

SOUNDING BOARD

Parshuram

C

Data Safety and Monitoring Boards

Arthur S. Slutsky, M.D., and James V. Lavery, Ph.D.

Data safety and monitoring boards, also known by other names (e.g., data monitoring committees), were first introduced in the 1960s as a mechanism for monitoring interim data in clinical trials in order to ensure the safety of the participating subjects. The key concept was to recruit board members who were expert in the field of interest but not otherwise intimately involved in the study (that is, not organizers, sponsors, or investigators), so that they could be objective in their assessment of issues that arose during the study. Over the past decade, the use of data and safety monitoring boards has increased substantially, for a number of reasons. These include the greater number of trials with end points such as death from any cause or death due to cardiovascular disease,¹ more trials funded or sponsored by government agencies that require such boards as part of their design, and greater awareness of methodologic issues in trial design, especially with respect to the potential bias that can result when trials are stopped early because of evidence that one treatment has greater efficacy or causes greater harm than another.

Data safety and monitoring boards are commonly composed of biostatisticians, scientists, bioethicists, and clinicians who are knowledgeable about the question being studied. Their principal role is to ensure the safety of patients, which they do by analyzing adverse events and by performing interim analyses of the clinical outcome data.² Although the boards may have other charges, their primary role is to protect the safety of trial participants; to do so, they may request further information from investigators, require that investigators change the informed-consent document, or recommend the discontinuation of a trial. The powerful mandate of data safety and monitoring boards — particularly their authority to recommend stopping trials early — can have far-reaching effects on clinical practice. Although there is a growing literature on the role and function of these boards,³⁻⁸ surprisingly few articles have focused explicitly on the

ethical dimensions and implications of their decisions, or on proposed approaches or solutions to the ethical challenges they face. In this article we address two central questions: What is the appropriate balance between the obligation of board members to protect study subjects, and their duty to patients and clinicians outside the trial, for whom a clear and convincing answer to the clinical question will have important implications? And should data safety and monitoring boards have a policy limiting disclosure of interim results, especially when such disclosure could influence the decision of a patient to enter the trial?

 THE OBLIGATION TO PROTECT
AND SERVE

Decisions by data safety and monitoring boards to stop trials early have potential implications for three distinct groups: research subjects currently enrolled in the trial, prospective subjects who may be recruited around or after the time of a decision to continue a trial after a board audit, and the population of patients, both current and future, who stand to benefit from clarification of an important clinical question. In research, the investigator's principal ethical obligations are to design the study so as to minimize risk, to ensure the adequate disclosure of the remaining risks to prospective subjects, and to protect individual subjects enrolled in the study. This much is uncontroversial. However, these obligations occasionally conflict with other compelling ethical obligations, particularly the investigator's responsibility to maximize the social value of his or her research by designing trials in a way that is most likely to provide a clear answer to the question under study. Such conflicts have been discussed in the literature on research ethics as it relates to the standard of care in clinical research,^{9,10} in international clinical trials,¹¹ and in studies of public health interventions.¹² Decisions by data safety and monitoring boards to

stop trials early, because the interim data indicate that one of the interventions being compared leads to greater benefit or harm, bring these tensions into sharp relief and demand an explicit account of how they can be balanced.

The critical challenge for the members of a data safety and monitoring board is to determine when the interim data cross some boundary and become "conclusive," rather than simply suggestive, thus justifying an early termination of the trial. This is where prespecified "stopping rules" come into play.¹³⁻¹⁵ By convention, the usual test of statistical significance at the end of a trial is that the probability of the expected findings' occurring by chance alone is less than 5 percent (i.e., $P < 0.05$). However, when interim data are analyzed multiple times, as occurs regularly in analyses by data safety and monitoring boards, the use of a P value of less than 0.05 as the criterion for stopping a trial can lead to a false conclusion that the difference between the study groups is significant, as a result of the effects of multiple tests of significance.¹⁶ Furthermore, the results of clinical trials that are stopped early are likely to exaggerate the magnitude of a treatment effect.¹⁷

As a result of these problems, statisticians have developed various rules to guide interim data analysis; these "stopping rules" can be less stringent with respect to evidence of harm than with respect to evidence of efficacy. One approach is to stop trials early only when there is overwhelming evidence that one intervention is better than another. An example of this is the suggestion by Peto et al.¹⁴ that trials should be stopped early only if the P value at any interim analysis is less than 0.001. Another approach is to develop a "stopping boundary," which changes as the trial gets closer to its predetermined sample size.¹⁶ The rationale is that as the end of the trial nears, a less stringent P value is required to indicate significance, since the results are less likely to change when additional patients are enrolled than was the case early on in the trial.

An instructive example of this situation is a trial comparing a recombinant tissue-factor pathway inhibitor (tifacogin) with placebo for the treatment of selected patients with severe sepsis.¹⁸ At the second interim analysis, when 722 patients had entered the trial (approximately half the projected number), mortality was 38.9 percent in the placebo group, as compared with 29.1 percent in the tifacogin group, with a P value of 0.006. This P value would have been significant at the end of the

trial but did not cross the predetermined boundary for stopping the trial at the time of the interim analysis. After considering the interim results, the board ultimately decided that the study should continue.

One could argue that at the interim analysis, given the relatively large treatment effect (relative reduction in mortality, about 25 percent), with a P value of 0.006, it was extremely likely that tifacogin was in fact better than placebo. The data safety and monitoring board had compelling (but not definitive) reason to believe that subjects enrolled after the interim analysis and randomly assigned to the placebo group were at greater risk of death than subjects assigned to the tifacogin group. Aside from the issue of tifacogin's status as a drug that had not yet been approved by the Food and Drug Administration, how should the board's decision to continue with the trial be characterized ethically?

The stringent early stopping rule used in the study had the effect of a "Ulysses contract," an agreement to stick to a particular course of action made in anticipation of a persuasive, but potentially misleading, argument or set of data (named after Ulysses' insistence, in advance of his encounter with the Sirens, that his men not untie him from the mast of his ship no matter how much he pleaded with them). It thus protected the board members against the seductive power of the early (and, as it turned out, misleading) evidence and permitted adequate time for the real findings of the study to emerge.

For the ethical analysis, it is important to recognize that the decision to specify a stringent stopping rule in the trial itself represents an ethical commitment — namely, to protect the scientific integrity (and hence the ultimate clinical value) of the trial by maximizing its capacity to resolve the important treatment question. The stringent stopping rule provided a secure basis on which the board members could determine not only whether the interim evidence had satisfied the statistical requirements, but also whether the evidence partway through the trial would be sufficient to convince the medical community to accept the investigational therapy. Thus, although the necessary P value may vary according to the nature of the clinical trial, stringent early stopping rules may be an effective way to maximize the social value of clinical trials. In the trial described above, recruitment was continued until the predetermined sample size was reached, and the very strong trend favoring tifacogin vanished: in the final results, mortality in the

tifacogin group was 34.2 percent, as compared with 33.9 percent in the control group — a statistically nonsignificant result.¹⁸

DISCLOSURE OF INTERIM DATA

The subjects who were enrolled after the interim analysis in the tifacogin trial were not told the details of this interim analysis, nor were the trial investigators. For many data safety and monitoring boards, this is standard practice. The rationale is that the evidence at an interim analysis may not be conclusive. Disclosure of suggestive interim results could seriously jeopardize the clinical trial by disturbing the assumption of clinical equipoise on the part of investigators that is necessary to justify the random assignment of subjects, it could make physicians hesitant to enroll any more of their patients, and it could adversely affect adherence to clinical trial protocols.¹⁹ In contrast, nondisclosure has been criticized by some commentators, who suggest that withholding the specific results of interim analyses from prospective subjects is unethical and may even be tantamount to fraud.⁷ Although not representative of previous trials of treatment protocols for patients with severe sepsis, in which interim trends persisted through the final analysis, the case of the tifacogin trial provides an excellent illustration of how disclosure of the interim data would not have clarified the true risk for potential subjects and, in fact, would have been misleading.

DISCLOSURE OF GENERAL RISKS AS PART OF INFORMED CONSENT

How, then, can the need of society for clarification of important clinical questions be satisfied while research subjects in clinical trials are protected? In cases in which data safety and monitoring boards do not stop trials in the face of an interim analysis that suggests a particular outcome, it is customary for the board not to disclose the interim data or the rationale for their decision to the investigators, to the subjects already enrolled, or to potential subjects, though there have been notable recent exceptions.²⁰ We propose that when a data safety and monitoring board is involved in a trial, prospective subjects should be informed about the board and told that any recommendation it makes to continue the trial, rather than stop it in the face of evidence suggestive of important differences in

effectiveness among the treatments, might prolong exposure to a therapy that could turn out to be suboptimal or even harmful and could delay their access to an intervention that could ultimately prove to be more effective. In addition, subjects should be informed that the interim results will not necessarily be revealed to them or their physicians, even if there is a suggestive trend in favor of one treatment or other, so as to maintain the scientific integrity of the study.

An explanation of the decision-making process of the data safety and monitoring board and its potential implications for research subjects is currently not built into the informed-consent process in most trials. Consent forms are already notorious for their length and complexity, and there is sensible resistance to expanding them further. But we believe that information about the board, how it assesses risks, and how it makes its recommendations to the investigators through the course of the trial is directly relevant to potential subjects and must be conveyed in order to meet even the basic regulatory requirements of disclosure. Therefore, we propose that such information be given to prospective subjects when their informed consent is being obtained. A standard section on the informed-consent form might be developed for use in many trials involving a data safety and monitoring board; this section could easily be modified to accommodate specific issues in any given trial, even before the board is constituted.

PUBLICIZING THE BOARD'S DECISIONS AND RATIONALE

Informed consent alone does not render the decisions of a data safety and monitoring board ethically acceptable. It is also necessary that the full rationale behind a decision either to terminate or to continue a trial (though not necessarily all the specific data) be publicly available. The preferred way to do this, we think, is through publication of the board's decisions after the trial has been completed. For example, Wheatley and Clayton recently reported their summary of the rationale behind the data safety and monitoring board's decision not to stop a chemotherapy trial funded by the Medical Research Council (MRC) in the United Kingdom in which four courses of chemotherapy for acute myeloid leukemia were compared with five courses.²¹ The target sample size was 1000 patients; an interim analysis performed after 480 pa-

tients had been enrolled showed that the five-course regimen was associated with lower mortality than was the four-course regimen (17.1 percent vs. 27.5 percent, $P=0.002$).

Although there were no formal rules for early termination of MRC-funded leukemia trials, the MRC suggested at the time that trials should be stopped early only when there was proof beyond a reasonable doubt and endorsed the general guideline that a P value of 0.002 be required to justify stopping a trial. The chief reason given by Wheatley and Clayton for the board's not stopping this trial even though this stringent P value was obtained was that the board members felt that the treatment effect was implausibly large.²¹ Their judgment proved to be correct. At the end of the trial, the mortality rate in the five-course group turned out to be slightly, but not significantly, higher (29.2 percent vs. 25.9 percent). Wheatley and Clayton also argue that fixed rules based simply on rigid predefined statistical boundaries should never be used to stop a trial. In this case, had the trial been stopped early, the clinical recommendation would have been that young patients with acute myeloid leukemia should receive five courses of chemotherapy instead of four, with substantial additional toxicity and cost.

It is a reasonable expectation that data safety and monitoring boards should be transparent about how they arrived at any decision on the early termination of a trial. Although technical criteria for stopping the trial may have been met, the board may make the additional assessment that the evidence that could be derived at the time of the interim analysis was insufficient to ensure that the finding would, indeed, affect policy, practice, or both. We believe these judgments, when made deliberately, conscientiously, and for reasons that are transparent and accessible, can be ethical and may even be critical for ensuring the highest respect for the subjects' contributions to the research (as long as the other necessary conditions are met).

CONCLUSIONS

Uncertainty is an unavoidable feature of interim data analyses. The critical challenge with respect to the ethics of decisions that data safety and monitoring boards make on the basis of these analyses is to find the appropriate balance between the obligation to maximize the scientific validity and val-

ue of the clinical trial, on the one hand, and the obligation to protect human subjects and provide them with meaningful information on which to base their decisions about whether to enroll in a trial, on the other. Data safety and monitoring boards must be rigorous in their analysis of the data so as to ensure the highest potential value of the science. Since there are no accepted standards of what constitutes conclusive evidence, the process by which board members arrive at their decisions must be transparent. One way to achieve this goal would be for the board, after a trial has been completed, to disclose publicly the full rationale for continuing the trial or for stopping it prematurely. If prospective subjects are fully informed about the process and the potential implications of the board's decisions, we believe the ethical obligation to disclose relevant risks adequately to human subjects can be satisfied, while stringent evidentiary standards are maintained that will maximize the scientific value and — more important — the clinical value of the trial.

Dr. Slutsky reports that he currently chairs two data safety and monitoring boards, one for a trial of surfactant for patients with acute lung injury (sponsored by Leo Pharmaceuticals) and the other for two trials of bronchial thermoplasty for the treatment of chronic asthma (sponsored by Asthmatx [formerly Broncus]) and is a member of two other boards, one for trials conducted by the Acute Respiratory Distress Syndrome Network investigators, examining various treatment options for acute lung injury (sponsored by the National Institutes of Health), and the other for a trial examining the efficacy of vasopressin in the treatment of septic shock (sponsored by the Canadian Institutes of Health Research). Dr. Lavery reports that he is a member of the data safety and monitoring board for the current phase 3 trial of a vaccine against human immunodeficiency virus type 1 (Aventis Pasteur and VaxGen) in Thailand, sponsored by the U.S. government.

From the Departments of Medicine and Critical Care (A.S.S.) and the Center for Global Health Research and Inner City Health Research Unit (J.V.L.), St. Michael's Hospital; and the Departments of Medicine, Surgery, and Biomedical Engineering (A.S.S.) and the Department of Public Health Sciences and the Joint Centre for Bioethics (J.V.L.), University of Toronto — both in Toronto. Address reprint requests to Dr. Slutsky at St. Michael's Hospital, Queen Wing, Rm. 4-042, 30 Bond St., Toronto, ON M5B 1W8, Canada, or at arthur.slutsky@utoronto.ca.

1. Guidance for clinical trial sponsors: on the establishment and operation of clinical trial data monitoring committees: draft guidance. Rockville, Md: Food and Drug Administration, November 2001. (Accessed February 19, 2004, at <http://www.fda.gov/cber/gdlns/clindatmon.pdf>.)

2. Ellenberg SS, Fleming IR, DeMets DL. Data monitoring committees in clinical trials: a practical perspective. Chichester, England: John Wiley, 2002.

3. Walters I. Data monitoring committees: the moral case for maximum feasible independence. *Stat Med* 1993;12:575-80.

4. Baum M, Houghton J, Abrams K. Early stopping rules — clinical perspectives and ethical considerations. *Stat Med* 1994;13:1459-69.

SOUNDING BOARD

5. Gordon VM, Sugarman J, Kass N. Toward a more comprehensive approach to protecting human subjects: the interface of data safety monitoring boards and institutional review boards in randomized clinical trials. *IRB* 1998;20:1-5
6. Dixon DO, Lagakos SW. Should data and safety monitoring boards share confidential interim data? *Control Clin Trials* 2000;21:1-6
7. Wells RJ, Gartside PS, McHenry CI. Ethical issues arising when interim data in clinical trials is restricted to independent data monitoring committees. *IRB* 2000;22:7-11.
8. Veatch RM. Longitudinal studies, sequential design, and grant renewals: what to do with preliminary data. *IRB* 1979;1:1-3.
9. Miller PB, Weijer C. Rehabilitating equipoise. *Kennedy Inst Ethics J* 2003;13:93-118
10. Miller FG, Rosenstein DL. The therapeutic orientation to clinical trials. *N Engl J Med* 2003;348:1383-6.
11. Angell M. Investigators' responsibilities for human subjects in developing countries. *N Engl J Med* 2000;342:967-9
12. Ross LF. In defense of the Hopkins Lead Abatement Studies. *J Law Med Ethics* 2002;30:50-7
13. Califf RM, Ellenberg SS. Statistical approaches and policies for the operations of data and safety monitoring committees. *Am Heart J* 2001;141:301-5.
14. Peto R, Pike MC, Armitage P, et al. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. I. Introduction and design. *Br J Cancer* 1976;34:585-612.
15. DeMets DL, Lan KK. Interim analysis: the alpha spending function approach. *Stat Med* 1994;13:1341-52
16. O'Brien PC, Fleming TR. A multiple testing procedure for clinical trials. *Biometrics* 1979;35:549-56.
17. Pocock SJ, Hughes MD. Practical problems in interim analyses, with particular regard to estimation. *Control Clin Trials* 1989;10: Suppl:209S-221S.
18. Abraham B, Reinhart K, Opal S, et al. Efficacy and safety of tifa-cogin (recombinant tissue factor pathway inhibitor) in severe sepsis: a randomized controlled trial. *JAMA* 2003;290:238-47.
19. Fleming TR, Ellenberg S, DeMets DL. Monitoring clinical trials: issues and controversies regarding confidentiality. *Stat Med* 2002; 21:2843-51
20. Rossouw JE, Anderson GI, Prentice RI, et al. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results from the Women's Health Initiative randomized controlled trial. *JAMA* 2002;288:321-33
21. Wheatley K, Clayton D. Be skeptical about unexpected large apparent treatment effects: the case of an MRC AML12 randomization. *Control Clin Trials* 2003;24:66-70

Copyright © 2004 Massachusetts Medical Society

CLINICAL PROBLEM-SOLVING SERIES

The *Journal* welcomes submissions of manuscripts for the Clinical Problem-Solving series. This regular feature considers the step-by-step process of clinical decision making. For more information, please see www.nejm.org/hfa/articles.asp.