	1.794	-5.039	1.899	1.938
4	-5.025	3.201	3.225	-5.063	3.660	3.723
5	-5.015	5.328	5.343	-4.956	5.821	5.865
6	-5.167	7.930	8.098	-5.239	7.921	8.161
7	-5.353	9.787	10.140	-5.372	10.015	10.387
8	-5.253	11.797	12.050	-5.169	12.011	12.181

2 further demonstrates the advantages of the proposed methods, and particularly the similarity weighted Pocock–Simon design that performs the best in terms of reducing the imbalance in both the sample size and covariates.

To explore the estimation of the treatment effect under the four designs, we considered a two-arm trial, where $I_i = 1$ indicates that the i th patient is allocated to arm 1, and $I_i = 0$ otherwise. We simulated 1,000 clinical trials with response Y_i generated as

$$Y_i = \mu I_i + \exp\left(\frac{\boldsymbol{\beta}^T \mathbf{X}_i}{2}\right) + \epsilon_i,$$

where the true parameter values were $\mu = -5$ and $\boldsymbol{\beta} = (p/3, \dots, p/3)^T$, ϵ_i was a zero-mean normal random error with standard deviation 0.1, and \mathbf{X}_i was generated as before. We chose $p = 1, \dots, 8$ and sample size $n = 30$. We allocated the first patient with equal probability to each arm, and started the adaptive allocation from the second patient.

Table 3 shows the estimated treatment effect $\hat{\mu} = \sum_{i=1}^n I_i Y_i / n_1 - \sum_{i=1}^n (1 - I_i) Y_i / n_2$, and its empirical standard deviation and mean squared error for $p =$

1, ..., 8. Here $\hat{\mu}$ is a consistent estimator of μ regardless of the regression form (Shao, Yu and Zhong (2010)). The biases of the estimates of μ are negligible under all four designs, while the empirical standard deviations and mean squared errors deteriorate as p grows. Again, the similarity weighted biased coin design and the similarity weighted Pocock–Simon design outperform the unweighted counterparts, in terms of the mean squared errors.

4.2. Data example

We applied the biased coin design, the Pocock–Simon design, and the corresponding similarity weighted versions, to the data from the AIDS trial A5175. To study the treatment effect, seven covariates were considered important, which should be balanced between the two arms at randomization: CD4 cell count and percentage (at screening), Karnofsky score, Hepatitis-B surface antigen reactivity, the laboratory test values including platelets, white blood cell count, absolute neutrophil count, and albumin. In the original trial, there were $n = 370$ patients with complete observations, and they were allocated to the two arms with equal probability. We took the standardized CD4 count at week 96 as the outcome, and transformed the standardized covariates via $2 \exp(x) / \{1 + \exp(x)\} - 1$ to ensure that all the covariate values were within $[-1, 1]$. Let \mathbf{X}_{iu} and Y_{iu} denote the covariates and response, respectively, for the i th patient in arm u , $u = 1, 2$. We built separate models for each arm,

$$Y_{iu} = \beta_{0u} + \beta_u^T \mathbf{X}_{iu} + e_{iu}, \quad u = 1, 2,$$

with $e_{iu} \sim N(0, \sigma_u^2)$. We obtained the least squared estimators $(\hat{\beta}_{0u}, \hat{\beta}_u^T)$ for each arm, and used these parameter estimates as the true values to generate the outcomes in different randomization procedures.

For illustration, we selected the first 50 samples to evaluate and compare the four designs. The observed difference of the mean outcomes between the two arms over these 50 samples is 0.39, which is substantially different from that using the full 370 samples, 0.22. Since the trial data are balanced in covariates for $n = 370$, we used 0.22 as a benchmark to approximate the true underlying mean difference between the two arms. Using each of the four randomization procedures, we re-randomized the 50 patients and each procedure is replicated 1,000 times to obtain the average effect. The means of $|n_1 - n_2|$ under the similarity weighted biased coin design, biased coin design, similarity weighted Pocock–Simon design and Pocock–Simon design were 1.30, 1.49, 0.17 and 0.32, respectively, and the corresponding means of the F statistics in (4.2), summing over all the covariates, were 3.24, 5.41, 2.29 and 2.56. The results show that the similarity weighted

procedures outperform the original counterparts in reducing both the sample size and covariate imbalance and, overall, the similarity weighted Pocock–Simon design performs the best among the four designs. In addition, the estimates of the difference of the mean responses were 0.306, 0.306, 0.306 and 0.313 using the similarity weighted biased coin design, biased coin design, similarity weighted Pocock–Simon design and Pocock–Simon design, respectively. Compared with the observed mean difference in the first 50 samples, the estimates from the four covariate–adaptive designs were closer to the benchmark value 0.22, indicating that covariates adaptation helps to improve the balance.

5. Discussion

To accommodate continuous covariates in the biased coin design and Pocock–Simon design, we develop a kernel-based similarity measure and its associated imbalance assessment criterion. We define the allocation probability function based on the new imbalance measure and show that the covariate equilibrium measure \mathbf{D}_{nu} of the proposed similarity weighted biased coin design asymptotically follows a normal distribution. We choose the continuous allocation function π instead of a discrete one, because discrete allocation functions can neither discriminate between large versus small values of $|g_{n1} - g_{n2}|$ nor discriminate between large versus small numbers of subjects, hence typically yield designs with poor small sample properties (Wei (1978); Smith (1984); Hu and Zhang (2004)). In terms of the bandwidth requirement, we find that as long as the bandwidth is chosen to be slightly larger than the covariate range, the results are not sensitive to the bandwidth choice. Not only does the asymptotic property of the covariate equilibrium \mathbf{D}_{nu} explain the covariate discrepancy between the arms, but it is also an essential component for analyzing the hypothesis testing procedures in the linear regression problem (Shao, Yu and Zhong (2010); Ma, Hu and Zhang (2015)). Our theoretical results are essential for constructing inference procedures under the similarity weighted biased coin design.

Supplementary Materials

The proof of Theorem 1 and several related lemmas and their proofs are presented in the online Supplementary Materials.

Acknowledgment

We thank the guest editor Raymond J. Carroll and Haolun Shi for many

constructive comments that have led to significant improvements of the article. Ma's research was supported in part by a grant (DMS-1608540) from the National Science Foundation, and Yin's research was supported in part by a grant (17326316) from the Research Grants Council of Hong Kong.

Appendix

A.1 Allocation probability function

Suppose that n samples have been enrolled in the trial. The allocation probability π_u is a function of the imbalance measure vector $\mathbf{g}_n = (g_{n1}, \dots, g_{n(m-1)})^\top$. Let $\boldsymbol{\pi} = (\pi_1, \dots, \pi_{m-1})^\top$. Furthermore, let $\boldsymbol{\kappa} = (\kappa_1, \dots, \kappa_{m-1})^\top$. We show that $\boldsymbol{\pi}$ drives \mathbf{g}_n towards $\boldsymbol{\kappa}$ under the following conditions. For notational simplicity, we suppress the subindex n in these conditions.

- (C1) $\pi_u(\mathbf{g})$ is a nonnegative and monotonically decreasing function with respect to the u th element g_u . Let $|\cdot|$ be the L_1 norm of a vector, the vector $\boldsymbol{\pi}(\mathbf{g})$ satisfies $|\boldsymbol{\pi}(\mathbf{g})| \leq 1$ for any component-wise nonnegative $m-1$ dimensional vector \mathbf{g} with $|\mathbf{g}| \leq 1$. Moreover, $\pi_m(\mathbf{g}) = 1 - |\boldsymbol{\pi}(\mathbf{g})|$, $g_m = 1 - |\mathbf{g}|$, $\kappa_m = 1 - |\boldsymbol{\kappa}|$. If $g_u \geq \kappa_u$, then $\pi_u(\mathbf{g}) \leq \kappa_u$, and if $g_u < \kappa_u$, then $\pi_u(\mathbf{g}) > \kappa_u$, $u = 1, \dots, m$.
- (C2) $\pi_u(\mathbf{g})$ is a twice continuously differentiable function of \mathbf{g} with a uniformly bounded Hessian matrix.

Let $\pi'_u(\mathbf{g}) = \partial \pi_u(\mathbf{g}) / \partial \mathbf{g}$, $\pi'_{ur}(\mathbf{g})$ be the partial derivative of $\pi_u(\mathbf{g})$ with respect to its r th argument, and π''_u be the $(m-1) \times (m-1)$ Hessian matrix.

Remark A1. Conditions (C1) and (C2) were used in Smith (1984) to establish the properties of the biased coin design. Condition (C1) implies $\pi_u(\boldsymbol{\kappa}) \leq \kappa_u$. If the inequality is strict for any u , summing both sides over $u = 1, \dots, m$, we obtain $1 < 1$, which is a contradiction. Therefore, we have $\pi_u(\boldsymbol{\kappa}) = \kappa_u$ for $u = 1, \dots, m$.

Remark A2. For an arbitrary δ ,

$$\pi_m(\kappa_1 + \delta, \kappa_2 - \delta, \kappa_3, \dots, \kappa_{m-1}) = \kappa_m + \delta\{\pi'_{m1}(\boldsymbol{\kappa}) - \pi'_{m2}(\boldsymbol{\kappa})\} + O(\delta^2).$$

Here $\pi_m(\kappa_1 + \delta, \kappa_2 - \delta, \kappa_3, \dots, \kappa_{m-1})$ is at most κ_m regardless of the sign of δ . As $\delta \rightarrow 0$, $O(\delta^2)$ goes to 0 faster than the leading terms. This gives $\delta\{\pi'_{m1}(\boldsymbol{\kappa}) - \pi'_{m2}(\boldsymbol{\kappa})\} \leq 0$ and $-\delta\{\pi'_{m1}(\boldsymbol{\kappa}) - \pi'_{m2}(\boldsymbol{\kappa})\} \leq 0$. Therefore, $\pi'_{m1}(\boldsymbol{\kappa}) = \pi'_{m2}(\boldsymbol{\kappa})$. Similarly, for each $u = 1, \dots, m-1$, $\pi'_{mu}(\boldsymbol{\kappa}) = \rho$, a constant that does not

depend on u . Following the same argument, for any $u < m, u \neq 1$,

$$\pi_u(\kappa_1 + \delta, \kappa_2, \dots, \kappa_{m-1}) = \kappa_u + \delta\pi'_{u1}(\boldsymbol{\kappa}) + O(\delta^2) \leq \kappa_u$$

for all δ , which implies $\pi'_{u1}(\boldsymbol{\kappa}) = 0$ and in turn $\pi'_{ur}(\boldsymbol{\kappa}) = 0$ for $r < m, r \neq u$.

Remark A3. Because $\sum_{u=1}^m \pi_u(\mathbf{g}) = 1$ for all \mathbf{g} , $\sum_{u=1}^m \pi'_{ur}(\mathbf{g}) = 0$ for any $r = 1, \dots, m - 1$, we have

$$\sum_{u=1}^m \pi'_{ur}(\mathbf{g}) = \pi'_{rr}(\mathbf{g}) + \sum_{u=1, u \neq r}^m \pi'_{ur}(\mathbf{g}) = \pi'_{rr}(\mathbf{g}) + \pi'_{mr}(\mathbf{g}) = 0,$$

which implies

$$\pi'_{rr}(\boldsymbol{\kappa}) = -\pi'_{mr}(\mathbf{g}) = -\rho \leq 0,$$

for $r = 1, \dots, m - 1$. The last inequality holds since $\pi_r(\mathbf{g})$ is non-increasing at $g_r = \kappa_r$.

Remark A4. Combining the results in Remarks 1 to 3, we have $\pi_u(\boldsymbol{\kappa}) = \kappa_u$ for all $u = 1, \dots, m$; $\pi'_{ur}(\boldsymbol{\kappa}) = 0$ for $u, r = 1, \dots, m - 1$ and $u \neq r$; and $\pi'_{mr}(\boldsymbol{\kappa}) = -\pi'_{rr}(\boldsymbol{\kappa}) = \rho \geq 0$ for $r = 1, \dots, m - 1$.

With the imbalance measure

$$g_{nu} = U_{nu}(\mathbf{X}_{(n+1)}) = \frac{\sum_{i=1}^n \prod_{k=1}^p K_{h_n}(X_{ik} - X_{(n+1)k}) I_{iu}}{\sum_{i=1}^n \prod_{k=1}^p K_{h_n}(X_{ik} - X_{(n+1)k})}, \tag{A.1}$$

if $\mathbf{U}_n = (U_{n1}, \dots, U_{n(m-1)})^T$, the allocation probability is $\pi_u(\mathbf{g}_n) = \pi_u\{\mathbf{U}_n(\mathbf{X}_{(n+1)})\}$, where π_u satisfies Conditions (C1) and (C2).

A.2. Asymptotic properties

We need additional conditions for the theoretical development.

(C3) In the kernel function $K_{h_n}(t) = K(t/h_n)/h_n$, K is a second order symmetric kernel function that satisfies $\int K(t)dt = 1$, $\int K(t)^2dt < \infty$, and $\int t^2K(t)^2dt < \infty$. h_n satisfies $nh_n^2 \rightarrow \infty$, and $nh_n^4 \rightarrow 0$.

(C4) The density function $f_k(X_k)$ is bounded away from zero and infinity almost surely on its support for all k .

(C5) X_k^2 is a uniformly integrable random variable.

(C6) Let $n_0 > \rho \geq 0$. If the first n_0 patients are randomized to arms $1, \dots, m$ with probabilities $\kappa_1, \dots, \kappa_m$, respectively, the adaptive allocation process starts from the $(n_0 + 1)$ th patient.

Let $I_{0u} = 0$. When the desired allocation ratio in the long run is κ_u , $u = 1, \dots, m$, we show that the covariate equilibrium of the similarity weighted

biased coin design, $\mathbf{D}_{nu} \equiv \sum_{i=1}^n (I_{iu} - \kappa_u) \mathbf{X}_i$, has mean zero and is asymptotically normally distributed. We first state the asymptotic property for similarity weighted biased coin design with one covariate.

Lemma A1. *Assume that Conditions (C1) – (C6) hold. Let $Q = E(X_i^2)$, $D_{nu} = \sum_{i=1}^n (I_{iu} - \kappa_u) X_i$,*

$$\mathbf{\Omega} = \begin{pmatrix} (1 + 2\rho)^{-1}(1 - \kappa_1)\kappa_1 & \dots & -(1 + 2\rho)^{-1}\kappa_1\kappa_{m-1} \\ \vdots & \ddots & \vdots \\ -(1 + 2\rho)^{-1}\kappa_1\kappa_{m-1} & \dots & (1 + 2\rho)^{-1}(1 - \kappa_{m-1})\kappa_{m-1} \end{pmatrix} Q$$

and $\mathbf{D}_n = (D_{n1}, \dots, D_{n(m-1)})^T$. Then the similarity weighted biased coin design with the allocation probability $\pi_u\{\mathbf{U}_n(X_{n+1})\}$, $n^{-1/2}\mathbf{\Omega}^{-1/2}\mathbf{D}_n$ converges to a standard multivariate normal distribution.

The proof of Lemma 1 is given in the Supplementary Material.

References

- Antognini, A. B. and Giovagnoli, A. (2004). A new ‘biased coin design’ for the sequential allocation of two treatments. *Journal of the Royal Statistical Society C Applied Statistics* **53**, 651–664.
- Atkinson, A. (1982). Optimum biased coin designs for sequential clinical trials with prognostic factors. *Biometrika* **69**, 61–67.
- Campbell, T. B., Smeaton, L. M., Kumarasamy, N., Flanigan, T., Klingman, K. L., Firnhaber, C., Grinsztejn, B., Hosseinipour, M. C., Kumwenda, J., Lalloo, U., Riviere, C., Sanchez, J., Melo, M., Supparatpinyo, K., Tripathy, S., Martinez, A. I., Nair, A., Walawander, A., Moran, L., Chen, Y., Snowden, W., Rooney, J. F., Uy, J., Schooley, R. T. De Gruttola, V., Hakim, J. G. and P. study team of the ACTG. (2012). Efficacy and safety of three antiretroviral regimens for initial treatment of HIV-1: a randomized clinical trial in diverse multinational settings. *PLOS Medicine* **9**, e1001290.
- Efron, B. (1971). Forcing a sequential experiment to be balanced. *Biometrika* **58**, 403–417.
- Epanechnikov, V. A. (1969). Non-parametric estimation of a multivariate probability density. *Theory of Probability and its Applications* **14**, 153–158.
- Frane, J. W. (1998). A method of biased coin randomization, its implementation, and its validation. *Drug Information Journal* **32**, 423–432.
- Hall, P. (1981). Laws of the iterated logarithm for nonparametric density estimators. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **56**, 47–61.
- Hall, P. (1984). Asymptotic properties of integrated square error and cross-validation for kernel estimation of a regression function. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **67**, 175–196.
- Hall, P. (1992). Effect of bias estimation on coverage accuracy of bootstrap confidence intervals for a probability density. *The Annals of Statistics* **67**, 675–694.
- Hall, P. (1993). On edgeworth expansion and bootstrap confidence bands in nonparametric

- curve estimation. *Journal of the Royal Statistical Society B Statistical Methodology* **55**, 291–304.
- Hall, P. and Heyde, C. (1980). *Martingale Limit Theory and its Application*. Probability and Mathematical Statistics. Academic Press, New York.
- Hall, P. and Horowitz, J. (2013). A simple bootstrap method for constructing nonparametric confidence bands for functions. *The Annals of Statistics* **41**, 1892–1921.
- Hall, P. and Marron, J. (1988). Choice of kernel order in density estimation. *The Annals of Statistics* **16**, 161–173.
- Hu, F. and Zhang, L.-X. (2004). Asymptotic properties of doubly adaptive biased coin designs for multitreatment clinical trials. *The Annals of Statistics* **32**, 268–301.
- Hu, Y. and Hu, F. (2012). Asymptotic properties of covariate-adaptive randomization. *The Annals of Statistics* **40**, 1794–1815.
- Ma, W., Hu, F. and Zhang, L. (2015). Testing hypotheses of covariate-adaptive randomized clinical trials. *Journal of the American Statistical Association* **110**, 669–680.
- Ma, Z. and Hu, F. (2013). Balancing continuous covariates based on kernel densities. *Contemporary Clinical Trials* **34**, 262–269.
- Pocock, S. J. and Simon, R. (1975). Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. *Biometrics* **31**, 103–115.
- Shao, J., Yu, X. and Zhong, B. (2010). A theory for testing hypotheses under covariate-adaptive randomization. *Biometrika* **97**, 347–360.
- Smith, R. L. (1984). Properties of biased coin designs in sequential clinical trials. *The Annals of Statistics* **12**, 1018–1034.
- Stigsby, B. and Taves, D. R. (2010). Rank-minimization for balanced assignment of subjects in clinical trials. *Contemporary Clinical Trials* **31**, 147–150.
- Su, Z. (2011). Balancing multiple baseline characteristics in randomized clinical trials. *Contemporary Clinical Trials* **32**, 547–550.
- Taves, D. R. (1974). Minimization: a new method of assigning patients to treatment and control groups. *Clinical Pharmacology and Therapeutics* **15**, 443.
- Wei, L. J. (1978). The adaptive biased coin design for sequential experiments. *The Annals of Statistics* **6**, 92–100.
- Yuan, Y., Huang, X. and Liu, S. (2011). A Bayesian response-adaptive covariate-balanced randomization design with application to a leukemia clinical trial. *Statistics in Medicine* **30**, 1218–1229.

Department of Statistics and Actuarial Science, University of Hong Kong, Pokfulam, Hong Kong.

E-mail: feijiang@hku.hk

Department of Statistics, Pennsylvania State University, University Park, PA 16802, USA.

E-mail: yanyuanma@gmail.com

Department of Statistics and Actuarial Science, University of Hong Kong, Pokfulam, Hong Kong.

E-mail: gyin@hku.hk

(Received December 2016; accepted April 2017)