Published in final edited form as:
Stat Med. 2014 December 30; 33(30): 5239–5248. doi:10.1002/sim.6266.

A better alternative to stratified permuted block design for subject randomization in clinical trials

Wenle Zhao
Department of Public Health Sciences, Medical University of South Carolina, Charleston, SC, 29425, USA

SUMMARY

Stratified permuted block randomization has been the dominant covariate-adaptive randomization procedure in clinical trials for several decades. Its high probability of deterministic assignment and low capacity of covariate balancing have been well recognized. The popularity of this sub-optimal method is largely due to its simplicity in implementation and the lack of better alternatives. Proposed in this paper is a two-stage covariate-adaptive randomization procedure that uses the block urn design or the big stick design in stage one to restrict the treatment imbalance within each covariate stratum, and uses the biased-coin minimization method in stage two to control imbalances in the distribution of additional covariates that are not included in the stratification algorithm. Analytical and simulation results show that the new randomization procedure significantly reduces the probability of deterministic assignments, and improves the covariate balancing capacity when compared to the traditional stratified permuted block randomization.

Keywords

clinical trial; randomization; permuted block; block urn design; big stick design; biased-coin minimization method

1. INTRODUCTION

Randomization in clinical trials provides broad comparability of treatment groups and validates the use of statistical methods for the analysis of results [1], [2]. Rosenberger and Lachin distinguish randomization procedures into four classes: complete randomization, restricted randomization, covariate-adaptive randomization, and response-adaptive randomization [3]. Both complete randomization and response-adaptive randomization are not commonly used [2]. Restricted randomization is employed when it is desired to have equal numbers of patients assigned to each treatment group [3]. The primary motivation is to maximize power of the trial [2]. However, it is well known that the power loss due to treatment imbalance under complete randomization is trivial, unless the sample size is extremely small [2][4]. Covariate-adaptive randomization is the most commonly used treatment assignment procedure for randomized controlled clinical trials. Imbalance in important baseline covariates can yield misleading results [2]. The motivation of using

Correspondence to: Wenle Zhao, Department of Public Health Sciences, Medical University of South Carolina, 135 Cannon Street, Suite 305H, Charleston, SC, 29425, U.S.A. zhaow@musc.edu.

covariate-adaptive randomization is to reduce the imbalances between treatment groups with respect to certain known covariates. To achieve this goal, three different approaches have been proposed. Stratified randomization controls treatment imbalances within each covariate stratum [5][6]. The minimization method proposed by Taves [7], and Pocock and Simon [8] controls treatment imbalances in covariate margins. The optimal allocation procedure proposed by Begg and Iglewicz [9], and Atkinson [10] attempts to minimize the variance of the estimate of the treatment effect in the presence of covariates. The use of the minimization method in its deterministic format has been controversial due to lack of randomness in the treatment allocation [11]. It is estimated that among all randomized controlled trials, less than 2% use the minimization method [12]. Trials using optimal allocation procedures are even rarer. The overwhelming majority of covariate-adaptive randomization procedures are based on stratified randomization, where a restricted randomization design is employed within each stratum independently.

Stratified permuted block randomization [13] is the most popular covariate-adaptive randomization procedure currently used in clinical trial practice, and is recommended by regulatory guidelines for multicenter trials [14]. It uses the permuted block design (PBD) to consistently control treatment imbalance within each stratum to a pre-specified maximal tolerated imbalance (MTI), which equals the half the block size. This procedure can be easily implemented in multicenter trials without requiring sophisticated technologies. Unfortunately the stratified permuted block randomization inherits disadvantages from both the PBD and the stratified randomization procedure. First, the PBD lacks treatment allocation randomness and is vulnerable to selection bias [15]. Its use in clinical trials has been strongly criticized [16]–[19]. Second, limited by the sample size and the total number of strata, a stratified randomization procedure usually balances no more than two or three covariates [3][14][20]. Nevertheless, primarily due to the lack of better alternatives, stratified permuted block randomization remains the most popular covariate-adaptive randomization procedure in clinical trial practice.

In section 2, a two-stage covariate-adaptive randomization procedure is proposed to reduce the probability of deterministic assignment and to allow balancing more covariates that are not included in the stratification algorithm. Stage one uses stratified randomization with the block urn design or the big stick design to restrict treatment imbalance within each stratum. Stage two uses the biased-coin minimization method to control imbalances in the distribution of additional covariates not included in the stratification algorithm. Stage two is triggered when the stage one algorithm yields a complete random assignment. Section 3 compares the proposed randomization procedure to the stratified permuted block randomization and the minimization method under different trial settings. Section 4 discusses the implementation and limitations of the proposed randomization procedure.

2. METHOD

The disadvantages associated with the stratified permuted block randomization can be overcome by the following two steps. First, replace the permuted block design by a restricted randomization design that has lower probability of deterministic assignment while preserving the same maximal tolerated imbalance (MTI). Second, when no balancing effort

needed within the stratum, employ a covariate-adaptive randomization procedure to control the imbalances in the distribution of additional covariates between the treatment arms. These two steps are further explored below.

2.1. Better alternatives to permuted block design

Stratified randomization procedures use a restricted randomization design within each stratum. The permuted block design (PBD) [21] is the most common restricted randomization design used in stratified randomization procedures. Several other restricted randomization designs have been proposed in the past few decades. Some of them, such as the random allocation rule [4], the truncated binomial design [22], and the maximal procedure [23], enforce perfect balance between the two treatment arms at the end of the trial. They require the number of total subjects (that is, the length of the treatment allocation sequence) be specified at the beginning of the study. These designs are not applicable for stratified randomization procedures, because the size of each stratum is usually unknown before the end of the study. Other designs, including Efron's biased coin design [24], Wei's urn design [25], and Smith's generalized biased coin design [26], do not restrict the treatment imbalance by the pre-specified maximal tolerated imbalance (MTI), and therefore are also not considered here. Soares and Wu's big stick design (BSD) [27], and Zhao and Weng's block urn design (BUD) [28] are potential alternatives to the PBD. Like the PBD, both the BSD and the BUD apply the MTI restriction throughout the study. Unlike the PBD which enforces perfect balance at the end of each block, the BSD and the BUD do not have the block issue. They are restricted only by the two boundaries formed by the MTI.

The randomization process for the PBD, the BSD, and the BUD can be illustrated by a model with two urns, one active and one inactive. The trial starts with an empty inactive urn and a full active urn, in which there are δ white balls for arm *A* and δ black balls for arm *B*. When a treatment assignment is requested, with the PBD and the BUD, a ball is randomly selected from the active urn. For the BSD, if a pair (one white and one black) of balls is available in the active urn, one ball from this pair is randomly selected. Otherwise, when all balls in the active urn are of the same color, one ball from the active urn is picked. The treatment assignment is made according to the color of the selected ball. Then, this ball is placed in the inactive urn. Under the BUD and the BSD, whenever a pair of balls (one white and one black) is available in the inactive urn, the pair of balls is returned to the active urn immediately. For the PBU, all balls (δ white and δ black) are returned to the active urn when the active urn is empty.

The treatment assignments for the PBD, the BSD, and the BUD can be made based on the conditional allocation probability. Consider a stratum in a two-arm trial with an equal allocation ratio, let δ = MTI, and $b = 2\delta$ be the block size for the PBD. For the i^{th} subject in the stratum, let $n_{i-1,A}$ and $n_{i-1,B}$ be the number of subjects previously assigned to arm *A* and *B* respectively, $k_{i-1} = \text{int}((i-1)/b)$, where function int($x$) rounds a number x down to the nearest integer, be the number of completed blocks and $k^*_{i-1} = \min\left(n_{i-1-A}, n_{i-1,B}\right)$ be the number of completed pairs (one *A* and one *B*) in the previous$(i-1)$ subjects. The conditional allocation probabilities for the PBD, the BUD, and the BSD can be defined as follows:

$$p_{i,A}(PBD)=\frac{\delta+\delta k_{i-1}-n_{i-1,A}}{2\delta+2\delta k_{i-1}-(i-1)}, \quad (1)$$

$$p_{i,A}(BUD)=\frac{\delta+k^{*}_{i-1}-n_{i-1,A}}{2\delta+2k^{*}_{i-1}-(i-1)}, \quad (2)$$

$$p_{i,A}(BSD)=\begin{cases} 0 & \text{if } n_{i-1,A}-n_{i-1,B}=\delta, \\ 0.5 & \text{if } |n_{i-1,A}-n_{i-1,B}|<\delta, \\ 1 & \text{if } n_{i-1,A}-n_{i-1,B}=-\delta. \end{cases} \quad (3)$$

The treatment assignment T_i is made by comparing $p_{i,A}$ to the value of a random number R_i with a uniform distribution on (0, 1). The subject is assigned to arm *A* if $R_i \le p_{i,A}$, otherwise to arm *B*. Treatment assignment T_i is defined as deterministic if $p_{i,A} = 1$ or $p_{i,A} = 0$, and is considered as complete random if $p_{i,A} = 0.5$.

The statistical properties of the PBD, the BSD, and the BUD have been well studied [13] [19] [27][31]. With the same value of the MTI, treatment allocation randomness is the focus of the comparison for the three randomization designs. Probability of deterministic assignments and correct guess probability are two commonly used measures for treatment allocation randomness. Deterministic assignment is defined based on the conditional allocation probability (1–3). Correct guess is defined based on the Blackwell and Hodges' convergence strategy [22], in which the next assignment is always guessed as the arm currently has enrolled fewer patients. In case of perfect balance, the guess is made completely at random. The analytical results of the probability of deterministic assignment and correct guess probability for the PBD, the BSD, and the BUD are provided by Matts and Lachin [13], Kundt and Chen [30][31], and Zhao and Weng [28] respectively, as shown in Table 1. The maximal procedure (MP) is included in the comparison. It assigns an equal probability for all possible treatment allocation sequences under the restriction of the MTI and the pre-specified allocation sequence length. The MP is not easy to be implemented in a stratified randomization procedure due to two reasons. First, the stratum size is usually unknown before the end of the study. Second, the MP does not have an analytical format for the conditional allocation probability. The MP is included in Table 1 for comparison purposes due to its excellent treatment allocation randomness. Data for the MP are obtained through computer simulation using the MP randomization sequence generation algorithm proposed by Salama et al. [32]. The simulation program uses a sample size of 300, in order to obtain stable assessments comparable to those obtained based on analytical formulas for the other three designs.

The BUD has the lowest probability of deterministic assignment and this probability decreases as the MTI increases. For example, when MTI = 3, the probability of deterministic assignment for the BUD is only 5.9%, compared to 25% for the PBD. When the MTI is greater than 3, the risk of selection bias caused by deterministic assignment becomes trivial for the BUD. The BSD has the lowest correct guess probability. The MP has a probability of

deterministic assignment and a correct guess probability between those of the BSD and the BUD. Among the four restricted randomization designs compared in Table 1, the PBD has both the highest probability of deterministic assignment, the highest correct guess probability, and is the most vulnerable design in terms of selection bias. Replacing the PBD with either the BUD or the BSD in a stratified randomization procedure can significantly enhance the treatment allocation randomness, and reduce the risk of selection bias.

2.2. Balancing covariates beyond stratification

In a traditional stratified randomization procedure, a restricted randomization design is employed independently within each covariate stratum. When the treatment imbalance occurs within the stratum, a biased coin or a deterministic assignment is used to reduce the imbalance. When the two arms are perfectly balanced within the stratum (for the PBD and the BUD), or when the treatment imbalance within the stratum is less than the MTI (for the BSD), the conditional allocation probability for the current subject will be $p_{i,A} = 0.5$. This complete random assignment makes no contribution to the control of the treatment imbalance within the stratum. However, it creates an opportunity to balance the distribution of other baseline covariates that are not included in the stratification algorithm. This led to the proposed two-stage covariate-adaptive randomization procedure. Stage one uses stratified BUD or BSD to restrict treatment imbalances within each stratum. A complete random assignment from the stage one triggers the stage two of the randomization procedure, using a covariate-adaptive randomization to balance the distribution of additional covariates. The covariate balancing capacity of stage two depends on the amount of complete random assignment left from stage one and the covariate-adaptive randomization design used in stage two. Table 2 lists the expected probability of complete random assignment for the PBD, the BSD, and the BUD based on formulas in [13][28][30][31].

The BSD has a higher probability of complete random assignment than the BUD and the PBD have. When MTI = 3, the probability of complete random assignment for the BSD, the BUD and the PBD is 83.3%, 26.5% and 36.7% respectively. As the MTI increases, the probability of complete random assignment increases in the BSD, but decreases in the PBD and the BUD. The selection between the BSD and the BUD as the within stratum randomization design depends on the tolerance levels on the probability of deterministic assignment, and the number of additional covariates to be balanced.

The second stage uses a covariate-adaptive randomization procedure to balance the distributions of covariates beyond the stratification algorithm. Possible candidates include Efron's biased coin design (BCD) [24], Chen's biased coin design with imbalance tolerance (BCDWIT) [33], and the biased-coin minimization method (BCM) [34]. The BCM is a better choice than the BCD and the BCDWIT, because it can simultaneously balance the distributions of multiple covariates. Let $n_{i-1,\,j,A}$ and $n_{i-1,\,j,B}$ be the number of subjects in the study with the category of the j^{th} ($j = 1, 2, \cdots, m$) covariate same as current subject (the i^{th} subject of his/her stratum). Let $F_{i,A}=\sum_{j=1}^{m} w_j \left|n_{i-1,j,A}-n_{i-1,j,B}+1\right|$ and $F_{i,B}=\sum_{j=1}^{m} w_j \left|n_{i-1,j,A}-n_{i-1,j,B}-1\right|$ be the weighted sums of absolute marginal imbalances

assuming the current subject is assigned to arm *A*, and *B* respectively. The conditional allocation probability for the BCM is defined as:

$$G_i = F_{i,A} - F_{i,B} = \sum_{j=1}^{m} w_j \left| n_{i-1,j,A} - n_{i-1,j,B} + 1 \right| - \sum_{j=1}^{m} w_j \left| n_{i-1,j,A} - n_{i-1,j,B} - 1 \right| \quad (4)$$

$$p_{i,A}(BCM) = \begin{cases} p_{bc} & \text{if } G_i < 0 \\ 0.5 & \text{if } G_i = 0 \\ 1 - p_{bc} & \text{if } G_i > 0 \end{cases} \quad (5)$$

Here $p_{bc} > 0.5$ is the biased coin probability; G_i is the difference between the weighted sums of the absolute marginal imbalances under the two treatment assignment assumptions.

3. RESULTS

To evaluate the performance of the proposed two-stage covariate-adaptive randomization procedure, and compare it to the traditional stratified permuted block randomization, a two-arm multicenter trial with an equal allocation ratio is considered. Potential confounding factors include clinical site and a few other covariates. In actual trials, covariates may be correlated to each other, and each has different distribution among study subjects. To simplify the simulation program, all covariates are assumed independent and have a binomial distribution of $B(1, 0.5)$ for each study subject. All covariates included in the randomization procedure are weighted equally. The chance the current subject is enrolled in any site is assumed equal.

The simulation program includes the following parameters:

- n: sample size of the trial
- s: number of clinical sites
- m_1: number of covariates included in the stratification algorithm in addition to site
- MTI: maximal tolerated imbalance within stratum
- m_2: number of covariates to be balanced in stage two
- p_{bc}: biased coin probability for the minimization method in stage two

Randomization procedures included in the simulation study for performance comparisons are:

P1) Complete randomization

P2) Stratified (by site only) permuted block randomization

P3) Stratified (by site and additional covariates) permuted block randomization

P4) Deterministic minimization method balancing multiple covariates

P5) Stratified (by site only) big stick block randomization plus biased-coin minimization method balancing additional covariates

P6) Stratified (by site only) block urn randomization plus biased-coin minimization method balancing additional covariates

The performances of these six randomization procedures are evaluated based on measures in three domains: treatment allocation randomness, treatment imbalances, and covariate imbalances. Treatment allocation randomness is measured by the probability of deterministic assignment *DA*, and the correct guess probability *CG*. They are estimated as follows:

$$DA=\frac{1}{n\times n_{simu}}\sum_{j=1}^{n_{simu}}\sum_{i=1}^{n}\begin{cases}1 & p_{iA}=0 \text{ or } p_{iA}=1\\ 0 & \text{Otherwise}\end{cases} \quad (6)$$

$$Guess_i=\begin{cases}A & \text{if } n_{i-1,site,A}<n_{i-1,site,B}\\ B & \text{if } n_{i-1,site,A}>n_{i-1,site,B}\\ null & \text{if } n_{i-1,site,A}=n_{i-1,site,B}\end{cases} \quad (7)$$

$$CG=\sum_{j=1}^{n_{simu}}\sum_{i=1}^{n}\begin{cases}1 & \text{if } n_{i-1,site,A}\neq n_{i-1,site,B} \text{ and } T_i=Guess_i\\ 0.5 & \text{if } n_{i-1,site,A}=n_{i-1,site,B}\\ 0 & \text{if } n_{i-1,site,A}\neq n_{i-1,site,B} \text{ and } T_i\neq Guess_i\end{cases} \quad (8)$$

Here n is the sample size, n_{simu} is the number of simulation runs, p_{iA} is the conditional treatment allocation probability for subject i, $n_{i-1,site,A}$ and $n_{i-1,site,B}$ are the numbers of subjects previously assigned to arm A and B respectively within the site of subject i, T_i is the treatment assignment for subject i, and $Guess_i$ is the guess of T_i, made based on the within site treatment imbalance.

Treatment imbalances include the overall imbalance and the within site imbalance. Let n_{jA} and n_{jB} be the total number of subjects assigned to arm A and B respectively at the end of the study in simulation j, $n_{j,k,A}$ and $n_{j,k,B}$ be the number of subjects in site k assigned to arm A and B respectively in simulation j, $d_j = n_{j,A} - n_{j,B}$ be the overall treatment imbalance observed in simulation j, and $d_{j,k} = n_{j,k,A} - n_{j,k,B}$ be the treatment imbalance within site k in simulation j. The overall imbalance $D_{overall}$ and the within site imbalance D_{site} are estimated by:

$$D_{overall}=\sqrt{\frac{1}{n_{simu}-1}\sum_{j=1}^{n_{simu}}\left(d_j-\bar{d}\right)^2} \quad (9)$$

$$D_{site}=\frac{1}{n_{simu}}\sum_{j=1}^{n_{simu}}\left[\max_{k=1\sim n_{site}}\left(|d_{j,k}|\right)\right] \quad (10)$$

The overall treatment imbalance is quantified by the standard deviation of the observed imbalance in each simulation run. With multiple sites involved in the study, the maximal

absolute within site imbalance from each simulation run is averaged as the measure for within site imbalance.

Covariate imbalance is measured by the difference of treatment arm sizes within covariate marginal. Because all covariates are considered independent and having the same binomial distribution $B(1,0.5)$ for all subjects, imbalances in all covariates controlled by the stage two will have the same distribution. Let $n_{j,x,A}$ and $n_{j,x,B}$ be the number of subjects in a covariate margin x assigned to arm A and B respectively in simulation j, $d_{j,x} = n_{j,x,A} - n_{j,x,B}$ be the imbalance within the covariate margin x in simulation j. The covariate imbalance is measured by:

$$D_X = \sqrt{\frac{1}{n_{simu}-1}\sum_{j=1}^{n_{simu}}\left(d_{j,x}-\bar{d}_x\right)^2} \quad (11)$$

It is noticed that the p-value of the baseline covariate distribution test has been used in practice as a measure of covariate imbalance. A small p-value is often considered as an indicator for serious covariate imbalances and potential selection bias. For example, Berger identified 30 trials with direct trial-level evidence of selection bias [35], 19 of them were identified primarily based on the p-values of covariate imbalance tests. In this simulation study, a Chi-square test is performed at the end of each simulation run to exam the distribution of covariate X between the two arms. The 1st percentile and the 5th percentile of the p-values of these tests from all simulation runs are calculated as measures of covariate imbalances, denoted as $pp1_X$ and $pp5_X$ respectively. For example, $pp1_X = 0.3$ indicates that there is a 1% chance the imbalance test for covariate X has a p-value less than 0.3; $pp5_X = 0.8$means that there is a 5% chance the imbalance test will have a p-value less than 0.8. It is important to remember that under complete randomization, there are $pp1_X = 0.01$ and $pp5_X = 0.05$. Table 3 shows the computer simulation results comparing the proposed randomization procedure to other commonly used randomization procedures under different trial settings.

Scenario 1 uses the complete randomization without any restrictions on the treatment assignment. Each subject has 50% chance being assigned to either arm A or arm B. When sample size n is large, the standard deviation of the overall treatment imbalance can be estimated based on the binomial distribution, $D_{Overall} = \sqrt{Var[n_A - n_B]} = \sqrt{2\,Var(n_A)} = \sqrt{n/2} = 15.81$, which is very close to the result observed in the simulation. As expected, the 1st and 5th percentiles of the covariate imbalance test p-value are approximately 0.05 and 0.01 respectively. With the highest level of treatment allocation randomness, the largest treatment imbalance, and lowest p-value percentile for covariate imbalance tests, this scenario is included as a reference point for other randomization procedures.

Scenarios 2 and 3 use the permuted block randomization stratified by site only, with the block size of 4 and 6 respectively. This design is aimed to control the overall treatment imbalance and the within site imbalance. While both imbalances are properly controlled, the

probability of deterministic assignment is high, and the covariate imbalance is not controlled.

Scenarios 4 and 5 use the permuted block design stratified by site and three additional covariates, with the objective of controlling covariate imbalances. Due to the inclusion of the three covariates, the total number of strata becomes $25 \times 2^3 = 200$, and the average stratum size is reduced to 2.5, smaller than the block size. In this case, most blocks are incomplete. The overall treatment imbalance control becomes weak. When block size of 6 (i.e. MTI=3) is used, the performance of the randomization procedure is close to that of the complete randomization.

Scenario 6 is the deterministic minimization method controlling the marginal imbalances in site and four additional covariates. Simulation results demonstrate that, the minimization method offers the most tighten control on multiple covariate imbalances. The 1st percentile of the covariate imbalance test *p*-value reaches 0.854, indicating that among the 10,000 simulation runs, only 1% time a covariate imbalance test yields a *p*-value less than 0.854. The cost paid for the tighten covariate balance is the 87.2% deterministic assignment, which is the primary reason the minimization method has been criticized for the concern of selection bias.

Scenarios 7 and 8 use the proposed two-stage covariate-adaptive randomization procedure; with the big stick design in stage one controlling treatment imbalances within site and the biased-coin minimization method in stage two controlling imbalances in the distributions of four covariates. This randomization procedure provides an effective control on treatment imbalances and covariate imbalances comparable to those of the deterministic minimization method, but has a lower probability of deterministic assignment, 22.9% for MTI = 2 and 13.8% for MTI = 3.

Scenarios 9 and 10 use the two-stage randomization procedure with the block urn design in stage one and the biased-coin minimization method in stage two. This procedure significantly reduces the probability of deterministic assignment while maintaining a sufficient control on treatment imbalances and covariate imbalances. The 5th percentile of covariate imbalance test *p*-value is 0.373 for MTI = 2, and 0.290 for MTI = 3 respectively. The chance covariate imbalances become a concern due to a small *p*-value in imbalance tests is practically eliminated.

Among the above six randomization procedures compared for a medium size trial with a sample size of 500 from 25 sites, the two-stage BUD+BCM procedure exhibits the best overall performance, and is therefore recommended to replace the stratified permuted block design when the controls of both treatment imbalances and covariate imbalances are needed.

A small trial with 100 subjects from 5 sites and 2 covariates, and a large trial with 1500 subjects from 50 sites and 4 covariates are included in the computer simulation. The MTI is set to 2 for the small trial and 3 for the large trial. Computer simulation results in Table 3 (scenarios 11–16) shown that performances of the BUD+BCM procedure and the BSD +BCM procedure are better than those of the stratified PBD for both the small and the large

trial, with regard to the treatment allocation randomness, the treatment imbalance control, as well as the covariate imbalance control.

The biased-coin probability use in stage two affects the performance of the biased-coin minimization procedure. Figure 1 indicated that as the biased-coin probability increase, both the overall treatment imbalance and the covariate imbalance decrease. It is suggested that the biased-coin probability be selected based on the 1st or the 5th percentile of the covariate imbalance test *p*-value. For example, with $p_{bc} = 0.75$, the 1st percentile for the imbalance test p-value is 0.211. It indicates that a Chi-square test for the imbalance of a covariate will have 99% chance being greater than 0.211. The chance any one of the three covariates having an imbalance test *p*-value less than 0.211 is about 97%. If $p_{bc} = 0.65$ or $p_{bc} = 0.85$ was selected, the 1st percentile of the *p*-value will be 0.106 and 0.277 respectively. It is suggested that $p_{bc} = 0.75$ be used when the number of covariates to be balanced is no more than 3. Otherwise, a slightly high biased-coin probability can be considered.

4. DISCUSSION

The implementation of the proposed two-stage covariate-adaptive randomization procedure in multicenter clinical trials requires a centralized subject randomization system. Based on the conditional treatment allocation probability functions (2–5), a real-time randomization algorithm can be developed without complex programming. The proposed method can be used in multicenter trials when preventing serious imbalances in the distributions of multiple covariates are desired.

It is also noticed that a statistical test for the distribution of baseline covariate is a controversial topic. The purpose of using the test *p*-value as a measure of covariate imbalance in this paper is to compare the effectiveness of covariate balancing for different randomization procedures. A small *p*-value could trigger challenges on suspicious selection bias, although it could simply be a small chance random phenomenon. To protect the trial, it is a natural choice to include those important covariates in the randomization algorithm, in order to reduce the chance of seeing a *p*-value less than 0.05 from 5% to a comfortable level like < 0.5%.

The intention of this paper is to present an alternative to the commonly used stratified permuted block randomization procedure. In fact, if the restriction of the maximal tolerated imbalance within each stratum is released, one can apply the biased-coin minimization method directly to the study to gain good covariates balancing without any deterministic assignment. When the imbalances in covariate margins are controlled, the overall treatment imbalance will be well controlled.

Like all covariate-adaptive randomization designs, the use of the proposed randomization procedure may affect the way the trial data be analyzed. Balancing of baseline covariates does not remove nor reduce the confounding impacts of these covariates on the estimation of the treatment affect. A conventional method is to adjust all covariates used in the randomization algorithm in the analysis model. It is clear that adjusting baseline covariates with confounding factors will increase the power of the trial, disregards whether or not they have been balanced in the randomization model. However, the inclusion of multiple

covariates in the analysis model may make the interpretation and acceptance of the trial results hard.

When design a randomization procedure for a clinical trial, it is necessary to make sure that the baseline covariate balancing does not create practical risk for selection bias. In trials with a medium or large sample size, the concern for the power loss due to treatment imbalance is baseless. Tighten covariate balancing is also not necessary for randomized controlled clinical trials. The focus is to ensure the treatment allocation randomness to prohibit potential selection bias and to prevent statistically significant covariate imbalances that could raise concerns on suspicious selection bias.

Acknowledgments

The author gratefully acknowledges the comments and edits of colleague Dr. Viswanathan Ramakrishnan at the Medical University of South Carolina, and the helpful suggestions and comments from the editor, and the two anonymous reviewers. This research is partly supported by the NIH/NINDS grants U01NS054630, U01NS0059041, and U01NS087748.

References

1. Armitage P. The role of randomization in clinical trials. Stat Med. 1982; 1:345–352. [PubMed: 7187102]
2. Friedman, LM.; Furberg, CD.; DeMets, DL. Fundamentals of Clinical Trials. 3. Springer Science; 1988.
3. Rosenberger, WF.; Lachin, JM. Randomization in Clinical Trials Theory and Practice. Wiley-Intersince; 2002.
4. Lachin JM. Properties of simple randomization in clinical trials. Controlled Clinical Trials. 1988; 9:312–326. [PubMed: 3203523]
5. Zelen M. The randomization and stratification of patients to clinical trilas. Journal of Chronic Disease. 1974; 28:365–375.
6. McEntegart DJ. The pursuit of balance using stratified and dynamic randomization techniques: an overview. Drug Information Journal. 2003; 37:293–308.
7. Taves DR. Minimization: a new method of assigning patients to treatment and control groups. Clinical Pharmacology and Therapeutics. 1974; 15:443–453. [PubMed: 4597226]
8. Pocock SJ, Simon R. Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. Biometrics. 1975; 31:103–115. [PubMed: 1100130]
9. Begg CB, Iglewicz B. A treatment allocation procedure for sequential clinical trials. Biometrics. 1980; 36:81–90. [PubMed: 7370375]
10. Atkinson AC. Optimal biased coin designs for sequential clinical trials with prognostic factors. Biometrika. 1982; 69:61–67.
11. Berger VW. Minimization, by its Nature, Precludes Allocation Concealment, and Invites Selection Bias. Contemp Clin Trials. 2010; 31(5):406. [PubMed: 20457277]
12. Taves DR. The use of minimizaiton in linical trilas. Comtemp Clin Trials. 2010; 31(2):180–4.
13. Matts JP, Lachin JM. Control Properties of permuted-block randomization in clinical trials. Clin Trials. 1988; 9:327–344.
14. ICH Harmonized Tripartite Guideline E9: Statistical Principals for Clinical Trials. 5 February 1998.
15. Zhao W, Weng Y, Wu Q, Palesch Y. Quantitative comparison of randomization designs in sequential clinical trials based on treatment balance and allocation randomness. Pharm Stat. 2012; 11:39–48. [PubMed: 21544929]
16. Berger VW. Do not use blocked randomization. Headache. 2006A; 46(2):343. [PubMed: 16492254]

17. Berger VW. Misguided precedent is not a reason to use permuted blocks. Headache. 2006B; 46(7): 1210–12. [PubMed: 16866731]
18. Berger VW. Varying block sizes does not conceal the allocation. Journal of Critical Care. 2006C; 21(2):229. [PubMed: 16769475]
19. Zhao W. Selection bias, allocation concealment and randomization design in clinical trials. Contemp Clin Trials. 2013; 36:263–265. [PubMed: 23871796]
20. Kernan WN, Viscoli CM, Makuch RW, Brass LM, Horwitz RI. Stratified randomization for clinical trials. J Clin Epidemiol. 1999; 52(1):99–26.
21. Hill AB. The Clinical trial. Br Med Bull. 1951; 71:278–282. [PubMed: 14879025]
22. Blackwell D, Hodges JL. Design for the control of selection bias. Annals of Mathematical Statistics. 1957; 28:449–460.
23. Berger VW, Ivanova A, Deloria-Knoll M. Minimizing Predictability while Retaining Balance through the Use of Less Restrictive Randomization Procedures. Statistics in Medicine. 2003; 22:3017–3028. [PubMed: 12973784]
24. Efron B. Forcing a Sequential Experiment to be Balanced. Biometrika. 1971; 58:403–417.
25. Wei LJ. A Class of Designs for Sequential Clinical Trials. Journal of the American Statistical Association. 1977; 72:382–386.
26. Smith RL. Sequential Treatment Allocation Using Biased Coin Designs. Journal of the Royal Statistical Society Series B (Methodological). 1984; 46:519–543.
27. Soares JF, Wu CF. Some restricted randomization rules in sequential designs. Communications in Statistics - Theory and Methods. 1983; 12:2017–2034.
28. Zhao W, Weng Y. Block urn design - a new randomization algorithm for sequential trials with two or more treatments and balanced or unbalanced allocation. Contemp Clin Trials. 201; 32:953–61. [PubMed: 21893215]
29. Zhao W, Weng Y. A simplified formula for quantification of the probability of deterministic assignment in permuted block randomization. J Stat Plan Inference. 2011; 141(1):474–478. [PubMed: 21057655]
30. Kundt G. A new proposal for setting parameter values in restricted randomization methods. Methods Inf Med. 2007; 46:440–9. [PubMed: 17694238]
31. Chen YP. Which Design Is Better? Ehrenfest Urn versus Biased Coin. Advances in Applied Probability. 2000; 32:738–749.
32. Salama I, Ivanova A, Qaqish B. Efficient generation of constrained block allocation sequences. Stat Med. 2008; 27(9):1421–8. [PubMed: 17729381]
33. Chen YP. Biased coin design with imbalance tolerance. Commun Stat Stochastic Models. 1999; 15:953–75.
34. Efron, B. Randomizing and balancing a complicated sequential experiment. In: Miller, RG.; Efron, B.; Brown, BW.; Moses, LE., editors. Biostatistics Casebook. Wiley; New York: 1980. p. 19-30.
35. Berger, VW. Selection bias and covariate imbalances in randomized clinical trials. John Wiley & Sons; 2005.

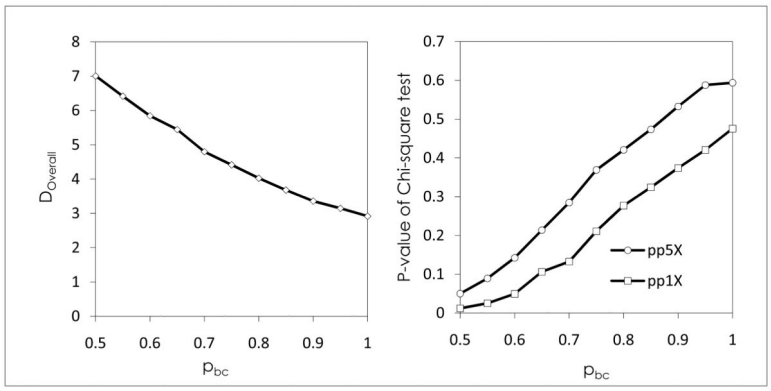

Figure 1.

Impact of the biased-coin probability on the performance of the Minimization Method

Sample size = 500; number of sites = 25; number of covariates = 3;

Stage one: Block urn design stratified by site;

Stage two: Biased-coin minimization balancing three covariates;

$D_{Overall}$: Standard deviation (over simulations) overall treatment imbalance.

pp1X, pp5X: The 1st, the 5th percentile of the p-value of the chi-square test for the distribution of the covariate X between the two arms.

Number of simulations: 5000 per scenario.

Table 1

Comparison of treatment allocation randomness

MTI	Probability of deterministic assignment				Correct guess probability			
	PBD	BSD	BUD	MP	PBD	BSD	BUD	MP
1	0.500	0.500	0.500	0.500	0.750	0.750	0.750	0.750
2	0.333	0.250	0.167	0.166	0.708	0.625	0.667	0.666
3	0.250	0.167	0.059	0.073	0.683	0.583	0.632	0.624
4	0.200	0.125	0.021	0.038	0.665	0.562	0.613	0.598
5	0.167	0.100	0.008	0.022	0.653	0.550	0.600	0.582
6	0.143	0.083	0.003	0.014	0.643	0.542	0.590	0.569
7	0.125	0.071	0.001	0.009	0.633	0.536	0.583	0.560
8	0.111	0.063	0.000	0.007	0.625	0.531	0.577	0.553

MTI: Maximal Tolerated Imbalance.

PBD: Permuted block design. Data obtained based on formulas in [4] and Hypergeometric distribution.

BSD: Big Stick Design. Data obtained based on formulas in [30,31].

BUD: Block Urn Design. Data obtained from [29].

MP: Maximal Procedure. Data obtained based on computer simulation with sample size = 300, using algorithm provided by Salama at el. in [32]. Ten thousand simulation runs are performed for each scenario.

Table 2

Proportion of Equal-Probability Assignment

Maximal Tolerated Imbalance	Permuted Block Design	Big Stick Design	Block Urn Design
1	0.500	0.500	0.500
2	0.417	0.750	0.333
3	0.367	0.833	0.265
4	0.332	0.875	0.225
5	0.306	0.9	0.199
6	0.294	0.917	0.180
7	0.270	0.929	0.166
8	0.256	0.937	0.154

Table 3

Performance comparison of six randomization procedures for a medium size trial

No.	Sample Size	# of Sites	Stage 1 (within stratum) Randomization			Stage 2 (across strata) Randomization			Allocation Randomness		Treatment Imbalance		Covariate Imbalance		
			Design	m_1	MTI	Design	m_2	P_{bc}	DA	CG	$D_{Overall}$	D_{Site}	D_X	$pp5_X$	$pp1_X$
1	500	25	Complete randomization						0	0.5	22.48	10.12	15.87	0.049	0.009
2			PBD	0	2				0.313	0.699	4.57	1.88	11.47	0.049	0.009
3			PBD	0	3				0.223	0.671	5.40	2.32	11.43	0.049	0.011
4			PBD	3	2				0.161	0.631	13.14	5.91	9.29	0.248	0.127
5			PBD	3	3				0.044	0.584	16.19	7.23	11.44	0.153	0.060
6			Minimization method				4	1.00	0.872	0.633	1.03	3.03	1.10	0.858	0.854
7			BSD	0	2	BCM	4	0.75	0.229	0.617	2.28	2.00	2.72	0.655	0.533
8			BSD	0	3	BCM	4	0.75	0.138	0.572	2.22	3.00	2.56	0.659	0.591
9			BUD	0	2	BCM	4	0.75	0.154	0.658	3.48	1.99	5.14	0.373	0.244
10			BUD	0	3	BCM	4	0.75	0.050	0.622	4.50	2.79	6.24	0.290	0.155
11	100	5	PBD	2	2				0.250	0.671	4.15	2.83	2.91	0.421	0.304
12			BSD	0	2	BCM	2	0.75	0.219	0.622	1.68	1.41	1.76	0.543	0.405
13			BUD	0	2	BCM	2	0.75	0.154	0.659	2.09	1.60	2.95	0.296	0.112
14	1500	50	PBD	4	3				0.022	0.566	30.53	10.89	21.74	0.119	0.039
15			BSD	0	3	BCM	4	0.75	0.148	0.575	2.31	3.00	2.61	0.797	0.757
16			BUD	0	3	BCM	4	0.75	0.053	0.626	5.24	2.95	6.94	0.502	0.354

Simulation run = 10,000 / scenario

m_1: Number of covariates included in the stratification algorithm.

MTI: Maximal tolerated imbalance within stratum.

m_2: Number of covariates included in the minimization method in stage 2.

P_{bc}: Biased coin probability for the minimization method in stage 2.

DA: Average (over simulations) proportion of deterministic assignment.

CG: Average (over simulations) probability of correct guess.

$D_{Overall}$: Standard deviation (over simulations) overall treatment imbalance.

DSite: Average (over simulations) maximal (over sites) within site imbalance.

DX: Standard deviation (over simulation) of marginal imbalance of covariate *X*.

pp1X, pp5X: The 1st, the 5th percentile of the p-value of the chi-square test for the distribution of the covariate X between the two arms.