as primary clinical trial endpoints. Although I fundamentally share his concerns, I would cite one important aspect of the AIDS context that is relevant to the use of such endpoints in trials designed to obtain marketing approval for new drugs. HIV-infected patients, unlike patients recovering from myocardial infarction or suffering from chronic granulomatous disease, will inevitably die of their disease within a short time relative to their otherwise expected remaining lifetime. The best we can hope for from current therapies is a modest to moderate prolongation of survival. In this circumstance, it does not seem inappropriate to accept a higher level of risk in deciding what therapies might be made available. Whether therapies that have only shown positive effects on early markers should be distributed in "expanded access" or "parallel track" programs, or whether the FDA should permit their manufacturers to market them, may be more of an economic than a scientific issue. Whatever mechanism is used, it will ultimately fall to federally funded research programs of the Public Health Service to mount trials that compare available regimens and move toward defining optimal treatment strategies for patients at various stages of disease. In these trials, it will be essential to study clinical efficacy—that is, physical rather than laboratory manifestations of disease—until and unless we discover markers that come much closer to meeting the Prentice criteria.

It is encouraging to learn of the innovative investigations by Fleming and colleagues of the potential use of the auxiliary information present in early markers of disease to strengthen evaluation of therapies when only limited long-term clinical data are available. As Fleming notes, the circumstances under which this type of approach will significantly add to our ability to assess treatments reliably are somewhat limited. Nevertheless, it would be of interest to test out such approaches in data sets in which the relationship between the surrogate and the "true" endpoint is fairly well characterized—for example, if S were blood pressure and T were heart attack or stroke. The problem is complicated in AIDS because there has been experience with relatively few treatments and therefore little data regarding the correlation between S and T in the presence of different therapies. If this correlation varies greatly according to the particular regimen being administered, it would be difficult to use this approach in any routine way.

In conclusion, I would like to congratulate Dr. Fleming for highlighting some of the issues biomedical statisticians are struggling with, and hope that his paper will inspire more statisticians to become actively involved in, and even leaders of, the process of planning and carrying out medical research programs.

Comment

Vern T. Farewell and Richard J. Cook

INTRODUCTION

In this paper, Dr. Fleming provides an excellent review of some current methodological problems facing health scientists involved in clinical trials. Some issues considered in detail are monitoring clinical trials, the analysis of equivalence trials, multiple endpoints and surrogate markers. We will remark on each of these in turn.

MONITORING

The examples cited clearly demonstrate the importance of a monitoring committee for moderate to large-scale sequential clinical trials. In particular, a specialized and centralized Data Monitoring Committee (DMC) for the AIDS Clinical Trial Group (ACTG) is discussed. Such a specialized monitoring committee has immediately obvious advantages. As more trials are passed through the DMC, the disease-specific knowledge gained from early trials can be applied to later studies.

In principle, there are a variety of other diseases that require DMCs. For fields with less trial activity and experience, it may be advantageous to provide access to less specialized DMCs. Although it may be necessary to supplement the available expertise for individual trials, this more general DMC could provide statistical expertise on monitoring and advice on termination to a wide range of clinical investigators. Such a committee, perhaps under the sponsorship of a funding agency, would help to make the most efficient use of available research funds.

Vern T. Farewell is Professor and Richard J. Cook is Ph.D. student, Department of Statistics and Actuarial Sciences, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1.