DIGITALCOMMONS @WAYNESTATE

Journal of Modern Applied Statistical Methods

Volume 2 | Issue 1 Article 8

5-1-2003

Randomization Technique, Allocation Concealment, Masking, And Susceptibility Of Trials To Selection Bias

Vance W. Berger
Biometry Research Group, National Cancer Institute, vb78c@nih.gov

Costas A. Christophi
George Washington University

Follow this and additional works at: http://digitalcommons.wayne.edu/jmasm

 Part of the Applied Statistics Commons, Social and Behavioral Sciences Commons, and the Statistical Theory Commons

Recommended Citation
Berger, Vance W. and Christophi, Costas A. (2003) "Randomization Technique, Allocation Concealment, Masking, And Susceptibility Of Trials To Selection Bias," *Journal of Modern Applied Statistical Methods*: Vol. 2 : Iss. 1 , Article 8.
DOI: 10.22237/jmasm/1051747680
Available at: http://digitalcommons.wayne.edu/jmasm/vol2/iss1/8

This Regular Article is brought to you for free and open access by the Open Access Journals at DigitalCommons@WayneState. It has been accepted for inclusion in Journal of Modern Applied Statistical Methods by an authorized editor of DigitalCommons@WayneState.

Journal of Modern Applied Statistical Methods
May, 2003, Vol. 2, No. 1, 80-86
Copyright © 2003 JMASM, Inc.
1538 – 9472/03/$30.00

Randomization Technique, Allocation Concealment, Masking, And Susceptibility Of Trials To Selection Bias

Vance W. Berger
National Cancer Institute
& University of Maryland-
Baltimore County

Costas A. Christophi
George Washington University

It is widely believed that baseline imbalances in randomized clinical trials must necessarily be random. Yet even among masked randomized trials conducted with allocation concealment, there are mechanisms by which patients with specific covariates may be selected for inclusion into a particular treatment group. This selection bias would force imbalance in those covariates, measured or unmeasured, that are used for the patient selection. Unfortunately, few trials provide adequate information to determine even if there was allocation concealment, how the randomization was conducted, and how successful the masking may have been, let alone if selection bias was adequately controlled. In this article we reinforce the message that allocation details should be presented in full. We also facilitate such reporting by identifying and clarifying the role of specific reportable design features. Because the designs that eliminate all selection bias are rarely feasible in practice, our development has important implications for not only the implementation, but also the reporting and interpretation, of randomized clinical trials.

Key words: Baseline imbalance, confounding, masking, randomized clinical trials, validity

Introduction

When lecturing on selection bias, we have addressed audience questions about how selection bias can occur in randomized clinical trials (RCTs). After all, it may be argued, if any subversion occurred, then the trial was not truly randomized. This statement implies that randomization confers absolute protection against any subversion, so that any covariate imbalances must be random. Similar abilities are often ascribed to allocation concealment or masking. Yet the effect of an action may differ from its objective; washed dishes, e.g., may remain dirty; cooked food may remain cold; and treated patients may remain sick.

Correspondence should be sent to Vance W. Berger, Ph.D., Biometry Research Group, National Cancer Institute, Executive Plaza North, Suite 3131, 6130 Executive Boulevard, MSC 7354, Bethesda, MD 20892-7354. Phone: (301) 435-5303. Fax: (301) 402-0816. E-mail: vb78c@nih.gov

It is in this light that we critically evaluate the ability of masking, allocation concealment, and randomization *as actually implemented* to produce treatment groups that differ only randomly. If they cannot do so, then observed covariate imbalances may be systematic, and may reflect selection bias. Observed treatment effects could then be attributable to biases, and not to the treatments themselves.

Selection bias can compromise the credibility of standard between-group comparisons, especially when the trial is conducted by a sponsor with a vested interest in the outcome (Hogel & Gaus, 1999). Yet details sufficient to assess the success of randomization, allocation concealment, and masking are rarely reported (Kyriakidi & Ioannidis, 2002).

This draws into question the reliability of the results of many RCTs that have been otherwise well conducted. In fact, if randomization is defined so as to eliminate the possibility of any subversion, then we question whether there has ever been a truly randomized trial. The irony is that until sufficient design details are routinely reported, it will be impossible to quantify the

extent to which selection bias actually occurs in RCTs, yet this lack of reporting is likely due to failure to appreciate the extent to which selection bias occurs in RCTs. Our development clarifies those details that should be presented in RCT reports. It is our hope that more RCT reports will provide these details, and test for selection bias explicitly (Berger & Exner, 1999).

What Are Randomization, Allocation Concealment, and Masking?

In a discussion of the distinction between a claim of masking and true masking, Oxtoby et al. (1989) pointed out that "the presumption that a plan to which one has aspired has come to fruition by virtue of aspiration alone is not science, and is particularly inapposite for a profession which should have a reputation for making clear distinctions between fantasy and reality". This profound remark highlights the distinction between an action and its effect. Masking may be defined as either the process (researchers not revealing treatment codes until the database is locked) or the result (complete ignorance of all trial participants as to which patients received which treatments). A masking claim indicates only the former; this may help to ensure the ignorance of some parties, but is unlikely to ensure the desired state of complete ignorance.

As the legal term "inevitable discovery" suggests, knowledge transfers by various mechanisms. It may be possible to fool all of the people some of the time, or some of the people all of the time, but it is not possible to fool all of the people all of the time. Just as a speed limit is a statement not about how fast drivers drive but rather about how fast they are *encouraged* to drive, so too is a policy of masking a statement not about who knew what (and when) but rather about a process.

Masking is often said to be possible only some of the time, while allocation concealment (Schulz, 1995a,b; 1996), which is essentially the masking of each allocation just until it is executed, is always possible. This confusion of the two definitions is a double-standard. If masking is possible only some of the time, then clearly reference is being made to the result, and not the process.

To be fair, then, one would have to ask if the *result* of allocation concealment is always possible. Sealed envelopes have been held to lights, phantom patients have been enrolled, and locked files have been raided to determine upcoming treatment allocations in successful subversions of allocation concealment (Schulz, 1995a). Also, it may be clear what a given patient would receive, if enrolled, if cluster randomization (Jordhoy et al., 2002) or minimization (Pocock & Simon, 1975) is used. Drug bottle numbers can also lead to prediction (Kuznetsova, 2002). So only the *process* of allocation concealment, but not its result, can be ensured. Without the result of allocation concealment, selection bias remains a concern.

Mechanisms for Selection Bias, and Specific Countermeasures

To focus ideas, we confine our attention to selection bias that interferes with internal validity (a fair comparison, Mark, 1997); we do not consider external validity. Groups of patients to be compared may differ in important ways even before any intervention is applied (Prorok, Hankes, & Bundy, 1981). These baseline imbalances cannot be attributed to the interventions, but they can interfere with and overwhelm the comparison of the interventions (Green & Byar, 1984).

If treatments are independent of patient characteristics, then any baseline imbalances (even if statistically significant) are due to chance variation only. This is one reason often cited for using randomization.

On the other hand, a systematic explanation for the imbalances, known or unknown, would constitute selection bias, even if the imbalances are not statistically significant, or even readily observed (Berger & Exner, 1999). We present a sequence of mechanisms by which selection bias may occur, starting with observational studies in Section A, and such countermeasures as randomization, allocation concealment, and masking (see Table 1).

Table 1: What to Report in Randomized Clinical Trials To Control Selection Bias

Concern	Report
Differential Allocation Discretion	Planned allocation proportions Number of screened and randomized patients by the group to which they were or would have been randomized had they been randomized
Deferred Enrollment	List patients who were screened twice or more, or that there were none
Allocation Concealment	Specific means of concealing the future allocations
Predicted Allocations	Specific restrictions on the randomization (including block sizes) Specific methods of concealing the past allocations (masking) Evidence of unmasking (including differential rates of observable adverse events, any emergencies requiring intentional unmasking, and rates of correct treatment group guesses at de-briefing)
Baseline Imbalances	Compare baseline covariates across treatment groups
Selection Bias	Graph key covariates against P{active}, as in Berger and Exner (1999) Graph response against P{active} within each treatment group, per Berger and Exner (1999). List stratification errors (if any), or that there were none

A. Selection Bias in Observational Studies or with Consumer Randomization

Investigators may assign treatments based on patient characteristics (Green & Byar, 1984; Rubin, 1977). Patients may select either their treatment or, with consumer randomization (Bird, 2001), their randomization probability, at least from among a given set of choices. Allocation discretion may be available to the patient, the investigator, both, or neither (dictated allocation). Those patients selecting one treatment or probability may differ systematically from those selecting another (Green & Byar, 1984), so dictated allocation (no freedom of choice) is a countermeasure to prevent patient characteristics from influencing the allocation sequence through either overt treatment assignment based on patient characteristics or self-selection.

B. Selection Bias with Dictated Allocation

If allocation is alternated, then either patients with even accession numbers or patients with odd accession numbers receive the active treatment. The others receive the control. This dictated allocation would prevent the type of selection bias considered in Section A. But with sequential accrual, knowledge of the upcoming treatment, and enrollment discretion (Chalmers, 1990), an investigator could deny enrollment to patients lacking the characteristics that would make them "suitable" to receive the upcoming treatment (Schulz, 1995a; Schulz & Grimes, 2002a).

The selection bias enabled by the predictable allocation sequence (Schulz & Grimes, 2002b) can be controlled by creating instead an unpredictable allocation sequence, or randomizing (Rosenberger & Lachin, 2002). The second countermeasure is the use of actual (not virtual, quasi-, or pseudo-) randomization (Berger & Bears, 2003) to prepare the allocation sequence.

C. Selection Bias with Dictated Allocation and Randomization

Urn randomization (Wei & Lachin, 1988) is conducted by tossing a (possibly biased) coin each time a patient is to be allocated. Heads indicates active treatment, and tails indicates control. There is no actual allocation discretion, yet having screened and evaluated a given patient, the investigator might exercise *de facto* allocation discretion to reject the toss and repeat until the preferred allocation is observed.

Another mechanism for selection bias with dictated allocation and randomization would be possible if minimization, or dynamic randomization (Pocock & Simon, 1975), were used to force balance with respect to certain covariates. The allocation is determined by minimizing an imbalance function, and randomization may be used to break the ties. So there is both dictated allocation and randomization. Yet because most allocations will be deterministic, it would be possible to determine the allocation to be made once a patient has been identified. A patient enrollment decision may be based on a combination of the treatment to be assigned and values of observed covariates that were not used to define the imbalance function. Randomization is *conventional* if the allocation sequence is generated in advance of screening any patients, and *unconventional* otherwise. Conventional randomization prevents the types of selection bias discussed in this section, and is our third countermeasure.

D. Selection Bias with Dictated Allocation and Conventional Randomization

As in Section B, selection bias may result from enrollment discretion and advance knowledge of the allocation sequence; the latter may be facilitated by conventional randomization, as the allocation sequence may be posted publicly before patients are screened (Schulz & Grimes, 2002a). A countermeasure to eliminate this advance knowledge is that each allocation be determined only after the patient to be enrolled is identified (Clarke, 2002), as occurs with minimization (Pocock & Simon, 1975). Either the allocation to be made or the patient to be enrolled has to be selected first; whichever it is may influence the other, and the biases possible with unconventional randomization (Section C) are at least as serious as the biases possible with conventional randomization.

Unconventional randomization may not be able to eliminate advance knowledge of patient characteristics, but one might hope to eliminate advance knowledge of the allocation sequence with conventional randomization and the fourth countermeasure, allocation concealment, which is often taken to mean precisely this lack of advance knowledge. But recall that allocation concealment signifies only that the allocation codes are not intentionally revealed. Even with steps to ensure that these codes cannot be observed, e.g. by holding an envelope to a light (Schulz, 1995a,b), it is not possible to enumerate, and rule out, all mechanisms by which allocations can be observed. We are not prepared to take the success of allocation concealment on faith in an actual trial; we do so for the purpose of this article to demonstrate that even in this unrealistically optimistic case, subversion is still possible.

E. Selection Bias with (D) and Allocation Concealment

In a randomized depression study of nurse telehealth care (Hunkeler et al., 2000), the initial 40:60 randomization to two groups later became 40:20 to those same two groups, with the remaining 40% allocated to a new third group. If the change in allocation proportions was planned (which need not be the case; see Lippman et al., 2001), then even with allocation concealment it may still be possible to *predict* (but not observe) future allocations. Knowing that more late patients than early patients would be allocated to the third group constitutes advance knowledge of the allocations which, though imperfect, allows for deferred enrollment (Schulz, 1996) of those subjects most "suitable" for the third group until after the new proportions took effect. The fifth countermeasure, then, is the fixed allocation proportions that prevent this.

F. Selection Bias with (E) and Fixed Allocation Proportions

Randomization is *unrestricted* (Schulz & Grimes, 2002b) if a patient's likelihood of receiving either treatment is independent of all previous allocations, and is *restricted* (ter Riet & Kessels, 1995) otherwise. The random allocation rule (Schulz & Grimes, 2002b), in which both

treatment groups must be assigned equally often, is one form of restricted randomization, as the final allocation would be determined by the prior ones. Even with allocation concealment and fixed allocation proportions, patterns created by restrictions on the randomization allow prediction of the allocation sequence. Berger and Exner (1999) quantified this extent of advance knowledge with the probability, P{active}, of a given patient being allocated to the active group given the previous allocations.

With 1:1 allocation, P{active}=0.5 for the first patient; with alternation (Section B), P{active} is always either 0 or 1. Note that P{active} reflects the restrictions on the allocation sequences, and becomes a patient characteristic only after that patient is randomized. With enrollment discretion, P{active} may be used, in conjunction with the estimated potential outcomes of each patient to each treatment, say **Y**={Y(A),Y(C)} for the active and control treatments, respectively, as a basis for enrollment decisions.

Gender, age, race, pre-existing medical conditions, or other baseline characteristics may be considered in deriving the value of **Y** for a given patient. Based on **Y**, the investigator might select a range of P{active} values for which the patient would be enrolled. If the P{active} value at the time this patient is screened happens to fall outside of this patient's P{active} range, then the patient will be denied enrollment, and another patient will be screened. Only when a patient is found with a P{active} range to match the actual P{active} value will the patient be enrolled.

Selection bias occurs if the P{active} range is restricted based on **Y**. It would be possible, e.g., to enroll patients only if P{active} and **Y** are both large (suppose that larger **Y** values indicate better responses) or both small, but not if they are discordant (Schulz, 1995a). This possibility is depicted in Table 2, using randomized blocks of size four to calculate P{active} (Berger & Exner, 1999). Notice that not only does treatment assignment for randomized patients depend upon the allocation sequence, but in fact Patients #S5, #S7, #S9, and #S10 may or may not be randomized depending on the allocation sequence, and Patient #S3 cannot get the control.

Discussion

Few RCT reports make any effort to address the potential for selection bias. Presumably, this is due to unrealistically optimistic definitions of randomization, allocation concealment, and masking. Unfortunately, even in combination, these design features *as implemented* cannot eliminate selection bias. One may argue that while selection bias is possible in theory, its mechanisms are implausible, especially when the main analyses have low p-values.

Unfortunately, history has demonstrated the fallibility of the plausibility test; at best low p-values rule out (probabilistically) chance events, but they do not rule out biases (Berger, 2000; Berger et al., 2000; Grimes and Schulz, 2002). Because of the one-sponsor problem (Hogel & Gaus, 1999) and the vested interest the one sponsor usually has in the outcome of the trial, the best way to offer a convincing argument that a trial was free of a certain bias is to eliminate the possibility of its occurrence. Hence, the burden needs to be on the researchers to demonstrate the reliability of their results. In this article we have presented a number of countermeasures, few combinations of which would eliminate the potential for selection bias. In most cases, then, it is unrealistically optimistic to believe that RCTs are insulated from severe bias (Schulz, 1996).

We are hopeful that the information presented in Table 1 will accompany reports of future trials, preferably in the text of the article, but possibly in an accompanying web site. Such transparency would enable readers to determine the extent to which various mechanisms for selection bias were possible in a given trial, and the extent to which it appears as though there actually was selection bias. The refined measures of trial quality could be used in determining the extent to which specific trials influence policy and meta-analyses. This would exert pressure on those who design trials to design better trials. We are hopeful that journal editors, regulators, and granting institutions will rely, in part, on this information to make their important decisions.

Table 2: Selection Bias with Randomization and Allocation Concealment.

		{(A C C A); (C C A A)}		{(A C A C); (C A A C)}	
S	P{active} Range*	P{active}	Randomized	P{active}	Randomized
S1	[0.50,1.00]	0.50	Active	0.50	Active
S2	[0.00,0.33]	0.33	Control	0.33	Control
S3	[1.00,1.00]	0.50	-	0.50	-
S4	[0.00,0.50]	0.50	Control	0.50	Active
S5	[0.50,1.00]	1.00	Active	0.00	-
S6	[0.00,0.50]	0.50	Control	0.00	Control
S7	[0.00,0.50]	0.67	-	0.50	Control
S8	[0.67,1.00]	0.67	Control	0.67	Active
S9	[0.67,1.00]	1.00	Active	0.50	-
S10	[0.00,0.50]	1.00	-	0.50	Active
S11	[0.33.0.67]	1.00	-	0.00	-
S12	[0.00,1.00]	1.00	Active	0.00	Control

*The range of P{active} values for which the patient gets randomized. P{active} computed according to the formula of Berger and Exner [3] using the randomized block procedure with a fixed block size of four. Not only does treatment assignment for randomized patients depend upon the allocation sequence, but in fact Patients #S5, #S7, #S9, and #S10 may or may not be randomized depending on the allocation sequence, and Patient #S3 cannot get the control.

References

Berger, V. W. (2000). Pros and cons of permutation tests in clinical trials. *Statistics in Medicine*, *19*, 1319-1328.

Berger, V. W., & Bears, J. (2003). When can a clinical trial be called 'randomized'? *Vaccine*, *21*, 468-472.

Berger, V. W, & Exner, D. V. (1999). Detecting selection bias in randomized clinical trials. *Controlled Clinical Trials*, *20*, 319-327.

Berger, V. W., Lunneborg, C., Ernst, M. D., & Levine, J.G. (2002). Parametric analyses in randomized clinical trials, *Journal of Modern Applied Statistical Methods*, *1*(1), 74-82.

Bird, S. M. (2001). Dissemination of decisions on interim analyses needs wider debate. *BMJ*, 323, 1424.

Chalmers, T. C. (1990). Discussion of biostatistical collaboration in medical research by Jonas H. Ellenberg. *Biometrics*, 46, 20-22.

Clarke, M. (2002). Last moment randomization and concealment. *British Medical Journal*, *323*, http://bmj.com/cgi/eletters/323/7310/446/

Day, S. (1998). Blinding or masking in the encyclopedia of biostatistics. P. Armitage & T. Colton (Eds.), Vol. *1*. Chichester: Johns Wiley and Sons, 410-417.

Green, S. B, & Byar, D. B. (1984). Using observational data from registries to compare treatments: The fallacy of omnimetrics. *Statistics in Medicine*, 3, 361-370.

Grimes, D. A, & Schulz, K. F. (2001). Randomized controlled trials in *contraception*: The need for "consort" guidelines. *Contraception*, *64*, 139-142.

Grimes, D. A, & Schulz, K. F. (2002). Bias and causal associations in observational research. *Lancet*, 359, 248-252.

Hogel, J., & Gaus, W. (1999). The procedure of new drug application and the philosophy of critical rationalism or the limits of quality assurance with good clinical practice. *Controlled Clinical Trials, 20*, 511-518.

Hunkeler, E. M., Meresman, J. F., & Hargreaves, W. A. (2000). Efficacy of nurse telehealth care and peer support in augmenting treatment of depression in primary care. *Archives of Family Medicine, 9*(8), 700-708.

Jordhoy, M. S., Fayers, P. M., Ahlner-Elmqvist, M., & Kaasa, S. (2002) Lack of concealment may lead to selection bias in cluster randomized trials of palliative care. *Palliative Medicine, 16*, 43-49.

Kuznetsova, O. M. (2002). Why permutation is even more important in ivrs drug codes schedule generation than in patient randomization schedule generation. *Controlled Clinical Trials, 22*, 69-71.

Kyriakidi, M., & Ioannidis, J. P. A. (2002). Design and quality considerations for randomized controlled trials in systematic sclerosis. *Arthritis Care and Research, 47*, 73-81.

Lippman, S. M., Lee, J. J., & Kurp, D. D. (2001). Randomized phase iii intergroup trial of isotretinoin to prevent second primary tumors in stage i non-small-cell lung cancer. *Journal of the National Cancer Institute, 93*, 605-618.

Mark, D. H. (1997). Interpreting the term selection bias in medical research. *Family Medicine, 29*(2), 132-136.

Oxtoby, A., Jones, A., & Robinson, M. (1989). Is your 'double-blind' design truly double-blind? *British Journal of Psychiatry, 155*, 700-701.

Pocock, S. J., & Simon, R. (1975). Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. *Biometrics, 31*, 103-115.

Prorok, P. C., Hankes, B. F., & Bundy, B. N. (1981). Concepts and problems in the evaluation of screening programs. *Journal of Chronic Diseases, 34*, 159-171.

Rosenberger, W., & Lachin, J. M. (2002). Randomization in clinical trials: Theory and practice. NY: John Wiley and Sons.

Rubin, D. B. (1977). Assignment to treatment group on the basis of a covariate. *Journal of Educational Statistics, 2*(1), 1-26.

Schulz, K. F. (1995a). Subverting randomization in controlled trials. *JAMA, 274*, 1456-1458.

Schulz, K. F. (1995b). Unbiased research and the human spirit: The challenges of randomized controlled trials. *Canadian Medical Association Journal, 153*, 783-786.

Schulz, K. F. (1996). Randomised Trials, Human Nature, and Reporting Guidelines. *Lancet*, 348, 596-598.

Schulz, K. F., & Grimes, D. A. (2002a). Allocation concealment in randomized trials: defending against deciphering. *Lancet, 359*, 614-618.

Schulz, K. F., & Grimes, D. A. (2002b). Generation of allocation sequences in randomized trials: chance, not choice. *Lancet, 359*, 515-519.

ter Riet, G., & Kessels, A. G. H. (1995). Restricted randomization in randomized controlled trials. *JAMA, 274*, 1835.

Wei, L. J., & Lachin, J. M. (1988). Properties of the urn randomization in clinical trials. *Controlled Clinical Trials, 9*, 345-364.