MONITORING FOR SAFETY, EFFICACY, AND FUTILITY

From the Therapy chapter for the 3rd edition of Clinical Epidemiology, by DL Sackett
17 April 2004 (day 108)

Scenarios:

1. In the RRPCE, three of us at the Methods Centre at McMaster carried out the monitoring for safety, efficacy and futility: Mike Gent, the Co-PI Biostatistician, Wayne Taylor, the Study Statistician, and I. None of the trial participants were my personal patients, and we kept the PI, Steering Committee, and participating neurologists blind to our analyses. Both study drugs had been in widespread use for years, but we nonetheless maintained a telephone "hotline" so that both anticipated (e.g., gastrointestinal) and unanticipated adverse drug reactions could be assessed at once. We continuously monitored for safety, but carried out a single analysis for efficacy at the scheduled end of the trial. It showed statistically significant efficacy for aspirin, and we then unblinded the Principal Investigator, who discussed the findings with us and his Steering Group and started an "orderly termination" of the trial.

2. Fifteen years later, when we began the North American Symptomatic Carotid Endarterectomy Trial (NASCET), our sponsor (the U.S. National Institutes of Health) insisted on naming, and sitting on, our Data Safety and Monitoring Board. We were dissatisfied with its procedures and with some of its members, who we thought had conflicts of interest. We had developed statistical warning rules for efficacy and futility, and obtained the DSMB's grudging agreement to remain blind to our efficacy analyses until those warning rules were triggered. Subsequent DSMB meetings were positive and placid during discussions of accrual, surgical performance, follow-up, and data quality. They became stormy and confrontational when we refused to show them unblinded data and limited our efficacy report to a statement that the warning rules had not been triggered. Tempers would flare, we would threaten to quit, and some DSMB members appeared eager to accept our resignations. Then things would settle down and our report would be accepted. However, this the cycle would recur at the next meeting. Our statistical warning rules (for patients with high-grade symptomatic stenosis) were triggered shortly before a regularly scheduled DSMB meeting. We presented the unblinded results to them, and they recommended stopping the trial. We showed the results to the PI and Steering Group that afternoon, and they decided to stop the trial. They announced the results to their clinical collaborators that same day, and a control patient underwent endarterectomy the next day.

3. In the interim, I had frequently provided a one-person, volunteer monitoring service for small (<100 patient) RCTs. I closely monitored their unblinded data for safety and periodically examined their unblinded data for efficacy. Based on my assessments, I recommended the continuation of these trials or the unblinding of their Principal Investigators.

4. Another fifteen years later, as this edition was being written, I was asked to chair the Trial Monitoring Committee for an RCT of drugs that might delay or prevent the onset of diabetes. With agreement from the Principal Investigator, I banned employees from any sponsor (government or industry) from the committee, and rejected any potential members who owned stock with the drugs' manufacturers. None of us will receive salaries or honoraria for serving on this Trial Monitoring Committee. Our committee will never be blind. The Data Center and I have established a system for alerting me within 24 hours of any unanticipated serious adverse event in any trial patient. Our Trial Monitoring Committee and the Principal Investigators have agreed upon and adopted statistical warning rules for efficacy and safety. It was understood from the outset that we are serving in an advisory (not executive) capacity to the Principal Investigators and their Steering Committee. That is, if we decide that the study treatment clearly works (or is clearly harmful), we will unblind the Principal Investigators, not stop the trial.

This section is about the external person(s) who serve the patients and investigators in RCTs by alerting them as soon as clear-cut evidence emerges about the safety or efficacy of the experimental treatment, or about the futility of continuing an indeterminate trial. These groups bear a wide array of names and acronyms, and Susan Ellenberg's collection[1] appears in Table 03-16-1. Some of the more common monikers are Data, Safety, and Monitoring Board (DSMB); Data Monitoring Committee (DMC); and Trial Monitoring Committee (TMC). In this section, I'll use the last term: TMC.

Table 03-16-1: Names assigned to monitoring groups

Subject	Function	Organization
Trial	Monitoring	Committee
Data	Review	Board
Safety	Advisory	Panel
Policy		
Efficacy		
Endpoint		
Ethics		

Adapted from Ellenberg SS. Independent data monitoring committees: rationale, operations and controversies. Stat Med 2001;20:2573-83.

If this section whets your interest in TMCs, or it becomes clear that you need one for your RCT, there are two very good resources where you can learn more about them. The first is a book by Susan Ellenberg, Thomas Fleming, and David DeMets[2], and the second is a report from the UK DAMOCLES project[3] that carried out a systematic review and several interviews on issues in data monitoring and the interim analysis of RCTs.

Monitoring Check List:

1 ❍	Can you justify **not** monitoring your RCT for safety, efficacy and futility?
If you decide to monitor your RCT:	
2 ❍	Specify the precise functions you want your monitor(s) to carry out
3 ❍	Establish requirements for monitor(s) in terms of expertise, experience, and freedom from conflicts of interest.
4 ❍	Recruit your monitor(s) and work with them to establish policies and procedures
5 ❍	Get on with it.

1. **Can you justify not monitoring your RCT for safety, efficacy and futility?**

Current opinion among trialists is that virtually every RCT needs external monitoring for safety, efficacy and futility (for example, the British MRC made monitoring of its trials mandatory in 1998). Current opinion also dictates that this monitoring be done by an individual or group who have no personal interest in its outcome. As shown in the third scenario above, this doesn't mean that every RCT needs a full-blown TMC. A single individual often can carry out these functions.

Exceptions to the need for monitoring are rare, but do occur. In some cases, monitoring is not necessary because no patients are at risk. An example here is an RCT that randomizes clinicians to receive the same efficacy information in relative or absolute terms, in order to determine whether these different formats lead to different conclusions about efficacy.

In other cases, monitoring is not feasible because all study patients have already been recruited and treated (and perhaps have experienced outcomes) before any monitoring function could be launched. One example here would be a single RCT of sufficient size to determine whether

marathoners' performance is importantly improved or worsened by different rehydration regimens. A second example would be an RCT into whether different vaccines administered today lead to different outcomes ten years from now (although even this RCT would need short-term monitoring for safety).

Finally, some trialists propose waiving monitoring in RCTs that test well-established treatments with minor hazards for their effects on trivial, reversible outcomes. I'd accept this view only if the criteria for "minor hazards" and "trivial, reversible outcomes" originated with the patients in these trials, and not from the investigators.

2. **Specify the precise functions you want your monitor(s) to carry out**

Table 03-16-2 lists some possible functions that monitors could carry out. For a more detailed list, see the report of the DAMOCLES project[4].

Table 03-16-2 Some Monitoring Functions

Before the RCT	Review, discuss with, and advise the Principal Investigator/Steering Group about the protocol and the logistics of its execution
During the RCT	Immediately review every serious unanticipated adverse event, taking necessary steps to protect patients in the trial.
	Periodically review accumulating safety data, taking necessary steps to protect patients in the trial.
	Review any emerging external evidence that might influence a recommendation to stop or continue the trial
	Periodically review (and provide feedback and advice about) actual vs. projected patient accrual; patients' compliance with treatments; investigators' adherence to the study protocol; the completeness, timeliness and accuracy of study data; and other measures of the quality of the conduct of the trial.
	Periodically review unblinded outcome data, applying previously established statistical warning rules for safety, efficacy, equivalence, and futility.
	Integrate this review with the totality of evidence about the target disorder and study regimens, and make a recommendation about continuing the study, enlarging it, or unblinding the PI so that a decision can be made about stopping it.
	Based on any of the foregoing, make other recommendations about changes to the protocol or conduct of the study, or in the analysis of its results.
After the RCT	Review and provide feedback on draft reports, presentations, and publications.
	Assist the Principal Investigator and Writing Committee in responding to comments and criticisms of the trial.
	Review the on-going care of study patients.

As in the 3rd scenario above, a single monitor can provide all these functions for small studies (say, less than 100 patients) with early outcomes (say, within 3 months of entry). For bigger, longer studies I recommend that you create a monitoring committee.

3. **Establish requirements for monitor(s) in terms of expertise, experience, and freedom from conflicts of interest.**

To begin with, you want monitors who are expert in the functions you listed in the previous step. Their collective expertise must extend from the clinical presentation and care of patients with the target disorder to the latest developments in trial design and statistical analysis.

Paradoxically, we usually exclude the most knowledgeable clinician, the Principal Investigator, from the monitoring process. In recognition of this paradox, Curt Meinert has challenged the notion that the PI must be kept blind to the emerging results[5]. Scenario #4 describes a partial solution to this paradox in which the triggering of a statistical warning rule leads monitors to unblind the PI, not stop the trial.

Monitors' experience in the conduct and monitoring of RCTs must be sufficient for them to have previously confronted all the problems you envisage in your trial, especially in the consideration of early, unstable trends in efficacy and safety. Stopping an RCT early for a trend in efficacy or safety that subsequently disappears is every trialist's nightmare.

Avoiding conflicts of interest (whether real or potential) among trial monitors is vital to both the validity and the credibility of your or anybody else's RCT. Some conflicts are obvious; others are subtle. Table 03-16-3 lists the ones the DAMOCLES gang and I could think of. Note that most of them apply as well to the authors of the letters, editorials, and guidelines that appear following the publication of the trial results.

Table 03-16-3: Real and potential conflicts of interest for monitors

Financial	Owns stock in the company that stands to gain from a positive trial of their product or process.
	Owns stock in a competing company that stands to lose from a positive trial of the product or process.
	Buys stock in the former company, or sells stock in the latter, based on unblinded data available only to monitors.
	Is a paid consultant or honorarium recipient of either company.
	Receives research or educational grants, fellowship support, or free travel-accommodation from either company.
	Receives payment beyond reimbursements for travel to, and accommodation at, monitoring meetings (Exception: income offsets for time spent by some university and government employees in monitoring, paid to their employers.)
Professional	Career success is tied to applying the product or process.
	Is (or will be) admitting patients to the RCT.
	Is involved in running any part of the RCT.
	Is a member of the regulatory agency that will approve/disapprove the product or process.
	Is a member of the funding agency whose prestige and budget will be affected by the outcome of the RCT.
Academic	Already on record as certain about the efficacy and safety (good or bad) of the experimental treatment.
	Would get credit for authorship of publications arising from the RCT.

The final entry in Table 03-16-3 stresses the "servant" role of monitors. They should seek neither fame nor publications for their efforts. They have to get all the satisfaction they need from working and learning behind the scenes, with nothing to show for it but an acknowledgement in small print at the end of a publication.

Even when monitors with conflicts of interest behave impeccably, their presence can detract from the credibility and acceptance of the trial result. Moreover, as described in the second scenario, their presence on a Trial Monitoring Committee can seriously impair its function. It was, in part, growing concern about these conflicts that has led to the organization and operation of Trial Monitoring Committees like the one described in the fourth scenario.

4. **Recruit your monitor(s) and work with them to establish policies and procedures**

For small, short, simple trials (say, of the next advance in the short-term treatment of a common condition), you can recruit a single monitor who can carry out the functions you require. For larger, longer, more complex trials you can begin by recruiting a Trial Monitoring Committee chair, and then work together to select the other members. The chair must have prior experience in trials and trial monitoring, plus the interpersonal skills required to resolve differences of opinion within the committee. The Trial Monitoring Committees I chair have six members, a number I find large enough both to provide the required range of expertise and to generate a quorum for our meetings and conference calls.

This recruitment process must include hammering out the policies and procedures for how the monitors will function and, crucially, how they will relate to you and your other investigators. The key issue here is whether they will act as advisors or executives. This distinction is best illustrated by how they behave when they conclude that the emerging results demonstrate efficacy, harm, or futility. An executive-style Trial Monitoring Committee would order the trial to stop. An advisory-style Trial Monitoring Committee would unblind the Principal Investigator, show her or him the data, and collaborate with them and any others they choose (such as the Steering Group) in deciding whether to stop or continue the trial.

By incorporating the Principal Investigator's (and Steering Group's) greater clinical and biological expertise, the expanded, advisory-style Trial Monitoring Committee can intelligently examine the totality of evidence. The ultimate termination decision is left to the Principal Investigator and Steering Group, and in third party mediators are called in if they disagree with the monitors. Writing from experiences on both sides of the Principal Investigator-monitor interaction, I vastly prefer the advisory approach from both perspectives. This advisory approach not only brings the most knowledgeable person (the Principal Investigator) into the decision to stop the trial. Also, the identification and solution of problems is so much more pleasant and productive when Principal Investigators and monitors work as collaborators, rather than as defendants and judges.

Another key policy decision to settle before the trial begins is how to monitor events that are both unanticipated and serious. In useful operational terms, events are designated "unanticipated" when they are not already specified in the trial protocol as determining the efficacy of the experimental treatment. These events are "serious" as well when their occurrence beyond chance would require unblinding the Principal Investigator and changing the protocol or even stopping the trial.

It is vital that clinicians looking after trial patients identify and manage every "serious unanticipated adverse event" (SUAE) as soon as it occurs. It is also vital that an unblinded monitor (who can combine this event with other, similar SUAEs) reviews this occurrence and determines whether it is happening at other centers with a combined frequency greater than chance. Thus, in scenario #4, the monitor is informed of any SUAE within 24 hours of its recognition.

5. **Get on with it.**

Once the monitor(s) have been named, you should meet with them face-to-face for a detailed discussion of your protocol, how you will execute it, how they will protect study patients, and how

they will help you achieve a result that is both valid and credible. This first meeting should also agree on the statistical warning rules and how everybody will act when they are triggered.

Subsequent monitors' meetings should be held when they can be most helpful to the trial. The second one could be held when patient recruitment is at a point where problems in accrual, eligibility, initial treatment, patient compliance, protocol adherence, data quality and timeliness, and the like can be identified and solutions suggested. Subsequent meetings might most sensibly be scheduled on the basis of study progression (such as when ½ the projected events have occurred, or when half of the projected follow-up has been reported) rather than on the passage of time (unless recruitment is lagging and the trial is bogging down).

Typical agenda for a monitoring meeting

The agenda for the Trial Monitoring Committee meetings that follow the fourth scenario observe the following sequence:

1. A "*closed*" session among just the monitors, to identify concerns and other issues for discussion later in the meeting.
2. An "*open*" session with the blinded Principal Investigator (perhaps accompanied by other members of the blinded Steering Group), the blinded Study Coordinator, and the unblinded Study Statistician(s). Patient accrual, data quality and timeliness, patient compliance and protocol adherence are usual topics, plus any other issues raised by, and appropriate for discussion among, blinded participants.
3. A "*semi-closed*" session between the monitors and the unblinded Study Statistician(s), to examine and discuss unblinded data on safety and efficacy, and to determine the results of applying the statistical warning rules. This session often generates specific requests and recommendations that apply to just the Study Statisticians, and these are transmitted on the spot.
4. A "*closed*" session among just the monitors, to discuss all the foregoing and to generate appraisals and recommendations for the study statisticians, the Principal Investigator, and the trial patients and staff.
5. A final "*open*" session with everyone, to present, explain, discuss and (if necessary) revise monitors' recommendations. The session closes with a decision on the timing and format (face-to-face or conference-call) for the next Trial Monitoring Committee meeting.

Following such a Trial Monitoring Committee meeting, its chair drafts 2 letters. The first is for general distribution to all the trial participants, makes comments on its (blinded) progress, offers praise where deserved, and concludes that it should continue as planned. The second is for the blinded Principal Investigator and Steering Group, and includes recommendations about proposed protocol changes, recruitment, follow-up, and the quality and timing of field data. The Principal Investigator decides whether to forward this second letter to the sponsors.

When the "semi-closed" session reveals that a statistical warning rule for safety, efficacy, or futility has been triggered, the "closed" session becomes a lengthy consideration of the totality of evidence for its completeness, consistency, sensibility, and coherence. If these criteria are met, the final "open" session unblinds the Principal Investigator. At that point the Principal Investigator assumes lead responsibility for deciding whether to stop the trial, involving Steering Group members and anyone else who could be helpful, and continuing to use the TMC as advisors.

63 **CHAPTER REFERENCES**

[1] Ellenberg SS. Independent data monitoring committees: rationale, operations and controversies. Stat Med 2001;20:2573-83.

[2] Ellenberg SS, Fleming TR, DeMets DL. Data Monitoring Committees in Clinical Trials; A Practical Perspective. Chicester: John Wiley & Sons, 2003.

[3] Altman D, Bebiker A, Campbell MK, Clemens F, Darbyshire J, Elbourne D, Grant AM, McLeer SK, Parma M, Pocock S, Spiegelhalter D, Sydes M, Walker A, Wallace S. Issues in Data Monitoring and Intermin Analysis of Trials. UK National Health Services Research & Development Programme. In Press.

[4] Altman D, Bebiker A, Campbell MK, Clemens F, Darbyshire J, Elbourne D, Grant AM, McLeer SK, Parma M, Pocock S, Spiegelhalter D, Sydes M, Walker A, Wallace S. Issues in Data Monitoring and Interim Analysis of Trials. UK National Health Services Research & Development Programme. In Press.

[5] Meinert CL. Clinical trials and treatment effects monitoring. Controlled Clin Trials 1998;19:515-22.