PRACTICAL TIPS FOR SURGICAL RESEARCH

Ensuring balanced groups in surgical trials

Forough Farrokhyar, MPhil, PhD*†
Sohail Bajammal, MBChB, MSc‡
Kamyar Kahnamoui, MD, MSc§
Mohit Bhandari, MD, MSc†¶

From the Departments of *Surgery and of
†Clinical Epidemiology and Biostatistics,
McMaster University, Hamilton, Ont., the
‡Department of Surgery, Umm Al-Qura
University, Saudi Arabia, and the
Divisions of §General Surgery and of
¶Orthopaedic Surgery, McMaster
University, Hamilton, Ont.

Correspondence to:
Dr. F. Farrokhyar
St. Joseph's Healthcare
Mary Grace Wing, Fl. 8, Rm. G845
50 Charlton Ave. E
Hamilton ON L8N 4A6
farrokh@mcmaster.ca

High-quality randomized controlled trials (RCTs) are the highest level of evidence in assessing the effectiveness of a treatment. It is random allocation that places RCTs in the highest level of evidence. The purpose of randomization is to create groups of patients that are comparable for known and unknown factors at the start of the trial so that any differences at the completion of the trial can be attributed to the treatment under investigation.[1] The purpose of this article is to discuss the processes that would help create balanced groups and maintain between-group comparability throughout the study period.

OBJECTIVES OF THE ARTICLE

This article will discuss the importance of balancing groups in RCTs of surgical treatments. By the end of the article, the reader will be able to appreciate the importance of having balanced groups in surgical trials and learn to apply appropriate strategies to ensure that the treatment groups are balanced for prognostic factors throughout the study period. The subject matter is divided into 5 sections:
- enrolment
- randomization
- concealment of allocation
- blinding
- complete follow-up
- intention-to-treat analysis

ENROLMENT

An earlier paper in this series[2] provided helpful tips on how to optimize patient recruitment in surgical trials. In this section, we will discuss how to maintain the comparability of groups once patients are recruited.

The inclusion and exclusion criteria of surgical trials have multiple implications; they are, therefore, deserving of time and attention. The inclusion criteria basically define the population of the research question. The exclusion criteria define populations of patients who will not help in answering the research question or might be harmed by research interventions. It is very important that a record is kept of all patients who were assessed for eligibility, identifying those who were excluded and stating the reason. This ensures that the risk of selection bias will be minimized (i.e., preferential exclusion of certain patients from joining the study).

Adherence to the trial protocol strengthens trial validity and credibility. Protocol violations can happen at the enrolment stage with the erroneous enrolment of an ineligible patient into the trial. The researcher should be prepared, in consultations with study investigators and biostatisticians, to deal with such potential protocol violations and ensure that eligibility criteria is diligently followed and inclusion of ineligible patients are avoided. The optimal method is

prevention through proper training. Thus, every effort should be made to adequately train research personnel to avoid protocol violations. In multicentre surgical trials, it is imperative to communicate the enrolment criteria to research collaborators across study centres. It is important that research collaborators understand and agree on these criteria to avoid any errors in enrolment. However, if this violation occurs, some researchers recommend "blinded postrandomization exclusion."[3] This basically means reviewing the eligibility of all randomly assigned patients and excluding those who are ineligible but who were enrolled in error. It is of paramount importance that this process is conducted blindly without the knowledge of patient group allocation to protect against preferential imbalance between groups. The number of these exclusions should be kept to minimum, preferably none, to avoid detrimental effects on credibility and validity.

Patient-entry forms are safeguards against enrolling ineligible patients. There are good resources to help researchers to design effective patient-entry forms.[4,5] Briefly, when designing a patient-entry form, include data points to:

1. uniquely identify and keep track of patients,
2. confirm that the patient has the disorder of interest,
3. ensure that all eligibility criteria are met, and
4. ensure that vital baseline characteristics and prognostic factors are collected.

RANDOMIZATION

Randomization is a process during which the patients have an equal chance of being allocated to either study treatment group. The goal is to produce comparable groups in a way that both known and unknown prognostic factors are balanced[1,6] and that any imbalance that might occur will be by chance rather than by choice. Randomization is the most optimal method to minimize selection bias and control for known and unknown confounding factors. A true randomization process eliminates selection bias.[1] The most robust and optimal method of randomization is computer-generated random numbers. Coin-tossing, dice-throwing or using random number tables (from statistical textbooks) represent reasonable approaches for the generation of simple randomization sequences, but might become nonrandom in practice. These methods do not provide concealment allocation. If, for example, using the coin-toss method, an investigator throws a series of "heads" with no "tails," he or she might be tempted to alter the results of a toss or a series of tosses.[7] Some researchers allocate patients to groups in a way that is not truly random (e.g., using the day of the week or alternate medical record number) and are called "quasirandom."[1] Although these methods might seem to generate comparable groups, they cannot provide concealment of allocation. This introduces a substantial risk of selection bias.[1,8] Using these systematic methods, the

study personnel can predict to which group the next patient will be assigned and might, consciously or unconsciously, exclude that patient from the study for different reasons. To minimize bias, patients should be assigned to study groups based on a truly random process. Timing of randomization is also very important in preventing "postrandomization exclusion." An eligible patient might become ineligible if the there is a lag time from randomization to surgical intervention. As in clinical trials, randomization should be performed very close to when the intervention is performed. If possible, patients' informed consent should be obtained preoperatively, but randomization occurs intraoperatively once there is certainty that the patient could receive either intervention.[1]

There are different methods of randomization. This paper will present the most commonly used of these methods.

Simple randomization

Simple or unrestricted randomization is the method in which each eligible patient will be randomly assigned to one of the treatments. This process will continue until the required numbers of patients are randomly assigned. Using this method, about one-half of the patients will be randomly assigned to the experimental and one-half to the conventional treatments. This method is easy to implement but has a major disadvantage for smaller trials. Although an equal number of patients in each group is expected, at any point, including the end, there could be a substantial sample size imbalance for small trials ($n < 100$).[9,10] Whereas such imbalances do not cause a statistical test to be invalid, they reduce the ability to detect the true differences and may lead to some loss of trial credibility.[11] For this reason, simple randomization it not often used even for large trials.

Blocked randomization

Blocked or permuted block randomization is a restricted randomization method frequently used in RCTs to ensure that roughly equal numbers of patients are randomly assigned to the treatment groups.[11] Blocked randomization increases the power as well as the credibility of the study. It means that within each block, patients are randomly assigned with equal probability to receive one of the study treatments. The block size could be determined by the researcher.[12] It is recommended to use smaller increments for small trials to ensure balanced groups.[9] After block size is determined, all possible combinations of assignment for that block size are calculated, and block combination is then randomly chosen to determine the patients' assignment into the treatment groups.[9,12] For a block size of 4, for example, 2 patients will be randomly assigned to treatment A and 2 to treatment B. There are only 6 blocks in

which 2 patients will receive treatment A and 2 will receive treatment B: AABB, ABAB, BABA, ABBA, BAAB and BBAA. One of these blocks is randomly chosen to determine the assignment. This process of random block selection is repeated until all patients are randomly assigned to one of the groups. For example, for a sample size of 120, 30 block sizes of 4 are needed. A potential risk with using a fixed blocked size, especially if small, is that the block size could be deciphered in unblinded trials.[7,9,13] This means that the assignment of the last person entered in each block is known before randomization. Let us assume, for example, patients randomly assigned to treatment A will receive laparoscopic surgery and those to treatment B will receive open surgery. If the blocking size is 4 and the first 3 assignments are ABB, then the next assignment must be treatment A or laparoscopic surgery. When treatment allocations become known, the sequence can be discerned from the pattern of the last assignment.[7] The personnel randomly assigning the patients to the treatments groups may decide (either intentionally or unintentionally) that the next patient's characteristics are not favourable for laparoscopic surgery (next assignment) and decide not to randomly assign the patient into the study or to randomly assign that patient later. Even if the allocation is effectively concealed, some future assignment could be guessed and selection bias introduced.[7] Thus, it is recommended to use blocks of variable sizes.[4,6,8] After the completion of each block, the size of the next block is determined at random. This method of allocation is called "random blocked randomization."

Stratified randomization

Before starting the randomization sequence, the researcher should assess whether there are major prognostic factors that are strongly associated with subsequent patient response or outcome.[9,11] Such factors should be considered for stratified randomization. Stratified randomization prevents an imbalance between treatment groups for factors that influence treatment responsiveness.[14] Stratified randomization requires the prognostic factor of interest to be measured a priori or at the time of randomization. Stratified randomization may be useful in small trials as some imbalances, for example age, might occur and complicate the interpretation of the results.[7,12] Within each stratum, the randomization process could be simple or restricted depending on the size of trial. In multicentre trials, centres may vary with respect to the type of patients, and the quality and type of care given to patients during follow-up. Thus, centre may be an important factor related to patient outcome, and the randomization process should be stratified accordingly.[12] By stratifying randomization within a centre (i.e., using separate randomization schedules at each centre), the extent to which major imbalances between treatment groups will occur across centres can be limited.[11]

Note that the factor of blocking and/or stratifying should be taken into consideration during data analysis. The purpose of blocking and/or stratifying is to ensure balance between treatment groups and increase the power of the study; therefore, ignoring blocking and/or stratifying factors in the data analysis may result in misleading conclusions.[11,15] There are other randomization methods, such as the adaptive randomization process[11] (i.e., minimization to avoid between-group imbalances) and the maximal procedure,[16] details of which can be found elsewhere.

CONCEALMENT OF ALLOCATION

Regardless of the method of randomization, the risk of selection bias and the ensuing imbalances in the characteristics of patients between the study groups is great when the allocation process is not appropriately concealed.[1] Concealment of allocation implies that individuals responsible for recruiting and assigning patients to treatment groups and patients themselves should remain unaware of the next assignment in the sequence. For example, if people making the decisions about patient eligibility are aware of the group to which patients will be allocated, they may systematically (either intentionally or unintentionally) enrol patients with favourable characteristics such as younger age and lower-grade tumours into the experimental treatment. This selective assignment would introduce bias and create treatment groups that are different with respect to age and tumour grade, thereby influencing trial credibility because patients with favourable factors are likely to have better outcomes regardless of the treatment. Thus it is important to conceal the allocation from the individuals who recruit and enrol patients in the trial. It is also crucial that the individual who generates the allocation scheme should not be involved in ascertaining eligibility, administering treatment or assessing outcome because such an individual will usually have knowledge of the allocation assignment and thus the opportunity to introduce bias.[7] It is possible to conceal the randomization sequence in every RCT. The optimal method is to have the randomization process independently administered using, for example, a 24-hour telephone randomization line, a web-based randomization service, a hospital pharmacy or a central office.[1] Many surgical trials randomly assign patients using envelopes containing the group allocation. Whereas this method theoretically provides concealed allocation, it is highly susceptible to corruption if investigators open multiple envelopes or are able to determine the group allocation without opening the envelope. The use of envelopes should be avoided. If envelopes must be used, the potential for abuse should be minimized. The envelopes should be
- opaque, sealed and serially numbered;
- opened sequentially and only after the participant's name and other details are written on the envelope; and
- kept in a locked and secure place.[1,13]

Note that allocation concealment relates to the process before randomization to prevent selection bias and is different from blinding mechanism after randomization.[1,13]

BLINDING

Blinding is an important methodological feature of surgical trials and needs to be considered more rigorously. Therefore, researchers should make every effort to incorporate blinding into their trial designs. In trials of surgical interventions, surgeons can usually not be blinded, but patients, health care providers, data collectors, outcome assessors and data analysts can often be blinded. Blinding or masking these individuals prevents systematic imbalances in effective concomitant interventions, outcome evaluations and between-group comparability for baseline characteristics. Randomized controlled trials of surgical interventions are often more difficult to blind than drug trials, which typically achieve blinding with placebos.[1] It is most problematic to blind allocation from patients and research personnel when comparing a surgical intervention to nonoperative management. Group imbalances in surgical trials could occur if the outcome assessors, care providers and patients are not blinded to the treatment allocation. The outcome assessors might assess the outcome differently if they are aware of the treatment allocation. Blinding outcome assessors protects the trial against the differential assessment of the outcomes. People who set up follow-up visits may (intentionally or unintentionally) make extra efforts for complete follow-up for patients who received experimental treatment than for those who received conventional treatment if they are not blinded for treatment allocation. This may create differential follow-ups between study groups and introduce attrition bias. When patients are aware of the treatment allocation, their attendance at follow-up visits are usually different than those who are blinded to the treatment allocation. The differential loss to follow-up is greater when surgical intervention is compared with conservative management and blinding is impractical. For example, Michaels and colleagues[17] compared surgery to conservative management for uncomplicated varicose veins. At 1 year follow-up, there was significant attrition owing to patients failing to attend follow-up visits or withdrawing from the trial (35% conservative arm v. 17% surgery arm). By the end of the third year, 52% of patients in the conservative arm had undergone surgery. To increase the internal validity of an RCT, researchers should blind as many involved individuals as possible and clearly state which individuals are blinded and how the blinding is achieved. When blinding of patients and health care providers is not feasible, to prevent group imbalances surgical researchers should ensure that the randomization process is independently administered and that people who randomly assign patients into the trial are not involved in patient care. To maintain group comparability, the surgical researchers should ensure that the study groups are, except for the intervention, treated equally (i.e.,

concomitant interventions) and that every effort is made for a complete follow-up for all participating patients. Another useful tip to avoid differential assessment of outcome measures and maintain comparability between the groups when blinding is not feasible is to have 2 or more individuals independently assess outcomes and resolve the disagreements with consensus.

COMPLETE FOLLOW-UP

Ideally, every patient should be followed until the completion of the study. Failure to account for all patients at the end of the study is another factor that risks introducing imbalances between treatment groups and losing the benefits conferred from randomization.[1] The imbalances become more prominent when there are systematic differences between comparison groups in the loss to follow-ups or drop-outs from the study. Patients who do not attend follow-up visits are usually different from the ones who do;[1] they may have died, experienced the outcome of interest or had a satisfactory outcome. Losses to follow-up are greater and differential when

- no treatment is required after surgery, especially when the follow-up time is long;
- concomitant treatments (e.g., rehabilitation or physiotherapy) are required for one treatment group but not the other; and
- the trial is not blinded and patients are aware of what treatment they have received.[1]

Loss to follow-up is a common problem in surgical trials. There are strategies that can be implemented to minimize the burden of this problem. The following strategies should be considered before randomly assigning patients into the trial:

- excluding patients unlikely to adhere,
- fully informing patients of both the burden of the study and of the harms and benefits of the treatment,
- establishing follow-up visits suited to patients' satisfaction,
- obtaining contact information to prompt patients to return for follow-up,
- keeping data collection to a minimum, and
- hiring a research person to accomplish all of the above.[1]

Despite these strategies, there will still most likely be patients lost to follow-up in each group. The proportion of losses to follow-up should be anticipated and accounted for in surgical research. Studies should be adequately powered to answer the research question at initiation, including allowance for losses to follow-up and attrition bias.[18] There are different approaches to handle losses to follow-up at the stage of data analysis, such as last observation carried forward, multiple imputation, mixed effect model and the best-case and worst-case scenarios.[19,20] These approaches do not eliminate the bias introduced from losses to follow-up and the conclusions drawn from these studies must be interpreted with caution.[1]

INTENTION-TO-TREAT ANALYSIS

In many trials, some patients invariably switch from one treatment to the other owing to side effects, apparent lack of effectiveness or a simple change in preference. If researchers analyze patients based on the treatment they receive (known as per protocol or analysis by treatment administered), they risk introducing prognostic imbalances between treatment groups and lose the benefits conferred by randomization. Alternatively, the intention-to-treat approach analyzes patients in the groups to which they were randomly assigned, regardless of the treatment they actually received, and provides the least biased assessment of the efficacy of the treatment.[1,20,21] Intention-to-treat analysis maintains prognostic balances in study groups. In surgical trials, adherence to protocol is not usually an issue when the treatment is a one-time irreversible process, but there may be a chance of conversion from new treatment to conventional treatment for technical reasons or owing to comorbidities. The intention-to-treat analysis does not eliminate bias introduced by conversion, losses to follow-up or withdrawals, but provides the best estimate of the effect size that can be expected for patients in whom the treatment is attempted (regardless of the need for conversion).[1]

We have summarized and listed the key considerations on this topic in Box 1 to help surgeons and surgical researchers sustain comparability between groups to enhance the validity and credibility of their research.

CONCLUSION

The goal of randomization is to create similar groups for the known and unknown confounding factors at the start of the trial. If necessary safeguards are used to maintain similarities between groups, any differences found at the completion of the study can be attributed to the treatment under investigation. Surgeons planning to conduct research need to consider these safeguards ahead of time and use appropriate methodology to avoid biases at different stages of their research and enhance the validity and credibility of their findings.

References

1. Farrokhyar F, Karanicolas PJ, Thoma A, et al. Randomized controlled trials of surgical interventions. *Ann Surg* 2010;251:409-16.

2. Thoma A, Farrokhyar F, McKnight L, et al. Practical tips for surgical research: how to optimize patient recruitment. *Can J Surg* 2010; 53:205-10.

Box 1. Tips for surgical researchers in optimizing between-group comparability: key considerations by study phase

Enrolment
- Clearly define inclusion and exclusion criteria.
- Adhere to the trial protocol, and communicate and train your research personnel and collaborators accordingly.
- Design a patient-entry form with key data points to identify and keep track of patients assessed for eligibility criteria.
- Design a patient-entry form with key data points to identify and keep track of enrolled patients, and ensure that all eligibility criteria are met.
- Pretest patient-entry forms.
- Predefine in the protocol how to deal with protocol violations.
- Check eligibility criteria diligently for every patient.
- Ensure that patients who meet all of the inclusion criteria and none of the exclusion criteria are enrolled.

Randomization
- Ensure that the individual who generates the allocation scheme is not involved in ascertaining eligibility, administering treatment or assessing outcomes.
- Use an appropriate randomization technique such as computer-generated random sequence to ensure true random allocation. Avoid using birth year, day of the week or alternate chart numbers as methods of randomization.
- Consider block randomization with blocks of variable sizes, if using block randomization, to avoid imbalance in the numbers of patients assigned in each group.
- Consider stratified randomization by factors that will likely cause group imbalances to ensure that these factors are evenly distributed between study groups.

Concealment of allocation
- Ensure that individuals responsible for recruiting and allocating patients to treatments remain unaware of the next assignment in the sequence.
- Consider using an optimal method for proper allocation concealment (e.g., a 24-hour telephone randomization line, web-based randomization service or hospital pharmacy).

Concealment of allocation (continued)
- When using envelopes, ensure that envelopes are
  - dark and opaque,
  - sealed and serially numbered,
  - opened sequentially and only after the patient's name and other details are written on the envelope, and
  - kept in a locked and secure place.

Blinding
- Blind as many individuals as possible.
- If blinding was not feasible
  - ensure that people who are randomly assigning patients are independent from the study and not involved in patient care,
  - ensure that study groups, except for the intervention, are treated equally and a complete follow-up is attempted for all participating patients, and
  - ensure that 2 or more individuals independently assess outcomes and resolve the disagreements with consensus.

Follow-up
- Implement strategies to minimize losses to follow-up in study groups, particularly when there is no concomitant treatment after surgical intervention.
- Anticipate losses to follow-up and define how these losses will be accounted for at the stage of study design and data analysis:
  - adjust your sample size to account for the losses to follow-up to maintain a well-powered study, and
  - implement appropriate methods (e.g., multiple imputations) to handle the losses to follow-up at data analysis.

Intention-to-treat analysis
- Analyze patients in the groups to which they were randomly assigned, regardless of the treatment they actually received.
- Minimize conversion rate from one treatment to another in surgical trials:
  - ensure that every enrolled patient is equally eligible to receive either treatment, and
  - ensure that the participating surgeons have passed the learning curve stage for the novel procedure.

3. Sackett DL. The tactics of performing therapeutic trials. In: Hanyes RB, Sackett DL, Guyatt GH, et al., editors. *Clinical epidemiology: how to do clinical practice research*. 3rd ed. Philadelphia (PA): Lippincott Williams and Wilkins; 2006. p. 66-172.

4. Meinert CL, Tonascia S. Data collection consideration. In: *Clinical trials: design, conduct, and analysis*. New York (NY): Oxford University Press; 1986. p. 119-37.

5. Spilker B. Preparing data collection forms. In: *Guide to clinical trials*. New York (NY): Raven Press; 1991. p. 262-71.

6. Li P, Mah D, Lim K, et al. Randomization and concealment in surgical trials: a comparison between orthopaedic and non-orthopaedic randomized trials. *Arch Orthop Trauma Surg* 2005;125:70-2.

7. Schulz KF, Grimes DA. Generation of allocation sequences in randomised trials: chance, not choice. *Lancet* 2002;359:515-9.

8. Chalmers TC, Celano P, Sacks HS, et al. Bias in treatment assignment in controlled clinical trials. *N Engl J Med* 1983;309:1358-61.

9. Kang M, Ragan BG, Park JH. Issues in outcomes research: an overview of randomization techniques for clinical trials. *J Athl Train* 2008;43:215-21.

10. Lachin JM. Properties of simple randomization in clinical trials. *Control Clin Trials* 1988;9:312-26.

11. Friedman LM, Furberg CD, DeMets AL. *The randomization process: fundamentals of clinical trials*. 3rd ed. Ann Arbor (MI): Edwards Brothers Inc.; 1998.

12. Altman DG, Bland JM. How to randomise. *BMJ* 1999;319:703-4.

13. Altman DG, Schulz KF. Statistics notes: concealing treatment allocation in randomised trials. *BMJ* 2001;323:446-7.

14. Kernan WN, Viscoli CM, Makuch RW, et al. Stratified randomization for clinical trials. *J Clin Epidemiol* 1999;52:19-26.

15. Matts JP, McHugh RB. Analysis of accrual randomized clinical trials with balanced groups in strata. *J Chronic Dis* 1978;31:725-40.

16. Berger VW, Ivanova A, Knoll MD. Minimizing predictability while retaining balance through the use of less restrictive randomization procedures. *Stat Med* 2003; 22:3017-28.

17. Michaels JA, Brazier JE, Campbell WB, et al. Randomized clinical trial comparing surgery with conservative treatment for uncomplicated varicose veins. *Br J Surg* 2006;93:175-81.

18. Fewtrell MS, Kennedy K, Singhal A, et al. How much loss to follow-up is acceptable in long-term randomised trials and prospective studies? *Arch Dis Child* 2008;93:458-61.

19. Hedden SL, Woolson RF, Carter RE, et al. The impact of loss to follow-up on hypothesis tests of the treatment effect for several statistical methods in substance abuse clinical trials. *J Subst Abuse Treat* 2009;37:54-63.

20. Montori VM, Guyatt GH. Intention-to-treat principle. *CMAJ* 2001; 165:1339-41.

21. Schulz KF, Grimes DA. Sample size slippages in randomised trials: exclusions and the lost and wayward. *Lancet* 2002;359:781-5.