Department of Health, Education, and Welfare, 1979:109-15. (NIH publication no. 79-1610.)
17. Goldman L, Cook F, Hashimoto B, Stone P, Muller J, Loscalzo A. Evidence that hospital care for acute myocardial infarction has not contributed to the decline in coronary mortality between 1973-1974 and 1978-1979. Circulation 1982; 65:936-42.
18. Goldberg R, Szklo M, Tonascia JA, Kennedy HL. Time trends in prognosis of patients with myocardial infarction: a population-based study. Johns Hopkins Med J 1979; 144:73-80.
19. Goldberg RJ, Kennedy HL, Gore JM. Absence of secular changes in the prognosis of patients with an initial myocardial infarction. Clin Cardiol 1982; 5:469-75.
20. Weinblatt E, Goldberg JD, Ruberman W, Frank CW, Monk MA, Chaudhary BS. Mortality after first myocadial infarction: search for a secular trend. JAMA 1982; 237:1576-81.
21. Gillum RF, Blackburn H, Feinleib M. Current strategies for explaining the decline in ischemic heart disease mortality. J Chronic Dis 1982; 35:467-74.

SPECIAL ARTICLE

BIAS IN TREATMENT ASSIGNMENT IN CONTROLLED CLINICAL TRIALS

THOMAS C. CHALMERS, M.D., PAUL CELANO, M.D., HENRY S. SACKS, PH.D., M.D., AND HARRY SMITH, JR., PH.D.

Abstract Controlled clinical trials of the treatment of acute myocardial infarction offer a unique opportunity for the study of the potential influence on outcome of bias in treatment assignment. A group of 145 papers was divided into those in which the randomization process was blinded (57 papers), those in which it may have been unblinded (45 papers), and those in which the controls were selected by a nonrandom process (43 papers). At least one prognostic variable was maldistributed ($P<0.05$) in 14.0 per cent of the blinded-randomization studies, in 26.7 per cent of the unblinded-randomization studies, and in 58.1 per cent of the nonrandomized studies. Differences in case-fatality rates between treatment and control groups ($P<0.05$) were found in 8.8 per cent of the blinded-randomization studies, 24.4 per cent of the unblinded-randomization studies, and 58.1 per cent of the nonrandomized studies. These data emphasize the importance of keeping those who recruit patients for clinical trials from suspecting which treatment will be assigned to the patient under consideration. (N Engl J Med 1983; 309:1358-61.)

CONTROLLED clinical trials employ a variety of methods for assigning treatment. It has been broadly accepted that randomization is the best method for obtaining comparability of groups in a clinical trial. Some have raised ethical objections to randomization[1,2] and have suggested other methods of treatment assignment.[3] These alternatives have been variably accepted[4,5] and have been criticized because they introduce the possibility of bias into the selection and assignment of patients in clinical trials.[6,7]

Differences in the outcomes of trials of similar therapies in which controls have been either randomized or chosen historically have been documented.[8] Trials that are historically controlled achieve more statistically significant results, whereas studies in which controls are selected at random obtain fewer significant differences between therapies tested.[9] There are many possible reasons for the differences in outcome between studies in which the controls are assigned at random and those in which they are selected by some other process, but data that might explain the differences are not available. This article explores a possible reason for the differences: "blinding" of the investigator with respect to treatment assignment at the time of selection of a patient for inclusion in the study.

Controlled trials of the management of acute myocardial infarction are an ideal setting for the study of this phenomenon for several reasons: the usual principal end point, death during hospitalization, is finite and reliable; there are prognostic variables that predict the outcome with various degrees of success; the number of reported trials in each category of treatment assignment is large enough to allow reasonably reliable analyses of the distribution of risk factors and of outcome; and finally, the rates of prognostic-variable occurrence and of death are large enough to allow conclusions to be drawn from the statistical analyses. This paper demonstrates a significant relation between aberrant distributions of prognostic variables and the differences in the death rates between experimental and control groups. The data suggest that the more frequently positive results found in studies in which assignment of controls is less blinded may be explained by bias in the selection or rejection of patients when the treatment to be given is known or suspected at the time of assignment.

From the Clinical Trials Unit and the Departments of Medicine and Biomathematical Sciences, Mount Sinai School of Medicine of the City University of New York. Address reprint requests to Dr. Chalmers at the Clinical Trials Unit, Mount Sinai Medical Center, One Gustave L. Levy Place, New York, NY 10029.

Supported in part by a grant (LM 03116) from the National Library of Medicine.

Portions of this work were submitted as partial fulfillment of the clerkship requirement in community medicine while Dr. Celano was a third-year student at Mount Sinai School of Medicine.

Portions of this work were presented at the meeting of the Society for Clinical Trials, San Francisco, May, 1981.

METHODS

Therapeutic trials of treatment for acute myocardial infarction were identified through a Medline search, *Current Contents*, and a review of the references listed in the more recently published studies. Only studies that used a control group were included. Each paper was then reviewed by two observers independently. Disagreements were resolved in conference.

The first tabulation of risk factors and case-fatality rates was unblinded — i.e., it was done with knowledge of how the treatment assignment was carried out. The second was blinded — i.e., it was

done after the pertinent assignment data had been obliterated in the photocopy. All risk factors and case-fatality-rate data were recorded, tabulated, and analyzed by chi-square and t-tests for each measurement.

Assignment to a treatment group was classified as having been according to one of three methods. The first, *blinded randomization*, was assignment prearranged at random and communicated to the investigator only after the patient had been accepted for the study and informed consent had been obtained. The usual methods of random assignment in the studies we examined involved opaque envelopes, a telephone call to a statistical center, or a prearranged order of blinded medications labeled consecutively by the pharmacy.

The second method, *unblinded randomization*, was assignment from an open table of random numbers, according to date of birth or chart number, or by some other variably random system in which patients could present for study in a chance order but be selected or rejected after the physician knew the treatment assignment. The obtaining of informed consent, a process highly susceptible to bias, thus could occur with knowledge of the treatment assignment.

The third method, *nonrandom assignment*, included the use of both simultaneously selected controls and historical controls. Assignment to treatment was made by a method more susceptible to clinical judgment than chance. Informed consent, if obtained at all, was obtained after assignment. An example of the use of nonrandomly assigned simultaneous controls was an evaluation of coronary-care units in which patients who were admitted were compared with those who were not admitted because the unit was full. In the case of historical controls, selection of patients has occurred from hospital charts, from reports in the literature, or from computerized data bases. Sometimes the treated patients are matched according to prognostic risk criteria — a process susceptible to bias. Selection for a special treatment and the obtaining of informed consent have rarely been employed to screen historical controls.

Table 1. Treatment Assignment According to Period of Years in 145 Published Studies of Treatment for Acute Myocardial Infarction.

Period	No. of Years	Blinded Randomization	Unblinded Randomization	Nonrandom Assignment	Totals
1946–1960	15	0	17	15	32
1961–1970	10	19	21	17	57
1971–1981	11	38	7	11	56
Totals		57	45	43	145

When the method of treatment assignment was not apparent in the published report, we inquired of the authors directly. The papers in which the method remained unknown were excluded from the analyses. In addition to the method of assigning treatment, the following data were tabulated for analysis: date of publication; case-fatality rate for experimental and control treatments; prognostic variables (risk factors); and the authors' conclusions about whether or not the therapies were of benefit (other factors than case-fatality rate were often influential). (The authors' conclusions were classified as follows: that there was a statistically significant difference favoring the experimental treatment; that there was a trend that did not reach statistical significance but was considered of enough interest by the authors to warrant continued study of the agent; that there was no trend of clinical interest; that there was a clinically interesting trend favoring the control therapy; or that there were statistically significant differences favoring the control treatment.) In each paper the control group was identified as the placebo or no-therapy group or as the group receiving the standard therapy with which the authors compared a new or less commonly used treatment. The validity of this identification was checked at a later date by the investigators who did not make the original identification, and no discrepancies were found. Three nonrandomized studies had more than one control group, and the raw data from these were combined to give single weighted means for the control groups in each study. Similarly, there were four studies with two treatment arms (two blinded randomization, one with unblinded randomization, and one with nonrandom assignment) for which the treatment data were combined.

Table 2. Treatment Assignment According to Types of Treatments Studied.

	Blinded Randomization	Unblinded Randomization	Nonrandom Assignment	Total No. of Studies (%)
Antithrombotic agents	20	26	24	70 (48.3)
Antiarrhythmic agents	9	6	1	16 (11.0)
Beta-blockers	9	2	1	12 (8.2)
Glucose, insulin, potassium	3	3	1	7 (4.8)
Coronary-care units	0	1	11	12 (8.2)
Early mobilization	2	3	3	8 (5.5)
Others	14	4	2	20 (13.8)
Total	57	45	43	145 (99.8)

Results

A group of 160 papers published between 1946 and 1981 formed the basis of the study. Fifteen were discarded because the method of assignment could not be determined. A list of the 145 papers analyzed for this report is available from the authors.

An increase over time in the proportion of studies having proper random assignment of treatments is illustrated by the data in Table 1. The spectrum of treatments evaluated is shown in Table 2. Antithrombotic agents (heparin, warfarin, and fibrinolytic agents) were the most frequently studied treatments and were evenly distributed among treatment-assignment methods. The more recently developed drugs — e.g., beta-blockers — were more frequently studied by means of blinded randomization.

Comparison of risk-factor tabulation by different observers revealed low levels of interobserver and intraobserver variability. Disagreements were uncommon and tended to be the result of oversights rather than differences in interpretation.

The risk factors reported in these analyses are

Table 3. Distribution of Various Risk Factors According to Treatment Assignment.

	Blinded Randomization	Unblinded Randomization	Nonrandom Assignment
Total no. of studies	57	45	43
Risk Factor	*no. of studies (%)*		
Age	25 (44)	21 (47)	22 (51)
Sex	44 (77)	39 (87)	25 (58)
History of previous myocardial infarction	32 (56)	24 (53)	14 (33)
History of congestive heart failure	7 (12)	5 (11)	1 (2)
Presence of congestive heart failure	29 (51)	11 (24)	12 (28)
Anterior infarct	16 (28)	11 (24)	11 (26)
Arrhythmia	18 (32)	7 (16)	7 (16)
Conduction defect	15 (26)	3 (7)	3 (7)
Shock	33 (58)	14 (31)	12 (28)
Prognostic index	12 (21)	10 (22)	24 (56)

Table 4. Distribution of Prognostic Variables According to Treatment Assignment.

	No. of Studies	Total No. of Variables Reported	No. of Variables per Study	Per Cent of Variables with P<0.05	No. (%) of Studies with at Least 1 Variable Maldistributed (P<0.05)	Per Cent of Variables Favoring Each Group		
						Control	Treatment	Neither
Blinded randomization	57	231	4.05±0.29	3.5	8 (14.0±4.6)	36.8	56.1	7.0
Unblinded randomization	45	145	3.22±0.29	7.0	12 (26.7±6.6)	20.2	77.6	2.2
Nonrandom assignment	43	131	3.05±0.32	34.4	25 (58.1±7.5)	16.3	81.4	2.2

shown in Table 3. The nonrandomized studies differ from those in the blinded-randomization group in the smaller number of variables reported, with the unblinded-randomization group falling between the two, and in the number of variables distributed significantly differently between patients in the treatment and control groups, again with the unblinded-randomization group falling between the other two groups (Table 4). In this tabulation the proportion of tests different in the blinded-randomization studies at the P<0.05 level — 3.5 per cent — is about what would be expected to occur by chance. Chance is less likely to explain the 7.0 per cent maldistribution in the unblinded-randomization group and is a very unlikely explanation for the 34.4 per cent of reported tests significantly different at the 5 per cent level in the nonrandomized studies. When the data were tabulated to show the number and percentage of studies in which at least one prognostic variable was maldistributed, the evidence for maldistribution was again apparent: 14, 26.7, and 58.1 per cent of the papers from the blinded-randomization, unblinded-randomization, and nonrandomized groups, respectively, had at least one prognostic variable maldistributed (chi-square with 2 degrees of freedom = 22.86, P<0.001). Tabulation of the prognostic variables with respect to whether they favored outcome in the control or treatment group also revealed a significant maldistribution. In the blinded-randomization studies, 56.1 per cent of the variables favored the treatment group; that is, the patients with better prognostic characteristics were assigned to the treatment group slightly more than half the time, whereas the equivalent percentages for the unblinded-randomization and nonrandomized groups were 77.6 and 81.4 per cent (chi-square with two degrees of freedom = 29.97, P<0.001). In this case, trends were tabulated without regard to their statistical significance.

The results of the trials in terms of case-fatality rate (Table 5) revealed a striking difference that has been noted before in comparisons of trials having randomized controls with those having historical controls.[7,8] Differences in case-fatality rate between treatment and control groups at P<0.05 were found in 8.8, 24.4, and 58.1 per cent, respectively (chi-square with 2 degrees of freedom = 22.86, P<0.001). The mean differences in case-fatality rates were 0.003±0.008 (not significant) for the blinded-randomization group, 0.052±0.016 (P<0.001) for the unblinded-randomization group, and 0.105±0.017 (P<0.001) for the nonrandomized group. The results favored the treatment group over the controls with statistical significance in 60, 100, and 93 per cent, respectively.

Finally, the conclusions of the papers about efficacy were tabulated to obtain a sense of results other than case-fatality rate, since other end points were considered in many of the studies (Table 6). The distributions were similar except that more of the authors of nonrandomized studies strongly favored the experimental therapy.

Discussion

Maldistribution of prognostic variables in pretreatment assignment has been demonstrated for controlled trials in which the investigator had a chance to suspect which treatment was likely to be assigned before the patient was accepted for the study. This maldistribution has been demonstrated to be associated with outcomes that were more frequently different at the 5 per cent level of significance. The data strongly suggest that bias in treatment assignment could be a more important determinant of outcome than the

Table 5. Results of Trials in Terms of Case-Fatality Rates (CFR).

	No.	No. (%) with Significant Result (P<0.05)	Control CFR	Treatment CFR	Difference	P Value
				mean ±S.E.		
Blinded randomization	57	5 (8.8)	0.156±0.017	0.152±0.016	0.003±0.008	NS
Unblinded randomization	45	11 (24.4)	0.227±0.020	0.175±0.014	0.052±0.016	<0.001
Nonrandom assignment	43	25 (58.1)	0.321±0.018	0.215±0.019	0.105±0.017	<0.001

Table 6. Conclusions of Authors about Efficacy of Treatment (Based on Clinical-Response Data in Addition to Case-Fatality Rate).

Type of Study	Total No. of Studies	Authors' Conclusions				
		Strongly Favor Treatment	Slightly Favor Treatment	No Preference	Slightly Favor Control	Strongly Favor Control
				per cent		
Blinded randomization	57	29.8	21.1	45.6	3.5	0
Unblinded randomization	45	31.1	15.6	48.9	4.4	0
Nonrandom assignment	43	55.8	16.3	20.9	4.7	2.3

treatments under investigation. The differences in the distribution of prognostic variables are smaller when randomization is not blinded, but the trend is there. There is little doubt about the evidence for bias when patients are assigned to treatments and the controls are selected or historical.

The results of the chi-square tests applied to our tables should be interpreted with caution, since use of the chi-square test assumes that the factors are independent, and this is probably not true of risk factors in acute myocardial infarction. In an attempt to obviate this problem, we looked at the numbers of studies that had one or more variables maldistributed, and the results were similar (Table 4). Many statisticians prefer a confidence-interval approach rather than hypothesis testing, with its frequently arbitrary decisions about significance or nonsignificance.[10] We agree with this for most applications, since reliance on significance testing may lead to incorrect conclusions.[11,12] This may partly explain why the number of studies that found significant differences in mortality was different from the number of studies whose authors concluded that the treatment was beneficial, but some of this difference may also be due to the fact that other end points were considered important.

The results of our study are pertinent to the various modifications of randomization that have been referred to as adaptive.[5] Blinding of assignment may be difficult to maintain when those modifications are employed. Also, when the relative ratios of treated to control patients are changed as a study progresses and the investigators are not kept blinded to the patient-selection process and the results, there is a definite opportunity for a changing bias in the selection of patients to be reflected in the treatment assignments. This danger has to be weighed against the advantages of monitoring equal groups or subgroups in the course of any study.

More recently it has been proposed by Zelen[3] that the randomization process should be changed to rectify the purported ethical problems of randomized clinical trials. He proposed that informed consent be obtained in half the patients with full knowledge of the treatment to be applied, thus giving bias full play in the determination of whether or not patients would participate in the study. We fear that the effect of this plan would be equivalent to that of unblinded randomization. Although the plan includes a recommendation that investigators follow the dropouts as carefully as those who consent, it is likely that such dropouts would be handled differently.

The data in this study confirm the necessity for blinding the randomization process in a controlled trial in such a way that chance alone determines the assignment of patients to treatment. If that process is not impenetrable to bias, the results of the study may not be creditable, and the effort devoted to the study will be wasted. A proper study will report all measurable prognostic variables that exist before randomization so that the reader may suspect a break in the randomization process if the prognostic variables are significantly maldistributed.

References

1. Gehan EA, Freireich EJ. Non-randomized controls in cancer clinical trials. N Engl J Med 1974; 290:198-203.
2. Bonchek LI. Are randomized trials appropriate for evaluating new operations? N Engl J Med 1980; 301:44-5.
3. Zelen M. A new design for randomized clinical trials. N Engl J Med 1979; 300:1242-5.
4. Horwitz RI, Feinstein AR. Advantages and drawbacks of the Zelen design for randomized clinical trials. J Clin Pharmacol 1980; 20:425-7.
5. Weinstein MC. Allocation of subjects in medical experiments. N Engl J Med 1974; 291:1278-85.
6. Chalmers TC. The control of bias in clinical trials. In: Shapiro SH, Louis TA, eds. Clinical trials: issues and approaches. New York: Marcel Dekker, 1983:115-27.
7. *Idem*. The clinical trial. Milbank Mem Fund Q 1981;0 59:324-39.
8. Sacks H, Chalmers TC, Smith H Jr. Randomized versus historical controls for clinical trials. Am J Med 1982; 72:233-40.
9. *Idem*. Sensitivity and specificity of clinical trials: randomized *v* historical controls. Arch Intern Med 1983; 143:753-5.
10. Rothman KJ. A show of confidence. N Engl J Med 1978; 299:1362-3.
11. Freiman JA, Chalmers TC, Smith H Jr, Kuebler RR. The importance of beta, the Type II error and sample size in the design and interpretation of the randomized control trial: survey of 71 "negative" trials. N Engl J Med 1978; 299:690-4.
12. Messer J, Reitman D, Sacks HS, Smith H Jr, Chalmers TC. Association of adrenocorticosteroid therapy and peptic-ulcer disease. N Engl J Med 1983; 309:21-4.

Not quite as random as I pretended

One of the cartoons drawn by the late Mel Calman for the Penguin edition of Darrell Huff's book *How to Lie with Statistics* shows an interviewer who has cast his clipboard aside to seduce a woman on a sofa. The caption reads "I'm afraid this sample isn't quite as random as I pretended it was". It would have been an apt illustration for Schulz's[1] recent account of his experience of discussing the conduct of randomised controlled trials at epidemiology workshops. Many people who had been investigators in such trials admitted that they had witnessed behaviour or had themselves behaved in a way that subverted the random allocation of treatment. Some had done no more than hold envelopes up to the light to reveal which treatment had been assigned; others had gone so far as to rifle filing cabinets in the principal investigator's office.

Why did they do it? Comments volunteered by the workshop participants suggest that their motives were naive. One person, for example, confessed that he had decoded the sequence of treatment allocation so that he could gain experience in vaginal rather than abdominal hysterectomies. Perhaps he was unaware of the scientific consequences of what he was doing. Like many of the people concerned with keeping trials running from day to day, he may not have appreciated why random allocation of patients to different treatments matters so much. It is not that randomisation ensures that the groups being compared will be identical. Indeed, the fact that treatments are allocated at random makes differences between groups inevitable; consequently, modifications of the randomisation procedure such as stratification or minimisation may be necessary to ensure that such differences are small. The overwhelming advantage of randomisation is that the allocation of patients to the treatments being compared is unbiased. This is true not only for factors that can be identified or measured—eg, age or severity of illness—but also (and equally important) for factors that cannot be so readily identified but that may nonetheless affect outcome. Randomisation eliminates the possibility that the whims and prejudices of the investigators affect who gets which treatment.

The theoretical need for randomisation and concealment of allocation of treatment is supported by empirical evidence.[1,2] Trials whose methodology is poor tend to produce more positive results. An assessment of 250 clinical trials showed that those in which concealment of treatment allocation was inadequate produced treatment effects that were around 40% greater than those in which treatment allocation was satisfactorily concealed. Bias in the way treatments are assigned can often be a more powerful determinant of the outcome of the trial than the treatments that are being investigated.

Most people are uncomfortable with the idea that important decisions should hinge on chance. If the available information is too incomplete to make a fully informed judgment, they fall back on intuition. Trialists are not immune to this tendency. They often find it hard to prevent their hopes that a new treatment might prove to be beneficial from evolving into a belief that it really is so. There is a conflict between their intellectual grasp of the necessity for randomised trials and their desire that the results will confirm what they already believe to be true. It is this conflict that simultaneously makes randomised controlled trials so necessary and drives people to subvert them.

The history of therapeutics contains many examples of treatments that we now know to be useless if not actually inimical. Randomised controlled trials may be a laborious way of determining the optimum treatment for our patients but they are the best we have. Our derision when we hear of fever being treated by venesection or peptic ulcer by milky diets should be tempered by the realisation that only the randomised controlled trial stands between us and similar acts of credulity. Schulz's account raises the possibility that investigators frequently kibitz the treatment allocation in trials that claim to be randomised. This is worrying. As Mark Twain said of a clock that struck thirteen, "it casts doubts not only on itself but on all that went before".

Remedies should not be hard to implement. Watertight ways have already been invented of preventing those who admit patients to a trial from knowing the treatment assignments. Trialists can reasonably be expected to use them, and journals can insist that the details of randomisation are described when the results of a trial are published. Principal investigators should make certain that colleagues who are collaborating in a trial understand the reasons why the allocation of treatment is randomised. And they must take responsibility for ensuring that it stays that way.

Christopher Martyn
MRC Environmental Epidemiology Unit, Southampton General Hospital, Southampton, UK

1 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: 408–12.
2 Chalmers TC, Celano P, Sacks HS, Smith H Jr. Bias in treatment assignment in controlled clinical trials. *N Engl J Med* 1983; 309: 1358–61.

Pathogenesis of sudden ischaemic cardiac death

The treatment of acute myocardial infarction (AMI) has been revolutionised by thrombolytic and antithrombotic therapy. Thrombolysis administered within 12 hours and aspirin initiated within 24 hours of initial symptoms are each associated with a reduction in AMI mortality of around 20%.[1] These treatments are based on the hypothesis that AMI is caused by occlusive coronary artery thrombosis interrupting the blood supply to the myocardium. Although coronary artery dissection, vasculitis, and embolism may exceptionally cause acute occlusion in the absence of atherosclerosis, coronary artery thrombosis does seem to have a central role in the precipitation of AMI.[2] DeWood et al[3] used coronary artery angiography to establish that 97·6% of 126 patients with AMI had total or subtotal occlusion of a major coronary artery within 4 hours of symptom onset; at 12–24 hours the proportion was only 80·7%. Many of these occlusions showed angiographic features of thrombus. DiSciascio et al[4] histologically identified thrombus in 69% of 48 patients undergoing coronary atherectomy up to 15 days after AMI, and Alfonso et al[5] directly viewed coronary artery lumina via an angioscope, identifying occlusive coronary artery thrombus in all but two of 21 patients undergoing angioplasty for total coronary occlusion. Muller et al[6] noted that AMI occurs with increased frequency after awakening, and suggested

Subverting Randomization in Controlled Trials

Kenneth F. Schulz, PhD, MBA

Recent empirical evidence supports the importance of adequate randomization in controlled trials. Trials with inadequate allocation concealment have been associated with larger treatment effects compared with trials in which authors reported adequate allocation concealment. While that provides empirical evidence of bias being interjected into trials, trial investigators rarely document the sensitive details of subverting the intended purpose of randomization. This article relates anonymous accounts of deciphering assignment sequences before allocation based on experiences acquired from epidemiologic workshops for physicians. These accounts run the gamut from simple to intricate operations, from transillumination of envelopes to searching for code in the office files of the principal investigator. They indicate that deciphering is something more frequent than a rare occurrence. These accounts prompt some methodological recommendations to help prevent deciphering. Randomized controlled trials appear to annoy human nature—if properly conducted, indeed they should.

(*JAMA*. 1995;274:1456-1458)

JAMA is stimulating increased rigor in the conduct and reporting of randomized controlled trials (RCTs).[1,2] First, it published reporting guidelines.[1] Then a subsequent Editorial[2] called for comments on those proposed guidelines and on criteria[3] published in another journal. That Editorial also endorsed the tenet of randomization being essential for reducing bias in controlled trials.[2] Is *JAMA* inflating the importance of adequate randomization?

I think not. Recent empirical evidence supports the necessity of adequate randomization.[4] We assessed the quality of randomization reporting in 250 controlled trials extracted from 33 meta-analyses and then analyzed the associations between those assessments and estimated treatment effects. Trials in which the allocation sequence had been inadequately concealed yielded larger estimates of treatment effects (odds ratios exaggerated, on average, by 30% to 40%) compared with trials in which authors reported adequate allocation concealment.[4] These results support other findings[5] and provide empirical evidence that inadequate randomization, particularly poor allocation concealment, contributes to bias in estimating treatment effects.

While we have empirical evidence of bias being interjected into trials, do investigators actually relate the delicate details of subverting the intended purpose of randomization? That has happened,[6] but given the obvious sensitivities involved, documented accounts are rare. In this article, I discuss the important elements of randomization and then present anonymous accounts of deciphering assignment sequences before allocation. Basically, since RCTs are anathema to the human spirit, we must acknowledge the human elements of this important scientific process. To help prevent deciphering, I provide a few methodological recommendations.

From the Division of Sexually Transmitted Disease Prevention, Centers for Disease Control and Prevention, Atlanta, Ga.

Reprint requests to Division of Sexually Transmitted Disease Prevention, Centers for Disease Control and Prevention, Mail Stop E-02, Atlanta, GA 30333 (Dr Schulz).

WHAT IS 'RANDOMIZATION'?

Randomization, if successfully accomplished, prevents bias in allocation of participants to comparison groups. Its success depends on two interrelated processes.[4] First, an unpredictable allocation sequence must be generated based on a random procedure. Second, strict implementation of that schedule must be secured through an assignment mechanism (allocation concealment process) that prevents foreknowledge of treatment assignment.[7] Crucially, allocation concealment shields those who admit patients to a trial from knowing the upcoming assignments. The decision to accept or reject a participant must be made and informed consent obtained without knowledge of the treatment to be assigned.[8]

Traditionally, many medical researchers mistakenly consider the sequence generation process as "randomization." They properly stress generation but frequently slight concealment. Without adequate allocation concealment, however, even random, unpredictable assignment sequences can be subverted.[4,9] For example, suppose that an investigator generates an adequate allocation sequence using a random number table. However, the investigator then posts that sequence on a bulletin board, which equates to basically no allocation concealment. Those responsible for admitting participants could detect the upcoming treatment allocations and then channel participants with a better prognosis to the experimental group and those with a poorer prognosis to the control group, or vice versa.[4] Bias could easily be introduced.

Allocation concealment should not be confused with blinding. Allocation concealment seeks to prevent selection bias,

protects the assignment sequence before and until allocation, and can always be successfully implemented.[7] In contrast, blinding seeks to prevent ascertainment bias, protects the sequence after allocation, and cannot always be implemented.[7] I do not address issues pertaining to blinding in this article.

PERSONAL ACCOUNTS OF DECIPHERING ASSIGNMENT SEQUENCES

During the last 8 years, a colleague and I have conducted more than 20 epidemiology workshops for medical residents and medical school junior faculty. Each workshop included 20 to 25 participants. When we discussed allocation concealment, we asked how many of the participants had deciphered, or had witnessed someone else decipher, an assignment sequence. Typically, they responded with apprehension and silence. However, once we assured them of our lack of interest in individual names and of our preservation of everyone's anonymity, a brave soul would relate her or his experiences. Thereafter, responses usually flowed freely. When queried, more than half of the participants at each workshop related at least one instance of deciphering. This should not be interpreted as representing more than half of all the trials, however. Many participants had been involved in more than one trial, some in more than 10 trials. We do not have an accurate denominator. Nevertheless, their responses indicate that deciphering is something more frequent than a rare occurrence.

The personal accounts of those decipherings ran the gamut from simple to intricate operations. The simple operations were the most frequent and usually involved taking advantage of inadequate allocation concealment schemes. One frequently mentioned approach, taking advantage of the posting of the allocation sequence on a bulletin board, required little effort. Workshop participants admitted to adjusting allocations based on preenrollment checks of the board. Other examples of simple operations included opening unsealed assignment envelopes, holding translucent envelopes up to a regular lightbulb, feeling the differential weight of envelopes, and opening many envelopes that were not sequentially numbered until a desired treatment was found.

More elaborate operations, however, were needed to circumvent more adequate allocation schemes. For example, we have now heard of a few accounts of taking sequentially numbered, opaque, sealed envelopes to the "hot light" (an intense incandescent bulb) in the radiology department for deciphering of the assignment scheme. Apparently that works!

Workshop participants rarely implicated the more impervious allocation concealment schemes in their personal accounts of deciphering operations. Even with the good schemes, however, slight faults could develop into fatal flaws. For example, in trials using sequentially numbered drug containers, someone reported deciphering the scheme based on the appearance of the tablets and another based on the appearance of the label on the containers. Also, in trials using central randomization, we have heard a couple of accounts of physicians ringing a central number for allocation and obtaining the next few allocations all at once.

Still another workshop participant had attempted to decipher a numbered container scheme but had given up after her attempts bore no success. One evening she noticed a light on in the principal investigator's office and dropped in to say hello. Instead of finding the principal investigator, she found an attending physician who also was involved in the same trial. He unabashedly announced that he was rifling the files for the assignment sequence because he had not been able to decipher it any other way. What materialized as most curious was her response. She admitted being impressed with his diligence and proceeded to help in rifling the files. Obviously, the assignment sequence should have been kept in a locked location.

We have not asked the workshop participants why they deciphered sequences. Based on some volunteered comments, however, I have gleaned a few scanty notions. Frequently, they simply lacked knowledge of the scientific ramifications of their actions. One said that he wanted experience in vaginal rather than abdominal hysterectomies. Other documented accounts also reflect lack of knowledge.[10] The persons involved do not necessarily have unscrupulous motives.

I can speculate on a general explanation for many subversions: RCTs are anathema to the human spirit. The need for unbiased research conducted by human beings on human beings embodies a volatile mix. Investigators intellectually grasp the need but have many contradictory interests once they are immersed in a trial. They perhaps "know" the more effective treatment, so they may want certain patients to benefit or may want the results of a study to reveal what they believe to be valid. Some aspects of properly conducted RCTs, then, annoy investigators, because trial procedures attempt to impede human inclinations.

As Oscar Wilde wrote, "The only way to get rid of a temptation is to yield to it."[11] For those conducting a trial that has not incorporated proper procedures for allocation concealment, the challenge of deciphering the allocation scheme may frequently become too great a temptation to resist. Succumbing to temptation may sometimes reflect deliberate acts to alter findings. At other times, succumbing may be an innocent reflection of human inquisitiveness and ingenuity rather than scientific malevolence. Whatever the motivation, however, the effects are the same when such actions undermine the validity of the trial.

RECOMMENDATIONS

Since many allocation decipherings emanate from a lack of scientific knowledge among those conducting trials, education in the rationale and importance of trial procedures would avert many problems. More important, investigators must acknowledge the vagaries of human nature; they should establish methodological safeguards that thwart attempts to contaminate trials with bias. Particular attention to allocation concealment will prevent or deflect attempts at subversion. Moreover, medical journals should insist on adequate reporting of randomization.[1-3,7]

Regarding the generation of assignment schedules, available texts[12,13] and an entire journal issue[14] comprehensively cover the details. One aspect of generation, however, deserves greater attention. If blocked randomization is used in an unblinded trial, the block size should be randomly varied to reduce the chances that the assignment schedule will be inferred by those responsible for recruiting participants. If the block size in such a trial is not randomly varied but fixed, particularly if the size is small (eg, six or fewer participants), the block size could be unraveled. With treatment assignments becoming known after allocation, a sequence can be discerned from the pattern of the past assignments. Some future assignments could then be accurately anticipated and selection bias introduced, regardless of the effectiveness of allocation concealment.

In previous research,[4,7,15] investigators have considered the following approaches to allocation concealment to be adequate: use of sequentially numbered, opaque, sealed envelopes; pharmacy control of allocation; use of numbered or coded containers; and central randomization (eg, by telephone to a trials office). These criteria describe minimal methodological standards, yet they are met by only about one fourth of recent

trials.[7,15] Realistically, those standards should be exceeded.

Methods using envelopes are more susceptible to manipulation through human ingenuity than are other approaches and are therefore considered a less than ideal method of concealment.[9] If investigators use envelopes, they must diligently develop and monitor the allocation process to preserve concealment. In addition to using sequentially numbered, opaque, sealed envelopes, they should ensure that the envelopes are opened sequentially, and only after the participant's name and other details are written on the appropriate envelope.[16] I also recommend using pressure-sensitive paper or carbon paper inside the envelope. That transfers such information to the assigned allocation and thus creates a valuable audit trail. Cardboard or aluminum foil placed inside the envelope further inhibits detection of assignments.

Reports in which the assignment was stated to have been made by the pharmacy have been classified as having used an acceptable allocation concealment mechanism.[4,7,15] The pharmacists' compliance with proper randomization methods in these trials is unknown, however, and the precautions taken should have been reported. I am aware of instances in which pharmacists have been responsible for gross distortions of assignment schedules. For instance, one large pharmacy charged a project $150 per participant for randomization. During one weekend in the course of the trial, the pharmacy ran out of one of the two drugs being compared. The pharmacists then allocated the other drug to everyone "to avoid slowing enrollment." Investigators should not assume that pharmacists, or others involved in their trials, for that matter, are knowledgeable in RCT methods. Investigators must ensure that their research partners follow proper trial procedures.

The use of numbered or coded containers prevents foreknowledge of treatment assignment, but only if investigators take proper precautions. Authors of trial reports should specify further details of the methods. Assurances that all of the containers were of equal weight and similar appearance and that some audit trail had been established, such as writing the names of participants on the empty bottles or containers, would help readers to assess whether randomization was likely to have been concealed successfully. Similarly, although central telephone randomization is an adequate approach to allocation concealment, effective trial procedures should have been established and followed. All these details should be addressed in the trial execution and in the trial report.[1,2,7]

Investigators and methodologists often neglect one other critical element of RCT design and reporting. With all approaches, the person or persons who prepared the randomization scheme should not be involved in determining eligibility, administering treatment, or assessing outcome. That is obviously important because, whatever the methodological quality of the randomization process, such an individual would usually have access to the allocation schedule and thus the opportunity to introduce bias. Faults in this critical trial element may indeed be the crack through which much of the bias seeps into controlled trials. Nevertheless, under some extraordinary circumstances, someone may have to prepare the scheme and be involved in the trial. In those instances, the investigators must make sure that the assignment schedule is unpredictable and locked away even from the person or persons who generated it.

These recommended randomization procedures are not onerous. While trial organizers may feel some minor additional complexities, trial implementers could actually experience simplified operations. Moreover, these inexpensive procedures would account for only a small fraction of a total trial budget. As for other specific recommendations on randomization, please forward your suggestions to me as well as any additional reports of deciphering.

CONCLUDING COMMENTS

Proper randomization provides our only hope for eliminating selection bias from investigations.[17] Juxtaposed to its value, however, is the challenge human inclinations pose to RCTs. The paradox is that the reason that trials are crucial is the very reason that they are problematic. Researchers need to realize that humans, if given the opportunity, frequently subvert the intended aims of randomization. Thus, opportunities for deciphering need to be eliminated, or at least constricted, with painstaking, assiduous attention to the design and conduct of randomization schemes. Adequate allocation concealment is a vital part of that process.

I gratefully acknowledge the helpful suggestions on an earlier draft provided by David A. Grimes, MD, Department of Obstetrics, Gynecology, and Reproductive Sciences, University of California, San Francisco (my colleague in the conduct of the epidemiology workshops); Iain Chalmers, MBBS, MSc, UK Cochrane Centre, Oxford, England; and Douglas G. Altman, Imperial Cancer Research Fund, London, England.

References

1. The Standards of Reporting Trials Group. A proposal for structured reporting of randomized controlled trials. *JAMA*. 1994;272:1926-1931. Correction: *JAMA*. 1995;273:776.
2. Rennie D. Reporting randomized controlled trials: an experiment and a call for responses from readers. *JAMA*. 1995;273:1054-1055.
3. 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.
4. 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.
5. Chalmers TC, Celano P, Sacks HS, Smith H Jr. Bias in treatment assignment in controlled clinical trials. *N Engl J Med*. 1983;309:1358-1361.
6. Carleton RA, Sanders CA, Burack WR. Heparin administration after acute myocardial infarction. *N Engl J Med*. 1960;263:1002-1004.
7. 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.
8. Chalmers TC, Levin H, Sacks HS, Reitman D, Berrier J, Nagalingam R. 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.
9. Pocock SJ. Statistical aspects of clinical trial design. *Statistician*. 1982;31:1-18.
10. Crewdson J. New breast cancer data flaws. *Chicago Tribune*. May 1, 1994: section 1:1.
11. Wilde O; Ackroyd P, ed. *The Picture of Dorian Gray*. New York, NY: Penguin Books; 1985:41.
12. Pocock SJ. *Clinical Trials: A Practical Approach*. Chichester, England: John Wiley & Sons; 1983.
13. Meinert CL. *Clinical Trials: Design, Conduct, and Analysis*. New York, NY: Oxford University Press; 1986.
14. Properties of randomization in clinical trials. *Control Clin Trials*. 1988;9:287-382. Theme issue.
15. Altman DG, Doré CJ. Randomization and baseline comparisons in clinical trials. *Lancet*. 1990;335:149-153.
16. Bulpitt CJ. *Randomized Controlled Clinical Trials*. The Hague, the Netherlands: Martinus Nijhoff; 1983.
17. Altman DG. Randomization: essential for reducing bias. *BMJ*. 1991;302:1481-1482.