Introduction to Case Studies Showing Harmful Effects of the Intervention

David L. DeMets
Curt D. Furberg
Lawrence M. Friedman

Based on the Declaration of Helsinki and established principles of ethics in research, one major objective of data monitoring is the protection of trial participants from being harmed by the study intervention. This section presents nine cases of clinical trials showing evidence of harm attributed to a trial intervention, eight of which were terminated earlier than planned. Three of the trials tested more than one active intervention (the Coronary Drug Project (CDP—Case 12), the Cardiac Arrhythmia Suppression Trial (CAST—Case 13). and the Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT—Case 18). In this section, we focus on the treatment arms associated with harmful effects. The complexities behind the decision to recommend trial termination are illustrated.

Eight of the nine trials were placebo-controlled, a design feature that facilitates determination of harm. The ninth trial, ALLHAT (Case 18), was an active-control trial designed to determine whether any of three newer and costlier antihypertensive drugs was superior to a generic diuretic. Due to the lack of a placebo group, this trial evaluated the treatment effect along the axis of superiority-indifference-inferiority rather than the axis efficacy-indifference-harm, which applies to placebo-controlled trials. In active-control trials, inferiority does not automatically mean harm, since an inferior intervention could be less beneficial or neutral. The magnitude of the inferiority may indirectly shed some light on the question of harm. The two-fold higher risk of congestive heart failure (CHF) in the doxazosin group of ALLHAT compared to the chlorthalidone group, in spite of only a small difference in systolic blood pressure reduction, suggests a harmful effect of doxazosin on CHF risk. All-cause mortality was the pre-specified primary outcome in four trials (CDP, the Prospective Randomized Milrinone Survival Evaluation Trial [PROMISE—Case 14], the Diaspirin Cross-Linked Hemoglobin for Emergency Treatment of Post-Traumatic Study [Baxter DMC—Case 16] and the Moxonidine

Congestive Heart Failure Trial [MOXCON—Case 19]) and a pre-specified secondary outcome in the other five trials (CAST, the Carotene and Retinol Efficacy Trial [CARET—Case 15], the Heart and Estrogen/progestin Replacement Study [HERS—Case 17], ALLHAT, and the Placebo-Controlled Trial of Daclizumab in Acute Graft-versus-Host Disease [ECOG—Case 20]). Five trials had a disease-specific primary outcome (CAST, CARET, HERS, ALLHAT, and ECOG).

The CDP (Case 12) may have been the first clinical trial with external monitoring, although the Data Monitoring Committee was not appointed until a couple of years after the trial was launched. Many of the methods for data monitoring that we use today were developed during the course of CDP. The other eight trials all had pre-specified monitoring guidelines or "stopping rules."

Two trials were designed with a one-sided hypothesis. The daclizumab trial (Case 20) tested whether active treatment was superior to placebo at a statistical significance level of 0.05. Interestingly, when the trial was terminated due to excess mortality in the actively treated group, there was no difference between the daclizumab and placebo groups for the primary outcome. CAST (Case 13) was also initially designed to determine benefit at the statistical significance level of 0.05. However, the monitoring committee changed the alpha level for benefit to $p = 0.025$ and added the same significance level for the testing of harm. The lesson from these two trials is that an adverse effect of the intervention, however promising, can never be ruled out.

Most trials employed symmetric boundaries during the monitoring process. In other words, they required similar strength of evidence for claim of benefit and harm. One exception was MOXCON (Case 19), which had an asymmetric boundary with stricter criteria for benefit. The monitoring of the moricizine–placebo comparison in CAST (Case 13) also relied on an asymmetric lower boundary.

Five of the trials (CDP [two treatment arms], CAST, PROMISE, CARET, and MOXCON) were terminated due to group differences in the pre-specified primary outcome. In two trials, the recommendations for early termination were based on observed group differences in secondary or other outcomes (CDP [one treatment arm] and ALLHAT). In HERS (Case 17), one of the components of the primary outcome (non-fatal MI plus CHD death) appeared early destined to cross the stopping boundary. An excess of early CHD deaths (a nominal $p = 0.02$) was observed in the hormone therapy group. For a variety of good reasons, the board voted to continue the trial. Later this trend reversed itself and the relative hazard at trial termination was 0.99. In the middle years, the risk of one of the pre-specified secondary outcomes, venous thromboembolic events (VTE), crossed the stopping boundary. Rather than

recommending trial termination, the board advised the Steering Committee to inform all participants of this risk, to modify the study protocol to reduce the future risk of thromboembolic complications and to publish the VTE data (*JAMA* 1997;278:477).

Futility was, in addition to harm, a consideration in the recommendation to terminate four trials (CDP [one treatment arm], CARET, ALLHAT, and MOXCON). In two trials (CARET and MOXCON) external scientific evidence was considered in the decision making.

Deciding to terminate a trial is difficult and it is very common that the monitoring committee votes will be split. This leaves the sponsor in a difficult position. One solution to this dilemma is consultation with a second advisory group. In fact, in three of the trials (CDP [one treatment arm], CARET, and ALLHAT), a second committee was formally consulted before a final decision to terminate was made.