$I_{(T|S,\,Z)}(\theta) = 0$—a highly unrealistic situation. Therefore, perfect surrogacy serves only as an interesting theoretical construct. Equation (1) structures an assessment of the cost/benefit of using a surrogate. If the term in square brackets is negative, then use of the surrogate will be more efficient for inferences on θ than using the ultimate endpoint. Even when this term is positive, the surrogate may still be attractive. Use of it will require additional patients (or events), but total trial duration and person-months on study may be shorter than having to wait for the ultimate endpoint.

Candidates for surrogates abound, but validation is usually elusive. In AIDS research, lab values such as CD4, neopterin and β_2 microglobulin, and disease status indicators such as Karnofsky score, weight and intermediate clinical events are contenders for surrogate status. CD4 currently has top billing, but several challenges remain. The measurement process produces considerable intra- and interlab variability. The true value reacts to short-term infections, can be influenced by smoking and has a pronounced circadian rhythm. In addition, we don't know the best method of using a CD4 *trajectory* to define a surrogate endpoint, and different classes of treatments can have differential effects on CD4, but equivalent therapeutic value. Accruing information from treatment studies linking potential surrogates to long-term follow-up will pin down their status.

Methods Development

A DMC must be ready for anything, and the challenges of monitoring have spawned a variety of methods. Most notable has been introduction of the alpha-spending function, which eliminates the need for specifying in advance the number of monitoring looks. Even with the spending function we can get into difficult and exceedingly unproductive deliberations about how many looks *have been* performed. Fortunately, monitoring boundaries based on a large number of looks (even after each observation) are only slightly broader than the usual, and the broader boundaries should be used.

Fleming presents other exciting recent developments resulting from the wide variety of analyses required for proper monitoring (e.g., multiple measures of treatment effect), the need to react to interesting leads and the need to increase precision (e.g., use of auxiliary variables). Their use in monitoring puts special importance on robustness of validity and efficiency.

CONCLUSION

I enthusiastically thank Professor Fleming for preparing his article. I have learned a great deal and have been energized to give careful thought to technical and broad issues related to clinical trials.

The exigencies of clinical trial design and conduct, especially those associated with monitoring, will continue to seed conceptual and methodologic research that crosses disciplinary and philosophical boundaries. Monitoring and other components of clinical trial design and analysis must balance robustness and efficiency; each trial gets stopped only once. Striking this balance will continue to challenge clinical trialists from all disciplines.

ACKNOWLEDGMENT

Partial support was provided by contract NO1-AI-05073 from the National Institute of Allergy and Infectious Diseases.

Rejoinder

Thomas R. Fleming

MONITORING CLINICAL TRIALS

I appreciate the comments, clarifications and extensions of my distinguished colleagues who have long provided extensive statistical scientific leadership to this area of evaluating therapeutic interventions. I thank the editors for this opportunity for further discussion of some issues related to their comments.

Data Monitoring Committees

The discussants uniformly endorse the concept of Data Monitoring Committees (DMCs), with Professor DeMets specifically advocating their use "for any comparative (Phase III) trial that is pivotal and has either mortality or irreversible morbidity as a primary outcome." With the increasing implementation of such committees pointed out by Professors Ellenberg and Louis, certain issues will need further attention. These include guidelines for membership in various settings and for financial compensation and procedures for expansion of the group of interested, qualified statisticians.

We have stated that DMCs should be "independent," specifically indicating that DMC members should be