Restricted Randomization in Studies With Unequal Allocation

Olga M. Kuznetsova [1]
[1]Merck & Co., Inc., 126 East Lincoln Ave, Rahway, NJ 07065-0900

Abstract

The allocation space of an unequal allocation permuted block randomization can be quite wide. The development of unequal allocation procedures with a narrower allocation space is complicated by the need to preserve the unconditional allocation ratio at every step (ARP property). When the allocation paths are depicted on the *K*-dimensional unitary grid, where allocation to the *j*-th treatment is represented by a step along the *j*-th axis, the ARP property can be expressed in terms of the center of the probability mass after i allocations. For an ARP procedure that randomizes subjects to *K* treatment groups in w_1: … :w_K ratio, $w_1+...+w_K=1$, the coordinates of the center of the mass are ($w_1 i$, …, $w_K i$). In this presentation the momentum with respect to the center of the probability mass (expected distance) is used to compare ARP procedures in how closely they approximate the target allocation ratio. It is shown that the 2-arm Brick Tunnel Randomization (BTR) by Kuznetsova and Tymofyeyev (2011) has the smallest momentum of all ARP procedures with the same allocation ratio. The resident and transition probabilities for 3-arm BTR are analytically derived.

Key Words: Brick Tunnel Randomization, Unequal Allocation, Large Block Size, Allocation Ratio Preserving Randomization, Restricted Randomization, Imbalance in Treatment Totals

1. Introduction

Unequal allocation is becoming more wide spread in clinical trials [1-5].

Equal allocation algorithms are usually symmetric with respect to the treatment arms and as a result, all patients are allocated with the same unconditional allocation ratio regardless of their place in the allocation sequence. For example, in a study with equal allocation to Control and Experimental treatments, the patients allocated first, second and third, all have 0.5 probability to be allocated to Experimental treatment.

This is not necessarily the case in studies with unequal allocation [6-10]. For example, in a study with 1:2 allocation to Control and Experimental treatments, the allocation procedure can be designed in such a way that the first three patients have unconditional probabilities of 0.2, 0.6, and 0.2 of being assigned Control treatment; the same pattern of unconditional probabilities is repeated for patients 4-6, 7-9, and so on. While the overall allocation ratio across large samples remains close to 1:2, such variations in the unconditional allocation ratio from allocation to allocation present a problem [6-11].

In single-center open-label studies with restricted randomization, where a sequence of all previous allocations is known to the investigator, selection bias is often a problem. When the investigator knows that, given the sequence of previous treatment assignments, the Experimental treatment is more likely to be assigned to the next patient than the Control treatment, he might choose to enroll a patient with a better prognosis; the patients with a worse prognosis are left to be enrolled when the Control treatment is more likely to be assigned. As a result, the Experimental group has patients with a better prognosis which leads to biased results.

Selection bias in studies without variations in the unconditional allocation can be typically prevented by proper blinding of the study. Not knowing the previous assignments, the investigator will not be able to predict which treatment is more likely to be assigned to the next patient. In studies with variations in the allocation ratio, however, blinding does not eliminate the selection bias. Indeed, if the investigator knows that the 1^{st}, 3^{rd}, 4^{th}, 6^{th}, 7^{th}, … patients have lower than average probability of being assigned to Control, while the 2^{nd}, 5^{th}, 8^{th}, … patients have higher than average probability of being assigned to Control, he can introduce selection bias without knowing the actual sequence of treatment assignments. The evaluation of the patients can also be biased by the knowledge of variations in the unconditional allocation ratio.

Thus, variations in the unconditional allocation ratio provide potential for selection and evaluation bias even in double blind studies [6-10]. Such variations can also result in an accidental bias – in particular, in multi-center studies with randomization stratified by center [11, 8-9] and lead to randomization test problems [11, 10]. As was demonstrated by Proschan et al. [11], the distribution of the randomization test statistic is shifted away from 0 when the unconditional allocation ratio varies from allocation to allocation, resulting in a lower power of the test. Kuznetsova and Tymofyeyev [10] derived the value of the shift from the sequence of the unconditional allocation ratios. Kaiser [12] demonstrated that such variations cause the treatment effect estimator to be biased from a randomization perspective and recommended to avoid allocation procedures that do not preserve the unconditional allocation ratio.

Kuznetsova and Tymofyeyev [10] call procedures that preserve the unconditional allocation ratio at every step Allocation Ratio Preserving (ARP) procedures.

Unfortunately, the need for ARP property in studies with unequal allocation was not recognized until recently and many unequal allocation procedures lack this property. Among non-ARP unequal allocation procedures are urn design [1, p.51], expansion of the maximal procedure by Salama et al. [13], biased coin randomization and minimization expansion by Han et al. [14], doubly adaptive biased coin design procedure by Hu and Zhang [15] applied to fixed unequal allocation as described by Sverdlov and Zhang [16], minimum quadratic distance constrained balance randomization by Titterington [17] (as demonstrated in [16]), adaptation of biased coin randomization by Frane [18], generalized method for adaptive randomization by Russel et al. [19], generalized multidimensional dynamic allocation method by Lebowitsch et al. [20].

Permuted Block Randomization (PBR) [21] – the most commonly used unequal allocation procedure – is an ARP procedure, and so are the Drop-the-Loser Urn Design by Ivanova [22] applied to fixed unequal allocation and the Urn Block design by Zhao and Weng [23]. However, when the unequal allocation ratio gives rise to a large block size, the PBR has a wide allocation space – and it is even wider for the procedures by

Ivanova [22] and Zhao and Weng [23]. As a result, in small cohorts of subjects the observed allocation ratio can considerably deviate from the target allocation ratio. Nevertheless, ARP procedures with the allocation space narrower than that of the PBR designed for arbitrary allocation ratio were lacking until the recent introduction of the Brick Tunnel Randomization (BTR) by Kuznetsova and Tymofyeyev [6-7].

BTR is an ARP randomization procedure that requires the allocation path to stay within the tunnel built of K-dimensional unitary cubes pierced by the allocation ray, which guarantees a certain closeness of the observed allocation ratio to the targeted allocation ratio. BTR is defined uniquely for randomization to two treatment arms but not necessarily uniquely for $K>2$ arms. For 1:m two-arm allocation, BTR coincides with the PBR with the block size $1+m$.

Kuznetsova and Tymofyeyev offered an iterative way to derive the resident and transition probabilities for BTR [6-7]. They did not prove that BTR can always be constructed in this way (that is the derivations can always be successfully executed to the end of the block), but that was the case in the numerous examples of the BTR they considered.

Narrow allocation space of the BTR guarantees that the observed allocation ratio is close to the target one throughout the enrollment; it also reduces the accidental bias associated with the time trend. This is helpful in adaptive design dose-ranging studies [5, 24, 25], studies with a time trend in response or baseline characteristics, multi-center studies [7, 25]. However, in a single-center open-label trial, where the investigator knows the sequence of previous treatment assignments, it makes the allocation procedure more predictable and thus, prone to selection bias. Predictability is most pronounced in a two-arm BTR, where all generations within a block have two nodes; the allocation from one of the two nodes is deterministic, while both treatments can be assigned from the other node. In this respect, the two-arm BTR is similar to the 1: m PBR. To reduce the predictability of the BTR in open-label trials, Kuznetsova and Tymofyeyev [26, 27] widened its allocation space to a strip of desired width narrower than the PBR allocation space while keeping the ARP property. The resulting allocation procedure, Wide Brick Tunnel, is less predictable and can be a better choice for open-label single-center trials. Kuznetsova and Tymofyeyev offered other ways to expand the allocation space of the BTR, not necessarily to a strip, while keeping the ARP property in [25].

In this paper, we will introduce the concept of the center of the probability mass of the i-th generation in the allocation space of a procedure. All ARP procedures with the same allocation ratio have the same sequence of the centers of the probability mass. We will use momentum of the probability mass with respect to its center (expected distance) to measure how close the probability mass is to the target allocation ratio. This measure will be used to compare unequal allocation ARP procedures in how well they approximate the target allocation ratio.

We will show how to build an ARP procedure with nodes closest to the center of mass for $K=2$ (Section 3) and K=3 (Section 4). For K=2, such procedure coincides with the BTR. We will show that the two-arm BTR is the minimum momentum two-arm ARP procedure. For K=3 such procedure is a version of the 3-arm BTR. The resident probabilities for the 3-arm BTR can thus be derived analytically for all generations at once, without the need for iterative calculations. It can be proven that respective transition probabilities always exist [28]. We will compare the momenta of the 2- and 3-arm BTR, PBR, and CR in Section 5. Discussion completes the paper.

2. Concepts and Notation

Let us introduce some notation that will be used throughout the paper.

Consider a study with $C_1:C_2: ...:C_K$ allocation to K treatment groups $G_1, G_2, \ldots, G_K$, where $C_1, \ldots, C_K$ are integers that have no common divisor. Let us use $S=C_1+C_2+...+C_K$ as the block size. Let us denote by $w_l, l=1, \ldots, K$, the allocation probabilities $w_l = C_l /S$; thus, $w_1+ w_2+ \ldots+ w_k =1$. We will denote the vector of allocation probabilities by $W = (w_1,...,w_k)'$.

For simplicity, we will also denote the treatments as Treatments A and B in 2-arm examples and as Treatments A, B, and C in 3-arm examples.

We will visualize an allocation sequence as a path along the integer grid in the K-dimensional space as described in [29]. Axis l, $l=1, \ldots, K$, represents allocation to treatment G_l. The allocation path starts at the origin and with each allocation moves one unit along the axis that corresponds to the assigned treatment. After i allocations, the allocation path ends up at the node with coordinates $(N_{1i}, N_{2i}, \ldots, N_{Ki})$, where N_{li} is the number of G_l allocations within the first i allocations.

The set of nodes that can be realized with a given allocation procedure forms its allocation space. For some allocation procedures, such as CR or biased coin randomization [30], the allocation space is equal to the whole non-negative sector of the K-dimensional space; for other allocation procedures, such as PBR, it is a subset of the non-negative sector. Specifically, for PBR with the permuted block size mS, the allocation space is a chain of K-dimensional parallelepipeds with the lowest corners at $(jmC_1, ..,jmC_K), j \geqslant 0$, and dimensions $mC_1 \times mC_2 \times ... \times mC_K$.

We will call the nodes that can be realized with the allocation procedure after i allocations the nodes of generation i. We will number them from 1 to m_i, where m_i is the number of nodes in generation i, and denote the j-th node in generation i and its coordinates as $X_{ij}=(x_{ij1}, x_{ij2}, \ldots, x_{ijK})$. All nodes of generation i belong to the $(K-1)$-dimensional plane $\pi_i = \{ (x_1, x_2, \ldots, x_K): x_1+ x_2+ \ldots+ x_K =i\}$. The origin $(0, 0, \ldots, 0)$ could be considered 0 generation – the start of every allocation sequence.

We will call the ray $AR=(C_1u, C_2u, \ldots, C_Ku)$, $u \geqslant 0$, the Allocation Ray (AR). We will call the intersection of the plane π_i and the AR the point of perfect balance of the i-th generation $B_i=(w_1i, w_2i, \ldots, w_Ki)$. The coordinates of the point of perfect balance are proportional to the allocation ratio. The point of perfect balance belongs to the unitary grid (and thus can represent observed treatment group totals) only when $i=mS$, $m \geqslant 1$.

We will call the probability for an allocation sequence to reside in the node X_{ij} after i allocations the resident probability of the node X_{ij} and denote it as $R(X_{ij})$ (or as R_{ij} for shortness). Let us denote by $R_i = (R_{i1},...,R_{im_i})'$ the m_i-dimensional vector of resident probabilities of the nodes in the i-th generation. The sum of the resident probabilities across the nodes of the same generation is 1: $\sum_{j=1}^{m_i} R_{ij} = 1$.

The probability of treatment G_l allocation from the i-th generation node $X_{ij}=(x_{ij1}, x_{ij2}, \ldots, x_{ijK})$ will be called the transition probability from X_{ij} to $(x_{ij1},\ldots, x_{ij(l-1)}, x_{ijl}+1, x_{ij(l+1)}, \ldots, x_{ijK})$ (or the transition probability from X_{ij} in the direction l) and denoted by p_{ilj}. If the restricted allocation procedure does not allow allocation to G_l from X_{ij}, $p_{ilj}=0$. If only one of the transition probabilities from the node X_{ij} is positive, the allocation from the node X_{ij} is deterministic. For example, the last allocation in a block of the PBR sequence is deterministic.

Let us denote by P_i the $K\times m_i$ -dimensional matrix of transition probabilities from generation i to generation $(i+1)$. The m_i columns of the matrix represent the m_i nodes $X_{i1},\ldots,X_{im_i}$ in the i-th generation. The K rows represent the transition along each of the K treatment axes with the $(i+1)$-th allocation. The element $0 \leqslant p_{ilj} \leqslant 1$, $l = 1,\ldots, K$, $j = 1,\ldots, m_i$, of the matrix P_i represents the transition probability from the node X_{ij} within the generation i in the direction l. For every node X_{ij} in the i-th generation the sum of the transition probabilities from X_{ij} is equal to 1:

$$\sum_{l=1}^{K} p_{ilj} = 1, \quad \text{for all } i \geqslant 0, j = 1,\ldots, m_i \qquad (1)$$

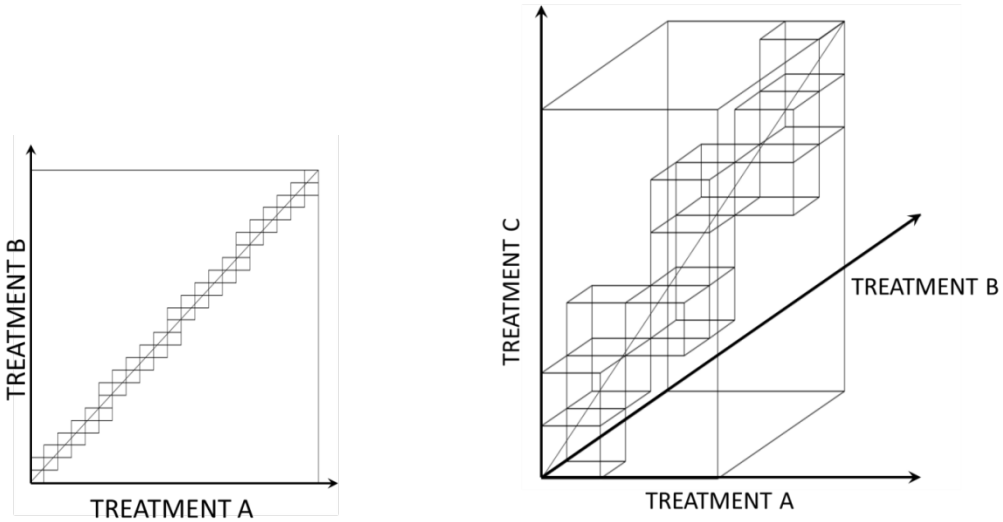
The ARP property can be expressed in terms of resident and transition probabilities in the following way [6-7]. The K-dimensional vector W_{i+1} of the unconditional probabilities of allocation to Treatments $l = 1, \ldots, K$, at the $(i+1)$-th allocation is equal to the product of the matrix of the transition probabilities from the i-th generation P_i and the vector of the resident probabilities in the i-th generation R: $W_{i+1} = P_{i\,[K,m_i]} \times R_i'$. Thus, the requirement of preserving the unconditional allocation ratio at every allocation step $W_i = W$ is

$$P_{i\,[K,m_i]} \times R_i = W \text{ , } i \geqslant 0. \qquad (2)$$

The matrix equation (2) can be written as a system of K linear equations, each corresponding to a respective row of the matrix P_i. Due to the constraint (1), the last equation in this system is redundant – it holds if all the previous equations hold.

Kuznetsova and Tymofyeyev [6-7] introduced K-dimensional BTR as an ARP randomization procedure that requires the allocation path to be constrained within the chain of the K-dimensional unitary cubes that are pierced by the allocation ray. This closeness of the allocation path to the allocation ray guarantees a certain closeness of the observed allocation ratio to the target allocation ratio even in a short cohort of patients.

Figure 1 illustrates the reduction of the allocation space with BTR compared to PBR with the examples of the 2-arm BTR with 21: 25 allocation ratio and the 3-arm BTR with 3:5:7 allocation ratio. The reduction of the PBR allocation space is more pronounced for large block sizes, like 46 for 21:25 allocation compared to small block sizes, like 5 for 2:3 allocation.



a) b)

Figure 1: The allocation space for two examples of the BTR. a) 2-arm BTR with 21:25 allocation ratio pictured within 21 × 25 block; b) 3-arm BTR with 3:5:7 allocation ratio, pictured within the 3 × 5 × 7 block.

Kuznetsova and Tymofyeyev [7] derived the transition and resident probabilities for BTR by solving (2) iteratively for i=0, 1, 2, …, S. They showed that for K>2 the tunnel can be further reduced for some generations by elimination of some of the nodes. They did not prove that the iterations can always be carried out to the end of the block, although this was the case in practice.

In this paper we explicitly derive the resident probabilities for 2- and 3- arm BTR using the concept of the center of the probability mass of the i-th generation: $Cent_i = (cent_{i1}, \ldots, cent_{iK})$, where $Cent_{il} = \sum_{j=1}^{m_i} R_{ij} x_{ijl}$.

It is easy to see that the allocation procedure is an ARP procedure if $Cent_i$ belongs to the allocation ray for all $i \geqslant 1$; specifically, $Cent_i = B_i$. Thus, all ARP procedures with the same allocation ratio have the same sequence of the centers of the probability mass $Cent_i$. These procedures can be compared in how tightly they distribute the probability mass around $Cent_i$.

For an allocation sequence that after i allocations ends up at the node $X_{ij}=(x_{ij1}, x_{ij2}, \ldots, x_{ijk})$, the imbalance in treatment assignments after i allocations describes how close the node X_{ij} is to the point of perfect balance of the i-th generation. Following Sverdlov and Zhang [16] we will use the Euclidean distance between X_{ij} and B_i as the measure of imbalance in treatment assignments after i allocations:

$$Imb(X_{ij}) = \sqrt{\sum_{l=1}^{K} \left(x_{ijl} - w_l i\right)^2} \qquad (3)$$

The absolute imbalance is 0 and the observed allocation ratio is equal to the target one if the node X_{ij} lies exactly on the allocation ray. For example, for PBR with the block size S, the imbalance is 0 at the end of each block, that is for $i=mS$, $m \geqslant 1$. For two-arm allocation the absolute imbalance in treatment assignments is commonly defined as

$I_2=|N_{2i} - N_{1i} \times C_2/C_1|$ (or proportional to this difference) [12, 13] – a quantity proportional to (3). For two-arm equal allocation ($C_2=C_1=1$) the absolute imbalance reduces to $|N_{2i} - N_{1i}|$.

Momentum of the probability mass $Mom_i = \sum_{j=1}^{m_i} R_j Imb(X_j)$ (the expectation of the imbalance in treatment assignments after i allocations) is a convenient measure of the closeness of the probability mass in the i-th generation to its center of mass. We will use this measure to compare ARP procedures in the balance they provide. We will call an unequal allocation procedure well balanced if it results in a low imbalance, that is approximates well the target allocation ratio.

3. How to Build an ARP Procedure With Nodes Closest to the Center of Mass for K=2

Figure 2 presents the centers of the probability mass for a 2-arm ARP procedure. To build the ARP procedure with the nodes closest to the centers of mass, take two nodes (D and E) closest to $Cent_i$ in generation i. These two nodes will be the corners of the unitary square that contains the center of mass.

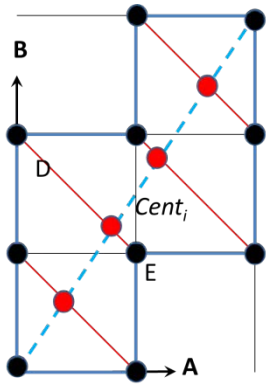


Figure 1: 2-Arm ARP Procedure with Centers of the Probability Mass

There is a unique way to place resident probabilities R_{i1} and R_{i2} in E and D so that $Cent_i$ is the center of the mass. Specifically, the resident probabilities should be inversely proportional to the distances from the nodes to the center of the mass. Accounting for $R_{i1} + R_{i2} =1$, it follows that

$$R_{i1} = w_1 i - [w_1 i] \qquad (5)$$
$$R_{i2} = w_2 i - [w_2 i]$$

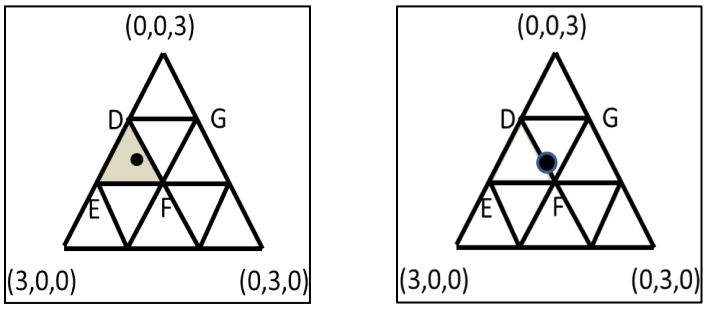
The transition probabilities can be derived and it can be easily proven that they always exist [28].

The described procedure is nothing else but the 2-arm Brick Tunnel randomization. It follows from the above that the 2-arm BTR is the minimum momentum ARP procedure, that is, for every generation the BTR has the momentum lower than any other ARP procedure with the same allocation ratio.

4. How to Build an ARP Procedure With Nodes Closest to the Center of Mass for K=3

The nodes of the 3-dimensional grid (x_1, x_2, x_3), $x_l \geqslant 0$ such that $x_1 + x_2 + x_3 = i$ form a triangular grid on the plane π_i. Figure 3 depicts the grid for i=3. The perimeter of the triangular grid is an equilateral triangle with (i+1) nodes on each side. If i is the last generation in a block ($i=mS$), the center of the BTR probability mass of the generation i falls on the node of the grid, since its coordinates $w_1 i$, $w_2 i$, or $w_3 i$ are all integers. Such node has the resident probability of 1.

For other generations, the center of the BTR probability mass of the generation i $Cent_i$ can either fall within one of the triangles on the grid (when none of the coordinates of $Cent_i$: $w_1 i$, $w_2 i$, or $w_3 i$ is an integer; Figure 3a) or belong to a segment connecting a pair of nodes (when one of the coordinates of $Cent_i$ is an integer, Figure 3b).



a) b)

Figure 3: The triangular grid formed by the nodes on the plane π_3 with the coordinates of the three vertices. The center of mass of the 3rd generation $Cent_3$ (black dot) belongs to the inner part of the triangle DEF on the grid.

If $Cent_i$ belongs to the inner part of triangle DEF (as in Figure 3a), the resident probabilities in the three nodes D, E, and F, that result in $Cent_i$ being the center of the probability mass are determined uniquely. Similarly, if $Cent_i$ belongs to segment DF, the resident probabilities in nodes D and F that result in $Cent_i$ being the center of the mass are also determined uniquely.

The resident and transition probabilities can be easily derived [28]; it is proven that the resident probabilities always exist [28]. Described derivation of the resident probabilities and corresponding transition probabilities is simple and can be easily programmed.

This procedure is a version of the 3-arm BTR where the nodes other the DEF (when the center of the mass belongs to an inner part of a triangular) or DF (when the center of the mass belongs to a segment connecting two nodes) are removed from the brick tunnel.

Although the 3-arm BTR with resident probabilities defined by Lemma 4 is not necessarily the minimum momentum ARP procedure, it is the ARP procedures with the nodes closest to the center of mass in each generation.

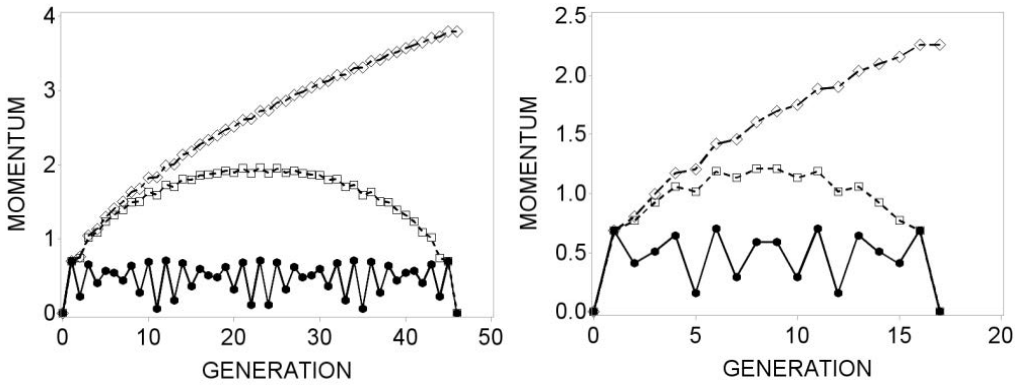
It should be noted that while $C_1, C_2, ..., C_K$ were defined above as integers following the PBR conventions, for BTR $C_1, C_2, ..., C_K$ can be any positive numbers. If S is not a

rational number, there will be no generation where BTR has a node on the allocation ray. For K=2 and K=3, the resident and transition probabilities are derived by the same formulae, regardless of the type of C_1, C_2, ...,C_K.

5. Comparison of the Momenta of the BTR, PBR, and CR

Formulae for the momenta of the 2- and 3-arm BTR, PBR, and CR are derived in [28]. For 2-arm BTR, $Mom_i \leqslant \sqrt{2}/2$ regardless of the block size or the allocation ratio. The maximum momentum of $\sqrt{2}/2$ is reached in the middle of the block when both C_1 and C_2 are odd. The average momentum across all generations in a block depends on the block size, but not on the allocation ratio. When the bloc size increases, the average momentum converges to $\sqrt{2}/3$.

Figure 4 provides the comparison of the momenta of the probability mass by generation for BTR, PBR and CR for 21:25 allocation ratio (a) and 7:10 allocation ratio (b). As Figure 4a shows, for small cohorts of 10-12 subjects the momentum of the PBR is almost the same as the momentum of the CR. This illustrates the point that for large block sizes PBR offers little advantage over CR in approximating the targeted allocation ratio in small cohorts. The highest momentum with the 21:25 PBR (reached in the middle of the block of 46 allocations) is 1.95883, while for BTR it is 0.70711. For smaller block size of 17 (7:10 allocation, Figure 4b), the highest momentum with the PBR is 1.20728, while for BTR it is 0.70466.



a) b)

Figure 4: Comparison of the momenta of the probability mass by generation for BTR (bottom line), PBR (middle line) and CR (top line) for a) 21:25 allocation ratio; b) 7:10 allocation ratio.

Since 2-arm BTR has the lowest momentum among all ARP procedures, it can be used to benchmark the momenta of other allocation procedures. Figure 5 illustrates this point by presenting the momenta of 21:25 PBR and three types of ARP Biased Coin Randomization plotted against the BTR momenta.

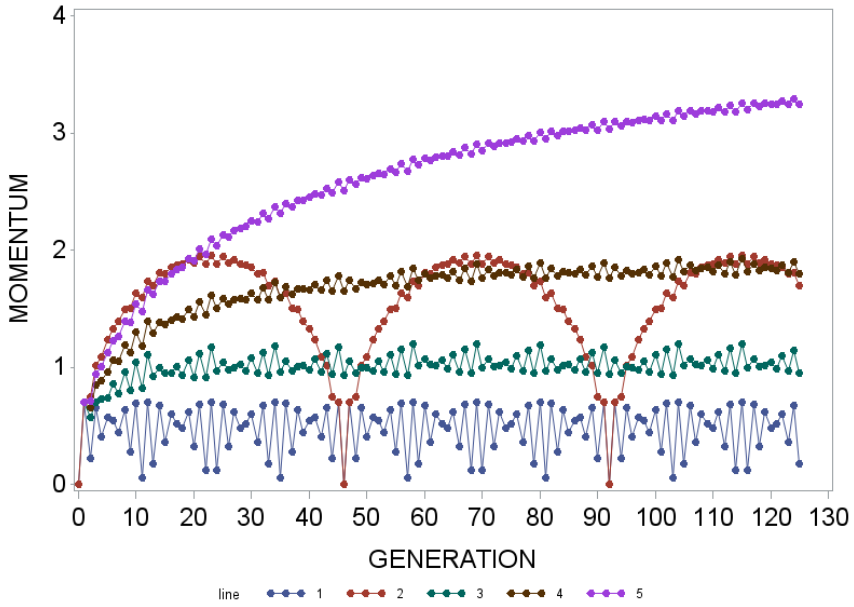


Figure 5: Momenta of the 21:25 BTR (line 1), PBR (line 2), and Three Types of ARP Biased Coin Randomization (lines 3-5).

For 3-arm BTR, maximum momentum of $\sqrt{2/3}$ is reached when $Cent_i$ is placed exactly at the center of the triangle formed by the nodes of the i-th generation. Thus, $Mom_i \leq \sqrt{2/3} = 0.817$ for any allocation ratio.

Figure 6 presents the comparison of the momenta of the BTR, PBR, and CR with 3:5:7 allocation ratio. Since there are only 3 nodes in the first generation for each procedure, all three procedures (as well as any other APR procedure) have the same momenta in the 1-st generation. (For PBR and BTR the same is true for the 14^{th} generation in the block). For other generations, the allocation space for PBR is wider than for the BTR which is reflected in the momentum of the PBR exceeding the momentum of the BTR (Figure 6). For CR, momenta are close to those of the PBR for a few generations, but continue to grow and diverge from the PBR momenta with generation.

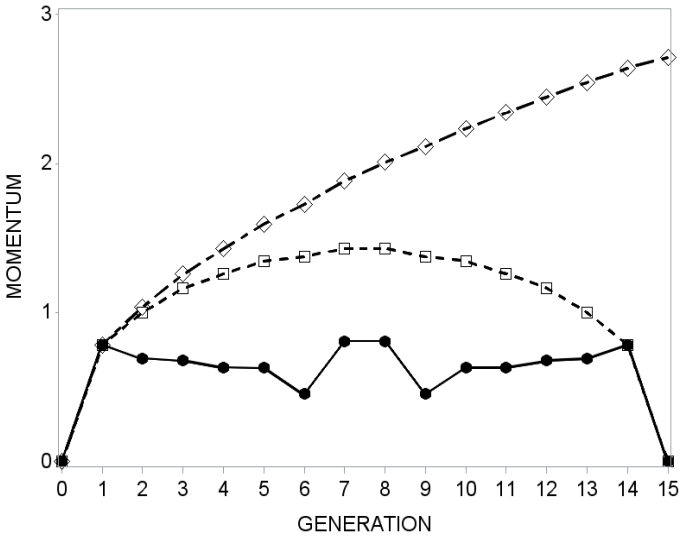


Figure 6: Momenta of the Probability Mass by Generation for 3:5:7 allocation: BTR (bottom line), PBR (middle line), CR (top line)

Similar to the 2-arm case, the advantage of the BTR over PBR and CR in approximating the targeted allocation ratio is even more pronounced for larger block sizes than the block size of 15 in the considered example.

6. Discussion

For equal allocation to two arms the allocation space can be as narrow as the space of the permuted block with the block size 2. Such allocation has one node in even generations and two nodes in odd generations and limits the absolute imbalance in treatment totals to 1. No similarly tight allocation space was available for ARP unequal allocation with arbitrary allocation ratio before the BTR was introduced by Kuznetsova and Tymofyeyev [6-7]. The permuted block design, the most common unequal allocation procedure, can have a wide allocation space. Unless $C_1, C_2, \ldots, C_K$ are small, such space can be too wide for many applications as it might result in the undesirable imbalance in treatment assignments.

Other allocation procedures developed for unequal allocation often lack an ARP property – the deficiency that was not well recognized. For example, to address the need for narrow allocation space, Salama et al. [13] offered an expansion of the maximal procedure for unequal allocation. They allowed all allocation sequences that fit within a strip of pre-specified width around the allocation ray. Unfortunately, they assigned equal probabilities to all permitted allocation sequences which lead to variations in the allocation ratio from allocation to allocation (a lack of an ARP property).

The unequal allocation procedures by Zhao and Weng [23] and Ivanova [22] are ARP procedures. They can be described as generated following the mapping principle [8-10] that ensures an ARP property. First, an equal allocation procedure to any number of arms symmetric with respect to the arms is described. Then, to generate an unequal allocation procedure to $K \geqslant 2$ treatment groups G_l, $l = 1,\ldots, K$ in $C_1{:}C_2{:} \ldots {:}C_K$ ratio, where S

$=C_1+C_2+ \ldots + C_K$, one first generates an equal allocation to S "fake" treatment arms F_1, F_2, … , F_s following the equal allocation procedure and then maps the groups of "fake" treatment arms to the actual treatment arms. Specifically, the first C_1 "fake" treatment arms $F_1 - F_{C_1}$ are mapped to treatment G_1; the next C_2 "fake" treatment arms $F_{C_1+1} - F_{C_1+C_2}$ are mapped to treatment G_2; …, ; and finally, the last C_K "fake" treatment arms $F_{C_1+\ldots+C_{k-1}+1} - F_S$ are mapped to treatment G_k. Due to symmetry, such procedure provides equal allocation to S "fake" treatment arms F_1, F_2, … , F_s and thus, C_1:C_2: … : C_K unconditional allocation ratio to actual treatment groups at every allocation. Other unequal allocation procedures can be developed following the mapping approach, but the allocation space for all such procedures will, by necessity, include the allocation space of the C_1:C_2: … :C_K permuted block allocation.

The ARP allocation that limits the set of permuted block sequences to the better balanced ones and thus reduces the PBR allocation space can be built following constrained randomization [31-35]. For small C_1, C_2, …, C_K, a reasonable constraint can often be built by a statistician, but for an arbitrary allocation ratio coming up with an acceptable constraint is a difficult task.

The BTR [6,7], an unequal allocation ARP procedure that works for any allocation ratio and has the space narrower than the permuted block with the minimal block size, provides a very tight balance in treatment assignments. It was followed by the wide brick tunnel [26, 27] – a two-arm ARP procedure that expands the BTR space to a strip around the allocation ray. The BTR allocation space for $K \geqslant 2$ can be also expanded by randomly permuting short segments of the BTR [25]. The allocation space wider than the BTR, but narrower than the permuted block space, can be useful in open-label studies (mainly two-arm ones) to reduce the potential for selection bias.

BTR closely approximates the targeted allocation ratio throughout the enrollment, in particular, for small samples. This makes BTR useful in adaptive design dose-ranging studies [5, 24, 25], unequal allocation studies with a time trend in response or baseline characteristics, multi-center studies [7, 25]. BTR application in two-arm response-adaptive randomization setting, where the allocation ratios are too inconvenient for PBR so that CR is used to allocate patients, should be further explored. There BTR can be used to randomize small cohorts of 5-6 patients with the same allocation ratio instead of allocating each patient using CR with his own allocation ratio. This approach would allow avoiding considerable deviations of the observed allocation ratio from the targeted one.

Complexity of the BTR algorithm for $K>2$, where the transition probabilities are derived iteratively for generation after generation using optimization techniques is the main obstacle to its implementation. This paper uses the concept of the center of probability mass in the i-th generation to explicitly derive, for $K=2$ and $K=3$, the resident and transition probabilities for all generations at once through simple formulae. It also proves that the required set of transition probabilities always exists for $K=2$ and $K=3$. It demonstrates that the two-arm and 3-arm BTR are the ARP allocation procedures with the tightest allocation space among all allocation procedures with the same allocation ratio. It further shows that the 2-arm BTR is the minimum momentum 2-arm ARP allocation procedure. Although the simplification of the algorithm for $K>3$ needs to be further explored, simplifying it for $K=2$ and $K=3$ provides an important step forward as

the 2- and 3-arm studies cover a large sector of clinical trials with unequal allocation. Moreover, the same algorithm applies when a study has $K>3$ arms, but only 2 or 3 distinct allocation ratios.

Comparisons of the BTR with the PBR and the CR in the momentum with respect to the center of the probability mass demonstrate the advantage of the BTR in the balance it provides. With better understanding of the BTR properties and easy generation for two and three arms, the BTR can find its way into clinical trials.

Contact: olga_kuznetsova@merck.com

References

1. Rosenberger WF, Lachin JM. Randomization in Clinical Trials. John Wiley & Sons, New York, 2002.
2. Dumville JC, Hahn S, Miles JNV, Torgerson DJ. The use of unequal randomisation ratios in clinical trials: A review. *Contemporary Clinical Trials* 2006; **27**:1-12. DOI: 10.1016/j.cct.2005.08.003
3. McEntegart D, Dawson R. Letter to the Editor re Dumville et al. *Contemporary Clinical Trials* 2006; **27**: 207-208, DOI: 10.1016/j.cct.2006.02.003.
4. Dragalin V. Adaptive designs: terminology and classification. *Drug Information Journal*. 2006; **40**: 425-435.
5. Kuznetsova OM. Randomization challenges in adaptive design studies. In: He W, Pinheiro J, and Kuznetsova OM (eds.). Practical Considerations for Adaptive Trial Design and Implementation: pp. 157–181. Springer, New York (2014).
6. Kuznetsova OM, Tymofyeyev Y. Brick tunnel randomization – a way to accommodate a problematic allocation ratio in adaptive design dose finding studies. ASA Proceedings of the Joint Statistical Meetings, 1356-1367. American Statistical Association (Alexandria, VA) (2009).
7. Kuznetsova O, Tymofyeyev Y. Brick Tunnel Randomization for unequal allocation to two or more treatment groups. *Statistics in Medicine* 2011a; **30:** 812-824 (2011a). DOI: 10.1002/sim.4167.
8. Kuznetsova OM, Tymofyeyev Y. Expansion of the modified Zelen's approach randomization and dynamic randomization with partial block supplies at the centers to unequal allocation. ASA Proceedings of the Joint Statistical Meetings, pp. 3494-3507. American Statistical Association, Miami Beach (2011b).
9. Kuznetsova OM, Tymofyeyev Y. Expansion of the modified Zelen's approach randomization and dynamic randomization with partial block supplies at the centers to unequal allocation. *Contemporary Clinical Trials* 2011c; **32**: 962–972. DOI:10.1016/j.cct.2011.08.006.
10. Kuznetsova OM, Tymofyeyev Y. Preserving the allocation ratio at every allocation with biased coin randomization and minimization in studies with unequal allocation. *Statistics in Medicine* 2012; **31**: 701-723, DOI: 10.1002/sim.4447.
11. Proschan M, Brittain E, and Kammerman L. Minimize the use of minimization with unequal allocation. *Biometrics* 2011; **67** (3): 1135-1141, DOI: 10.1111/j.1541-0420.2010.01545.x.
12. Kaiser LD. Dynamic randomization and a randomization model for clinical trials data. *Statistics in Medicine*. 2012; **31**:3858-3873. DOI: 10.1002/sim.5448.
13. Salama I, Ivanova A, Qaqish B. Efficient generation of constrained block allocation sequences. *Statistics in Medicine* 2008; **27**: 1421-1428, DOI:10.1002/sim3014.

14. Han B, Enas NH, McEntegart D. Randomization by minimization for unbalanced treatment allocation. *Statistics in Medicine* 2009; **28:** 3329-3346, DOI:10.1002/sim.3710
15. Hu F, Zhang LX. Asymptotic properties of doubly adaptive biased coin designs for multitreatment clinical trials. *The Annals of Statistics* 2004; **32(1)**, 268-301.
16. Sverdlov O, Zhang Y. (2014) Statistical properties of covariate-adaptive randomization procedures for trials with unequal treatment allocation ratios. ASA Proceedings of the Joint Statistical Meetings, American Statistical Association, Boston, MA, 2014.
17. Titterington DM. On constrained balance randomization for clinical trials. *Biometrics* 1983; **39(4):** 1083-1086 ()
18. Frane JW. A method of biased coin randomisation, its implementation and its validation. *Drug Information Journal* 1998; **32**: 423-432.
19. Russell D, Hoare ZSJ, Whitaker Rh, Whitaker CJ, Russell IT. Generalized method for adaptive randomization in clinical trials. *Statistics in Medicine* 2011; **30**: 922-934.
20. Lebowitsch J, Ge Y, Young B, Hu F. Generalized multidimensional allocation method, *Statistics in Medicine* 2012; **31**: 3537-3544.
21. Zelen M. The randomization and stratification of patients to clinical trials. *Journal of Chronic Disease* 1974; **27:** 365-375.
22. Ivanova A. A play-the-winner-type urn design with reduced variability. *Metrika* 2003; **58**: 1-13.
23. Zhao W, Weng Y. Block urn design — A new randomization algorithm for sequential trials with two or more treatments and balanced or unbalanced allocation. *Contemporary Clinical Trials*, 2011, **32**, 6, 953-961.
24. He W, Kuznetsova OM, Harmer MA, Leahy CJ, Anderson KM, Dossin DN, Li L, Bolognese JA, Tymofyeyev Y, Schindler JS. Practical considerations and strategies for executing adaptive clinical trials. *Drug Information Journal* 2012; **46**: 160-174, DOI: 10.1177/0092861512436580.
25. Kuznetsova OM, Tymofyeyev Y. Brick Tunnel and Wide Brick Tunnel Randomization for studies with unequal allocation. Sverdlov A. (ed.). Modern Adaptive Randomized Clinical Trials: Statistical, Operational, and Regulatory Aspects: pp.83-114. CRC Press, Taylor & Francis Group (2015)
26. Kuznetsova OM, Tymofyeyev Y. Expanding Brick Tunnel Randomization to allow for larger imbalance in treatment totals in studies with unequal allocation. Proceedings of the Joint Statistical Association 2013 Meetings, Montreal, Canada, 8/4/13 - 8/8/13.
27. Kuznetsova OM, Tymofyeyev Y. Wide Brick Tunnel Randomization – an unequal allocation procedure that limits the imbalance in treatment totals. *Statistics in Medicine* 2014; 33: 1514–1530, DOI: 10.1002/sim.6051.
28. Kuznetsova OM. Brick Tunnel Randomization and the Momentum of the Probability Mass. Statistics in Medicine 2015 DOI: 10.1002/sim.6601
29. Berger VW, Ivanova A, Knoll M. Minimizing predictability while retaining balance through the use of less restrictive randomization procedures. *Statistics in Medicine* 2003; **22**, 3017-3028. DOI:10.1002/sim.1538
30. Efron B. Forcing a sequential experiment to be balanced. Biometrika 1971; **58**: 403–417.
31. Youden WJ: Inadmissible random assignments. *Technometrics* 1964; **6:** 103-104.
32. Youden, W.J.: Randomization and experimentation. *Technometrics* 1972; **14**: 13-22.

33. Song C, Kuznetsova OM. Implementing constrained or balanced across the centers randomization with SAS v8 Procedure PLAN. PharmaSUG 2003 proceedings, pp. 473-479, Miami FL, May 4-7, 2003.
34. Kuznetsova O, Ivanova A. Allocation in Randomized Clinical Trials. In: Dmitrienko, A., Chuang-Stein, C., D'Agostino, R. (eds.) Pharmaceutical Statistics Using SAS, pp.213-236. SAS Press: Cary, NC (2006)
35. Kuznetsova OM. Randomization Schedule. In: D'Agostino, R., Sullivan, L., Massaro, J. (eds.). Wiley Encyclopedia of Clinical Trials. Hoboken: John Wiley &Sons., Inc; Published Online: 19 SEP 2008 DOI: 10.1002/9780471462422.eoct314.