

The Clinical Trial

THOMAS C. CHALMERS

*Mount Sinai School of Medicine,
City University of New York*

THE CLINICAL TRIAL IS THE MEETING PLACE OF the practice of medicine and clinical research. Webster's *Third New International Dictionary (Unabridged)* defines research as a "careful or diligent search," a "studious inquiry or examination" aiming at "the discovery of new facts and their correct interpretation, the revision of accepted conclusions . . . or the practical application of such new or revised conclusions." The way in which this search is conducted in clinical medicine forms the subject of this paper.

Traditionally, sharp distinctions have always been drawn between medical practice and therapeutic research using human subjects. When a physician goes through the process of diagnosing an illness in a patient, he is testing his knowledge both of biomedical phenomena and of the patient as an individual. Keeping specific disease possibilities in mind, he takes down the patient's medical history, does a careful physical examination, and orders laboratory tests. He formulates a diagnosis on the basis of his research, and then tries a therapy with varying degrees of confidence and enthusiasm. If the patient improves, the diagnosis has probably been correct and the therapy effective, although the patient could also improve even when

the medical evaluation is incorrect. But the conscientious clinician asks the right questions, obtains his answers accurately, and draws the proper conclusions, both from the individual patients he sees and treats and from the medical literature. His subsequent practice is guided by the experience he gathers from both sources.

The practice of medicine is in effect the conduct of clinical research, in which questions are asked and new facts are obtained, synthesized, analyzed, and acted upon. Every practicing physician conducts clinical trials daily as he is seeing patients. The research discipline known as the "clinical trial" is the formalization of this daily process; clinical trials provide the means for obtaining the most reliable answers about alternative medical therapies and their most appropriate application.

In the last thirty years the randomized control trial has become the ultimate means of applying the scientific method to the practice of medicine. A major purpose of this dissertation is to demonstrate that, whenever there is uncertainty about proper diagnosis and therapy, scientific clinical trials are the most ethical way to benefit both the individual patient and all others. Good practice is ethical research. Poor research is unethical practice.

It is customary to consider four phases of the clinical trial of new therapies in man:

Phase I: The first time a therapy is administered to a human being. Extensive earlier research has almost always been conducted in animals. With research on drugs, the subjects are usually normal people, but operations are tried for the first time on human patients who have the specific disease to be treated.

Phase II: Early trials in patients with a specific disease. Traditionally these have been uncontrolled, although a case is made below for having controls from the very beginning of new therapies.

Phase III: Large-scale comparative trial of a new therapy with the old.

Phase IV: The use of the therapy in practice with some monitoring of its outcome.

Phase I is the only stage of this process that is not a part of the practice of medicine, i.e., that does not involve sick people who need to be treated. In phases II and III, the patient requires treatment and volunteers are needed to provide controlled comparisons. In phase IV a more systematic recording of data is added to ordinary practice.

History of the Clinical Trial

In the documented history of the clinical trial, two designs dating back to the eighteenth century have been contrasted to demonstrate the right and the wrong way to try out a new therapy (Thomas, 1969). In 1747 Sir James Lind divided twelve sailors with scurvy into six groups:

Two of these were ordered each a quart of cyder a-day. Two others took twenty-five gutts of elixir vitriol three times a-day, upon an empty stomach; using a gargle strongly acidulated with it for their mouths. Two others took two spoonfuls of vinegar three times a-day, upon an empty stomach. . . . Two of the worst patients . . . were put under a course of sea-water. . . . Two others had each two oranges and one lemon given them every day. . . . The two remaining patients, took the bigness of a nutmeg three times a-day . . . made of garlic, mustard-seed, rad. raphan, balsam of Peru, and gum myrrh.

The answer was clear-cut: the two sailors who received citrus fruit were promptly cured, and the others were not. Nevertheless, there was a forty-year lag between the publication of this paper and the routine addition of lemons and limes to the ships' holds of the Royal Navy. In the same century Benjamin Rush, a famous physician in

TABLE 1
Types of Controls Used in Therapeutic Trials Reported in the *New England Journal of Medicine*

Year	Total No. of Trials	Type of Controls in Percent of Trials			
		None	Historical	Simultaneous	Randomized
1953	29	55%	10%	34%	0%
1963	35	54	6	31	9
1973	45	40	18	18	24
1974	55	44	22	7	27
1975	47	34	13	13	40
1976	36	31	0	25	44
1977	36	17	0	25	58
1978	46	35	2	17	46
1979	35	29	0	14	57

Philadelphia, treated yellow fever with extensive purges and bleeding and probably killed large numbers of people. But because he had no controls for comparison, he concluded from his clinical trial that he was benefiting his patients.

The first extensive use of the randomized controlled trial applied methods worked out by Fisher (1935) in horticulture to the trials of new drugs for tuberculosis after World War II (Medical Research Council, 1948). Since that time an increasing number of trials reported in clinical journals have employed randomization, as shown in Table 1. In 1962 the Food and Drug Administration required that new drugs must demonstrate efficacy by adequate and well-controlled investigations (Crout, 1977). Lately, trials have reached a peak of sophistication and, unavoidably, expense.

The Influence of Clinical Trials on the Practice of Medicine

An evaluation of the place of clinical trials in the practice of medicine requires data on how physicians respond to trials, especially when the conclusions differ somewhat from their own opinions. A disturbing picture was obtained in a review of four instances (Chalmers, 1974). In an effort to prevent spontaneous abortions and perinatal complications, stilbestrol was administered to hundreds of thousands of pregnant women every year for at least ten years after five trials with simultaneous controls had found the drug to be ineffective. Use of the drug did not stop until after carcinoma of the vagina was described in the offspring of women who had been so treated. In that case, poorly controlled trials were believed more than the well-controlled ones; an element of the "it can't hurt" philosophy also prevailed. Doctors, in effect, were responding positively to advertisements by the drug companies even though the textbooks had concluded that the drug lacked efficacy.

There has been no pharmacologic treatment for acute infectious hepatitis, but because patients felt fatigued physicians prescribed rest—even keeping patients in bed for long periods of time. Years after a total of three trials have shown that bed rest lacks efficacy, some physicians still play it safe and prescribe bed rest, even though bed rest itself has its deleterious side effects.

A number of clinical trials have shown that composition of the diet has no effect on the healing rate of peptic ulcer. Yet a number of physicians continue to prescribe modifications of the "sippy" diet, which consists of frequent feedings of milk and cream, despite its possible deleterious effect on the arteries.

The most notorious clinical trial with regard to its impact on the public sector is the study of the University Group Diabetes Program (UGDP) (1970), which (Cornfield, 1971) showed that oral hypoglycemic agents, administered to approximately one million diabetics per year to normalize their blood sugar and decrease the complications of their disease, actually resulted in an increased frequency of death from cardiovascular disease. This study met with such strong opposition from diabetologists, who had found the drugs useful in their practice, that their lawyers have blocked the inclusion of any reference to the study in the package insert required by the Food and Drug Administration. Fifty-five experts have written opinions of the study since its publication in 1970, 27 generally against it and 28 in favor. It is of interest to note that the possibility of conflicts of interest (bias) exists in the interpretation of these different opinions. Those opposed to the study and in favor of the drugs were significantly more likely to have had their past research supported by the manufacturers of the oral agents. No one has as yet repeated the UGDP trial, possibly because of the costs and the difficulties that would be encountered obtaining human subject approval and informed consent. Yet doctors continue to attack the study design of the UGDP, and to prescribe the drugs in vast quantities.

It is of some interest that the original study showing that anti-hypertensive drugs reduced the death rate from the complications of high blood pressure was of no better quality than the UGDP study, but was accepted universally by physicians and has been followed by many more studies confirming the effect. In this case a preconceived notion was confirmed.

The reluctance of physicians to accept the results of clinical trials when the conclusions are contrary to conventional wisdom, no matter how well the trials have been designed, may gradually disappear as physicians become better educated in clinical trial methodology. The physician who is going to act on the result of a clinical trial needs to be able to evaluate the quality of its methodology.

Criteria of a Good Trial

To appreciate the clinical, ethical, and fiscal considerations of a reliable clinical trial, it is important to summarize the requisite criteria, which may be listed under three major headings.

The Control of Bias

The major advance of clinical trial methodology in this century has been the recognition that unconscious bias may distort the outcome of a trial.

The traditional method of countering the effects of bias is the "double-blind" approach—disguising the therapy so that neither the patient nor the observer knows what therapy is being given, the experimental or the traditional. About 50 percent of trials are still not carried out in double-blind fashion, sometimes because the investigators are intellectually lazy or don't know enough to do a study properly, but often because it is impossible to conceal the therapy from either the patient or the observer. However, it is still important to attempt to disguise some of the interpretations in order to minimize observer bias.

Observer bias in medical research is the unconscious distortion of observations or data as a result of preconceived notions. Conscious distortion, of course, is fraud, and should be treated accordingly. To understand how important observer bias may be one must appreciate the potential magnitude of observer error. The classical example of this was the demonstration many years ago that when two skilled radiologists read the same X-ray their interpretations differed significantly around 15 percent of the time. In fact, the error was almost as large when the same physician reread the same X-ray several months later. This 15 percent frequency of observer error has been found to occur with surprising regularity whenever it has been sought: in observations of the skin, endoscopic observations, electrocardiographic interpretations, and in the reading of tissue biopsy preparations, for example.

Observer error and observer bias can be illustrated by a simple clinical experiment. When a heart is beating fast, as occurs in a syndrome called atrial fibrillation, the rate is irregular; consequently,

the magnitude of the pulse also varies because of the variable time available for the heart to fill with blood between beats. When the heart is beating about 200 times a minute, different skilled observers will obtain rates from 180 to 220 by counting through the stethoscope. If the pulse is taken at the wrist, however, these observers will find a rate of 80 to 120. Now divide the observers into two groups and have one count the apical rate first and the other the rate at the wrist. The difference between the mean rates at the two locations will be half as large for the group who listen to the apical rate first; observer bias leads them to anticipate that the wrist should show a faster heart rate. Conversely, the group who take the pulse at the wrist first will hear a slower apical rate afterwards, but the difference will be less because observer error is less at the apex, so that there is less chance for observer bias to occur.

In biological research observer error and bias can never be eliminated; they can only be minimized and recorded. In the clinical trial more than simply the outcomes of comparative therapies have to be blinded if this is to be accomplished.

In clinical trials in which patients are not randomly assigned, the difference between the experimental and the control treatment is almost invariably larger as a result of uncontrollable observer bias (Chalmers et al., 1980). The physician assigning a patient to a new or standard therapy is caught in the moral dilemma of having to conduct research while doing what he thinks is best for the individual patient at the moment. Ethical justification for the randomized assignment of patients is discussed in detail below. Suffice it to say here that the process is essential if misleading results of the experiment are to be avoided. Nor is it sufficient merely to employ a system of randomization such as a list of random numbers or the flipping of a coin, because knowledge about which treatment is about to be assigned can influence whether or not a patient is deemed suitable to enter a trial, as well as how hard the physician tries to persuade a reluctant patient to volunteer. The randomization process must be properly blinded by the blind envelope technique, or by the pre-assignment of numbers in the pharmacy that prepares the drugs to be tried.

Rarely do all the patients entered into a trial complete the whole course. Some patients are withdrawn because one of the prescribed therapies is no longer considered suitable; some withdraw because

they are tired of the trial and want to try something else; and some just disappear. Patient data are discarded because some reason is found after the trial to disqualify certain subjects. All of these contingencies require decision-making by the investigators, and offer the opportunity for additional biases to distort the results of the trial. The decision about how vigorously to pursue a lost patient, to remove one, or persuade one not to drop out should be made by someone who does not know to which therapeutic group the patient has been assigned. Only in this fashion can bias be reasonably controlled.

The tentative results of an ongoing study can distort or ruin a trial if it influences the clinician investigators with regard to the admission of new patients or the removal of patients, especially in the 50 percent of situations when adequate blinding of therapies is impossible. For this reason it is important that the trends in a trial be hidden from the physicians who are making the ongoing clinical decisions. Data should be monitored closely by skilled people who are not involved in the recruiting and care of the participating patients.

In effect, a good clinical trial is quadruple-blinded rather than merely double-blinded: randomization, therapy (both to the patient and the physician), and the ongoing results.

Recording and Analysis of Data

The second major distinction between the good practice of medicine and the scientific clinical trial is the increased need within the trial for detailed recording of data. It is probable that most clinical trials waste a lot of money by recording data that are never used. But the investigators never know when the protocol is written that some commonplace observation might not be crucial in detecting or interpreting the results of the trial. So the error is acceptably made on the side of recording too much rather than too little.

Analysis of the data, as well as the design of the study on which the validity of the analyses depends, requires the expert skills of a biostatistician. Few clinical investigators are sufficiently versed to handle the statistics alone. Ability to perform a χ^2 or a t test without a thorough knowledge of the pitfalls and principles on which the tests are based creates a real likelihood for distortion of the results of a trial. In an extensive evaluation of the quality of trials of the treatment of duodenal ulcer there was a statistically significant correlation with

the presence of a biostatistician as an author or as a credited helper (Silverman et al., 1978).

The Numbers Problem

Many clinical trials are inconclusive or suggest the wrong conclusion because the numbers involved are too small, or the wrong question was asked in view of the numbers available for study. All clinical investigators and most practitioners are familiar with the concept of Type I error, or alpha, the likelihood that an observed difference is due to chance. But many clinicians miss the importance of beta, or Type II error, that is, the chance that a difference of interest might be missed because not enough subjects were studied. They almost universally ignore it in the protocol section of their papers and in the interpretation of so-called negative results (Freiman et al., 1978). Estimating beta requires a knowledge of the event rate to be expected in the control group as well as the acceptance of a difference not to be missed. When the rate in the control group is very small or very large, and the difference of interest relatively small, the numbers required to be sure that that difference is not being missed will be very large. Conversely, if the control rate is around 50 percent and the difference of interest about 50 percent of that, i.e., 25 or 75 percent, the numbers involved will be small. Of the 71 so-called negative papers in which data were recalculated, only 14 had sufficient numbers to warrant the negative statement of the authors, i.e., to rule out a 25 percent improvement.

The problem of Type II error is a real handicap in the conduct of clinical trials (see the discussion of cost considerations, below). For example, the rate of recurrent myocardial infarction each year is about 4 percent. These are patients who have had one previous attack. A reduction of only 10 percent in that rate is very worthwhile, because the 4 percent figure represents several hundred thousand people each year. But a properly sized study to detect that kind of reduction requires over 10,000 patients in each group. More common endpoints, such as subclinical changes in electrocardiograms or coronary blood flow, or exercise tolerance tests, reduce the number of subjects required, but the "softness" of these endpoints when compared with death reduces their utility.

Ethical Considerations

One possible explanation for the fact that so many therapies are used in practice despite the lack of adequate documentation of their efficacy, while relatively few patients are included in randomized control trials, is a prevalent belief that there may be something unethical about the latter. The decision to employ a treatment on the basis of chance rather than clinical judgment could be considered as an anathema by the conscientious physician if he is not critical enough about the validity of the information on which he ordinarily relies.

A number of procedures designed to handle satisfactorily the ethical issues related to clinical trials have evolved through the years. In general, these issues—which are reviewed below—have prompted those responsible for supporting and conducting research to try to protect the patient participating in that research from an overeager investigator who might be biased with regard to known efficacy-toxicity ratios.

Peer Review

The techniques of peer review of protocols for therapeutic trials have developed into highly efficient procedures as a result of the national concern for the protection of human subjects involved in biomedical research. The responsibilities of peer review committees include the assurance that the research methods are appropriate to the objectives of the research and the field of study.

At the Mount Sinai Medical Center 350 protocols are approved each year. The parent committee has 17 members of the faculty and 5 laymen (currently including a rabbi, a lawyer, and 3 other members of the community with past training in social work, psychiatry, and child care). There are six subcommittees that do the first review of applications—one at each of two affiliate hospitals and four at Mount Sinai—with expertise appropriately divided among those four. Although the department chairmen must approve all protocols for human safety and scientific quality before they are submitted for institutional review, the subcommittees look at both processes with great care. In the course of last year 82 percent were approved without qualifications, 11 percent had questions to be resolved before final approval, and 7 percent were deferred for a total rewrite. All research conducted in

the institution, whether funded by the Department of Health, Education, and Welfare or by another outside donor, is reviewed by the research office. Projects that do not involve humans are also reviewed for conformance to safety rules because of the possibility that they may create some hazard to technicians or bystanders. Studies involving animals must conform to nationally approved animal care policies.

Informed Consent

Much is written about how detailed the process of obtaining informed consent should be, and opinions vary from the requirement of most lawyers and ethicists that every single possible alternative outcome be presented to the patient, in writing, to settling for evidence that a dialogue on efficacy and safety has occurred between investigator and patient. Some practicing physicians participating in research feel that the patient is often made more sick by details included in the effort to obtain permission for the performance of clinical trials. This is especially applicable to patients with neoplastic diseases. At Mount Sinai the pros and cons of the protocol are presented verbally to a patient by the investigator and one professional witness, who takes into consideration the therapeutic needs of the patient. The patient then signs a form stating that the procedure has been explained and has been fully understood. The consent form and a brief description of what will be said to the patient are included in each protocol approved by the committees, and a report is filed with the patient's chart.

When Should Randomization Begin?

It is customary to carry out the first phases of an investigation of a new therapy in sick patients without using randomized controls. There is, however, a serious ethical issue involved with this custom.

Once a therapy reaches a stage in which it is administered with the hope that it may help the patient, it is still possible that the patient could be harmed rather than helped. The eventual abandonment of many new therapies is testimony to this. Therefore, if the patients need therapy, it would be more ethical to begin randomization at this stage and give them an equal opportunity of receiving the standard therapy in view of the possibility that it could turn out to be superior to the new one, especially when the exact dosage or other details of the regimen may have not yet been worked out. Most

investigators do not want to randomize early because they fear that in the preliminary stages of the investigation the new treatment may not be perfected enough to fare well in comparison with the standard treatment. What are the ethical considerations of that decision, and how is the informed consent requested?

The classical example of this dilemma is the introduction of a new operation for a serious disease. Surgeons don't like to randomize at that stage because of the fear that the operation has not yet been adequately perfected (Bonchek, 1979). The surgeon usually does a consecutive series of patients before submitting the new operation to a randomized trial. How do the peer review committee and the volunteering patient handle the information that randomization is being avoided because the operation has not been perfected well enough to compare favorably with the standard form of treatment? If the procedure is the same as that required for randomized trials, the study could not be performed without a consent form detailed enough to dissuade all patients from volunteering for new operations. The only answer to this ethical dilemma is to begin the randomization process as soon as the new therapy is initiated for the purpose of evaluating the response of the patient (Chalmers, 1975).

When to Stop a Trial

This problem has ethical implications fully as serious as those occurring at the start of a trial. Randomization is justified solely on the basis that it is not known whether the therapy in question might be better, as good as, or worse than the standard in a given patient. In the trial in which one therapy turns out to be better than the other, the uncertainty becomes less and less likely (Chalmers et al., 1976). Suppose, then, that patients are admitted in pairs and the endpoint being measured is only the question of survival. If one of the treatments has won most of the pair comparisons, and the P value is .06, so that only one or two more pairs are necessary to prove by acceptable standards that one treatment is better, what should the patients be told at this stage? If they could know the data, they would obviously prefer not to be in the study but to receive the "better" treatment. At that stage there is practically no chance that the therapy could be significantly worse. In those circumstances the study could not be continued until the customary 5 percent probability of error has been achieved. When should it be stopped? Actually, the probability is

no longer 50–50, once the first pair has been decided one way or the other. Should the trial be stopped after that? There is no answer to this ethical dilemma other than referral to the phenomenon in protocols and informed consent procedures, and deferring decision about when to stop studies to third parties, such as data-monitoring committees or policy advisory boards. Such boards are now routine in large studies and include biostatisticians and ethicists as well as specialists in the disease process being studied.

The data must be looked at in an objective and unbiased way during the course of the trials because unexpected phenomena may appear or because other studies may be reported that can influence the conduct of the ongoing one. Procedures for discontinuance should have been considered in detail in the protocol so that the board does not have to make critical decisions in the middle of the study. But important outcomes can never be entirely anticipated. For instance, lack of decided therapeutic advantage in the presence of an unanticipated side effect may lead the board to decide to abort the study or one arm of the study before conventional biostatistical significance is achieved.

The Placebo

Much is written, some of it misinformed, about the ethical aspects of using a placebo in biomedical research. The placebo is used in a clinical trial in order to facilitate the blinding and control of bias as detailed above. It never implies depriving a patient of a known effective therapy. The placebo disguises the new or the standard treatment and consequently diminishes bias.

Another ethical issue occasionally intrudes on the decision-making process: when to abandon no-treatment controls or standard therapy as a control when a new treatment has clearly been found to be better. This is a complex issue because it involves clinical judgment about when data results become clear-cut and when they are still equivocal. The determination depends on the controversial technique of combining the results of diverse clinical trials. The opportunities for bias in these decisions are easy.

The best example of this problem is the clinical trial of antibiotics in preparation of the colon for surgery (Anish et al., 1980). In 1971 it was apparent that antibiotics reduced postoperative infections; by 1975 it was certain. Yet investigators continued to use no-treatment

controls, sometimes disguised as placebo, in papers reported as late as 1979 and presumably started after 1975. Possible explanations for this phenomenon involve a lack of confidence in or knowledge of previous results, as well as a need to demonstrate more positive differences. In this case it could be said that the placebo was properly used to disguise an improper therapy.

Cost Considerations

It took few patients and little money to demonstrate that penicillin was effective in lobar pneumonia, vitamin B¹² in pernicious anemia, and insulin in diabetic coma. Other instances of dramatic efficacy have also been determined in a small number of patients without adequate controls. But the great majority of new therapies are not dramatically effective and may have only minor advantages over standard treatments or no therapy at all. A good example is the use of antiplatelet drugs by patients who have had one myocardial infarction, in the hope that they will prevent a second and perhaps even a first attack in susceptible patients. A number of very expensive trials have been done, and there is a suggestion of a very small effect. A recent trial of aspirin and persantine, involving 2,000 such patients, cost \$5 million and gave an equivocal answer (Research Group, 1980b). A repeat trial with larger numbers, where therapy is administered earlier after the original infarct, is about to begin and will cost \$8 million. In this case a pharmaceutical firm is paying for the study; if it should happen to be positive, the company will get its investment back within a year or two. This is the first time, however, that a pharmaceutical firm has supported a large-scale therapeutic trial without having any influence whatever on the decision-making process. Most clinical trials have had to be supported, at great expense, by the federal government. A recent trial of aspirin alone cost \$17 million and revealed the drug to be ineffective, in contrast to six or eight other studies that have shown a slight benefit (Research Group, 1980a). All too often, results are unclear and costs are enormous. Such studies are no longer popular when government funds available for biomedical research are so limited.

However, this is bad economic thinking. If antiplatelet drugs should happen to reduce the rate of recurrent myocardial infarction by 10

percent, the number of infarcts would be approximately 20,000 less in the United States each year. Considering that \$10,000 in direct and indirect costs are accrued for each attack, \$20 million would be saved per year. One year's savings would finance two large-scale studies.

If large-scale clinical trials are to be carried out to discover meaningful but small therapeutic differences, the funds have to be found from somewhere other than the limited moneys available for fundamental biomedical research. The American public spent nearly \$200 billion on health care in the year 1979. All good businesses spend at least 5 percent of their gross income on developmental research. The costs of medical care could be reduced if research and development funds, specifically money for clinical trials, came from the general medical care budget rather than from the 2 percent that is devoted to all biomedical research.

The Future of Clinical Trials

The great cost and the equivocal results of many trials have led to a resurgence of interest in the use of historical controls and other less scientifically valid methods of obtaining useful information. How strange that some should think that the smaller the difference expected, the less elaborate the control procedures required, when exactly the opposite prevails. Large-scale as well as small trials must be conducted properly; when they are well designed, the payoff in financial benefit as well as improvement of health will be enormous.

References

- Anish, D.S., Baum, M.L., Chalmers, T.C., et al. 1980. Prophylactic Antibiotics in Colon Surgery: The Use and Misuse of No-Treatment Controls (Abstract). *Clinical Research* 28:540A.
- Bonchek, L. 1979. Are Randomized Trials Appropriate for Evaluating New Operations? *New England Journal of Medicine* 301:44-45.
- Chalmers, T.C. 1974. The Impact of Clinical Trials on the Practice of Medicine. *Mount Sinai Journal of Medicine* 41:753-759.
- . 1975. Randomization of the First Patient. *Medical Clinics of North America* 59:1035-1038.

- , Lasagna, L., et al. 1976. How To Turn Off an Experiment. In Cooper, J.D., and Ley, H.L., eds., *Ethical Safeguards in Research on Humans*, 119–143. Washington, D.C.: Interdisciplinary Communications Associates.
- , Sacks, H.S., and Smith, H., Jr. 1980. Historical vs Randomized Controls for Clinical Trials. *Controlled Clinical Trials* 1:177.
- Cornfield, J. 1971. The University Group Diabetes Program: A Further Statistical Analysis of the Mortality Findings. *Journal of the American Medical Association* 217:1676.
- Crout, J.R. 1977. The Nature of Regulatory Choices. Presented at annual meeting of the Association of Military Surgeons of the U.S. Washington, D.C., November 27.
- Fisher, R.A. 1935. *The Design of Experiments*. London: Oliver and Boyd.
- Freiman, J.A., Chalmers, T.C., Smith, H., Jr., et al. 1978. 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. *New England Journal of Medicine* 299:690–694.
- Medical Research Council. 1948. Streptomycin Treatment of Pulmonary Tuberculosis. *British Medical Journal* 2:769–782.
- Research Group, Aspirin Myocardial Infarction Study, 1980a. A Randomized Controlled Trial of Aspirin in Persons Recovered from Myocardial Infarction. *Journal of the American Medical Association* 243:661–669.
- Research Group, Persantine-Aspirin Reinfarction Study. 1980b. Persantine and Aspirin in Coronary Heart Disease. *Circulation* 62:II, 1–42.
- Silverman, B.L., Shareck, E.P., and Ambroz, A. 1978. Duodenal Ulcer Treatment: An Analysis of the Randomized Control Trials (Abstract). *Gastroenterology* 74:1174.
- Thomas, D.P. 1969. Experiment versus Authority: James Lind and Benjamin Rush. *New England Journal of Medicine* 281:932–934.
- University Group Diabetes Program. 1970. Mortality Results. *Diabetes* 19 (Supplement 2):1.

The work in this paper was supported in part by NIH Grant LM 03116 from the National Library of Medicine.

Address correspondence to: Thomas C. Chalmers, M.D., Office of the Dean, Mount Sinai School of Medicine, City University of New York, Fifth Avenue and 100th Street, New York, N.Y. 10029.