SPECIAL REPORT

The Quest for Unbiased Research: Randomized Clinical Trials and the CONSORT Reporting Guidelines

Kenneth F. Schulz, PhD, MBA

Editor's Note

The large and increasing number of randomized clinical trials in all of medicine has prompted the introduction of guidelines that are intended to improve the quality of the research and of the published papers derived from such investigations. These guidelines for reporting clinical trials, the Consolidated Standards of Reporting Trials (the CONSORT statement) were set forth by a group of journal editors, statisticians, biomedical investigators, and editors. The CONSORT statement has been published in several major medical journals and is likely to have a major impact on editorial policies as well as the design of such investigations. I have invited Dr Kenneth F. Schulz of the Centers for Disease Control and Prevention, a leading investigator in the evaluation of clinical trials and a participant in the design of the CONSORT guidelines, to prepare this brief summary for the *Annals*. The CONSORT statement is expected to have a substantial effect on the peer review of clinical trials, and it promises to elevate the quality of such investigations. Future submissions to the *Annals* will be evaluated in the context of these guidelines.

Robert A. Fishman, MD
Editor

Assessing treatments can be misleading unless investigators ensure unbiased comparison groups. Random allocation to those groups remains the only method that eliminates selection and confounding biases. Thus, randomized controlled trials (RCTs) serve as the foundation for advancing medical science. Deservedly, medical researchers consider them the "gold standard."

The design and implementation of an RCT, however, takes more knowledge and effort than investigators seem to fathom. They often encounter unanticipated, and sometimes unnoticed, difficulties. In particular, trial implementers sometimes subvert the intended aims of randomization [1–3]. While proper attention to design issues prevents or deflects attempts at subversion, investigators often fail to attain that requisite level of attention.

The poor reporting of RCTs reflects an inadequate attention to methodology. An article published in 1990 detailed the poor reporting of RCTs in the premier general medical journals in the world [4]. That article spawned another analysis that yielded similar results for specialty journals [5]. Regrettably, the most recent analysis indicates persistent poor reporting [6]. These and other recent reports [7, 8] buttressed earlier analyses that found inadequate reporting [9–11]. Justifiably, the medical community frets that shoddy reporting reflects shoddy methods, and that with shoddy methods come biased results.

With the aim of directly addressing shoddy reporting, and with the hope of also dealing with shoddy methods, a group recently published reporting guidelines for RCTs, the Consolidated Standards of Reporting Trials (CONSORT) statement [12]. In this special report, I discuss the important elements of randomization, review the recent research on bias and allocation methods, refer to some instances in which investigators have provided anonymous accounts of subverting randomization, and address recent findings on bias and double-blinding and exclusions after entry. I then introduce and discuss the CONSORT guidelines.

Important Elements of Randomization

The success of randomization depends on two interrelated processes [1, 13]. The first process entails generating a sequence by which the participants in a trial are allocated to treatment and control groups. To ensure unpredictability of that allocation sequence, investigators should generate it by a random process. The sec-

From the Division of Sexually Transmitted Disease Prevention, National Center for HIV, STD, and TB Prevention, Centers for Disease Control and Prevention, Atlanta, GA.

Received Jan 2, 1997, and in revised form Feb 6. Accepted for publication Feb 6, 1997.

Address correspondence to Dr Schulz, Division of Sexually Transmitted Disease Prevention, National Center for HIV, STD, and TB Prevention, MS-E02, Centers for Disease Control and Prevention, 1600 Clifton Road, Atlanta, GA 30333.

ond process involves meticulous implementation of that allocation sequence through an assignment mechanism (allocation concealment process) that prevents foreknowledge of the treatment assignments to those implementing the allocation sequence [5]. Many medical researchers mistakenly regard the process of generating an allocation sequence as "randomization" and overlook the process of allocation concealment [1, 2]. Without adequate allocation concealment, however, even random, unpredictable assignment sequences can be subverted [13, 14]. For example, suppose that an investigator creates an adequate allocation sequence using a random number table. However, the investigator then affixes the list of that sequence to a bulletin board, thereby essentially providing no allocation concealment. Those responsible for admitting participants could ascertain the upcoming treatment allocations and then route participants with better prognoses to the experimental group and those with poorer prognoses to the control group, or vice versa. Bias could easily be introduced. Crucially, allocation concealment shields those who admit participants to a trial from knowing the upcoming assignments [1, 2]. The decision to accept or reject a participant must be made and informed consent obtained without knowledge of the treatment to be assigned [15].

Investigators must not confuse blinding with allocation concealment. Substantial differences between the two exist. In particular, allocation concealment focuses on preventing selection and confounding biases, safeguards the assignment sequence *before and until* allocation, and can *always* be successfully implemented [1, 3]. In contrast, blinding focuses on preventing ascertainment bias, safeguards the sequence *after* allocation, and cannot always be implemented [1, 3].

Recent Research on Bias and Allocation Methods

As with all studies, bias jeopardizes even RCTs if investigators improperly execute them. For example, in a study of 250 controlled trials from 33 metaanalyses, we found that alleged RCTs with inadequate allocation concealment yielded larger estimates of treatment effects than did trials in which the authors reported adequate concealment [13]. Odds ratios were exaggerated, on average, by 30 to 40%. Furthermore, two other recent analyses in different subject matter areas also revealed exaggerated treatment effects related to inadequate allocation concealment [16, 17]. These exaggerated estimates of treatment effects likely reflect methodological problems. They also reveal meaningful levels of bias. If a medical researcher aspires to find a decrease in mortality of 35% from a particular treatment, certainly potential biases of 30 to 40% would overwhelm estimates of the treatment effect. Indeed, the elimination of bias becomes especially crucial in trials designed to detect moderate treatment effects.

Randomization and the Human Spirit

In the previous section, I referred to studies that discovered empirical evidence of bias. In practice, do investigators actually subvert the intended purpose of randomization? Have they divulged the details of deciphering allocation sequences? While that has happened [18], documented accounts are rare. Nevertheless, when physicians responded to queries during epidemiological workshops, more than half related instances of deciphering an allocation concealment scheme [1]. This should not be interpreted as representing more than half of all the trials [1]; indeed, most published RCTs probably estimate treatment effects reliably. Allocation breaches, however, seem to be something more frequent than a rare occurrence.

The personal accounts of those decipherings run the gamut from simple to intricate operations [1]. One frequently mentioned approach, the posting of an allocation sequence on a bulletin board, takes little effort to decipher. Other examples of simple deciphering operations include opening unsealed assignment envelopes and holding translucent envelopes up to a regular lightbulb. Workshop participants also related more involved efforts. For example, a few took sequentially numbered, opaque, sealed envelopes to the "hot light" (an intense incandescent bulb) in the radiology department for deciphering of the assignment scheme. Another admitted to searching for code in the office files of the principal investigator. The stories frequently reflected ingenious efforts.

The stories also frequently reflected naive, dubious judgment [1–3]. Assignment manipulations by those in trials reveal larger underlying conflicts. The need for unbiased research struggles with the inherently biased sources of that information—human beings. While investigators understand the need for unbiased research in the abstract, they may have difficulty maintaining an impartial posture once they are engaged in a trial. They may want certain patients to benefit from one of the treatments that they believe to be better or may want the results of a study to show what they believe to be valid [2, 3]. Aspects of properly conducted RCTs, then, annoy investigators, because trial procedures attempt to impede human inclinations. Unfortunately, RCTs are an anathema to the human spirit [1–3].

For those conducting a RCT with inadequate allocation concealment, the challenge of deciphering the allocation scheme may become too great a temptation to resist. Those who thwart assignment mechanisms do not necessarily have corrupt motives. Frequently, they are unaware of the scientific ramifications of their actions [1]. Whether their motives reflect innocent or

pernicious intents, however, their actions undermine the validity of a trial.

Given the human ingredients in RCTs, we must realize that trial implementers, if provided the opportunity, sometimes subvert randomization. Fortunately, trialists can usually prevent subversions with diligent attention to design and implementation issues [1–3]. Trial designers and implementers should adopt the methods and standards of proper randomization and understand the rationale behind them [1, 3].

Bias and Nonallocation Methodological Elements

Methodological elements in RCTs other than proper allocation also reduce bias. We found that trials that were not double-blinded yielded larger estimates of treatment effects compared with trials in which authors reported double-blinding (odds ratios exaggerated, on average, by 17%) [13]. While double-blinding appears to prevent bias, its effect appears weaker than allocation concealment. Another recent analysis also denoted the importance of double-blinding [17].

Absence of bias persists throughout a trial only if the analysis includes all properly randomized participants in the originally assigned groups. Thus, we had expected to find some evidence of bias in those trials that reported having excluded participants after randomization. However, those trials did not yield exaggerated estimates of treatment effects compared with trials in which the reports gave the impression of no exclusions [13]. Some of the trials with reports that gave the impression of no exclusions probably had exclusions [13].

Reporting and CONSORT

While investigators must execute RCTs properly to minimize bias, they also must communicate those efforts to the reader. Without proper reporting, readers cannot discern the trials with valid results from those with questionable results. The use of proper methodology should be transparent to editors, reviewers, and readers. They should not have to assume or guess the methods employed; they should be informed explicitly. Yet assessments of the reporting quality of published trials have consistently found major flaws [4–11].

Violations of randomization probably happen more frequently than suspected. Only 9% of reports in the specialist journals and 15% in the general journals reported both an adequate method of generating random sequences and an adequate method of allocation concealment [4, 5, 19]. Of the reports on studies considered double-blind, only 45% described similarity of the treatment and control regimens and only 26% provided information on the protection of the allocation schedule [20]. Investigators reported testing the efficacy of blinding in only 2 of 31 trials, and found substantial unblinding of assignments in both [20].

Perhaps the most dismal reporting stems from authors neglecting to address, or incompletely addressing, exclusions of participants after randomization. Two studies found that trials with no apparent reported exclusions may in fact have made exclusions during the trial but ignored them in the report [13, 20]. Moreover, two trials with documented exclusions have published reports indicating no exclusions [21]. Readers of RCT reports should warily regard the potentially misleading information provided. Published information on exclusions may frequently provide a misleading impression of trial quality [20].

To address the prevailing flaws in reporting, two groups independently met and separately published guidelines for the reporting of RCTs [22, 23]. A subsequent editorial suggested that the two groups meet to generate unified guidelines [24]. This flurry of activity culminated in the CONSORT statement [12]. It includes a checklist of 21 items (Table) that pertain to a RCT report. Those items reflect the fundamental information necessary to evaluate accurately the internal and external validity of a trial. The CONSORT statement also includes a flow diagram (Fig) that depicts the progress of participants throughout a two-group parallel-design RCT, the type of trial most commonly reported. Modifications will need to be made in reports of trials with a larger number of groups or trials using a different design [12].

While this guideline avalanche may seem precipitous, it really is long overdue [3]. Instigated by the pioneering work of A. Bradford Hill, the *British Medical Journal (BMJ)* published the first randomized trial almost 50 years ago [25]. For all these years, calls for improved reporting published in medical and statistical journals have languished on library shelves.

The CONSORT statement, however, represents a change. A number of high quality and high profile journals are adopting the guidelines, including *JAMA, Lancet,* and *BMJ.* The journals encourage authors to follow the statement immediately, but from January 1, 1997, they will be required to do so. After almost 50 years, the age of adequate reporting of RCTs approaches.

With such a long legacy of poor reporting, can the research community hope that these guidelines will produce a positive impact? While surely many factors will affect events, at least three factors justify hope. First, the greater number of methodologically astute journal editors constitutes perhaps the most substantial change. The editors contributing to the CONSORT guidelines portray a symptom of that trend. They represent a large number of editors worldwide who share similar views. Second, empirical evidence indicates that many improperly implemented RCTs may indeed be

Table. Consolidation of Standards for Reporting Trials—CONSORT

Heading	Subheading	Descriptor	Was It Reported?	On What Page No.?
Title		Identify the study as a randomized trial.		
Abstract		Use a structured format.		
Introduction		State prospectively defined hypothesis, clinical objectives, and planned subgroup or covariate analyses.		
Methods				
	Protocol	Describe Planned study population, together with inclusion/exclusion criteria. Planned interventions and their timing. Primary and secondary outcome measure(s) and the minimum important difference(s), and indicate how the target sample size was projected. Rationale and methods for statistical analyses, detailing main comparative analyses and whether they were completed on an intention-to-treat basis. Prospectively defined stopping rules (if warranted).		
	Assignment	Describe Unit of randomization (e.g., individual, cluster, geographic). Method used to generate the allocation schedule. Method of allocation concealment and timing of assignment. Method to separate the generator from the executor of assignment.		
	Masking (Blinding)	Describe mechanism (e.g., capsules, tablets); similarity of treatment characteristics (e.g., appearance, taste); allocation schedule control (location of code during trial and when broken); and evidence for successful blinding among participants, person doing intervention, outcome assessors, and data analysts.		
Results				
	Participant Flow and Follow-up	Provide a trial profile (see Fig) summarizing participant flow, numbers and timing of randomization assignment, interventions, and measurements for each randomized group.		
	Analysis	State estimated effect of intervention on primary and secondary outcome measures, including a point estimate and measure of precision (confidence interval). State results in absolute numbers when feasible (e.g., 10/20, not 50%). Present summary data and appropriate descriptive and inferential statistics in sufficient detail to permit alternative analyses and replication. Describe prognostic variables by treatment group and any attempt to adjust for them. Describe protocol deviations from the study as planned, together with the reasons.		
Comment		State specific interpretation of study findings, including sources of bias and imprecision (internal validity) and discussion of external validity, including appropriate quantitative measures when possible. State general interpretation of the data in light of the totality of the available evidence.		

biased. Journal editors and readers of scientific reports could understandably be less concerned with poor reporting if they believed that poor reporting did not indicate biased results. However, with evidence indicating an association between poor reporting and exaggerated estimates of treatment effects [13, 16, 17], they

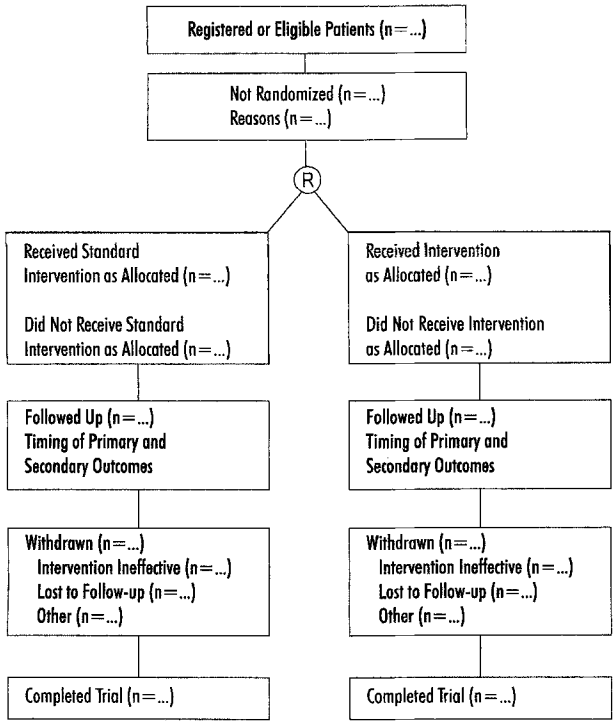

Fig. Progress through the various stages of a trial, including flow of participants, withdrawals, and timing of primary and secondary outcome measures. R = randomization. (Reproduced with permission from Begg C, Cho M, Eastwood S, et al. Improving the quality of reporting and randomized controlled trials: the CONSORT statement. JAMA 1996;276:638; copyright 1996, the American Medical Association.)

will be less comfortable simply presuming insulation from bias. Third, journals adopting the CONSORT guidelines do not merely suggest cooperation. They mandate adherence to the guidelines as a precondition for publication.

References

1. Schultz KF. Subverting randomization in controlled trials. JAMA 1995;274:1456–1458
2. Schulz KF. Unbiased research and the human spirit: the challenges of randomized controlled trials. Can Med Assoc J 1995; 153:783–786
3. Schulz KF. Randomized trials, human nature, and reporting guidelines. Lancet 1996;348:596–598
4. Altman DG, Doré CJ. Randomization and baseline comparisons in clinical trials. Lancet 1990;335:149–153
5. Schulz KF, Chalmers I, Grimes DA, Altman DG. Assessing the quality of randomization from reports of controlled trials published in obstetrics and gynecology journals. JAMA 1994;272: 125–128
6. Moher D, Fortin P, Jadad AR, et al. Assessing the completeness of reporting of trials published in languages other than English: implications for the conduct and reporting of systematic reviews. Lancet 1996;347:363–366
7. Williams DH, Davis CE. Reporting of assignment methods in clinical trials. Control Clin Trials 1994;15:294–298
8. Sonis J, Joines J. The quality of clinical trials published in *The Journal of Family Practice.* J Fam Pract 1994;39:225–235
9. Mosteller F, Gilbert JP, McPeek B. Reporting standards and research strategies for controlled trials: agenda for the editor. Control Clin Trials 1980;1:37–58
10. DerSimonian R, Charette LJ, McPeek B, Mosteller F. Reporting on methods in clinical trials. N Engl J Med 1982;306: 1332–1337
11. Tyson JE, Furzan JA, Reisch JS, Mize SG. An evaluation of the quality of therapeutic studies in perinatal medicine. J Pediatr 1983;102:10–13
12. Begg C, Cho M, Eastwood S, et al. Improving the quality of reporting of randomized controlled trials: the CONSORT statement. JAMA 1996;276:637–639
13. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. JAMA 1995;273:408–412
14. Pocock SJ. Statistical aspects of clinical trial design. Statistician 1982;31:1–18
15. Chalmers TC, Levin H, Sacks HS, et al. Meta-analysis of clinical trials as a scientific discipline. I: control of bias and comparison with large co-operative trials. Stat Med 1987;6:315–325
16. Khan KS, Daya S, Jadad AR. The importance of quality of primary studies in producing unbiased systematic reviews. Arch Intern Med 1996;156:661–666
17. Khan KS, Daya S, Collins JA, Walter SD. Empirical evidence of bias in infertility research: overestimation of treatment effect in crossover trials using pregnancy as the outcome measure. Fertil Steril 1996;65:939–945
18. Carleton RA, Sanders CA, Burack WR. Heparin administration after acute myocardial infarction. N Engl J Med 1960;263: 1002–1004
19. Schulz KF, Chalmers I, Altman DG, et al. The methodologic quality of randomization as assessed from reports of trials in specialist and general medical journals. Online J Curr Clin Trials [serial online] 1995 August 26;4(Doc No 197)
20. Schulz KF, Grimes DA, Altman DG, Hayes RJ. Blinding and exclusions after allocation in randomised controlled trials: survey of published parallel group trials in obstetrics and gynaecology. BMJ 1996;312:742–744
21. Gøtzsche PC. Multiple publication of reports of drug trials. Eur J Clin Pharmacol 1989;36:429–432
22. The Standards of Reporting Trials Group. A proposal for structured reporting of randomized controlled trials. JAMA 1994; 272:1926–1931
23. Working Group on Recommendations for Reporting of Clinical Trials in the Biomedical Literature. Call for comments on a proposal to improve reporting of clinical trials in the biomedical literature. Ann Intern Med 1994;121:894–895
24. Rennie D. Reporting randomized controlled trials: an experiment and a call for responses from readers. JAMA 1995;273: 1054–1055
25. Medical Research Council. Streptomycin treatment of pulmonary tuberculosis. BMJ 1948;2:769–782