Chapter 6
The Randomization Process

The randomized controlled clinical trial is the standard by which all trials are judged. In the simplest case, randomization is a process by which each participant has the same chance of being assigned to either intervention or control. An example would be the toss of a coin, in which heads indicates intervention group and tails indicates control group. Even in the more complex randomization strategies, the element of chance underlies the allocation process. Of course, neither trial participant nor investigator should know what the assignment will be before the participant's decision to enter the study. Otherwise, the benefits of randomization can be lost. The role that randomization plays in clinical trials has been discussed in Chap. 5 as well as by numerous authors [1–12]. While not all accept that randomization is essential [10, 11], most agree it is the best method for achieving comparability between study groups, and the most appropriate basis for statistical inference [1, 3].

Fundamental Point

Randomization tends to produce study groups comparable with respect to known as well as unknown risk factors, removes investigator bias in the allocation of participants, and guarantees that statistical tests will have valid false positive error rates.

Several methods for randomly allocating participants are used [6, 9, 12–14]. This chapter will present the most common of these methods and consider the advantages and disadvantages of each. Unless stated otherwise, it can be assumed that the randomization strategy will allocate participants into two groups, an intervention group and a control group. However, many of the methods described here can easily be generalized for use with more than two groups.

Two forms of experimental bias are of concern. The first, *selection bias*, occurs if the allocation process is predictable [5, 15–18]. In this case, the decision to enter a

participant into a trial may be influenced by the anticipated treatment assignment. If any bias exists as to what treatment particular types of participants should receive, then a selection bias might occur. All of the randomization procedures described avoid selection bias by not being predictable. A second bias, *accidental bias*, can arise if the randomization procedure does not achieve balance on risk factors or prognostic covariates. Some of the allocation procedures described are more vulnerable to accidental bias, especially for small studies. For large studies, however, the chance of accidental bias is negligible [5].

Whatever randomization process is used, the report of the trial should contain a brief, but clear description of that method. In the 1980s, Altman and Doré [15] reported a survey of four medical journals where 30% of published randomized trials gave no evidence that randomization had in fact been used. As many as 10% of these "randomized" trials in fact used non-random allocation procedures. Sixty percent did not report the type of randomization that was used. In one review in the 1990s, only 20–30% of trials provided fair or adequate descriptions, depending on the size of the trial or whether the trial was single center or multicenter [18]. More recently, a review of 253 trials published in five major medical journals after the release of the Consolidated Standards for Reporting Trials (CONSORT) [19] recommendations found little improvement in reports of how randomization was accomplished [20]. Descriptions need not be lengthy to inform the reader, publications should clearly indicate the type of randomization method and how the randomization was implemented.

Fixed Allocation Randomization

Fixed allocation procedures assign the interventions to participants with a prespecified probability, usually equal, and that allocation probability is not altered as the study progresses. A number of methods exist by which fixed allocation is achieved [6, 9, 12, 14, 21–25], and we will review three of these—simple, blocked, and stratified.

Our view is that allocation to intervention and control groups should be equal unless there are compelling reasons to do otherwise. Peto [7] among others, has suggested an unequal allocation ratio, such as 2:1, of intervention to control. The rationale for such an allocation is that the study may slightly lose sensitivity but may gain more information about participant responses to the new intervention, such as toxicity and side effects. In some instances, less information may be needed about the control group and, therefore, fewer control participants are required. If the intervention turns out to be beneficial, more study participants would benefit than under an equal allocation scheme. However, new interventions may also turn out to be harmful, in which case more participants would receive them under the unequal allocation strategy. Although the loss of sensitivity or power may be less than 5% for allocation ratios approximately between 1/2 and 2/3 [8, 21], equal allocation is the most powerful design and therefore generally recommended. We also believe that equal allocation is more consistent with the view of indifference or equipoise

toward which of the two groups a participant is assigned (see Chap. 2). Unequal allocation may indicate to the participants and to their personal physicians that one intervention is preferred over the other. In a few circumstances, the cost of one treatment may be extreme so that an unequal allocation of 2:1 or 3:1 may help to contain costs while not causing a serious loss of power. Thus, there are tradeoffs that must be considered. In general, equal allocation will be presumed throughout the following discussion unless otherwise indicated.

Simple Randomization

The most elementary form of randomization, referred to as simple or complete randomization, is best illustrated by a few examples [9, 12]. One simple method is to toss an unbiased coin each time a participant is eligible to be randomized. For example, if the coin turns up heads, the participant is assigned to group A; if tails, to group B. Using this procedure, approximately one half of the participants will be in group A and one half in group B. In practice, for small studies, instead of tossing a coin to generate a randomization schedule, a random digit table on which the equally likely digits 0 to 9 are arranged by rows and columns is usually used to accomplish simple randomization. By randomly selecting a certain row (column) and observing the sequence of digits in that row (column) A could be assigned, for example, to those participants for whom the next digit was even and B to those for whom the next digit was odd. This process produces a sequence of assignments which is random in order, and each participant has an equal chance of being assigned to A or B.

For large studies, a more convenient method for producing a randomization schedule is to use a random number producing algorithm, available on most computer systems. A simple randomization procedure might assign participants to group A with probability p and participants to group B with probability $1 - p$. One computerized process for simple randomization is to use a uniform random number algorithm to produce random numbers in the interval from 0.0 to 1.0. Using a uniform random number generator, a random number can be produced for each participant. If the random number is between 0 and p, the participant would be assigned to group A; otherwise to group B. For equal allocation, the probability cut point, p, is one-half (i.e., $p = 0.50$). If equal allocation between A and B is not desired ($p \neq 1/2$), then p can be set to the desired proportion in the algorithm and the study will have, on the average, a proportion p of the participants in group A.

This procedure can be adapted easily to more than two groups. Suppose, for example, the trial has three groups, A, B and C, and participants are to be randomized such that a participant has a 1/4 chance of being in group A, a 1/4 chance of being in group B, and a 1/2 chance of being in group C. By dividing the interval 0 to 1 into three pieces of length 1/4, 1/4, and 1/2, random numbers generated will have probabilities of 1/4, 1/4 and 1/2, respectively, of falling into each subinterval. Specifically, the intervals would be <0.25, 0.25–0.50, and ≥0.50. Then any

participant whose random number is less than 0.25 is assigned A, any participant whose random number falls between 0.25 and 0.50 is assigned B and the others, C. For equal allocation, the interval would be divided into thirds and assignments made accordingly.

The advantage of this simple randomization procedure is that it is easy to implement. The major disadvantage is that, although in the long run the number of participants in each group will be in the proportion anticipated, at any point in the randomization, including the end, there could be a substantial imbalance [23]. This is true particularly if the sample size is small. For example, if 20 participants are randomized with equal probability to two treatment groups, the chance of a 12:8 split (i.e., 60% A, 40% B) or worse is approximately 50%. For 100 participants, the chance of the same ratio (60:40 split) or worse is only 5%. While such imbalances do not cause the statistical tests to be invalid, they do reduce ability to detect true differences between the two groups. In addition, such imbalances appear awkward and may lead to some loss of credibility for the trial, especially for the person not oriented to statistics. For this reason primarily, simple randomization is not often used, even for large studies. In addition, interim analysis of accumulating data might be difficult to interpret with major imbalances in number of participants per arm, especially for smaller trials.

Some investigators incorrectly believe that an alternating assignment of participants to the intervention and the control groups (e.g., $ABABAB\ldots$) is a form of randomization. However, no random component exists in this type of allocation except perhaps for the first participant. A major criticism of this method is that, in a single-blind or unblinded study, the investigators know the next assignment, which could lead to a bias in the selection of participants. Even in a double-blind study, if the blind is broken on one participant as sometimes happens, the entire sequence of assignments is known. Therefore, this type of allocation method should be avoided.

Blocked Randomization

Blocked randomization, sometimes called permuted block randomization, was described by Hill [4] in 1951. It avoids serious imbalance in the number of participants assigned to each group, an imbalance which could occur in the simple randomization procedure. More importantly, blocked randomization guarantees that at no time during randomization will the imbalance be large and that at certain points the number of participants in each group will be equal [9, 12, 26]. This protects against temporal trends during enrollment, which is often a concern for larger trials with long enrollment phases.

If participants are randomly assigned with equal probability to groups A or B, then for each block of even size (for example, 4, 6 or 8) one half of the participants will be assigned to A and the other half to B. The order in which the interventions are assigned in each block is randomized, and this process is repeated for

consecutive blocks of participants until all participants are randomized. For example, the investigators may want to ensure that after every fourth randomized participant, the number of participants in each intervention group is equal. Then a block of size 4 would be used and the process would randomize the order in which two A's and two B's are assigned for every consecutive group of four participants entering the trial. One may write down all the ways of arranging the groups and then randomize the order in which these combinations are selected. In the case of block size 4, there are six possible combinations of group assignments: $AABB$, $ABAB$, $BAAB$, $BABA$, $BBAA$, and $ABBA$. One of these arrangements is selected at random and the four participants are assigned accordingly. This process is repeated as many times as needed.

Another method of blocked randomization may also be used. In this method for randomizing the order of assignments within a block of size b, a random number between 0 and 1 for each of the b assignments (half of which are A and the other half B) is obtained. The example below illustrates the procedure for a block of size four (2As and 2Bs). Four random numbers are drawn between 0 and 1 in the order shown.

Assignment	Random number	Rank
A	0.069	1
A	0.734	3
B	0.867	4
B	0.312	2

The assignments then are ranked according to the size of the random numbers. This leads to the assignment order of $ABAB$. This process is repeated for another set of four participants until all have been randomized.

The advantage of blocking is that balance between the number of participants in each group is guaranteed during the course of randomization. The number in each group will never differ by more than $b/2$ when b is the length of the block. This can be important for at least two reasons. First, if the type of participant recruited for the study changes during the entry period, blocking will produce more comparable groups. For example, an investigator may use different sources of potential participants sequentially. Participants from these sources may vary in severity of illness or other crucial respects. One source, with the more seriously ill participants, may be used early during enrollment and another source, with healthier participants, late in enrollment [3]. If the randomization were not blocked, more of the seriously ill participants might be randomized to one group. Because the later participants are not as sick, this early imbalance would not be corrected. A second advantage of blocking is that if the trial should be terminated before enrollment is completed, balance will exist in terms of number of participants randomized to each group.

A potential, but solvable problem with basic blocked randomization is that if the blocking factor b is known by the study staff and the study is not double-blind, the assignment for the last person entered in each block is known before entry of that person. For example, if the blocking factor is 4 and the first three assignments are

ABB, then the next assignment must be *A*. This could, of course, permit a bias in the selection of every fourth participant to be entered. Clearly, there is no reason to make the blocking factor known. However, in a study that is not double-blind, with a little ingenuity the staff can soon discover the blocking factor. For this reason, repeated blocks of size 2 should not be used. On a few occasions, perhaps as an intellectual challenge, investigators or their clinic staff have attempted to break the randomization scheme [27]. This curiosity is natural but nevertheless can lead to selection bias. To avoid this problem in the trial that is not double-blind, the blocking factor can be varied as the recruitment continues. In fact, after each block has been completed, the size of the next block could be determined in a random fashion from a few possibilities such as 2, 4, 6, and 8. The probabilities of selecting a block size can be set at whatever values one wishes with the constraint that their sum equals 1.0. For example, the probabilities of selecting block sizes 2, 4, 6, and 8 can be 1/6, 1/6, 1/3, and 1/3 respectively. Randomly selecting the block size makes it very difficult to determine where blocks start and stop and thus determine the next assignment.

A disadvantage of blocked randomization is that, from a strictly theoretical point of view, analysis of the data is more complicated than if simple randomization were used. Unless the data analysis performed at the end of the study reflects the randomization process actually performed [26, 28–30] it may be incorrect since standard analytical methods assume a simple randomization. In their analysis of the data most investigators ignore the fact that the randomization was blocked. Matts and Lachin [26] studied this problem and concluded that the measurement of variability used in the statistical analysis is not exactly correct if the blocking is ignored. Usually the analysis ignoring blocks is conservative, though it can be anticonservative especially when the blocks are small (e.g. a block size of two). That is, the analysis ignoring blocks will have probably slightly less power than the correct analysis, and understate the "true" significance level. Since blocking guarantees balance between the two groups and, therefore, increases the power of a study, blocked randomization with the appropriate analysis is more powerful than not blocking at all or blocking and then ignoring it in the analysis [26]. Also, the correct treatment of blocking would be difficult to extend to more complex analyses. Being able to use a single, straightforward analytic approach that handles covariates, subgroups, and other secondary analyses simplifies interpretation of the trial as a whole. Performing the most correct analysis is even more problematic for adaptive designs, as discussed in the next section.

Stratified Randomization

One of the objectives in allocating participants is to achieve between group comparability of certain characteristics known as prognostic or risk factors [12, 31–44]. These are baseline factors which correlate with subsequent participant response or outcome. Investigators may become concerned when prognostic factors are not

evenly distributed between intervention and control groups. As indicated previously, randomization tends to produce groups which are, on the average, similar in their entry characteristics, known or unknown, or unmeasured. This is a concept likely to be true for large studies or for many small studies when averaged. For any single study, especially a small study, there is no guarantee that all baseline characteristics will be similar in the two groups. In the multicenter Aspirin Myocardial Infarction Study [45] which had 4,524 participants, the top 20 cardiovascular prognostic factors for total mortality identified in the Coronary Drug Project [43] were compared in the intervention and control groups and no major differences were found (Furberg CD, unpublished data). However, individual clinics, with an average of 150 participants, showed considerable imbalance for many variables between the groups. Imbalances in prognostic factors can be dealt with either after the fact by using stratification in the analysis (Chap. 18) or can be prevented by using stratification in the randomization. Stratified randomization is a method which helps achieve comparability between the study groups for those factors considered.

Stratified randomization requires that the prognostic factors be measured either before or at the time of randomization. If a single factor is used, it is divided into two or more subgroups or strata (e.g., age 30–34 years, 35–39 years, 40–44 years). If several factors are used, a stratum is formed by selecting one subgroup from each of them. The total number of strata is the product of the number of subgroups in each factor. The stratified randomization process involves measuring the level of the selected factors for a participant, determining to which stratum she belongs and performing the randomization within that stratum.

Within each stratum, the randomization process itself could be simple randomization, but in practice most clinical trials use some blocked randomization strategy. Under a simple randomization process, imbalances in the number in each group within the stratum could easily happen and thus defeat the purpose of the stratification. Blocked randomization is, as described previously, a special kind of stratification. However, this text will restrict use of the term blocked randomization to stratifying over time, and use stratified randomization to refer to stratifying on factors other than time. Some confusion may arise here because early texts on design used the term blocking as this book uses the term stratifying. However, the definition herein is consistent with current usage in clinical trials.

As an example of stratified randomization with a block size of 4, suppose an investigator wants to stratify on age, sex and smoking history. One possible classification of the factors would be three 10-year age levels and three smoking levels.

Age (years)	Sex	Smoking history
1. 40–49	Male	Current smoker
2. 50–59	Female	Ex-smoker
3. 60–69		Never smoked

Table 6.1 Stratified randomization with block size of 4

Strata	Age	Sex	Smoking	Group assignment
1	40–49	M	Current	*ABBA BABA...*
2	40–49	M	Ex	*BABA BBAA...*
3	40–49	M	Never	etc.
4	40–49	F	Current	
5	40–49	F	Ex	
6	40–49	F	Never	
7	50–59	M	Current	
8	50–59	M	Ex	
9	50–59	M	Never	
10	50–59	F	Current	
11	50–59	F	Ex	
12	50–59	F	Never	
	(etc.)			

Thus, the design has $3 \times 2 \times 3 = 18$ strata. The randomization for this example appears in Table 6.1.

Participants who were between 40 and 49 years old, male and current smokers, that is, in stratum 1, would be assigned to groups *A* or *B* in the sequences *ABBA BABA* .... Similarly, random sequences would appear in the other strata.

Small studies are the ones most likely to require stratified randomization, because in large studies, the magnitude of the numbers increases the chance of comparability of the groups. In the example shown above, with three levels of the first factor (age), two levels of the second factor (sex), and three levels of the third factor (smoking history), 18 strata have been created. As factors are added and the levels within factors are refined, the number of strata increase rapidly. If the example with 18 strata had 100 participants to be randomized, then only five to six participants would be expected per stratum if the study population were evenly distributed among the levels. Since the population is most likely not evenly distributed over the strata, some strata would actually get fewer than five to six participants. If the number of strata were increased, the number of participants in each stratum would be even fewer. Pocock and Simon [41] showed that increased stratification in small studies can be self-defeating because of the sparseness of data within each stratum. Thus, only important variables should be chosen and the number of strata kept to a minimum.

In addition to making the two study groups appear comparable with regard to specified factors, the power of the study can be increased by taking the stratification into account in the analysis. Stratified randomization, in a sense, breaks the trial down into smaller trials. Participants in each of the "smaller trials" belong to the same stratum. This reduces variability in group comparisons if the stratification is used in the analysis. Reduction in variability allows a study of a given size to detect smaller group differences in response variables or to detect a specified difference with fewer participants [22, 26].

Sometimes the variables initially thought to be most prognostic and, therefore used in the stratified randomization, turn out to be unimportant. Other factors may be identified later which, for the particular study, are of more importance. If randomization is done without stratification, then analysis can take into account those factors of interest and will not be complicated by factors thought to be important at the time of randomization. It has been argued that there usually does not exist a need to stratify at randomization because stratification at the time of analysis will achieve nearly the same expected power [7]. This issue of stratifying pre- versus post-randomization has been widely discussed [35–38, 42]. It appears for a large study that stratification after randomization provides nearly equal efficiency to stratification before randomization [39, 40]. However, for studies of 100 participants or fewer, stratifying the randomization using two or three prognostic factors may achieve greater power, although the increase may not be large.

Stratified randomization is not the complete solution to all potential problems of baseline imbalance. Another strategy for small studies with many prognostic factors is considered below in the section on adaptive randomization.

In multicenter trials, centers vary with respect to the type of participants randomized as well as the quality and type of care given to participants during follow-up. Thus, the center may be an important factor related to participant outcome, and the randomization process should be stratified accordingly [33]. Each center then represents, in a sense, a replication of the trial, though the number of participants within a center is not adequate to answer the primary question. Nevertheless, results at individual centers can be compared to see if trends are consistent with overall results. Another reason for stratification by center is that if a center should have to leave the study, the balance in prognostic factors in other centers would not be affected.

One further point might need consideration. If in the stratified randomization, a specific proportion or quota is intended for each stratum, the recruitment of eligible participants might not occur at the same rate. That is, one stratum might meet the target before the others. If a target proportion is intended, then plans need to be in place to close down recruitment for that stratum, allowing the others to be completed.

Adaptive Randomization Procedures

The randomization procedures described in the sections on fixed allocation above are non-adaptive strategies. In contrast, adaptive procedures change the allocation probabilities as enrollment progresses. Two types of adaptive procedures will be considered here. First, we will discuss methods which adjust or adapt the allocation probabilities according to imbalances in numbers of participants or in baseline characteristics between the two groups. Second, we will briefly review adaptive procedures that adjust allocation probabilities according to the responses of participants to the assigned intervention.

Baseline Adaptive Randomization Procedures

Two common methods for adaptive allocation which are designed to make the number of participants in each study group equal or nearly equal are biased coin randomization and urn randomization. Both make adaptations based only on the number of participants in each group, though they can be modified to perform allocation within strata in the same way as blocked randomization, and operate by changing the allocation probability over time.

The Biased Coin Randomization procedure, originally discussed by Efron [46], attempts to balance the number of participants in each treatment group based on the previous assignments, but does not take participant responses into consideration. Several variations to this approach have been discussed [47–63]. The purpose of the algorithm is basically to randomize the allocation of participants to groups A and B with equal probability as long as the number of participants in each group is equal or nearly equal. If an imbalance occurs and the difference in the number of participants is greater than some prespecified value, the allocation probability (p) is adjusted so that it is higher for the group with fewer participants. The investigator can determine the value of the allocation probability. The larger the value of p, the faster the imbalance will be corrected, while the nearer p is to 0.5, the slower the correction. Efron suggests an allocation probability of $p = 2/3$ when a correction is indicated. Since much of the time p is greater than $1/2$, the process has been named the "biased coin" method. As a simple example, suppose n_A and n_B represent the number of participants in groups A and B respectively. If n_A is less than n_B and the difference exceeds a predetermined value, D, then we allocate the next participant to group A with probability $p = 2/3$. If n_A is greater than n_B by an amount of D, we allocate to group B with probability $p = 2/3$. Otherwise, p is set at 0.50. This procedure can be modified to include consideration of the number of consecutive assignments to the same group and the length of such a run. Some procedures for which the allocation probability also depend on differences in baseline characteristics, as discussed below, are sometimes also called "biased coin" designs.

Another similar adaptive randomization method is referred to as the Urn Design, based on the work of Wei and colleagues [64–67]. This method also attempts to keep the number of participants randomized to each group reasonably balanced as the trial progresses. The name Urn Design refers to the conceptual process of randomization. Imagine an urn filled with m red balls and m black balls. If a red ball is drawn at random, assign the participant to group A, return the red ball, and add one (or more than one) black ball to the urn. If a black ball is drawn, assign the participant to group B, return that ball, and add one (or more than one) red ball to the urn. This process will tend to keep the number of participants in each group reasonably close because, like the biased coin procedure it adjusts the allocation probability to be higher for the smaller group. How much imbalance there might be over time depends on m and how many balls are added after each draw.

Since the biased coin and urn procedures are less restrictive than block randomization, they can be less susceptible to selection bias, but by the same token they do

not control balance as closely. If there are temporal trends in the recruitment pool during enrollment, imbalances can create difficulties. This happened in the Stop Atherosclerosis in Native Diabetics Study (SANDS), a trial comparing intensive intervention for cholesterol and blood pressure with less intensive intervention in people with diabetes [68, 69]. Randomization was done using a stratified urn design, but partway through the trial there was in imbalance in the intervention groups at the same time new and more aggressive guidelines regarding lipid lowering treatment in people who had known coronary heart disease came out. The participants in SANDS who met those guidelines could no longer be treated with the less intensive regimen and no new participants with a history of prior cardiovascular events could be enrolled. Not only was there a possibility of imbalance between study groups, the sample size needed to be reconsidered because of the lower average risk level of the participants.

The most correct analysis of a randomized trial from a theoretical point of view is based on permutation distributions modeling the randomization process. For adaptive procedures this requires that the significance level for the test statistic be determined by considering all possible sequences of assignments which could have been made in repeated experiments using the same allocation rule, assuming no group differences. How well population models approximate the permutation distribution for adaptive designs in general is not well understood [6, 14, 70]. Efron [46] argues that it is probably not necessary to take the biased coin randomization into account in the analysis, especially for larger studies. Mehta and colleagues [71] compared analyses ignoring and incorporating biased coin and urn procedures and concluded that the permutation distribution should not be ignored. Smythe and Wei [30, 46] and Wei and Lachin [46, 66] indicate conditions under which test statistics from urn designs are asymptotically normal, and show that if this randomization method is used, but ignored in the analyses, the p-value will be slightly conservative, that is, slightly larger than if the strictly correct analysis were done. Thus the situation for analysis of biased coin and urn designs is similar to that for permuted block designs. Ignoring the randomization is conservative, though not likely to be excessively conservative. Unlike the permuted block design, however, strong temporal trends can create problems for adaptive randomization, and make the permutation-based analysis more important. Although the biased coin method does not appear to be as widely used, stratified urn procedures have been used successfully, as in the multicenter Diabetes Control and Complication Trial [72, 73].

Minimization

In the Enforcing Underage Drinking Laws (EUDL) randomized community trial, 68 communities in five states were selected to receive either an intervention or a control condition. Matched pairs were created using community characteristics including population size, median family income, percentage of the population currently in college, and percentages that were black, Hispanic and spoke Spanish.

The specific set of pairings used was determined by sampling from all possible pairings and selecting the set of pairs with the smallest Mahalanobis distance measure. One community in each pair was then randomly assigned to receive the intervention [74]. In this situation, all the communities to be randomized and the key prognostic covariates are known in advance. The treatment and control groups are guaranteed to be well-balanced, and randomization provides a foundation for later statistical inference using standard population models. This type of a priori matching is a common feature of group-randomized trials [75].

Unfortunately, this is almost never possible in a clinical setting, where patients typically arrive sequentially and must be treated immediately. To accommodate the sequential nature of participant enrollment, some compromise between manipulation of allocation to achieve balance of prognostic covariates and a less restrictive treatment allocation must be made. Stratified block designs can balance a small number of selected prognostic covariates, and randomization will tend to balance unselected as well as unmeasured covariates, but such methods do not perform well when it is important to balance a large number of prognostic covariates in a small sample. For such settings, procedures which adapt allocation to achieve balance on prognostic covariates have been developed.

The biased coin and urn procedures achieve balance in the number of randomizations to each arm. Other stratification methods are adaptive in the sense that intervention assignment probabilities for a participant are a function of the distribution of baseline covariates for participants already randomized. This concept was suggested by Efron [46] as an extension of the biased coin method and also has been discussed in depth by Pocock and Simon [41], and others [47, 48, 51, 52, 59, 63, 76, 77]. In a simple example, if age is a prognostic factor and one study group has more older participants than the other, this allocation scheme is more likely to randomize the next several older participants to the group which currently has more younger participants. Various methods can be used as the measure of imbalance in prognostic factors. In general, adaptive stratification methods incorporate several prognostic factors in making an "overall assessment" of the group balance or lack of balance. Participants are then assigned to a group in a manner which will tend to correct an existing imbalance or cause the least imbalance in prognostic factors. Proschan and colleagues [70] distinguish between minimization procedures which are deterministic [59, 68], as 'strict minimization', reserving the term *mimimization* for the more general procedure described by Pocock and Simon [41] [see Appendix]. Generalization of this strategy exists for more than two study groups. Development of these methods was motivated in part by the previously described problems with non-adaptive stratified randomization for small studies. Adaptive methods do not have empty or near empty strata because randomization does not take place within a stratum although prognostic factors are used. Minimization gives unbiased estimates of treatment effect and slightly increased power relative to stratified randomization [68]. These methods are being used, especially in clinical trials of cancer where several prognostic factors need to be balanced, and the sample size is typically 100–200 participants.

The major advantage of this procedure is that it protects against a severe baseline imbalance for important prognostic factors. Overall marginal balance is maintained in the intervention groups with respect to a large number of prognostic factors. One disadvantage is that minimization is operationally more difficult to carry out, especially if a large number of factors are considered. Although White and Freedman [63] initially developed a simplified version of the minimization method by using a set of specially arranged index cards, today any small programmable computer can easily carry out the calculations. Unlike blocked, biased coin and urn procedures, however, the calculations for minimization cannot be done in advance. In addition, the population recruited needs to be stable over time, just as for other adaptive methods. For example, if treatment guidelines change during a long recruitment period, necessitating a change in the inclusion or exclusion criteria, the adaptive procedure may not be able to correct imbalances that developed beforehand, as with the SANDS example cited above.

For minimization, assuming that the order of participant enrollment is random and applying the allocation algorithm to all permutations or the order can provide a null distribution for the test statistic [14, 70]. Considerable programming and computing resources are required to do this, and biostatisticians prefer to use conventional tests and critical values to determine significance levels. Unfortunately, for minimization there are no general theoretical results on how well the standard analysis approximates the permutation analysis [6, 14, 70], though there are some simulation-based results for specific cases [78].

General advice for stratified block randomization and minimization is to include the baseline variables used to determine the allocation as covariates in the analysis [51, 79]. This seems to produce reliable results in most actual trials using stratified block randomization, and in most trials using minimization, though trials using minimization designs rarely examine the permutation distribution. Proschan et al. [70] however, report an example of an actual trial using minimization for which conventional analysis greatly overstated the significance of the intervention effect relative when compared to the permutation distribution. The use of unequal allocation contributed to the discrepancy in this case, but the Proschan et al. recommend that the permutation test be used to control type 1 error whenever allocation is done using minimization. Several regulatory guidelines make the similar recommendations [80–83].

Despite the appeal of improved balance on more prognostic covariates, most biostatisticians approach minimization and other dynamic allocation plans with caution. As conditions vary considerably from trial to trial, it is expected that the best choice for method of allocation also varies, with the primary goal of avoiding a method which is poorly suited for the given situation.

Response Adaptive Randomization

Response adaptive randomization uses information on participant response to intervention during the course of the trial to determine the allocation of the next participant. Examples of response adaptive randomization models are the Play the

Winner [84] and the Two-Armed Bandit [85] models. These models assume that the investigator is randomizing participants to one of two interventions and that the primary response variable can be determined quickly relative to the total length of the study. Bailar [86] and Simon [87] reviewed the uses of these allocation methods. Additional modifications or methods were developed [88–94].

The *Play the Winner* procedure may assign the first participant by the toss of a coin. The next participant is assigned to the same group as the first participant if the response to the intervention was a success; otherwise, the participant is assigned to the other group. That is, the process calls for staying with the winner until a failure occurs and then switching. The following example illustrates a possible randomization scheme where S indicates intervention success and F indicates failure:

	Participant								
Assignment	1	2	3	4	5	6	7	8	...
Group *A*	S	F				S	F		
Group *B*			S	S	F			S	

Another response adaptive randomization procedure is the *Two Armed Bandit* method which continually updates the probability of success as soon as the outcome for each participant is known. That information is used to adjust the probabilities of being assigned to either group in such a way that a higher proportion of future participants would receive the currently "better" or more successful intervention.

Both of these response adaptive randomization methods have the intended purpose of maximizing the number of participants on the "superior" intervention. They were developed in response to ethical concerns expressed by some clinical investigators about the randomization process. Although these methods do maximize the number of participants on the "superior" intervention, the possible imbalance will almost certainly result in some loss of power and require more participants to be enrolled into the study than would a fixed allocation with equal assignment probability [92]. A major limitation is that many clinical trials do not have an immediately occurring response variable. They also may have several response variables of interest with no single outcome easily identified as being the one upon which randomization should be based. Furthermore, these methods assume that the population from which the participants are drawn is stable over time. If the nature of the study population should change and this is not accounted for in the analysis, the reported significance levels could be biased, perhaps severely [93]. Here, as before, the data analysis should ideally take into account the randomization process employed. For response adaptive methods, that analysis will be more complicated than it would be with simple randomization. Because of these disadvantages, response adaptive procedures are not commonly used.

One application of response adaptive allocation can be found in a trial evaluating extra-corporeal membrane oxygenator (ECMO) in a neonatal population suffering from respiratory insufficiency [95–99]. This device oxygenates the blood to compensate for the inability or inefficiency of the lungs to achieve this task. In this trial,

the first infant was allocated randomly to control therapy. The result was a failure. The next infant received ECMO which was successful. The next ten infants were also allocated to ECMO and all outcomes were successful. The trial was then stopped. However, the first infant was much sicker than the ECMO-treated infants. Controversy ensued and the benefits of ECMO remain unclear. This experience does not offer encouragement to use this adaptive randomization methodology.

Mechanics of Randomization

The manner in which the chosen randomization method is actually implemented is very important [100]. If this aspect of randomization does not receive careful attention, the entire randomization process can easily be compromised, thus voiding any of the advantages for using it. To accomplish a valid randomization, it is recommended that an independent central unit be responsible for developing the randomization process and making the assignments of participants to the appropriate group [27, 101]. For a single center trial, this central unit might be a biostatistician or clinician not involved with the care of the participants. In the case of a multicenter trial, the randomization process is usually handled by the data coordinating center. Ultimately, however, the integrity of the randomization process will rest with the investigator.

Chalmers and colleagues [102] reviewed the randomization process in 102 clinical trials, 57 where the randomization was unknown to the investigator and 45 where it was known. The authors reported that in 14% of the 57 studies, at least one baseline variable was not balanced between the two groups. For the studies with known randomization schedules, twice as many, or 26.7%, had at least one prognostic variable maldistributed. For 43 non-randomized studies, such imbalances occurred four times as often or in 58%. The authors emphasized that those recruiting and entering participants into a trial should not be aware of the next intervention assignment.

In many cases when a fixed proportion randomization process is used, the randomization schedules are made before the study begins [103–107]. The investigators may call a central location, and the person at that location looks up the assignment for the next participant [103]. Another possibility, used historically and still sometimes in trials involving acutely ill participants, is to have a scheme making available sequenced and sealed envelopes containing the assignments [106]. As a participant enters the trial, she receives the next envelope in the sequence, which gives her the assignment. Envelope systems, however, are more prone to errors and tampering than the former method [27, 101]. In one study, personnel in a clinic opened the envelopes and arranged the assignments to fit their own preferences, accommodating friends and relatives entering the trial. In another case, an envelope fell to the bottom of the box containing the envelopes, thus changing the sequence in which they were opened. Many studies prefer web-based or telephone systems to protect against this problem. In an alternative procedure that

has been used in several double-blind drug studies, medication bottles are numbered with a small perforated tab [105]. The bottles are distributed to participant in sequence. The tab, which is coded to identify the contents, is torn off and sent to the central unit. This system is also subject to abuse unless an independent person is responsible for dispensing the bottles. Many clinical trials using a fixed proportion randomization schedule require that the investigator access a website or call the central location to verify that a participant is eligible to be in the trial before any assignment is made. This increases the likelihood that only eligible participants will be randomized.

For many trials, especially multicenter and multinational trials, logistics require a central randomization operations process. Web-based approaches to randomization and other aspects of trial management predominate now [108]. In some cases, the clinic may register a participant by dialing into a central computer and entering data via touchtone, with a voice response. These systems, referred to as Interactive Voice Response Systems or IVRS, or Interactive Web Response Systems, IWRS, are effective and can be used to not only assign intervention but can also capture basic eligibility data. Before intervention is assigned, baseline data can be checked to determine eligibility. This concept has been used in a pediatric cancer cooperative clinical trial network [109] and in major multicenter trials [110, 111].

Whatever system is chosen to communicate the intervention assignment to the investigator or the clinic, the intervention assignment should be given as closely as possible to the moment when both investigator and participant are ready to begin the intervention. If the randomization takes place when the participant is first identified and the participant withdraws or dies before the intervention actually begins, a number of participants will be randomized before being actively involved in the study. An example of this occurred in a non-blinded trial of alprenolol in survivors of an acute myocardial infarction [112]. In that trial, 393 participants with a suspected myocardial infarction were randomized into the trial at the time of their admission to the coronary care unit. The alprenolol or placebo was not initiated until 2 weeks later. Afterwards, 231 of the randomized participants were excluded because a myocardial infarction could not be documented, death had occurred before therapy was begun, or various contraindications to therapy were noted. Of the 162 participants who remained, 69 were in the alprenolol group and 93 were in the placebo group. This imbalance raised concerns over the comparability of the two groups and possible bias in reasons for participant exclusion. By delaying the randomization until initiation of therapy, the problem of these withdrawals could have been avoided.

Problems of implementation can also affect the integrity of the randomization procedure. Downs and colleagues [101] relate their experiences with problems caused by errors in programming, incomplete and missing data for stratification variables, and other problems. They also recommend testing of the proposed procedure before the trial begins, and monitoring of the allocation after it begins.

Recommendations

For large studies involving more than several hundred participants, the randomization should be blocked. If a large multicenter trial is being conducted, randomization should be stratified by center. Randomization stratified on the basis of other factors in large studies is usually not necessary, because randomization tends to make the study groups quite comparable for all risk factors. The participants can still, of course, be stratified once the data have been collected and the study can be analyzed accordingly.

For small studies, the randomization should also be blocked, and stratified by center if more than one center is involved. Since the sample size is small, a few strata for important risk factors may be defined to assure that balance will be achieved for at least those factors. For a larger number of prognostic factors, the adaptive stratification techniques should be considered and the appropriate analyses performed. As in large studies, stratified analysis can be performed even if stratified randomization was not done. For many situations, this will be satisfactory.

Appendix: Adaptive Randomization Algorithm

Adaptive randomization can be used for more than two intervention groups, but for the sake of simplicity only two will be used here. In order to describe this procedure in more detail, a minimum amount of notation needs to be defined. First, let

x_{ik} = the number of participants already assigned intervention k
$(k = 1, 2)$ who have the same level of prognostic factor i
$(i = 1, 2, \ldots, f)$ as the new participant.

and define

$$x_{ik}^t = x_{ik} \qquad \text{if } t \neq k$$
$$= x_{ik} + 1 \quad \text{if } t = k$$

The x_{ik}^t represents the change in balance of allocation if the new participant is assigned intervention t. Finally, let

$B(t)$ = function of the x_{ik}^t's, which measures the "lack of balance" over all prognostic factors if the next participant is assigned intervention t.

Many possible definitions of $B(t)$ can be identified. As an illustrative example, let

$$B(t) = \sum_{i=1}^{f} w_i \, Range \left(x_{i1}^t, x_{i2}^t \right)$$

where w_i = the relative importance of factor i to the other factors and the range is the absolute difference between the largest and smallest values of x_{i1}^t and x_{i2}^t.

Table 6.A1 Fifty randomized participants by group and level of factor $(x_{ik}$'s$)^a$

Factor	1		2			
Level	1	2	1	2	3	
Group						Total
1	16	10	13	9	4	26
2	14	10	12	6	6	24
	30	20	25	15	10	50

[a]After Pocock and Simon [41]

The value of $B(t)$ is determined for each intervention ($t=1$ and $t=2$). The intervention with the smaller $B(t)$ is preferred, because allocation of the participant to that intervention will cause the least imbalance. The participant is assigned, with probability $p > 1/2$, to the intervention with the smaller score, $B(1)$ or $B(2)$. The participant is assigned, with probability $(1-p)$, to the intervention with the larger score. These probabilities introduce the random component into the allocation scheme. Note that if $p=1$ and, therefore, $1-p=0$, the allocation procedure is deterministic (no chance or random aspect) and has been referred to by the term "minimization" [51, 59].

As a simple example of the adaptive stratification method, suppose there are two groups and two prognostic factors to control. The first factor has two levels and the second factor has three levels. Assume that 50 participants have already been randomized and the following table summarizes the results (Table 6.A1).

In addition, the function $B(t)$ as defined above will be used with the range of the x_{ik}'s as the measure of imbalance, where $w_1 = 3$ and $w_2 = 2$; that is, the first factor is 1.5 times as important as the second as a prognostic factor. Finally, suppose $p = 2/3$ and $1-p=1/3$.

If the next participant to be randomized has the first level of the first factor and the third level of the second factor, then this corresponds to the first and fifth columns in the table. The task is to determine $B(1)$ and $B(2)$ for this participant as shown below.

1. Determine $B(1)$

   (a)   Factor 1, Level 1

	K	x_{1k}	x^I_{1k}	Range (x^I_{11}, x^I_{12})
Group	1	16	17	$\|17-14\|=3$
	2	14	14	

   (b)   Factor 2, Level 3

	K	x_{2k}	x^I_{2k}	Range (x^I_{21}, x^I_{22})
Group	1	4	5	$\|5-6\|=1$
	2	6	6	

   Using the formula given, $B(1)$ is computed as $3 \times 3 + 2 \times 1 = 11$.

2. Determine $B(2)$

   (a) Factor 1, Level 1

	K	x_{1k}	x^2_{1k}	Range (x^2_{11}, x^2_{12})
Group	1	16	16	$\lvert16-15\rvert=1$
	2	14	15	

   (b) Factor 2, Level 3

	K	x_{2k}	x^1_{1k}	Range (x^1_{21}, x^1_{22})
Group	1	4	4	$\lvert4-7\rvert=3$
	2	6	7	

Then $B(2)$ is computed as $3 \times 1 + 2 \times 3 = 9$.

3. Now rank $B(1)$ and $B(2)$ from smaller to larger and assign with probability p the group with the smaller $B(t)$.

t	$B(t)$	Probability of assigning t
2	$B(2)=9$	$p=2/3$
1	$B(1)=11$	$1-p=1/3$

   Thus, this participant is randomized to Group 2 with probability 2/3 and to Group 1 with probability 1/3. Note that if minimization were used ($p=1$), the assignment would be Group 2.

References

1. Armitage P. The role of randomization in clinical trials. *Statist Med* 1982;1:345–352.
2. Brown BW. Statistical Controversies in the Design of Clinical-Trials—Some Personal Views. *Control Clin Trials* 1980;1:13–27.
3. Byar DP, Simon RM, Friedewald WT, et al. Randomized Clinical Trials. *N Engl J Med* 1976;295:74–80.
4. Hill AB. The Clinical Trial. *Brit M Bull* 1951;7:278–282.
5. Lachin JM. Statistical properties of randomization in clinical trials. *Control Clin Trials* 1988;9:289–311.
6. Lachin JM, Matts JP, Wei LJ. Randomization in clinical trials: Conclusions and recommendations. *Control Clin Trials* 1988;9:365–374.
7. Peto R. Clinical trial methodology. *Biomedicine* 1978;28:24–36.
8. Peto R, Pike MC, Armitage P, et al. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. I. Introduction and design. *British Journal of Cancer* 1976;34:585.
9. Pocock SJ. Allocation of Patients to Treatment in Clinical Trials. *Biometrics* 1979; 35:183–197.
10. Royall RM. Ethics and statistics in randomized clinical trials. *Statistical Science* 1991; 6:52–62.

11. Weinstein MC. Allocation of Subjects in Medical Experiments. *N Engl J Med* 1974;291:1278–1285.

12. Zelen M. The randomization and stratification of patients to clinical trials. *J Chronic Dis* 1974;27:365–375.

13. Bather JA. On the allocation of treatments in sequential medical trials. *International Statistical Review/Revue Internationale de Statistique* 1985;1–13.

14. Kalish LA, Begg CB. Treatment allocation methods in clinical trials: a review. *Statist Med* 1985;4:129–144.

15. Altman DG, Dore CJ. Randomisation and baseline comparisons in clinical trials. *Lancet* 1990;335:149–153.

16. Stigler SM. The use of random allocation for the control of selection bias. *Biometrika* 1969;56:553–560.

17. Wei LJ. On the Random Allocation Design for the Control of Selection Bias in Sequential Experiments. *Biometrika* 1978;65:79–84.

18. Williams DH, Davis CE. Reporting of assignment methods in clinical trials. *Control Clin Trials* 1994;15:294–298.

19. Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *The Lancet* 2001;357:1191–1194.

20. Mills EJ, Wu P, Gagnier J, Devereaux PJ. The quality of randomized trial reporting in leading medical journals since the revised CONSORT statement. *Contemp Clin Trials* 2005;26:480–487.

21. Brittain E, Schlesselman JJ. Optimal Allocation for the Comparison of Proportions. *Biometrics* 1982;38:1003–1009.

22. Kalish LA, Harrington DP. Efficiency of Balanced Treatment Allocation for Survival Analysis. *Biometrics* 1988;44:815–821.

23. Lachin JM. Properties of simple randomization in clinical trials. *Control Clin Trials* 1988;9:312-326.

24. Louis TA. Optimal allocation in sequential tests comparing the means of two Gaussian populations. *Biometrika* 1975;62:359–369.

25. Louis TA. Sequential Allocation in Clinical Trials Comparing Two Exponential Survival Curves. *Biometrics* 1977;33:627–634.

26. Matts JP, Lachin JM. Properties of permuted-block randomization in clinical trials. *Control Clin Trials* 1988;9:327–344.

27. Schulz KF, Grimes DA. Allocation concealment in randomised trials: defending against deciphering. *The Lancet* 2002;359:614–618.

28. Kalish LA, Begg CB. The impact of treatment allocation procedures on nominal significance levels and bias. *Control Clin Trials* 1987;8:121–135.

29. Matts JP, McHugh RB. Analysis of accrual randomized clinical trials with balanced groups in strata. *J Chronic Dis* 1978;31:725–740.

30. Smythe RT, Wei LJ. Significance tests with Restricted Randomization Design. *Biometrika* 1983;70:496-500.

31. Ducimetiere P. Stratification; in Boissel JP, Klimt CR (eds): Multi-Center Controlled Trials, Principles and Problems. Paris, Institut National de la Sante et de la Recherche Medicale, 1979.

32. Feinstein AR, Landis JR. The role of prognostic stratification in preventing the bias permitted by random allocation of treatment. *J Chronic Dis* 1976;29:277–284.

33. Fleiss JL. Multicentre clinical trials: Bradford Hill's contributions and some subsequent developments. *Statist Med* 1982;1:353–359.

34. Green SB, Byar DP. The effect of stratified randomization on size and power of statistical tests in clinical trials. *J Chronic Dis* 1978;31:445–454.

35. Grizzle JE. A note on stratifying versus complete random assignment in clinical trials. *Control Clin Trials* 1982;3:365–368.

36. Mantel N, McHugh R, Matts J. Pre-Stratification or Post-Stratification. *Biometrics* 1984;40:256–258.
37. McHugh R, Matts J. Post-Stratification in the Randomized Clinical Trial. *Biometrics* 1983;39:217–225.
38. Meier P. Stratification in the design of a clinical trial. *Control Clin Trials* 1981;1:355–361.
39. Palta M, Amini SB. Consideration of covariates and stratification in sample size determination for survival time studies. *J Chronic Dis* 1985;38:801–809.
40. Palta M, Amini SB. Magnitude and likelihood of loss resulting from non-stratified randomization. *Statist Med* 1982;1:267–275.
41. Pocock SJ, Simon R. Sequential Treatment Assignment with Balancing for Prognostic Factors in the Controlled Clinical Trial. *Biometrics* 1975;31:103–115.
42. Simon R. Restricted Randomization Designs in Clinical Trials. *Biometrics* 1979;35:503–512.
43. The Coronary Drug Project Research Group: Factors influencing long-term prognosis after recovery from myocardial infarctionΓÇöThree-year findings of the Coronary Drug Project. *J Chronic Dis* 1974;27:267–285.
44. Zelen M. Aspects of the planning and analysis of clinical trials in cancer. A Survey of Statistical Design and Linear Models Amsterdam: North-Holland 1975.
45. Aspirin Myocardial Infarction Study Research Group: A randomized, controlled trial of aspirin in persons recovered from myocardial infarction. *JAMA* 1980;243:661–669.
46. Efron B. Forcing a Sequential Experiment to be Balanced. *Biometrika* 1971;58:403–417.
47. Atkinson AC. Optimum biased coin designs for sequential clinical trials with prognostic factors. *Biometrika* 1982;69:61–67.
48. Begg CB, Iglewicz B. A Treatment Allocation Procedure for Sequential Clinical Trials. *Biometrics* 1980;36:81–90.
49. Begg CB. On Inferences from Wei's Biased Coin Design for Clinical Trials. *Biometrika* 1990;77:467–478.
50. Efron B. Randomizing and balancing a complicated sequential experiment. Biostatistics Casebook 1980;19–30.
51. Forsythe AB, Stitt FW. Randomization or minimization in the treatment assignment of patient trials: validity and power of tests. Health Sciences Computing Facility, University of California, 1977.
52. Freedman LS, White SJ. On the Use of Pocock and Simon's Method for Balancing Treatment Numbers over Prognostic Factors in the Controlled Clinical Trial. *Biometrics* 1976;32:691–694.
53. Halpern J, Brown BW. Sequential treatment allocation procedures in clinical trials-with particular attention to the analysis of results for the biased coin design. *Statist Med* 1986;5:211–229.
54. Hannigan JF, Brown BW. Adaptive Randomization Biased Coin-Design: Experience in a Cooperative Group Clinical Trial. 74. 1982. Department of Statistics, Stanford University.
55. Klotz JH. Maximum Entropy Constrained Balance Randomization for Clinical Trials. *Biometrics* 1978;34:283–287.
56. Raghavarao D. Use of distance function in sequential treatment assignment for prognostic factors in the controlled clinical trial. *Calcutta Statist Assoc Bull* 1980;29:99–102.
57. Smith RL. Sequential Treatment Allocation Using Biased Coin Designs. *J R Stat Soc Series B Stat Methodol* 1984;46:519–543.
58. Soares JF, Jeff Wu CF. Some restricted randomization rules in sequential designs. *Commun StatTheory Methods* 1983;12:2017–2034.
59. Taves DR. Minimization: a new method of assigning patients to treatment and control groups. *Clinical pharmacology and therapeutics* 1974;15:443.
60. Wei LJ. A class of treatment assignment rules for sequential experiments. *Commun Stat Theory Methods* 1978;7:285–295.
61. Wei LJ. The Adaptive Biased Coin Design for Sequential Experiments. *Ann Stat* 978; 6:92–100.
62. Wei LJ. A class of designs for sequential clinical trials. *J Am Stat Assoc* 1977;72:382–386.
63. White SJ, Freedman LS. Allocation of patients to treatment groups in a controlled clinical study. Br J Cancer 1978;37:849.

64. Wei LJ. An application of an urn model to the design of sequential controlled clinical trials. *J Am Stat Assoc* 1978;73:559–563.
65. Wei LJ, Smythe RT, Smith RL. K-treatment comparisons with restricted randomization rules in clinical trials. *Ann Stat* 1986;265–274.
66. Wei LJ, Lachin JM. Properties of the urn randomization in clinical trials. *Control Clin Trials* 1988;9:345–364.
67. Wei LJ, Smythe RT, Lin DY, Park TS. Statistical inference with data-dependent treatment allocation rules. *J Am Stat Assoc* 1990;85:156–162.
68. Birkett NJ. Adaptive allocation in randomized controlled trials. *Control Clin Trials* 1985;6:146–155.
69. Russell M, Fleg JL, Galloway W, et al. Examination of lower targets for low-density lipoprotein cholesterol and blood pressure in diabetes—the Stop Atherosclerosis in Native Diabetics Study (SANDS). *American Heart Journal* 2006;152:867–875.
70. Proschan MA, Brittain E, Kammerman L. Minimize the Use of Minimization with Unequal Allocation. *Biometrics* 2011;67:1135–1141.
71. Mehta CR, Patel NR, Wei LJ. Constructing Exact Significance Tests with Restricted Randomization Rules. *Biometrika* 1988;75:295–302.
72. The DCCT Research Group: The Diabetes Control and Complications Trial (DCCT): Design and Methodologic Considerations for the Feasibility Phase. *Diabetes* 1986;35:530–545.
73. The DCCT Research Group: Diabetes Control and Complications Trial (DCCT): Results of Feasibility Study. The DCCT Research Group. *Diabetes Care* 1987;10:1–19.
74. Reboussin BA, Preisser JS, Song EY, Wolfson M. Sample size estimation for alternating logistic regressions analysis of multilevel randomized community trials of under-age drinking. *R Stat Soc Ser A Stat Soc* 2012;175:691–712.
75. Murray DM. Design and Analysis of Group-randomized Trials. Oxford University Press, 1998.
76. Gail MH, Wieand S, Piantadosi S. Biased Estimates of Treatment Effect in Randomized Experiments with Nonlinear Regressions and Omitted Covariates. *Biometrika* 1984; 71:431–444.
77. Begg CB, Kalish LA. Treatment Allocation for Nonlinear Models in Clinical Trials: The Logistic Model. *Biometrics* 1984;40:409–420.
78. Aickin M. A Simulation Study of the Validity and Efficiency of Design-Adaptive Allocation to Two Groups in the Regression Situation. *Int J Biostat* 2009;5:Article 19.
79. Kahan BC, Morris TP. Improper analysis of trials randomised using stratified blocks or minimisation. *Statist Med* 2012;31:328–340.
80. Committee for Proprietary Medicinal Products (CPMP: Committee for Proprietary Medicinal Products (CPMP): points to consider on adjustment for baseline covariates. *Statist Med* 2004;23:701.
81. FDA: International Conference on Harmonization—Efficacy: Statistical principles for clinical trials. U S Food and Drug Administration.
82. Food and Drug Administration: Guidance for industry: Adaptive design clinical trials for drugs and biologics. Washington DC, USA: Food and Drug Administration 2010.
83. Green H, McEntegart DJ, Byrom B, et al. Minimization in crossover trials with non-prognostic strata: theory and practical application. *J Clin Pharm Ther* 2001;26:121–128.
84. Zelen M. Play the Winner Rule and the Controlled Clinical Trial. *J Am Stat Assoc* 1969;64:131–146.
85. Robbins H. Some aspects of the sequential design of experiments. *Bulletin of the American Mathematical Society* 1952;58:527–535.
86. Bailar JC. Patient assignment algorithms: an overview; in Proceedings of the 9th International Biometric Conference, Boston, August 22-27, 1976: Invited Papers: Biometric Society, 1976, pp 189–206.
87. Simon R. Adaptive Treatment Assignment Methods and Clinical Trials. *Biometrics* 1977;33:743–749.

88. Armitage P. The search for optimality in clinical trials. *International Statistical Review/Revue Internationale de Statistique* 1985;15–24.

89. Bather JA. Randomized Allocation of Treatments in Sequential Experiments. *J R Stat Soc Series B Stat Methodol* 1981;43:265–292.

90. Berry DA. Modified Two-Armed Bandit Strategies for Certain Clinical Trials. *J Am Stat Assoc* 1978;73:339–345.

91. Nordbrock E. An Improved Play-the-Winner Sampling Procedure for Selecting the Better of Two Binomial Populations. *J Am Stat Assoc* 1976;71:137–139.

92. Simon R, Weiss GH, Hoel DG. Sequential Analysis of Binomial Clinical Trials. *Biometrika* 1975;62:195–200.

93. Simon R, Hoel DG, Weiss GH. The use of covariate information in the sequential analysis of dichotomous response experiments. *Commun Stat Theory Methods* 1977;6:777–788.

94. Wei LJ. Exact two-sample permutation tests based on the randomized play-the-winner rule. *Biometrika* 1988;75:603–606.

95. Bartlett RH, Roloff DW, Cornell RG, et al. Extracorporeal Circulation in Neonatal Respiratory Failure: A Prospective Randomized Study. *Pediatrics* 1985;76:479–487.

96. O'Rourke PP, Crone RK, Vacanti JP, et al. Extracorporeal Membrane Oxygenation and Conventional Medical Therapy in Neonates With Persistent Pulmonary Hypertension of the Newborn: A Prospective Randomized Study. *Pediatrics* 1989;84:957–963.

97. Paneth N, Wallenstein S. Extracorporeal Membrane Oxygenation and the Play the Winner Rule. *Pediatrics* 1985;76:622–623.

98. Ware JH, Epstein MF. Extracorporeal Circulation in Neonatal Respiratory Failure: A Prospective Randomized Study. *Pediatrics* 1985;76:849–851.

99. Ware JH. Investigating Therapies of Potentially Great Benefit: ECMO. *Stat Sci* 1989;4:298–306.

100. Pocock SJ, Lagakos SW. Practical experience of randomization in cancer trials: an international survey. *Br J Cancer* 1982;46:368–375.

101. Downs M, Tucker K, Christ-Schmidt H, Wittes J. Some practical problems in implementing randomization. *Clin Trials* 2010;7:235–245.

102. Chalmers TC, Celano P, Sacks HS, Smith H. Bias in Treatment Assignment in Controlled Clinical Trials. *N Engl J Med* 1983;309:1358–1361.

103. Beta-Blocker Heart Attack Trial Research Group. A randomized trial of propranolol in patients with acute myocardial infarction: I. mortality results. *JAMA* 1982;247:1707–1714.

104. CASS Principle Investigators and Their Associates. Coronary artery surgery study (CASS): a randomized trial of coronary artery bypass surgery. Survival data. *Circulation* 1983;68:939–950.

105. Collaborative Group on Antenatal Steroid Therapy. Effect of antenatal dexamethasone administration on the prevention of respiratory distress syndrome. *Am J Obstet Gynecol* 1981;141:276–287.

106. Hypertension Detection and Follow-up Program Cooperative Group: Five-year findings of the hypertension detection and follow-up program: I. reduction in mortality of persons with high blood pressure, including mild hypertension. *JAMA* 1979;242:2562–2571.

107. Multiple Risk Factor Intervention Trial Research Group. Multiple risk factor intervention trial: Risk factor changes and mortality results. *JAMA* 1982;248:1465–1477.

108. Reboussin D, Espeland MA. The science of web-based clinical trial management. *Clin Trials* 2005;2:1–2.

109. Krischer JP, Hurley C, Pillalamarri M, et al. An automated patient registration and treatment randomization system for multicenter clinical trials. *Control Clin Trials* 1991;12:367–377.

110. Kjekshus J, Apetrei E, Barrios V, et al. Rosuvastatin in Older Patients with Systolic Heart Failure. *N Engl J Med* 2007;357:2248–2261.

111. SPORTIF Executive Steering Committee for the SPORTIF: Ximelagatran vs warfarin for stroke prevention in patients with nonvalvular atrial fibrillation: A randomized trial. *JAMA* 2005;293:690–698.

112. Ahlmark G, Saetre H. Long-term treatment with ß-blockers after myocardial infarction. *Eur J Clin Pharmacol* 1976;10:77–83.