CASE 2

Data and Safety Monitoring in the Beta-Blocker Heart Attack Trial: Early Experience in Formal Monitoring Methods

Lawrence M. Friedman
David L. DeMets
Robert Hardy

ABSTRACT

The Beta-Blocker Heart Attack Trial (BHAT) compared the beta-blocker propranolol against placebo in 3,837 people who had recently had a myocardial infarction. The primary outcome was total mortality. The trial ended nine months ahead of schedule because of clear benefit from propranolol. The independent monitoring committee considered several newly developed statistical approaches in recommending early stopping, as well as other factors, including what had been communicated in the consent form to the participants.

INTRODUCTION AND BACKGROUND

In the 1970s, it was thought that blockade of the beta-adrenergic receptors might be beneficial for patients with myocardial infarction. This led to the conduct of several clinical trials. Some of these trials treated patients with intravenous beta-blockers at the time of the acute MI;[1–3] others began treatment intravenously at the time of the acute event and continued with oral beta-blockers after hospital discharge;[4] still others began long-term oral treatment of patients after the acute recovery phase.[5,6,7] Relevant to the development of BHAT were concerns that the long-term trials that had been conducted were inconclusive. In particular, some were underpowered, one used a beta-blocker that had unexpected serious toxicity, and some may have used inadequate doses of medication.[8] Therefore, a workshop, conducted by the National Heart, Lung, and Blood Institute (NHLBI) recommended that another long-term trial with a sufficiently large sample size and using appro-

priate doses of a beta-blocker with which there was considerable experience and a known toxicity profile, such as propranolol, be conducted.[9]

PROTOCOL DESIGN

The design of BHAT, which was sponsored by the NHLBI, called for enrollment of 4,020 patients, aged 30–69 years, who had had a myocardial infarction 5–21 days prior to randomization. The primary objective of the study was to determine if long-term administration of propranolol would result in a difference in all-cause mortality. The alpha level was set at two-tailed 0.05, with 90% power to detect a 28% relative change in mortality, from a three-year rate of 18% in the control (placebo) group to 12.96% in the intervention group. This projected benefit was derived from the earlier beta-blocker trials. It was also assumed that over the three-year average follow-up, 26% of patients assigned to propranolol would discontinue the study drug, and 21% of patients assigned to placebo would begin taking a beta-blocker.[9] Thus, after taking into account non-adherence, the adjusted estimated control group event rate was 17.46% and the adjusted estimated treatment group event rate was 13.75%. The adjusted relative benefit was 21.25%, rather than 28%.

Participants were randomly assigned to either daily propranolol or placebo. Initial dosing was propranolol, 40 mg, three times a day or matching placebo. Depending on the serum drug level at one month, the dose was changed to either 60 mg three times a day or 80 mg three times a day. Approximately 80% of the participants randomized to propranolol were on the 60-mg regimen. Participants assigned to placebo also had their dose formulation changed in order to preserve the double-blind. Participant accrual was planned for two years, with follow-up for a minimum of two years and a maximum of four years (average follow-up of three years).

Participant enrollment began in 1978; a total of 3,837 participants were enrolled, instead of the planned 4,020. This reduced the power from the planned 90% only a small amount (to 89%), assuming all other factors remained unchanged.

As noted, several studies of beta-blockers had been conducted prior to BHAT. In addition, other studies were ongoing simultaneously. One, a trial of timolol, which was similar in many respects to BHAT, was published in April 1981.[10] This trial of 1,884 survivors of an acute myocardial infarction showed a statistically significant reduction in all-cause mortality, from 16.2% to 10.4%, during a mean follow-up of 17 months.[10] At this point, BHAT was no longer enrolling patients, but follow-up was continuing.

Six months later, in October 1981, the independent Policy and Data Monitoring Board (PDMB), which was advisory to the NHLBI, recommended

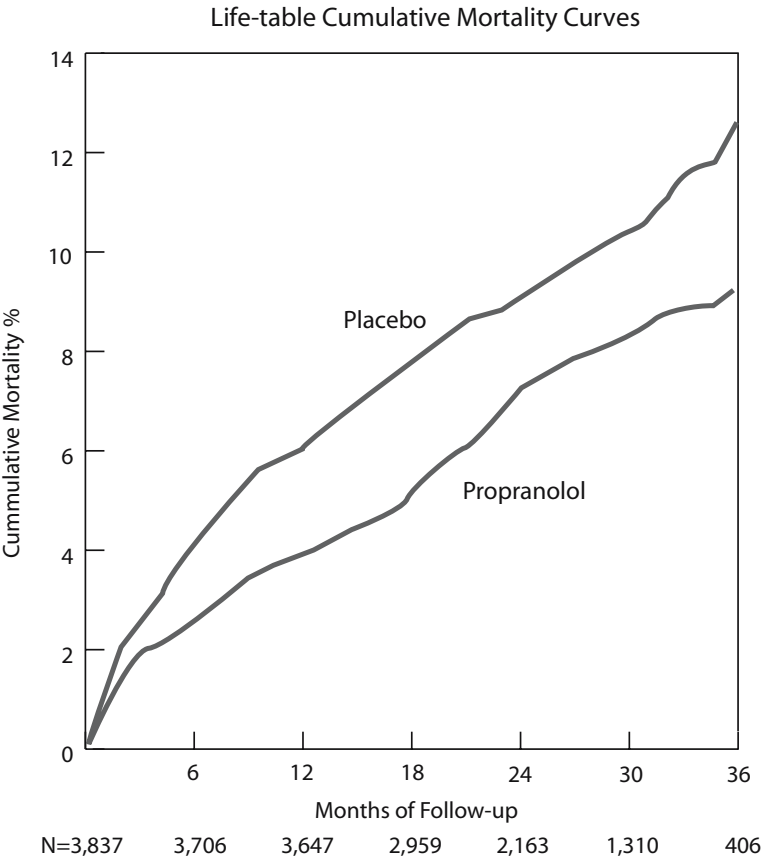

Figure 1 Life-table Cumulative Mortality Curves. Reprinted from BHAT[11] with permission from *JAMA*.

that BHAT be stopped, nine months ahead of schedule, because of a significant reduction in mortality in the propranolol group (Figure 1).[11]

DATA MONITORING EXPERIENCE

Early in the trial, the PDMB considered several monitoring boundaries. These included the ones suggested by Pocock[12] and Peto.[13] However, the PDMB selected the then recently published O'Brien–Fleming procedure for establishing monitoring boundaries.[14] The reasons for selecting this procedure were that (1) it protects the overall alpha; (2) it is quite conservative early in the study when small numbers and enrollment of participants who

are perhaps not representative of the final study sample could lead to misleading conclusions; (3) the final critical value is close to the nominal critical value, so that the power and sample size are not affected and communication of the outcome to the medical community is more straightforward; and (4) the decreasing boundary over time appropriately reflects confidence in the accumulating data.

The PDMB first reviewed the BHAT data in May 1979. Subsequent data reviews were to occur approximately every six months, until the scheduled end of the trial in June 1982.The logrank z-value exceeded the conventional 1.96 critical value for a nominal p of 0.05 at the October, 1979 meeting of the PDMB. However, because of the conservative nature of the O'Brien–Fleming boundaries early in the study, this was far from significant. At the regularly scheduled meeting in April 1981, the PDMB reviewed not only the accumulating BHAT data, but the results of the timolol trial that had just been published.[10] The PDMB recommended that BHAT continue, primarily because, despite the timolol findings, the BHAT data did not show convincing evidence of benefit. Not only had the monitoring boundary not been crossed, but the long-term effect on mortality and possible adverse events was unknown. Importantly, all patients in BHAT had been in the trial for at least six months post-infarction, and there was no evidence that beta-blockers started after that time produced benefit. Thus, there was not an ethical concern about leaving the participants on placebo off treatment. The PDMB advised that the study investigators be informed of the timolol results. However, it also advised that because there had been conflicting results from other beta-blocker trials, the positive results of the timolol trial should not preclude the continuation of BHAT. Furthermore, timolol was not then available for sale in the United States, where BHAT was being conducted.

At its October 1981 data review, the PDMB noted that the upper O'Brien–Fleming boundary had been crossed.[14] The normalized logrank statistic was then 2.82, which exceeded the boundary value of 2.23. (At the prior meeting of the PDMB, in April, 1981, the logrank statistic was 2.34, which was just short of the then boundary value of 2.44.) Figure 2 shows the logrank statistics at each time, along with the upper monitoring boundary.[15]

The PDMB considered a number of factors in addition to the monitoring boundaries in its recommendation to stop early. One was conditional power; that is, the likelihood that the observed results would remain significant if BHAT were to continue to its scheduled end.[15–17] Based on prior control group data, several estimates of the number of future events were made. If there were no additional benefit from propranolol (i.e., if the null hypothesis were to hold for the next nine months), the conditional probability of seeing a significant benefit at the end of the trial was calculated for these

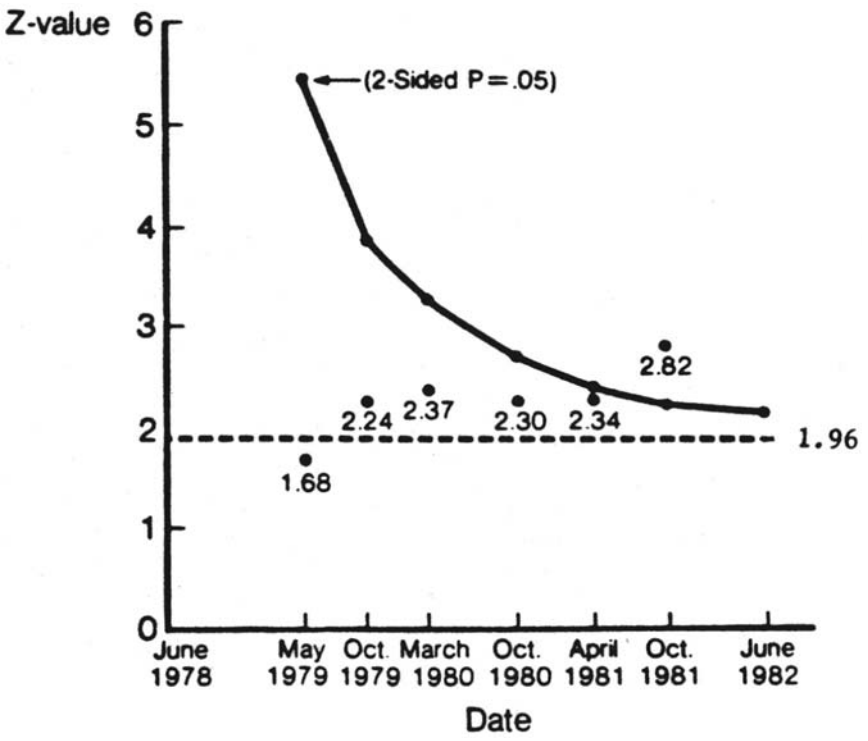

Figure 2 Beta-Blocker Heat Attack Trial Monitoring Boundary. Reprinted from DeMets et al.[15] with permission from *Control Clin Trials*.

different numbers of control group events. Under the most likely estimate, the error rate would at most be 5.5%, or only 0.5% more than the original type I error of 5%.[16,17]

The PDMB also looked at the additional precision that would derive from the added events. All participants had already been followed for one year, and only a few remained to be seen for their second annual visit. Therefore, the results for those years were complete, or essentially so. The additional precision for year 2 would have been minor. The year 3 data would have been somewhat improved by additional follow-up, as only about half of the participants had been seen for their third year visit. But even here, the increase in precision, as reflected by the narrowing of the standard error in the propranolol group from 0.0079 to 0.0068, and in the placebo group from 0.0130 to 0.0082, would have been modest. Very few participants had completed a four-year visit, so additional follow-up would have been helpful in estimating benefit at that point.[15]

The PDMB discussed whether the practicing medical community would be less likely to accept the BHAT results if the study were stopped early than if it were to continue to its scheduled end. Because the BHAT results were consistent with the recently published trial of timolol, this was not thought to be a serious problem. Ethical considerations were also raised. Although all of the control group participants were well past the time after their MI when propranolol was started, some might suffer a repeat MI. If so, it would be important for them to be aware of the BHAT results. For patients in the general public, knowledge of the BHAT outcome would be important to their medical care.

The PDMB reviewed a checklist of items to be considered when possibly recommending early termination. This checklist had been developed by one of the members of the PDMB.[18] In addition to the factors mentioned above, the list included examination of comparability of baseline variables and subsequent management of patients between the groups whether outcome ascertainment was sufficiently complete and equal in the groups consistency of subgroup results and overall benefit-to-risk, taking into account multiple outcomes and adverse events. None of these factors suggested that the observed outcome was due to anything other than the administration of propranolol or that the validity of the reported results would be seriously challenged.

A further consideration was the consent that had been signed by the study participants. The consent stated that "if propranolol proves to be beneficial for heart attack patients, the study will be stopped as soon as this is known. If, on the other hand, it proves to be harmful, the study will also be stopped, or those who have a tendency to be harmed will be removed from the study." Because the monitoring boundary had been crossed, it was argued that this "contract" with the patients required stopping the study.

In summary, the points in favor of early stopping were—

1. The pre-specified monitoring boundary had been crossed and propranolol was clearly beneficial.
2. Conditional power calculations indicated that there was little likelihood that the conclusions of the study would be changed if follow-up were to continue.
3. The gain in precision of the estimated results for the first two years would be tiny, and only modest for the third year.
4. The results were consistent with those of another beta-blocker trial.
5. There would be potential medical benefits to both study participants on placebo and to heart attack patients outside the study.
6. Other factors, such as subgroup examinations and baseline comparability, confirmed the validity of the findings.

7. The consent form clearly called for the study to end when benefit was known.

The points in favor of continuing until the scheduled end were—

1. Even though slight, there remained a chance that the conclusions could change.
2. Because therapy would be continued indefinitely, it would be important to obtain more long-term (4 year) data.
3. It would be important to obtain more data on subgroups and secondary outcomes.
4. The results of a study that stopped early would not be as persuasive to the medical community as would results from a study that went to completion, particularly given the mixed results from earlier trials.

The PDMB considered these issues and, in a closely divided vote, recommended early stopping. The NHLBI accepted this recommendation, and the investigators were informed of the decision.

As noted earlier, the sample size estimate assumed a three-year mortality rate of 18% in the control group. The mortality at one year was 5.99%. However, the two-year mortality was 9.15% and the three-year mortality (with a relatively small number of deaths) was 12.52%. At the time BHAT was stopped, the average follow-up was 25 months, with a control group mortality of 9.8%.[11] Thus, except for the first year, which included the high-risk early post-MI period, the observed mortality was considerably less than expected. However, the mortality in the propranolol group after the average follow-up of 25 months was 7.2%, an observed relative benefit of 26.5%, rather than the estimated relative benefit (after adjustment for nonadherence) of 21.25%.

LESSONS LEARNED

1. BHAT was one of the first major trials to use the O'Brien–Fleming approach to sequential boundaries. It proved particularly helpful in fostering a cautious attitude with regard to claiming significance prematurely. Even though conventional significance was seen early in the study, the use of sequential boundaries gave the study added credibility and probably helped make it persuasive to the practicing medical community.

2. The use of conditional power added to the persuasiveness of the results, by showing the extremely low likelihood that the conclusions would change if the trial were to continue to its scheduled end.

3. The decision-making process involves many factors, only some of which are statistical. Confidence that the data being observed are correct,

reasonably complete and current, and are not confounded by baseline or subsequent treatment imbalances provides assurance that the conclusions are due solely to the random assignment of the intervention. Use of a checklist of these factors helps ensure that they are adequately considered.

4. The lower than expected event rate in the control group is another demonstration of the need for randomized trials to assess treatment benefit or harm.

5. Ethical issues are paramount. If a study similar to the one being conducted presents results while the study is ongoing, the implications must be faced fully and honestly. The effect of the completed study on participant medical care and safety needs to be considered, as does the question as to whether the ongoing study remains important and ethical. The investigators need to be fully informed as to the data and relevance of the reported study, as do Institutional Review Boards. Study participants should also be informed of information pertinent to their medical care and continued involvement in the trial. During any discussion about continuation or early termination, the monitoring committee must be aware of the "contract" that was made with the subjects, namely, what was said during the informed consent process.

6. In the planning stages of a long-term trial, it is rare that all issues that might affect early termination can be anticipated. Because statistical considerations are only part of the deliberations, members of monitoring committees must always use their best judgment. The trial data themselves usually will not provide clear answers to key questions such as whether the results will be sufficiently persuasive to change practice, or the overall balance of benefits and risks. Judgment from a monitoring committee that contains members with diverse backgrounds and experience must come into play. Recommendations to stop or continue a trial are almost always accepted by the study sponsor, whose responsibility it is to implement those recommendations. Particularly when a recommendation involves a close vote, as in the case of BHAT, the study sponsor must also use judgment in its decision to accept or reject the recommendation. In BHAT, the recommendation to stop was accepted. But in situations where the recommendation is not accepted, the sponsor must fully and openly explain why it made its decision.

REFERENCES

1. Norris RM, Clarke ED, Sammel NL, et al. 1978. Protective effect of propranolol in threatened myocardial infarction, *Lancet* 2:907–909.
2. Sloman G, Stannard M. 1967. Beta-adrenergic blockade and cardiac arrhythmias. *BMJ* 4:508–512.
3. Waagstein F, Hjalmarson AC. 1976. Double-blind study of the effect of cardioselective beta-blockade on chest pain the acute myocardial infarction. *Acta Med Scand* (suppl) 587:201–208.

4. Andersen MP, Bechgaard P, Frederiksen J, et al. 1979. Effect of alprenolol on mortality among patients with definite or suspected acute myocardial infarction: Preliminary results. *Lancet* 2:865–868.
5. Ahlmark G, Saetre H, Korsgren M. 1974. Reduction of sudden deaths after myocardial infarction, *Lancet* 2:1563.
6. Multicentre International Study. 1977. Supplementary report: Reduction in mortality after myocardial infarction with long-term beta-adrenoceptor blockade. *BMJ* 2:419–421.
7. Wilhelmsson C, Vedin JA, Wilhelmsen L, et al. 1974. Reduction of sudden deaths after myocardial infarction by treatment with alprenolol. Preliminary results. *Lancet* 2:1157–1160.
8. Furberg CD, Friedewald WT. The effects of chronic administration of beta-blockade on long-term survival following myocardial infarction. 1978. In Braunwald E (ed.): *Beta-Adrenergic Blockade: A New Era in Cardiovascular Medicine*, Excerpta Medica, Amsterdam.
9. Byington RP for the Beta-Blocker Heart Attack Trial Research Group: Beta Blocker Heart Attack Trial. 1984. design, methods, and baseline results. *Control Clin Trials* 5:382–437.
10. The Norwegian Multicenter Study Group. 1981. Timolol-induced reduction in mortality and reinfarction in patients surviving acute myocardial infarction. *N Engl J Med* 304:801–807.
11. β-Blocker Heart Attack Trial Research Group. 1982. A randomized trial of propranolol in patients with acute myocardial infarction. 1. Mortality results. *JAMA* 247:1707–1714.
12. Pocock SJ. 1977. Group sequential methods in the design and analysis of clinical trials. *Biometrika* 64:191–199.
13. Peto R, Pike MC, Armitage P, et al. 1976. Design and analysis of randomized clinical trials requiring prolonged observations of each patient. I. Introduction and design, *Br J Cancer* 34:585–612.
14. O'Brien PC, Fleming TR. 1979. A multiple testing procedure for clinical trials, *Biometrics* 35:549–556.
15. DeMets DL, Hardy R, Friedman LM, Lan KKG. 1984. Statistical aspects of early termination in the Beta-Blocker Heart Attack Trial. *Control Clin Trials* 5:362–372.
16. Lan KKG, Simon R, Halperin M. 1982. Stochastically curtailed tests in long-term clinical trials. *Comm Stat* C1:207–219.
17. Halperin M, Lan KKG, Ware JH, Johnson NJ, DeMets DL. 1982. An aid to data monitoring in long-term clinical trials. *Control Clin Trials* 3:311–323.
18. Canner PL. 1983. Monitoring of the data for evidence of adverse or beneficial treatment effects. *Control Clin Trials* 4:467–483.