Comment

Stephen Lagakos

As Professor Fleming notes, leadership from statistical scientists is crucial in the design, conduct and interpretation of therapeutic intervention trials in major diseases such as AIDS, cancer and cardiovascular disease. For a disease such as AIDS, where many of the natural history and pathogenesis issues are still not well understood, yet where there is extreme pressure to develop effective therapies, this leadership is especially important. The statistical scientist must on one hand be familiar with the practical issues surrounding the conduct of a therapeutic trial, yet also have the technical skills to develop creative designs and methods of analysis that can address the variety of challenges present in these studies. The number of statisticians who are currently filling this dual role is very limited. Fortunately, Professor Fleming is one such individual and hence is ideally suited to relate some of the key issues in this area to his statistical colleagues.

I will focus on three of the issues raised in Professor Fleming's paper—data monitoring committees, group sequential analyses and surrogate markers—and attempt to raise additional points related to their roles in therapeutic trials, especially in AIDS, where I have had the most experience.

DATA MONITORING COMMITTEES

In the AIDS Clinical Trials Group (ACTG), which has conducted the largest number of federally sponsored trials in persons infected with HIV, a data monitoring committee is used in virtually all of the comparative (Phase III) trials. This committee is composed of clinicians, ethicists and statisticians. The only other individuals present at meetings (and with access to the interim results of a trial) are statisticians from the ACTG Statistical Center who conducted the analysis and representatives from the National Institutes of Health. Representatives from the pharmaceutical companies whose drugs are being tested, as well as the clinical investigators who are heading the trial, do not have access to the interim efficacy results.

When I first became involved with AIDS clinical trials, I was unconvinced about the need for and value of data monitoring committees. Furthermore, the restriction of information to a very small group that excluded the medical head of the study struck me as unnecessary and rather silly. However, over the years my attitude on the value and composition of a data monitoring committee has changed dramatically, and I now strongly support both the concept and current ACTG model because I believe that these help to ensure the scientific success of studies.

In my view, the value of a data monitoring committee is that it can provide an independent and unbiased "second opinion" about whether and how a study should be modified. In the ACTG, this advice is taken very seriously and has on many occasions led to improvements in studies that had either not occurred to those conducting the study or were a result of the data monitoring committee's support of a change recommended made by the investigators conducting the study.

The issue of who should have access to the interim results of a study is more complicated and can depend on many factors, including the disease area being studied. One rationale for excluding the medical head of a clinical trial from the interim results is that if this individual is treating patients in the trial, knowledge of the interim results may compromise his or her ability to unbiasedly evaluate the patient's response or may cause an ethical dilemma if the interim results show a possible "trend" favoring one of the treatments. Another reason is that this individual may be involved as a consultant with the pharmaceutical company that produces the drug used in the trial; clearly, this could make it difficult for him/her to effectively advise the company and at the same time not compromise the ongoing study.

Similar arguments about access to interim results apply to representatives from the pharmaceutical companies who produce the drugs being evaluated in the trial. Here, however, the company's desire for access to interim results may be due to their larger plans for development of this and other drugs. Because our ability to ultimately benefit individuals who are afflicted with the diseases being studied in these trials depends heavily on the licensure of drugs shown to be effective in trials, it becomes critical to explore ways of possibly modifying the current use of data monitoring committees that might be more responsive to the needs of sponsors, yet do not compromise our ability to successfully complete these trials. I note, however, that what might work best in one disease area, such as HIV/AIDS, may not be best in another, such as cancer.

Stephen Lagakos is Professor, Department of Biostatistics, Harvard School of Public Health, 677 Huntington Avenue, Boston, Massachusetts 02115.

GROUP SEQUENTIAL ANALYSES

In addition to the need for methods of monitoring multiple measures of therapeutic effect, more research is needed on methods of analyzing trials with three or more treatment arms. In a traditional two-arm trial, the determination that one arm is inferior to the other is tantamount to terminating the trial. Yet in a trial involving three or more treatment arms, the possible decisions to be made at each interim analysis are greater in number and in complexity, and will depend on the overall goals of the trial (e.g., to exclude clearly inferior treatments or to identify the single best treatment). Professor Michael Hughes, a colleague in my department, has made some important initial steps in this area, but additional approaches are needed.

SURROGATE MARKERS

In the field of HIV/AIDS, there is great interest in identification of "surrogate markers." Usually, this refers to a laboratory marker that can be used as a surrogate for a clinical outcome such as the development of an AIDS-defining opportunistic infection. The potential value of a valid marker is obvious, yet, as Professor Fleming notes, use of an invalid marker could lead to the widespread use of ineffective drugs and/or the non-use of effective drugs. Thus, the validation of markers becomes very critical. Some of the early investigations of surrogate markers in HIV/AIDS have attempted to determine whether a treatment's effect on a particular marker can fully explain its effect on clinical endpoints. And, in all cases, only part of the clinical effect could be explained by the effect on markers. In retrospect, this is not surprising because it is unrealistic to expect that any single laboratory marker could fully explain all of an AIDS drug's beneficial effect because of the complex nature of this disease and how it can be affected by intervention. Thus, it may be that a battery of several markers needs to be determined that collectively can explain most of a drug's beneficial clinical effect. Given the high cost of the assays that are needed to evaluate some of the virological and immunological markers in HIV/AIDS, the design of studies to assess the "surrogate marker" question becomes critical. More statistical research on this topic is urgently needed.

In closing, I would like to thank Professor Fleming for his excellent article.

Comment

Thomas A. Louis

INTRODUCTION

Professor Fleming has considerable experience in conducting clinical trials and serving on Data Monitoring Committees (DMCs). We are fortunate that he has prepared a debriefing. It reinforces the impact of statistical science potentiated by subject area expertise and of both technical and broad viewpoints. In complex applications, relevant disciplines must be represented, and statistics is central to the enterprise. As Fleming notes, we must provide strong and effective leadership. To do so, we must educate collaborators on the role of statistics and be educated on a study's scientific and clinical basis. Statistical philosophies, principles and methods (frequentist/Bayes, multiple comparisons, choice of tests and estimators) need to guide deliberations, but in the complex world of clinical trials absolute dictums are seldom appropriate.

Thomas A. Louis is Professor and Head, Division of Biostatistics, University of Minnesota School of Public Health, Box 197 Mayo, Minneapolis, Minnesota 55455.

Statisticians and other DMC members are truly on the line. Stopped trials are very difficult to restart, and the decision to terminate can essentially freeze out other, similar trials. Continuing a trial beyond what many think is a reasonable stopping point puts study participants at unnecessary risk and delays dissemination of important information. Ware (1989) and related discussion show the heat generated by these issues. Contrast this situation to analysis of a stable data base: investigators can analyze, reanalyze, critique other analyses and sustain the give and take for years or decades. A DMC must make important decisions in an acute time frame.

GENERAL DISCUSSION

Data Quality

Building trust with patients and clinicians being recruited for a trial often is, and should be, a sensitive negotiation. Of course, all stake-holders need to be convinced that the question is clinically relevant. Of equal importance is assurance that everyone's interests