I N T H E F I V E D E C A D E S S I N C E T H E C O M P L E T I O N O F T H E G R E E N B E R G R E P O R T recommendations in 1967 (which were later published\(^1\)), independent groups of experts have monitored the progress of many clinical trials for early definitive evidence of benefit, convincing evidence of harm, or sufficient evidence of no potential benefit to render continuation of the trial to be futile. Such monitoring is motivated primarily by an ethical imperative; for trials of treatments intended to prevent or delay serious outcomes, one would want to stop the trial and make the superior treatment available as soon as the evidence was definitive. The assessment of risk versus benefit throughout the course of the trial requires frequent access to accumulating data on efficacy and safety. To allow investigators and sponsors to manage the trial without bias, as well as to ensure an objective assessment of the accumulating data, independent committees, known as data monitoring committees (DMCs), data and safety monitoring boards, or other aliases,\(^2\) are established to perform this monitoring function and to make recommendations regarding trial modifications, including early termination, to investigators and sponsors. These committees require multidisciplinary expertise and experience, including knowledge of statistical methods for interim data monitoring; they also pay attention to recruitment progress and the general quality of the trial with respect to adherence to the protocol and completeness of data collection and follow-up.

Data are typically reviewed by DMCs according to the intervention groups as randomized; these groups should be identified fully, perhaps as “group A and group B,” in the printed report, in which the identities of groups A and B are known by the DMC.\(^2\)\(^4\) After the data have been reviewed, DMCs make a recommendation to the trial steering committee that the trial continue as planned or be terminated early for the reasons noted above (see Table 1 for a summary of DMC responsibilities). The Coronary Drug Project, an early, large cardiovascular trial sponsored by what was at that time known as the National Heart Institute, was probably the first trial to incorporate interim review by independent experts\(^5\) and provided many lessons regarding the DMC process.\(^6\)\(^7\) This process continued to evolve and became the general approach for most federally sponsored randomized trials during the 1970s\(^2\)\(^8\)\(^9\) and eventually for the increasing number of phase 3 industry-sponsored trials,\(^10\) with some modification of the National Institutes of Health (NIH) model.\(^11\) The evolution of DMCs continues as the art and science of clinical trial performance itself evolves.

In most cases, trials are completed as planned, with no DMC recommendation for early termination or major modification, but some trials have raised challenging issues for DMCs.\(^12\)\(^13\)\(^19\) In this article, we briefly summarize a few classic cases of early termination and then address some recent cases to further illustrate the complexity of the responsibilities of the DMCs.
† Assessment of efficacy outcomes can be a risk–benefit analysis as well as a consideration of early termination on the basis of efficacy.

The data assessed can include the dropout rate, incomplete data, and the timeliness of data.

Assessment of safety outcomes Always or almost always

Assessment of efficacy outcomes† Always or almost always

Recommendation regarding trial continuation Always or almost always

Review of trial presentations and manuscripts Sometimes

Approval of presentations and manuscripts Sometimes

It is important to protect the confidentiality of the interim data reviewed by DMCs. Although sharing of data across DMCs for separate trials with similar interventions is not common, it has on occasion proved useful. This was especially true for three randomized, double-blind trials evaluating ezetimibe, a drug for treating high cholesterol. The Improved Reduction of Outcomes: Vytorin Efficacy International Trial (IMPROVE-IT) evaluated treatment with ezetimibe combined with simvastatin versus simvastatin alone to assess whether combined therapy would be associated with a lower incidence of cardiovascular death, nonfatal myocardial infarction, or hospitalization for unstable angina or revascularization in patients with an acute coronary syndrome. The other two trials compared the combined regimen against placebo. The Study of Heart and Renal Protection (SHARP) assessed whether this combination could safely reduce the risk of coronary heart disease, non-hemorrhagic stroke, or the need for revascularization procedures in patients with chronic kidney disease; the Simvastatin and Ezetimibe in Aortic Stenosis (SEAS) trial involved patients with mild-to-moderate asymptomatic aortic stenosis. The SEAS trial, which was completed while the other two trials were still ongoing, did not show any benefit but did report a significant increase in cancer among the patients treated with simvastatin–ezetimibe, which raised concern about the continuation of the ongoing trials.

Concerned that external pressures might force the release of interim results, the steering committees of these three trials, in collaboration with their DMCs, proposed that interim cancer data for IMPROVE-IT be shared with the SHARP statisticians, who would conduct a meta-analysis to test the hypothesis generated by the SEAS trial that ezetimibe increased the risk of cancer. (Because there was no evidence that simvastatin,

---

**Table 1. Responsibilities of Data Monitoring Committees.**

<table>
<thead>
<tr>
<th>Responsibility</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Review of protocol</td>
<td>Often</td>
</tr>
<tr>
<td>Approval of protocol</td>
<td>Sometimes</td>
</tr>
<tr>
<td>Assessment of adequacy of recruitment progress</td>
<td>Always or almost always</td>
</tr>
<tr>
<td>Assessment of data quality*</td>
<td>Always or almost always</td>
</tr>
<tr>
<td>Assessment of safety outcomes</td>
<td>Always or almost always</td>
</tr>
<tr>
<td>Assessment of efficacy outcomes†</td>
<td>Always or almost always</td>
</tr>
<tr>
<td>Recommendation regarding trial continuation</td>
<td>Always or almost always</td>
</tr>
<tr>
<td>Review of trial presentations and manuscripts</td>
<td>Sometimes</td>
</tr>
<tr>
<td>Approval of presentations and manuscripts</td>
<td>Sometimes</td>
</tr>
</tbody>
</table>

* The data assessed can include the dropout rate, incomplete data, and the timeliness of data.

† Assessment of efficacy outcomes can be a risk–benefit analysis as well as a consideration of early termination on the basis of efficacy.

---

**Selected Classic Cases of Early Termination**

The NIH-sponsored Beta-Blocker Heart Attack Trial (BHAT) compared propranolol with placebo in a population of patients who had had a myocardial infarction. The DMC found a large early benefit with regard to mortality and recommended early termination, using formal sequential monitoring methods to guide their assessment. In contrast, the Cardiac Arrhythmia Suppression Trial (CAST), which evaluated a class of widely used drugs that suppressed cardiac arrhythmias, was terminated very early because of higher mortality in the treatment groups than in the placebo group, a result that was contrary to expectation. The Metoprolol CR/XL Randomized Intervention Trial in Congestive Heart Failure (MERIT-HF) evaluated the addition of a beta-blocker, metoprolol, to best care in chronic heart failure, although there were a priori safety concerns about the use of this drug class in this population. These concerns were resolved by the rapidly emerging large and highly significant mortality benefit attributed to metoprolol, which led to early termination of the trial.

Although most trials are not terminated early, trends may emerge and then disappear as the trial progresses, requiring DMCs to be very cautious. Balancing an early estimate of substantial benefit against the value of continued observation to confirm the early trend is challenging. It can be even more challenging when emerging evidence suggests harm.

The following cases illustrate additional challenging issues faced by DMCs. We also address concerns about the availability of knowledgeable clinical trialists for DMC service.
a drug that had long been in use, increased cancer risk, any excess cancer risk in the combination group would presumably be due to ezetimibe.) The meta-analysis did not confirm any increased cancer risk, a result that was reported with an accompanying editorial supporting the conclusions. The meta-analysis report defused the concerns about the potential risk of cancer sufficiently, and IMPROVE-IT and SHARP were able to continue to completion. IMPROVE-IT ultimately showed a small beneficial effect of the combination as compared with the statin alone with regard to the primary outcome (hazard ratio, 0.94; P=0.02), with a similar risk of cancer in both treatment groups. In SHARP, the rate of atherosclerotic events (the primary outcome) was 17% lower in association with simvastatin–ezetimibe than with placebo, a result that was highly significant, and there was no excess cancer risk associated with the drug combination. If interim results for the two ongoing trials had been shared publicly, these trials might not have been completed, and the opportunity for clarifying the benefits and risks of this combination might have been missed.

In general, the public sharing of interim data that are not yet conclusive may introduce bias into the conduct of the trial and threaten the validity of the results. Limiting the availability of interim data to DMCs protects the integrity of the trial and promotes the completion of trials unless interim results are overwhelmingly convincing. In this special case, only the cancer data from IMPROVE-IT were shared, and only with the SHARP statistician, so that only summary cancer results of IMPROVE-IT and SHARP were reported.

**Liability**

DMC members can become embroiled in litigation. The Adenomatous Polyposis Prevention on Vioxx (APPROVe) trial evaluated rofecoxib (Vioxx), a Food and Drug Administration (FDA)–approved pain reliever, as a possible agent for the prevention of colon cancer. The DMC recommended early termination of the trial because of convincing evidence of increased cardiovascular risk that appeared to emerge after approximately 18 months of treatment. (Later, additional follow-up data showed that the cardiovascular risk actually began well before 18 months.)

After the APPROVe report was published, lawsuits were filed on behalf of study participants who had had cardiovascular events after taking rofecoxib for the approved indication of arthritis pain. Although they were not litigation targets themselves, the DMC members were required to provide depositions and court testimony. Before the legal proceedings, the DMC members had no guarantee of indemnification or legal support, but that was ultimately provided by the pharmaceutical manufacturer before any depositions were made (Neaton J and Konstam M: personal communication). This example points to a more general problem: DMCs have substantial challenges in monitoring interim data, and their members bear some legal risk. DMC members should not have to worry about litigation costs when they are properly fulfilling their monitoring responsibilities; that is not in the best interest of the trial, sponsors, investigators, or trial participants. The issue of DMC indemnification remains difficult. Currently, many commercial sponsors provide indemnification, but they often require that their own corporate lawyers represent the DMC and its members. Federally sponsored trials offer no indemnification for DMC members. DMCs should ideally have independent insurance coverage included as part of their contract.

**Differing Perspectives on Evidence Required for Early Termination of a Trial**

Statistical monitoring plans allow for the early termination of a trial that has extreme emerging results while preserving the ability to claim significance, but there must be consensus among all involved with regard to the criteria for early termination before the trial begins. DMC members who do not agree with the monitoring plan should not participate. For some trials — for example, those comparing two or more treatments that are in widespread use — the criteria might be quite stringent, because clinical practice is unlikely to change without substantial evidence of the superiority of one treatment. In other trials, such as those testing a promising new treatment for a life-threatening disease, there may be a desire to allow termination for benefit more readily, because the benefits of more rapid access to a substantially improved treatment may outweigh the potential risks.
The trial investigators proposed a conservative Haybittle–Peto statistical stopping boundary that would permit early termination only if an interim P value fell below 0.001.\textsuperscript{36,37} They argued that very strong findings would be needed to support the costs of providing immediate treatment to all infected infants. Many members of the DMC preferred a boundary of the O’Brien–Fleming type, for which the boundary becomes less stringent as more data become available.\textsuperscript{39} Both types of boundaries provide appropriate control of the false positive error rate, but the Haybittle–Peto boundary makes termination more difficult in the later portion of the trial when the data are more precisely estimated. The committee understood the need for very substantial data to support a policy change requiring major new resources, but balancing that against accepting more deaths among infants who received placebo was very difficult. Ultimately, the committee did accept the Haybittle–Peto monitoring plan.

As the trial progressed, the DMC observed increasing advantages for the two immediate-treatment groups (which differed from one another only in duration of therapy) and became more uncomfortable with the Haybittle–Peto stopping boundary. After approximately 18 months, the P values were 0.008 and 0.01 for the comparisons between each immediate-treatment group and the delayed-treatment group, with almost all the observed events being death. These results led the DMC to request a comparison between the combined immediate-treatment groups and the delayed-treatment group. This comparison resulted in a P value of less than 0.001, which led the DMC to unanimously recommend termination of the trial (Fig. 1). The recommendation was accepted by the study team and the NIAID; the treatment guidelines for infected infants were changed shortly thereafter. This example illustrates the fact that statistical monitoring plans are only guidelines and that ultimately the DMC should exercise its best judgment based on trial data and the external context.\textsuperscript{2,6,7}

The Children with HIV Early Antiretroviral Therapy (CHER) trial, which was sponsored by the National Institute of Allergy and Infectious Diseases (NIAID), was a three-group trial that was started in 2005 to evaluate strategies for initiating antiretroviral treatment for South African infants with human immunodeficiency virus (HIV) infection, as well as to evaluate the duration of treatment.\textsuperscript{42} The trial end point was death or failure of first-line antiretroviral therapy. At the time the trial was launched, treatment guidelines worldwide recommended the initiation of treatment when the CD4 count fell below 350 cells per cubic millimeter or when clinical symptoms appeared. However, in the United States, the guidelines added that treatment should be considered for all infants with HIV infection immediately on diagnosis, as was increasingly the standard practice in Western countries.

**Figure 1. Probability of Death or Clinical Failure in the Children with HIV Early Antiretroviral Therapy (CHER) Trial.**

The figure is adapted from Violari et al.\textsuperscript{42} Clinical failure was defined as a Centers for Disease Control and Prevention (CDC) stage C or severe stage B event.

**MULTIREGIONAL TRIALS**

Industry-sponsored trials are increasingly global, and troublesome issues can arise if the interim treatment effect appears to differ across countries or regions. DMC members typically view subset results with some skepticism, but they
may not be knowledgeable about regional differences in medical practice that could plausibly lead to different effects of treatment. Further, regional differences in regulatory standards or national policies about adopting new treatments may lead sponsors to propose stringent monitoring plans, which can create ethical dilemmas about continuing studies when the overall emerging data appear strong enough to warrant early termination but not strong enough to support practice changes in some regions.

Despite the overwhelmingly positive mortality results for the primary outcome in the MERIT-HF trial overall,\textsuperscript{43} the results in the U.S. population were neutral,\textsuperscript{43} which led to the rejection by the FDA of a mortality indication for metoprolol. Whether one believes that the DMC terminated the trial too early or that the FDA was too conservative in rejecting the mortality indication, it is clear that such scenarios pose dilemmas for a DMC (and ultimately for regulatory agencies). Such observed regional differences that do not have strong biologic plausibility must be interpreted with great caution, because they can occur with a surprising frequency as a result of chance alone.\textsuperscript{44}

\textbf{PROVISION OF RELEVANT INFORMATION TO THE DMC}

In most cases, selecting the information to be presented to the DMC is straightforward; the information generally includes enrollment rates, data quality, demographic characteristics, emerging safety information, and efficacy summaries. Interim efficacy data should always be available to the DMC; otherwise, the DMC will not be able to weigh potential benefits and risks. Controlling the type I error can be addressed by implementing alpha spending function methods.\textsuperscript{41}

The DMC must also be notified of other information available to sponsors or investigators that might inform their assessment of the trial data. For this, they must rely on the judgment of the trial steering committee, which usually includes the trial statistician working with the clinical investigators on the trial design and conduct, the sponsor, and the independent statistician preparing the DMC reports. Without the relevant information, the ability of the DMC to protect trial participants will be hampered, as in the following example.

The study in question, HIP Impact Protection Program (HIP PRO), was an NIH-funded investigation of the use of hip pads worn by nursing home residents to prevent hip fractures.\textsuperscript{45} Trial participants were randomly assigned to have a pad inserted into special underwear on either their left or their right side, with the primary question being whether fewer fractures occurred on the “protected” side. As the trial progressed, the emerging results appeared to favor the unprotected side, with regard to both the number of fractures and the number of falls. These data were presented to the DMC, along with information that may have mitigated their effect — many falls were not observed, and therefore the data on falls were unreliable; there was a lag in adjudicating fractures; information on adherence to wearing the pads was not available; no previous study had shown any harm in association with hip pads; and in any case a new, improved hip pad was about to be introduced. The DMC permitted the trial to continue with the new hip pad; trial participants were told only that the previous pad had been ineffective and that a new pad would be substituted. However, the concern about the imbalance in falls led the investigators to review the data on falls from a previously completed but incompletely analyzed pilot study, which showed a highly significant substantial excess of falls on the “protected” side, a finding consistent with the trend in the ongoing trial. The investigators discussed whether to provide the pilot data to the DMC but decided that they would not; e-mail records obtained by the Office for Human Research Protections (OHRP) of the Department of Health and Human Services showed that the investigators were not convinced that the pilot data were reliably informative and were concerned that these data could lead the DMC to recommend that the study be terminated.

These issues came to light after a complaint was filed with the OHRP by a hip-pad manufacturer who objected to the assertion in the authors’ publication of the study that hip pads were ineffective in preventing hip fractures.\textsuperscript{46,47} The OHRP determined that after the investigators became aware of the excess falls on the protected side, they should have informed the trial participants, as well as the institutional review boards and others responsible for research oversight, of this potential risk. This example
serves to raise consciousness among those who are developing reports to DMCs. Second-guessing the recommendations made by DMCs is difficult, and whether the added information from the pilot study would have altered their discussion is speculative, but clearly the DMC should at least have been informed about the pilot data. This example also shows the difficulties that arise when investigators have access to interim data; the investment these investigators had in their study surely influenced their judgment with regard to the presentation of data to the DMC.

### STRETCHING DMC RESOURCES

Initially, DMCs were used in a very limited number of high-profile NIH-sponsored trials; a few leading academic researchers regularly served on these committees, and a small number of statistical centers reported data analyses to them. Today, there are hundreds, perhaps thousands, of DMCs monitoring government-sponsored and industry-sponsored trials. The number of scientists who have experience with and training for DMC activity has not kept pace. As a result, many people are asked to serve on numerous DMCs; however, some DMCs are operating with less-experienced members. There is currently no plan to address this gap in the workforce. In addition, the number of statistical centers that are prepared to meet the needs of a DMC is inadequate, and their competency and experience are critical for the proper functioning of the DMC. Strategic training efforts have been recommended both for members of DMCs and for statistical analysis support teams; such efforts would include teaching, simulation of DMC processes on the basis of data from completed trials, and apprenticeships in ongoing trials.

### CONCLUSIONS

DMCs play a critical role in the conduct of clinical trials by assessing the risks and benefits of an intervention as data accumulate, thereby enhancing the safety of trial participants. Although the fundamental principles of DMCs have been described, trials can present challenges for the DMC that are not anticipated or covered in the DMC statistical analysis plan. DMCs have to be prepared for the unexpected, relying in large part on their experience and collective wisdom. We have shared a few examples to illustrate the complexity of DMC activity; however, many other scenarios have plagued DMCs, and new issues will regularly arise.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

### REFERENCES


19. Effect of metoprolol CR/XL in chronic heart failure: Metoprolol CR/XL Randomised Intervention Trial in Congestive Heart
et al. Intensive lipid lowering with simvasta

The Journal welcomes consideration of new submissions for Images in Clinical Medicine. Instructions for authors and procedures for submissions can be found on the Journal’s website at NEJM.org. At the discretion of the editor, images that are accepted for publication may appear in the print version of the Journal, the electronic version, or both.