Current problems and future challenges in randomized clinical trials

ALVAN R. FEINSTEIN, M.D.

AS HISTORIANS review the advances of therapeutic science in the second half of the 20th century, one of the outstanding methodologic achievements will surely be the use of evidence from randomized trials to show the efficacy of new therapeutic agents. With this type of evidence, clinicians and patients can now be reasonably reassured, for the first time in medical history, that a new treatment actually does what is claimed for it.

Scientific problems in comparing therapy. Although the concept of randomized assignment was developed in the 20th century, the idea of controlled comparisons of therapy dates back to antiquity. A comparative trial of two diets was reported in the Book of Daniel; citrus fruits were compared with five other agents when James Lind treated scurvy in 1747; early and late bloodletting were contrasted in 1835 by Pierre Louis; and many of the new therapeutic agents introduced in the first half of the 20th century received documented comparison with older agents.

The comparisons, however, usually had four features that were scientifically unattractive. One obvious problem was that the comparisons were often nonconcurrent. Although the results would be presented as a contrast of newer treatment vs older treatment, the newer treatment was often given in a newer era to patients identified with newer diagnostic methods and receiving newer supportive treatment. When the results of this combination of new factors were compared against the effects of the diagnostic methods, main treatment, and supportive treatment used in the older era, the additional factors were not "controlled." Consequently, the better results attained with new treatment might have been due to earlier diagnosis or better supportive treatment.

A second problem was the failure to use similar admission criteria for the patients included in the contrasted groups. For example, when surgical treatment given judgmentally to "operable" patients is compared against nonsurgical treatment given to "inoperable" patients, the subsequent triumphs of surgery are often attributable not to the surgery but to the better pretherapeutic prognosis of the "operable" patients.

A third problem was produced by the often inadvertent use of different methods or criteria for decisions about "success" or "failure" in observing and recording the outcome of treatment in the compared groups. Furthermore, when the outcome was a subjective event, such as relief of symptoms, a clinician (and patient) who expected a particular treatment to be effective could often convert the expectation into a self-fulfilling prophecy when the response to treatment was experienced and noted.

A fourth problem was really statistical, rather than scientific. A success rate of 50% for a new treatment might be hailed as a major achievement in doubling the success rate of 25% for the old treatment, but the actual numerical values under contrast could have been as small as 2/4 vs 1/4. For these small groups, the observed distinctions could easily arise by chance alone if the two treatments were actually equivalent.

Advantages of randomization. The use of randomization as a method for allocating treatments in comparative trials was not specifically intended to address any of these four problems. The problems can be solved by

Dr. Feinstein is Professor of Medicine and Epidemiology; Director, Clinical Epidemiology Unit and Robert Wood Johnson Clinical Scholars Program, Yale University School of Medicine, New Haven; Senior Biostatistician, Cooperative Studies Program Coordinating Center, Veterans Administration Medical Center, West Haven, CT.

Supported in part by grant HS 04101 from the National Center for Health Services Research, OASH; by The Robert Wood Johnson Foundation Grant 6309; and by a grant from the Andrew W. Mellon Foundation.

This "Perspective" is developed from a larger chapter on clinical trials in Dr. Feinstein's new book Clinical Epidemiology, W. B. Saunders Co.

Address for correspondence: Alvan R. Feinstein, M.D., Yale University, I 456 SHM, P.O. Box 333, New Haven, CT 06510-8025.

Vol. 70, No. 5, November 1984
any type of research plan that arranges for treatments to be compared concurrently, in groups of adequate size, comprising subjects who fulfill similar criteria for admission to the study, and who are observed thereafter with similar objective procedures. By making the assignment of treatment unpredictable, the role of randomization was to prevent a fifth problem: the additional bias that might occur when treatments are selected judgmentally in the basic context of a research plan that calls for concurrent comparison, adequate size, similar eligibility, and similar follow-up observations.

If the other four problems are taken care of, the impact of this fifth problem will vary with the types of compared treatments. If the principal treatment under investigation has some strong contraindications — such as the co-morbidity or clinical severity that may militate against the use of surgery, anticoagulants, or other powerful agents — the contraindicating conditions will usually be accompanied by a worse prognosis in patients who are denied the principal treatment. This source of pretherapeutic “susceptibility bias” can be reduced or removed by the demand that all patients under comparison be equally eligible to receive all of the compared treatments. If there are no compelling clinical mandates to give or avoid any of the compared treatments, the use of similar eligibility criteria will have less importance in removing susceptibility bias.

Regardless of the effects of similar eligibility criteria, randomization has the vital role of removing or reducing the susceptibility bias that can arise when treatments are actually assigned. Thus, although medical vs surgical treatment might be compared only in patients who are all “operable,” randomization would prevent surgery from being selectively chosen for operable patients with the best prognoses.

The majestic achievements of randomization, although often ascribed to the stochastic virtues of chance assignments, really arise from the prophylaxis of the fifth problem. By necessitating the advance planning of an experiment, a randomized trial can solve the first four problems almost as a by-product. In a randomized trial the treatments are compared concurrently, with special efforts made to eliminate unauthorized “contamination”; the patients are diagnosed with similar methods and admitted with similar pretherapeutic criteria; the follow-up observations can be arranged with objectivity and, if necessary, with “double blind” procedures; and the statistician can calculate the group sizes needed for the results to sustain claims about “significance.” With these four problems managed as part of the routine plans, randomization makes its majestic contribution by helping avoid the fifth problem.

**Early development of randomized trials.** In historical sequence, randomization was first proposed by R. A. Fisher in 1923 for agricultural experiments, first used by Amberson et al. in 1926 to assign clinical treatment (with the toss of a coin) to two preformed groups, and first employed by Diehl et al. in 1938 to make individual assignments of two treatments. The “landmark” that inaugurated the modern era of biostatistically planned randomized trials occurred in 1947, when streptomycin was evaluated in the treatment of tuberculosis. The striking demonstration of efficacy in that study supported the merits of randomized trials both as a scientific method for evaluating therapy and as a procedure for deciding whether new drugs warranted major industrial development.

The success of the streptomycin trial, however, was achieved without the challenge of many difficult problems that would subsequently arise in the evaluation of other therapeutic agents. Because the supply of streptomycin was limited, enough was available for only those patients who would receive active treatment. No ethical quandaries were encountered in using placebo or in denying sick patients a potentially curative drug. The short supply of streptomycin also restricted the scope of admission criteria to the trial. It was limited to patients with “far advanced tuberculosis,” thus providing a relatively homogeneous group of clinical conditions, rather than a heterogeneous spectrum of different clinical types of patients with different degrees of severity of the same disease.

The main outcome event in the trial — a reduction in the size of pulmonary lesions — could be appraised by radiologists who were kept unaware of therapy. There was no need to use double-blind tactics for evaluations made subjectively by patients and clinicians. Since the changing size of the radiographic lesions was an easily measured remedial effect, statistically significant decrements in the lesions could be shown with small numbers of patients, thus avoiding the huge groups needed in prophylactic trials when the outcome event to be prevented has a low rate of attack. Since the radiographic changes occurred relatively promptly after initiation of therapy, the trial also had a short duration. A large number of people did not have to be followed for a protracted period of time. Finally, since the limited supply of streptomycin allowed it to be given only briefly, there was no opportunity to observe any of the adverse effects of its long-term use.

As randomized trials subsequently became well es-
tablished and widely employed, however, the investigators began to encounter all the problems that had been avoided in the streptomycin study: ethical quandaries, heterogeneous patients, subjective observations, huge sample sizes for prophylactic end points, protracted periods of observation, and long-term adverse side effects.

Distinctions of remedial vs prophylactic therapy. When a therapeutic agent has a remedial goal — such as reducing the size of a lesion, relieving pain, lowering blood sugar, or altering any other target condition that exists at the start of treatment — a randomized trial is relatively easy to do. The patients are reasonably homogeneous, because all of them have the same initial target condition; the measurable change in that condition usually allows “statistical significance” to be achieved with a small sample size; and the change occurs within a relatively short period of time. Because most pharmaceutical agents are developed and tested for remedial actions, randomized trials have been particularly successful in demonstrating the remedial effects of new (or old) agents.

The clinical situation is much more difficult, however, when the agent is used prophylactically to prevent a target condition that is not present when treatment begins. In primary prevention, such as vaccination against poliomyelitis, the prophylaxis is intended to thwart the future development of a disease. In secondary prevention the prophylaxis is aimed at avoiding or retarding the adverse consequences of an established disease. For this prophylactic role, an agent having an established remedial effect as its short-term primary action is used to prevent a secondary long-term effect. Examples of this type of secondary prophylaxis are the attempt to prevent vascular complications by altering hemostatic mechanisms or by lowering blood sugar, blood pressure, or blood lipids in patients with a variety of ailments including coronary artery disease, cerebrovascular disease, diabetes mellitus, and hypertension. Another common example of treatment aimed at secondary prevention is the effort to avoid or retard death by giving various antineoplastic therapies to patients with cancer.

If we recall some of the major ambiguities, disappointments, controversies, and confusions that have occurred after randomized trials in the past two decades, we find that almost all the problems have been associated with studies of prophylactic agents, most often in circumstances of secondary prevention. Cardiologists need no reminder about the controversies that have raged, and that still flourish, in coronary disease alone after randomized trials for the prophylactic bene-

fits of anticoagulants, clofibrate, vasodilators, sulfinpyrazone, aspirin, dipyridamole, intracoronary streptokinase, bypass surgery, and multiple risk factor intervention.

Problems produced by duration and sample size in trials of prophylactic therapy. Because the outcome event is something to be prevented, trials of prophylactic therapy have two major handicaps that do not occur in studies of remedial agents. The first handicap is that the patients must be followed for whatever duration is required to see whether the outcome event occurs. This duration is usually short in studies of treatment for far-advanced cancer because substantial proportions of the patients die within a year. The duration can be quite long, however, for therapeutic trials in which the individual observation periods are 5 years or more for such outcome events as recurrent myocardial infarction or other cardiovascular and cerebrovascular complications.

Aside from the logistic difficulties of maintaining people under observation for a protracted period, two main problems can arise when a trial has a prolonged duration. The first problem is the “contaminated” protocol. The longer the duration of the trial in each patient, the greater the opportunity for its basic plans to be violated. If scheduled tests are not carried out at the right times, if assigned treatments are altered or taken improperly, if many patients drop out of the study, or if other important violations occur, the ultimate results will be difficult to analyze and interpret, regardless of what statistical methods are used.

The second major problem is the hazard that new technologic developments, arising while the trial is in progress, will make the results of the trial obsolete before it is finished. (This problem can arise not only when the trial has a long duration for each patient but also when the total study takes a long time, despite relatively short durations of individual treatment and follow-up, to accrue enough patients with a relatively rare condition.) An example of this problem was the development of the bypass graft operation while a large-scale trial was in progress to evaluate the surgical tunnel-implant procedure for coronary disease. The tunnel-implant trial was eventually abandoned and the clinical investigators began a new trial of bypass grafting. As the latter trial was in progress, cold cardioplegia became developed as a method of improving the bypass operation. Consequently, the postoperative fatalities noted early in the trial could no longer be used to indicate what might be expected as postoperative mortality for the improved operative procedure. Today, before all the results have been fully appraised for
the early or later parts of the bypass trials, percutaneous transluminal angioplasty has arrived on the scene and is in need of evaluation. The difficulties of doing this evaluation with randomized trials have already been recognized as formidable. In an age of rapidly changing new technologic developments, this type of problem can be expected to occur repeatedly for both medical and surgical forms of treatment.

A second prominent handicap in trials of prophylactic therapy is the low rate of occurrence for many of the chosen outcome events. To get statistically significant evidence that a low attack rate has been further lowered, such trials will require a very large sample size. To get enough patients, multiple collaborating institutions must join and cooperate in following a common protocol and pooling their data. Aside from the logistic difficulties of arranging to coordinate and standardize the protocol at diverse institutions, a major problem of the large-size handicap is that the trials can be extremely costly. Among the listed costs for some well-known clinical trials are the following:

University Group Diabetes Program (UGDP) $10,000,000
Coronary Drug Project (CDP) $40,000,000
Hypertension Detection and Follow-up Program (HDFP) $70,000,000
Multiple Risk Factor Intervention Trial (MRFIT) $116,000,000

The vast expense of these activities becomes particularly distressing when the results do not provide the anticipated clarification of the issues under investigation.

Management of problems in cost, obsolescence, and contaminated protocols. None of the major problems just cited is easy to manage. Some of the expenses can be reduced by emulating the model developed by the Veterans Administration, which has pioneered in demonstrating that large-scale cooperative trials need not be conducted as fiscal colossi. The VA has achieved this desideratum by using an established network of hospitals, investigators who are salaried staff physicians at those hospitals, biostatistical coordinating centers that are part of the VA system, and an efficient VA committee mechanism for evaluating new studies and monitoring progress. This arrangement of hospitals, investigators, coordinating centers, and supervision is internal to the VA system and has avoided the large additional expenses needed for the ad hoc investigative personnel, special coordinating centers, university "overhead" costs, and project site-visit teams in trials conducted under other auspices. In exchange for the relatively low costs and high efficiency, however, the VA studies have been limited to topics that could be adequately investigated in VA patients. Although many large-scale randomized trials will have to be conducted outside the VA system, and although certain high expenses are inevitable, some of the costs might be substantially reduced if the sponsoring and investigative agencies were able to adopt some of the economies developed in the VA cooperative studies program.

A second major problem — the hazard of technologic obsolescence — is unavoidable in any form of research, as well as in randomized trials. For the relatively simple activities of laboratory research, the investigator can easily make and carry out the decision to stop a particular study when it seems unlikely to be productive. This decision is much more difficult, however, if the study is a large-scale trial that has required years of planning and that involves multiple institutions and investigators. If the investigators have committed a great deal of time, effort, and ego to the trial, they may not be coldly dispassionate in approaching the decision to stop it. To allow this type of decision to be made objectively, many large-scale trials today have established an external policy committee, from which the participating investigators are excluded. Although such committees were instituted mainly for deciding whether the accrued results of the treatments compared in an ongoing trial warranted a premature ending or postmature extension of the planned study, the committees might be given several additional assignments.

The committees could be asked to decide whether a trial should be stopped when the arrival of a powerful new technologic agent threatens the value of the results. A different type of "early-termination" decision might be considered when the early results of a trial indicate that the original objective of the study cannot be achieved because the intended plans have been too altered by reality. For example, in both the Coronary Drug Project and in the WHO primary prevention trial of clofibrate, relatively early results showed that patients' serum lipid levels were not being lowered to the extent anticipated or desired in the research plans. If continued, the trials would test the effects of prescribing clofibrate, but the results would not answer the original question of whether a substantial lowering of lipid levels will prevent myocardial infarction. An early cessation of those trials might have been disappointing and frustrating for the investigators, but the process might ultimately have been less disappointing and frustrating, as well as less expensive, than the inconclusive, controversial results that eventually emerged from the completed trials. An analogous ear-
ly-termination decision might have saved enormous amounts of time, effort, money, and confusion in the MRFIT study, when most of the “control” group was found to be receiving the same interventions as the “treated” group.

The difficulties just noted in the clofibrate and MRFIT studies are examples of the most thorny of the cited problems: the management of a contaminated protocol. This type of hazard will continue to arise in long-term studies when the assigned therapeutic agents do not act or are not maintained in the intended manner. No matter what solution is offered for the problem, the results will not satisfy everyone because the analysis of contamination requires a choice between conflicting policy options that can create and perpetuate inevitable controversies in randomized clinical trials.

Pragmatic vs fastidious policies in design and analysis. Regardless of whether a randomized trial is aimed at remedial or prophylactic therapy, every aspect of the design and analysis of the trial is affected by certain basic strategic decisions. The decisions arise from attitudes about why a trial is done, at whom it is aimed, what makes it pertinent, what makes it credible, and how its results are to be used. The attitudes behind these decisions often involve a choice of one of two fundamental but conflicting policies, each of which is reasonable and readily justified. Because only one of the two policies can be used, however, the subsequent results will often be unacceptable to adherents of the opposing policy. The controversies that arise in many randomized clinical trials are thus inevitable.

The two conflicting policies have been described in diverse ways and have received diverse names. Although the full details are beyond the scope of this essay, the conflict will be outlined here under the title of pragmatic vs fastidious policies. A more extensive discussion is presented elsewhere.

In the pragmatic policy, a therapeutic trial is intended to ask and answer questions that are directly cogent for the “messy” realities of clinical practice. The groups under study represent the heterogeneous patients seen in practice; the treatments are given in the manners employed by practicing physicians; and the results are analyzed according to what actually happened. In the fastidious policy, a therapeutic trial is aimed at getting a “clean,” unbiased answer to the research question. The groups under study may therefore be chosen to be relatively homogeneous; the treatments may be modified to facilitate a “double-blind” examination; and the results are analyzed according to the “intention-to-treat” plan created by the randomized assignment, regardless of how the assignment may have been altered during the trial. Beyond these bare outlines, the two policies will have extensive ramifications in every aspect of what is done in a randomized trial to choose patients, decide on the compared treatments, select the outcome events, acquire data, retain ineligible or “lost” patients, analyze contaminated therapy, and cope with unexpected outcome events.

An example of the problems occurred in the celebrated controversy over the UGDP trial of treatment for patients with diabetes mellitus. The trial was planned and analyzed with a fastidious approach. The patients were chosen to have non-insulin-dependent diabetes; the oral agents were given in fixed dosages to allow “double-blinding”; the main outcome event was cardiovascular mortality, which had not been anticipated in the original design of the study; and the results were analyzed according to the originally assigned treatments, although many of the treatments had not achieved good regulation of glucose and although many of the patients had changed to other treatments during the course of the trial. All of these decisions were unacceptable to adherents of the pragmatic viewpoint, who wanted to determine the effect of careful glucose regulation in any type of diabetes, using “titration” of therapy, examining clinical morbidity and total mortality as outcome events, and analyzing results according to the treatments that had actually been received, not just those that were initially assigned.

Each of the fastidious and pragmatic policies can readily be defended and justified with both passionate rhetoric and dispassionate logic, but the same single trial can seldom be conducted with both policies and will seldom satisfy proponents of the opposing viewpoints. The fastidious policy has a scientific and statistical appeal because of its attention to a “clean,” unbiased set of plans and analyses, but the results may not be directly cogent (and may be rejected) for clinical practice. The pragmatic policy is attractive for its attention to the needs of clinical science, but the results contain the hazard of a biased analysis and may be rejected for their statistical “messiness.” Although certain compromises can be achieved, particularly if the most adamant proponents of the two viewpoints were more tolerant of the opposing position, many conflicts will be inevitable because the main demands of the two policies can seldom be attained in a single trial. Consequently, the attempt to prevent controversy by satisfying everyone would require that many therapeutic questions be resolved not with a single, large-
scale, expensive randomized trial, but with several such trials — each conducted according to the appropriate policy choices.

Because the costs of such multiple activities would be prohibitive, the controversies and confusions that have occurred after many prominent randomized trials in the past two decades are destined to recur and to continue without resolution. The situation is reminiscent of the definition of tragedy as the destructive collision of two opposing protagonists, both of whom are right.

Observational approximations of randomized trials. Because randomized trials have developed such high scientific stature and acceptance, they are now accorded an almost religious sanctification; the accusation of antiscientific heresy may be given to any suggestion that cause-effect relationships can be evaluated observationally without randomization. This type of fervent belief in the irreplaceable virtues of randomized trials can be maintained, however, only by ignoring two cogent scientific realities. The first reality is that randomized trials have not been (and cannot be) applied for many of the cause-effect relationships that must be evaluated in modern medical science. To study the etiologic effects of smoking, exercise, diets, and other factors that may cause disease in healthy people and to appraise suspicions about the adverse pathogenetic consequences attributed to clinical therapy with sexual hormones, diuretics, or other pharmaceutical substances, we have had to rely on observational epidemiologic research rather than randomized trials. The second reality, as discussed in this essay and elsewhere, is that randomized trials are too difficult, too expensive, or too controversial for routine use in answering all the clinical questions that will arise in the future for a burgeoning diagnostic and therapeutic technology.

Whether we like it or not, most of our future decisions about medical practice, health care, and scientific technology will have to be made without evidence from randomized trials. To acknowledge this reality requires no loss of reverence, allegiance, or respect for the primacy of randomized trials as a “gold standard” in scientific research. Furthermore, we commit neither sacrilege nor disloyalty by recognizing that randomized trials cannot always be done, that they do not always yield unequivocal answers when done, and that alternative scientific methods must be developed to get satisfactory answers to questions for which randomized trials are either impossible or inadequate.

In a similar acknowledgment of reality, surgeons who prefer to operate in a suitably antiseptic environ-

ment realize that they must sometimes work in battlefield or other conditions where perfect antisepsis cannot be achieved. Accordingly, thoughtful surgeons can use a scientific knowledge of antisepsis to guide the creative efforts that set standards when surgery must be done under less-than-perfect conditions. Analogous efforts are now needed to develop scientific standards for all the evaluative “operations” that must be performed, without the intellectual “antisepsis” of randomization, in the “battlefield” of modern technology.

This major challenge in modern medical science has received relatively little creative thought or consideration. A preoccupation with the structure, conduct, and analysis of randomized trials has diverted suitable attention from being given to scientific principles rather than mathematical procedures in observational research, and no general strategies have been developed for using the scientific structure of randomized trials as a guide to design research in circumstances where randomization is impossible or unfeasible.

In a curious paradox of double standards for cause-effect evaluations, the benefits of modern therapeutic agents are usually appraised with the high scientific demands of randomized trials, but the risks of therapy or of etiologic agents are appraised without similar scientific principles. The observational cohort, retrospective case-control, and other nonexperimental studies used in epidemiologic investigations have evolved as a separate set of isolated research activities. They are not usually planned as scientific substitutes for randomized trials, and they are not usually evaluated with demands for fulfillment of the basic scientific principles noted earlier.

In many observational studies, mathematical models are used to analyze the data, and the results are required to be “statistically significant”; but the scientific methods of the research may not involve concurrent comparison of etiologic or therapeutic maneuvers, suitably similar admission criteria in groups receiving the compared maneuvers, and suitably similar methods for detecting the outcome events. These three scientific principles are seldom given specific attention when randomized trials are planned because the principles are easily achieved, almost inadvertently, as a by-product of the experimental design. Since the principles cannot be easily achieved or ensured in observational research, however, the fulfillment of these scientific standards requires adequate recognition and suitable efforts. With such efforts, observational studies could be suitably designed to cope with the four main scientific and statistical problems cited earlier.
The fifth problem, produced when treatment is assigned without randomization, is more difficult to solve, but reasonably effective approaches might be developed if thoughtful attention were given to the development of cogent prognostic analysis for patients' pretherapeutic status.\textsuperscript{7, 26}

In current policies and attitudes about clinical research, however, the idea of a nonrandomized investigation usually evokes such horror that the scientific challenges are neither acknowledged nor considered. One way to mitigate this horror is to determine whether similar results were obtained when the same topic was studied in the same way with randomized and nonrandomized trials. Unfortunately, almost no data are available for such comparisons. Because various differences in the protocol may lead to conflicting results even in two apparently similar randomized trials,\textsuperscript{27} experimental and observational studies may often disagree because they asked different questions in different ways, not because one trial was randomized and the other was not. Even if the research protocols had similar components, however, the randomized and nonrandomized results may not be comparable because the studies were not done concurrently or because the cited scientific principles were not adequately applied in the nonrandomized research.

In two recent instances where randomized and nonrandomized trials were devoted to the same topic, approached with similar protocols, and carried out with all of the three main scientific principles, the results of the two types of research were remarkably similar. One of these examples occurred when Hammermeister et al.\textsuperscript{28} used observational data to simulate a trial of medical vs surgical bypass treatment for coronary disease and obtained results similar to those of randomized studies. The other example was produced in a landmark investigation in which Paradise et al.\textsuperscript{29} simultaneously conducted randomized and nonrandomized trials of medical vs surgical therapy for severe tonsillitis. The two parallel trials showed almost identical results. Like other serendipitous events, these phenomena warrant careful methodologic examination to determine what aspects of the clinical topics and research architecture led to the concordant results.

The challenge is to learn what we can from the scientific architecture of randomized trials and to use that knowledge to improve the structure of nonrandomized research. The splendid scientific contributions of randomized trials have made them the "gold standard" of cause-effect research — but when gold is too expensive or too difficult to obtain, we need to have good substitutes.

Since this essay is intended to identify problems and challenges, not to offer solutions, I shall not discuss details of the strategies and tactics\textsuperscript{7, 26, 30} with which scientific quality can be improved in nonrandomized research. The main points to bear in mind are that these improvements must be developed as a vital necessity of clinical and epidemiologic science, that the improvements will enhance the quality of randomized as well as nonrandomized studies, that the work offers a fundamental intellectual challenge in basic clinical investigation, and that the work will not be quick or easy. Since more than three decades of intensive effort were needed to establish randomized trials as a routine investigative activity and to recognize their limitations, we cannot expect the current problems to be solved overnight.

Because numbers are "the language of science,"\textsuperscript{31} the work will need new forms of biostatistical creativity. Because any language, however, depends on basic information and reasoning, the fundamental activities will require that clinical investigators expand their creative horizons to give that scientific language a proper derivation, a sensible grammar, and a meaningful content.

References
1. Daniel Ev 1–15
3. Louis PCA: Researchers on the effects of bloodletting in some inflammatory diseases and on the influence of tartarized antimony and vesication in pneumonitis. (Trans. Putnam CG). Boston, 1836, Hillard, Gray

Vol. 70, No. 5, November 1984
Current problems and future challenges in randomized clinical trials.
A R Feinstein

Circulation. 1984;70:767-774
doi: 10.1161/01.CIR.70.5.767

Circulation is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 1984 American Heart Association, Inc. All rights reserved.
Print ISSN: 0009-7322. Online ISSN: 1524-4539

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circ.ahajournals.org/content/70/5/767.citation

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Circulation can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Circulation is online at:
http://circ.ahajournals.org//subscriptions/