

## DESIGN AND MANAGEMENT OF LARGE MULTICENTER CLINICAL TRIALS

Richard D. Remington

*School of Public Health  
University of Michigan  
Ann Arbor, Michigan 48109*

By the last half of the 1960s, clinical scientists and administrators were agreed that pharmacologic intervention in hypertension had progressed through its developmental stages. A number of relatively safe and individually effective drugs had been developed—effective in the sense that they produced a substantial and relatively consistent reduction of blood pressure in hypertensive patients. These conclusions had been established through animal studies, small human trials, and larger controlled studies in individual clinics and laboratories.

It remained to be shown, however, that intervention with blood-pressure-lowering drugs could reduce the major late effects of hypertension: premature mortality and increased risk of cerebral vascular disease, congestive heart failure, and renal disease. That elevated blood pressure carried with it the hazard of such late consequences had been well established from actuarial statistics, from longitudinal community surveys, and from clinical observation. There was a strong presumption that drug treatment was effective in this respect, but no proof. Clinicians were beginning to comment upon the disappearance of malignant hypertension, for example.

The United States Veterans Administration, with its network of medical care facilities extending across the U.S.A., had a long history of success in the design and conduct of cooperative, multicenter clinical trials, dating from its landmark investigations of chemotherapy for pulmonary tuberculosis. The VA's relatively tight administrative apparatus made it a natural locus for such development. Its Cooperative Studies Evaluation Committee provided a central review mechanism applicable to a broad array of proposed studies. Thus, it was natural that the Veterans Administration would decide to establish a tightly designed, controlled, multicenter trial of the late effects of antihypertensive medication. Dr. Edward Freis assumed major responsibility for the organization and development of protocols, and the applications passed internal review. The rest is history. Scientists will be discussing the results of this landmark investigation for years to come. The trials established that combination therapy for moderate and severe hypertension produced a substantial reduction in all-cause mortality and major morbidity, at least in middle-aged male U.S. war veterans. Remaining at issue were the effects in women, in mild hypertensives, and in less highly selected segments of the population. Furthermore, effects on coronary heart disease were unclear and remain controversial.

At a critical stage in this history, a decision had been made that the cooperative, multicenter clinical trial was the investigative tool of choice for answering an important question. Why and how are such decisions made? Although the Veterans Administration, the National Cancer Institute, the National Heart Institute, the armed forces, and other groups here and abroad had developed and applied this methodology in the 1950s, to this date little attention has been

given to the important public policy issues involved. Such trials are expensive. They demand substantial financial resources. They preempt the skills and energies of large numbers of investigators over relatively long periods of time. They require attention to a central theme, protocol, and detailed manual of operations, thereby placing constraints on individual investigative initiative. They are highly visible, in success and failure. They can produce conclusions or controversies, the latter sometimes extending over a period longer than that required to conduct the trial itself. Agencies, legislative bodies, and the general public need a better specification of the public policy issues involved in the decision to invest in such major undertakings.

At a minimum, a cooperative trial should not be launched in the absence of a substantial question of therapy or intervention. While many of us would agree with Dr. Chalmers' admonition that we should "randomize the first case," reminding us that as therapeutic habit becomes entrenched, it is more and more difficult to conduct valid scientific evaluation, this admonition does not apply to the cooperative, multicenter trial. Before such a trial is seriously contemplated, we should find ourselves at a therapeutic crossroads. Our need to know the efficacy and safety of a particular mode of intervention should be acute and should be impeding further scientific advance. It should be clear that a positive versus a negative answer would result in a marked change in investigative direction and therapeutic management of the typical patient. Major ethical issues should have been solved and the treatment should be acceptable for patient management.

The cooperative multicenter trial provides the opportunity to reach an answer to a substantive clinical question over a reduced period of time. Pooling of resources makes possible the enrollment of larger numbers of patients than could occur within any single clinical setting. While standardization and uniformity of approach are essential, conduct of therapeutic investigations in a variety of settings broadens the scope of scientific inferences. The involvement of a number of investigators working toward answers to the same questions strengthens the likelihood of a favorable result, while diminishing the opportunity for oversight and error. Nevertheless, this multiplicity of investigators also produces an averaging effect resulting from negotiation and compromise. Thus, the cooperative trial becomes a relatively blunt instrument, most appropriate for answering broad rather than highly detailed questions. While disease mechanisms may be elucidated in such trials, these results most often flow from ancillary studies off the main stream of the cooperative investigation. Finally, a collaborative trial should probably be contemplated only when investigating a disease process with major sequelae such as mortality or major morbidity.

How many cooperative, multicenter trials are needed? This question cannot be answered in any detail because of the public policy hiatus noted earlier. Will the conduct of such trials act as a stimulus or as a deterrent to other biomedical research? Is the "full bucket" model, which assumes a stable number of dollars for the entire national health research effort, appropriate? This model states that allocation of a major number of dollars for one purpose results in fewer dollars for other purposes. Does such research have a stultifying effect on investigative initiative? Does such research constitute bureaucratic "targeting" in which investigators have little or no role in following scientific leads of their own? Or alternatively, does such research have a stimulating effect on the entire national biomedical research effort, increasing public awareness and, in

turn, producing a greater emphasis upon this sector of the national economy? Do agencies adequately involve broad scientific opinion before launching such investigations and in their design? Do such investigations attract young people who might otherwise devote themselves entirely to the private practice of medicine? It seems curious that these important issues have received so little systematic attention.

What is the cooperative multicenter trial? First of all, it is a substantial scientific question demanding a relatively rapid answer in a variety of settings. In addition, it is an organization, a group of people and facilities oriented toward a common goal. Here, I will attempt to describe a minimal organizational pattern. Specific trials may require additional organizational components.

A senior operating or executive committee, composed of investigators with a direct stake in the research, is essential. These individuals should be empowered, within broad limits established by the funding agency, to make operating decisions. While agency and other outside individuals may be members of this body, its primary focus should be on getting the work done, and, therefore, its primary membership should be workers.

A coordinating and data-collecting center should be established. In some trials, administrative coordination and data storage and maintenance are separated in two locations. However, all study data should eventually arrive at a central point for editing, storage, retrieval, monitoring, and analysis. At this coordinating center, senior statistical and data-processing personnel should be available. This unit will be the sensory apparatus of the study, determining that the agreed upon goals and detailed operating procedures are being followed uniformly. The center will have multiple reporting responsibilities. For primary operational considerations, it will report to the steering or executive committee, but it will also bear responsibilities directly to the funding agency and to monitoring bodies.

A group of clinical centers, directed by strong individuals who nevertheless agree to pursue common objectives using common methods is essential to the success of any cooperative trial. At each center, the principal investigator will be responsible for producing a set of observations that can be pooled with similar data from other centers. He or she will be responsible for recruiting an adequate number of patients, for supervising local personnel, and for operating the study in a manner consistent with agreed procedures.

Standardization is essential, and the opportunity for lack of standardization multiplies as the number of clinical centers increases. Biochemical observations should be controlled through the use of blinded reference laboratory techniques. Often, a central biochemical laboratory will be used to receive specimens, to produce individual determinations, and to certify through external monitoring the accuracy of its own methods. In addition, clinical observations such as electrocardiograms and x-ray readings may require central control facilities.

A cooperative trial requires an external data monitoring and review structure, i.e., a mechanism for accountability. This monitoring group should determine that toxic or other adverse effects are being carefully watched and are controlled within acceptable limits. This body should also review endpoint data to determine whether the trial should be stopped prematurely, either for ethical or scientific reasons. This committee will report directly to the funding agency, which of course has final authority to stop the flow of funds. The committee will receive appropriate data from the coordinating center, creating another organizational link. Conclusion of any trial is a difficult matter and a variety

of technologies must be brought to bear. These include formal statistical methodology, sometimes sequential or quasi-sequential in character and sometimes based on a fixed sample size. Other components of a stopping rule include ethical issues, issues of relationship between the study and the community in which it is being conducted, and issues of extension of the study beyond the planned completion date.

A number of specific design features must be agreed upon before the study is initiated. This will require a planning or pilot phase in which field testing of proposed questionnaires and investigative methods can be conducted. In the case of the Hypertension Detection and Follow-up Program, this period included a decision whether to use an automated blood pressure measuring device. In addition, sample size must be determined and the procedures for data flow completely specified.

Determination of sample size in a scientifically defensible manner is essential before the study gets under way. If the trial is to involve a fixed sample size rather than a purely sequential design (the latter a relative rarity among cooperative therapeutic trials), then several issues are at stake. A primary endpoint or endpoints must be established for design purposes. The frequency of this endpoint in the control or comparison group should be estimated. The amount of effect on the primary endpoint as a result of treatment should be considered carefully. The investigators should decide how large an effect would be important to the subsequent scientific history of management of the disease. For example, would a 25% reduction in mortality be important and worth pursuing? Operating probability levels should then be determined: the level of significance or probability of finding a difference of the desired magnitude when, in fact, no such difference exists in the population from which the patients are drawn, and the power of the investigation, i.e., the probability of detecting a true difference of the magnitude specified. In addition, the dilution effect of loss to follow-up or other inability to determine endpoints should be estimated. The greater the dilution, the greater the sample size. It should be decided whether separate subdesigns for efficacy and safety will be used. In some instances, it may be possible to design a quasi-sequential scheme for detecting adverse effects. In this event, however, many of the same decisions concerning event frequency and operating probability levels must be made.

Throughout this discussion, we have assumed that patients will be allocated to one of several groups, differentiated by treatment status. In the classical trial, patients receive either active drugs or apparently identical but inactive placebo medication, the allocation of patient to specific group being at random, i.e., with the use of a table of random numbers. Randomization and monitoring to see that such randomization has been adequately completed is the responsibility of the statistical coordinating center. In the classical trial, patients and physicians are blinded with respect to the specific treatment group. Provision is made for breaking the blind for reasons of patient safety according to a standard set of provisions described in the manual of operations.

Advanced planning for the cooperative trial is relatively more important than for the single-center study. Objectives must be relatively fully specified and priorities among those objectives rigidly set. While one or more endpoints may be classified as major design endpoints, other outcome measures will ordinarily also be investigated. Assessment of all endpoints in as objective and reproducible a manner as possible is critical to the successful outcome of the study. It is important, wherever possible, to specify hypotheses concerning

outcomes in patient subgroups in advance of the trial. Strong and largely unsolvable statistical problems exist if an attempt is made to "dredge" data at the conclusion of the study, in the absence of advanced specification of problems of interest. This is not to say that incomplete data analysis is desirable. However, the investment in a cooperative trial makes it mandatory to specify scientific questions in advance in order that findings will be dependable.

Because a trial is so costly, it is desirable, whenever feasible, and without compromising the major objectives of the trial, to seek to accomplish multiple purposes. Although, as noted earlier, the cooperative trial is a relatively blunt investigative tool, this bluntness is no excuse for failing to realize the maximum potential from a large investment. The initial pilot phase can assist in maximizing study potential.

The collaborative trial requires cooperation on a wide variety of levels. Individual investigators must to some degree subjugate their own initiative to the demands of the design. Workers from different disciplines—clinical, epidemiologic, statistical—must all collaborate, sharing their special expertise in the conduct of the trial. While democratic processes are essential to make sure that every voice is heard, a final decision-making apparatus is necessary in order that the study may proceed smoothly toward its conclusion. The objectives of any such trial are primarily in the area of patient management and care. If experienced statisticians and computer technologists are available, there is seldom a reason for conflict between the quantitative demands of design and analysis and the reasonable desires of the clinical investigators. Trial ethics is a highly individualized matter, and investigators active in the trial must be convinced that they are embarked on a fully defensible ethical course.

Because of the ponderous nature of a clinical trial and the amount of inertia and momentum it develops, it is a relatively poor vehicle for major technical innovation in data processing or management. Such innovation has in the past occasionally produced unconscionable delays in releasing the data trapped inside large computer files. Again, the primary objective of such a trial is the improvement of our understanding of a therapeutic or intervention process, and technical or manipulative aspects must not interfere with this objective. Even so, when placed in proper perspective, such trials can and should provide important insights into improved data-management techniques for subsequent trials. In addition, the professional interests of statisticians and computer scientists must be recognized if high quality collaboration is to be expected. Ancillary technical studies can be helpful in this regard.

One of the administrative challenges of such investigations is the amount of "academic credit" available to investigators devoting a substantial portion of time to the study. Publications will often be produced at a relatively slow rate during the course of a cooperative trial, although it is important to plan the publication program concurrently with initiation of the trial. Rights of authorship and individual citation are important to individuals, particularly young scientists whose future is at stake. The devotion of a young scientist's full-time activity to collaborative research seems questionable, because of prevailing attitudes of administrators and faculty peers considering promotion and advancement. These are examples of administrative problems attendant to the management of clinical trials. Too often, a trial is begun without careful consideration of such issues.

In many respects, the most difficult part of collaborative research is ending it. There always seems to be just one more question that could be answered. It

seems a terrible waste to dismantle an efficiently functioning administrative structure. For this reason, cooperative trials tend to end with a whimper rather than a bang, or like old soldiers, simply fade away. As we move into trials involving direct participation by community groups and direct intervention into the public health, we can expect even more problems in phasing out such studies. When an improved process for delivering health care to the population is established as the result of a trial, particularly difficult questions of discontinuing delivery of that improved process to participants arise. While we have not completely solved such problems, it is important to involve community groups, investigators, and the funding agency in such considerations from the beginning of the program. Abrupt withdrawal of an effective mode of intervention will leave justifiably angry people and communities in the wake of the study.

Finally, the trial must end with an intellectual question mark. It should pose this question mark as carefully as possible in order that other investigators in other places can pick up the investigative thread and continue progress toward full understanding of the disease process and its management. While that distant goal will probably never be reached, the cooperative trial has a particular investigative obligation strongly to suggest next steps.

Not every question requires cooperative study methodology. However, as we develop a variety of safe and effective treatments, we can expect more and more cooperative trials, particularly ones designed to improve delivery of improved care to the general public. In this latter area, we are truly in our infancy, but trials like the Hypertension Detection and Follow-up Program constitute important first efforts. As we move into these areas of direct concern for the public health and its improvement, we can expect the public policy questions mentioned earlier to increase in importance. We should move promptly to establish a dialogue leading to an operating consensus on these issues.