PERSPECTIVE | Annals of Internal Medicine

Ethical Issues in Stopping Randomized Trials Early Because of Apparent Benefit

Paul S. Mueller, MD, MPH; Victor M. Montori, MD, MSc; Dirk Bassler, MD, MSc; Barbara A. Koenig, PhD; and Gordon H. Guyatt, MD, MSc

Stopping randomized trials early because of an apparent benefit is becoming more common. To protect and promote the interests of trial participants, investigators may feel obligated to stop a trial early because of the apparent benefit of a study treatment (compared with placebo or other treatment). There are, however, serious ethical problems with doing so. Truncated trials systematically overestimate treatment effects; in cases where the number of accrued outcome events is small, the overestimation may be very large. Generating seriously inflated estimates of treatment effect violates the ethical research requirement of scientific validity. Subsequent use of inflated estimates to inform clinical decision making and practice guidelines violates the ethical requirements of social value and a favorable risk–benefit ratio. Researchers should ensure that a large number of outcome events accrues before stopping a trial and then continue recruitment to assess whether positive trends continue. This can balance the need to protect research participants with the ethical requirements of scientific validity, social value, and a favorable risk–benefit ratio.

Ann Intern Med. 2007;146:878-881. www.annals.org

For author affiliations, see end of text.

Stopping randomized trials early because of an apparent benefit is a growing phenomenon. A recent systematic review found that the number of randomized trials stopped early for benefit has more than doubled since 1990 (1). To protect and promote the interests of trial participants, investigators may feel ethically obligated to stop a trial early because of the unexpected harm or apparent benefit of a study treatment. If a study treatment's benefit far outweighs its adverse effects, is it not unethical to continue enrolling patients in a trial in which, as is typically the case, patients have a 50% chance of receiving a placebo or an inferior treatment?

In this article, we argue that stopping a randomized trial early for apparent benefit is often unethical and can be justified only under restricted circumstances. If the scientific community were to accept our arguments, then the approach that investigators, institutional review boards, and data monitoring committees take to the practice of stopping trials early for apparent benefit would substantially change.

Ethical Considerations

Emanuel and colleagues (2) describe a framework of 7 requirements for determining whether clinical research is ethical. We use this framework to identify and assess the ethical issues raised by stopping trials early because of apparent benefit (**Table**).

See also:

Print
Related article 882

Web-Only
Conversion of table into slide

Scientific Validity

The purpose of a trial of alternative interventions is to generate an estimate of treatment effect that closely approximates the true effect and is not misleading. This requires application of scientific procedures that yield valid and reliable data and thus minimize both systematic and random error.

A systematic review of randomized trials stopped early for apparent benefit (1) found that many of the trials yielded implausibly large treatment effects; the median relative risk was 0.53. Apparent large treatment effects occurred much more frequently when trials accrued only a small number of events. The odds of a treatment effect larger than the overall median relative risk of 0.53 was 28 times greater (95% CI, 11 to 73) among trials in which fewer than the median of 66 events accrued than among trials in which more events accrued. These results, which are consistent with predictions from statistical theory (3), suggest that stopping trials early for apparent benefit will systematically overestimate treatment effects.

The scientific validity of trials that are stopped early is further compromised when trials yield inconclusive data about outcomes that did not influence trial truncation but are nonetheless important to patients, such as disease-free survival, symptom control, quality of life, and adverse effects of treatment. For example, a trial of vitamin E supplementation in premature newborns that was stopped early because of an apparent reduction in intracranial hemorrhage (4) failed to detect the increase in sepsis associated with vitamin E supplementation that subsequent trials identified (5).

Social or Scientific Value and Favorable Risk–Benefit Ratio

It is understandable that investigators focus their ethical obligations on research participants. Such focus, however, risks neglecting obligations to society. The tendency of truncated trials to overestimate the effect of a treatment

Table. **Ethical Violations Resulting from Stopping a Trial Early for Apparent Benefit**

Ethical Requirement	Potential Violation of the Ethical Requirement
Scientific validity	Statistical theory and empirical evidence suggest that trials stopped early overestimate treatment effects
Social or scientific value and favorable risk–benefit ratio	Overestimates of benefit will lead to misguided practice recommendations and suboptimal clinical practice
Participant consent and respect for potential and enrolled participants	Informing patients of early results compromises a trial's scientific validity and provides no assurance that potential patients will receive the putatively beneficial therapy or choose to do so, in the face of unreliable benefit and insufficient risk information
Independent review	Investigators, funding agencies, journals, and patients have incentives for stopping early; informed ethical review and data monitoring committee oversight are therefore crucial
Fair participant selection	None

on the end point that resulted in trial truncation and to yield insufficient data about other important outcomes endangers the wider community to whom the results will be applied (6). On reviewing the results of a truncated trial, astute clinicians might appropriately conclude that the benefits of the intervention remain uncertain. However, less skeptical clinicians might assume that the results are true and inappropriately expose patients to the intervention and its unknown harms.

Consider the results of a trial in which the investigators continued to enroll patients even though prespecified criteria for early stopping were met. Two interim analyses of a randomized trial of 5 versus 4 courses of chemotherapy in patients with acute myeloid leukemia (7) found apparent large benefits to the 5-course regimen (relative odds reduction of 53% [CI, 23% to 71%; $P = 0.003$] in the first analysis and 45% [CI, 20% to 62%; $P = 0.0002$] in the second analysis). Finding these results too good to be true, the data monitoring committee recommended continuing the trial, which ultimately showed a trend in favor of the 4-course regimen. Had the investigators terminated the trial in accordance with their stopping rule, subsequent patients with leukemia may have experienced the toxicity of an additional course of chemotherapy without benefit.

Harm resulting from the misleading findings of truncated trials can be compounded if the findings influence the recommendations of clinical practice guideline panels. Investigators conducting a trial that involved patients undergoing vascular surgery (8) stopped the trial early when 2 of 53 patients randomly assigned to receive the β-blocker bisoprolol and 18 of 59 control patients had major cardiovascular events (relative risk reduction, 90% [CI, 59% to 98%]). These results contributed to recommendations by the American Heart Association and the American College of Cardiology favoring administration of β-blockers to patients with cardiac risk factors who were undergoing noncardiac surgery (9). However, these results contradict those of 2 much larger subsequently published trials, neither of which suggested that β-blockers reduce cardiac risk in patients undergoing noncardiac surgery (10, 11).

Further social detriment may occur when clinicians compromise the ability of others to conduct more definitive studies by placing undue confidence in the results of a truncated trial. Investigators (including 2 contributors to this article) who obtained funding for a trial of β-blockers in noncardiac surgery with an enrollment target of 10 000 patients (12) faced challenges in persuading clinicians that the question remained unanswered.

Participant Consent and Respect for Participants

Key prerequisites for informed consent include the participant's decision-making capacity and voluntariness and whether he or she had received adequate information to decide that participation in the research was in alignment with his or her values and goals. However, informed consent is not a single event, but it is an ongoing collaboration between participants and investigators. When important changes occur during a trial, investigators should inform participants of the changes.

One justification for stopping a trial early for benefit is to inform study participants of the preliminary results and offer them the superior treatment. According to this argument, uncertainty about the relative merits of alternative interventions (equipoise) has been lost and informed clinicians and patients will overwhelmingly choose the superior treatment (13). However, as we have pointed out, the astute clinician or patient may remain skeptical about a treatment's apparent benefits if the findings come from a truncated trial. Unfortunately, many clinicians and even more patients probably will not have the knowledge and understanding to appropriately interpret the results. Disclosing interim results to study participants may therefore prove misleading.

Furthermore, if investigators were to continue a trial after informing patients of the interim results, patients would be unblinded and may cross over or leave the trial. These behaviors create problems in interpreting trial results by further weakening inferences about the efficacy and safety of the intervention and compromising the ethical requirement of scientific validity.

Finally, stopping a trial early does not guarantee that current and potential trial participants will receive the apparently beneficial treatment (assuming that one believes they should). Studies of dissemination of new treatments reveal that long delays, such as those between reports of randomized trials and recommendations of experts in re-

view articles and textbooks, are common (14). Continuing a 2-group trial gives participants at least a 50% chance of receiving the experimental treatment, whereas if the trial is stopped early, the probability that participants will receive the treatment due to rapid dissemination is likely to be considerably less than 50%.

Independent Review

Trials may have stopping rules that allow early termination because of genuine (although misguided) ethical concerns. However, investigators, trial sponsors, journals, and patients may all have additional motives for stopping trials early for apparent benefit. For example, truncated trials that report a large treatment effect tend to be published in the most prestigious medical journals (1), which enhances the careers of the investigators and increases the likelihood that they will receive grants. Funding agencies have an interest in stopping trials early to minimize research costs. Pharmaceutical and for-profit sources that financially support trials are interested not only in controlling costs but also in the publicity and market share that result from reporting a trial stopped early for apparent benefit. Medical journals are interested in these trials because of publicity and citations, which result in increased journal impact factor, prestige, and advertising revenue. And patients and their advocates are motivated to stop a trial early when the experimental intervention is promising in order to hasten delivery of the intervention to clinical practice. All of these motives may affect investigators' decisions and encourage an inappropriately early stop to a trial. These considerations mandate that institutional review boards and data monitoring committees understand the principles outlined in this article and insist on appropriate standards for stopping a trial early for apparent benefit to maintain the ethical integrity of clinical trials.

STOPPING TRIALS EARLY FOR BENEFIT: POTENTIAL SOLUTIONS

We have described the dangers of stopping trials early for apparent benefit and the ethical arguments for continuing trials to their planned conclusion. This issue, however, is not 1-sided. If a treatment truly has large benefits, one would want rapid dissemination of knowledge of that treatment. Indeed, if the treatment is associated with minimal harm, there is a legitimate ethical argument for terminating a trial early and thus facilitating rapid dissemination of the results.

Consider a randomized trial that tested the effect of intensive lifestyle modification on the incidence of diabetes (15). The trial initially planned a sample size of 3000 participants. After enrolling 3234 participants and observing a total of 436 events, the investigators found a plausible relative risk reduction of 42% (CI, 34% to 52%) and terminated follow-up of the enrolled trial participants 1 year earlier than planned. Other large trials (16, 17) confirmed the accuracy of this estimate of benefit.

Is it appropriate to stop a trial early if the apparent treatment benefit is unlikely to largely overestimate the true benefit? How can one be certain that is the case? Application of currently published stopping rules do not achieve this goal (3). Investigators may set a relatively rigorous *P*-value threshold (say, 0.001) but may examine their findings early when few events have accrued. If their trial crosses the prespecified threshold at this early point, the apparent treatment benefit almost certainly represents a very large overestimate. This was the case in the randomized trial of bisoprolol in patients undergoing noncardiac surgery (8): The investigators took a single look at their data according to an O'Brien–Fleming stopping rule with a *P*-value threshold of 0.001. However, if β-blockers do reduce vascular events in patients undergoing noncardiac surgery, the effect is far more modest than the 90% relative risk reduction that the trial results suggested.

The solution is to demand that a large number of events accrue before investigators or data monitoring committees examine interim data. However, just how many events are required to avoid large overestimates of effect remains to be determined. Without this knowledge, we have outlined the ethical considerations that mandate a conservative approach. Requiring that 200 to 400 events accrue before invoking a stopping rule associated with a low *P* value (1), such as 0.001, and then continuing enrollment and follow-up for a further period to ensure that the trend continues should minimize seriously misleading results.

The approach we describe should be included in trial protocols, reported when registering a trial, and conveyed to trial participants. Furthermore, data monitoring committees should not deviate from these rules (18).

In addition to more stringent early stopping rules, the ethical issues that we have discussed suggest further recommendations. First, journals should require that reports of truncated trials describe the rationale for early stopping, including the statistics and rules applied. These descriptions are not currently routine (1). Second, investigators and data monitoring committees should consider the context of the illness and the treatment being studied when deciding to stop a trial early. The adverse consequences of overestimating a treatment effect are greater if the treatment is toxic and expensive, making the decision to stop early grave. Even when the treatment seems to incur minimal toxicity and cost, investigators and data monitoring committees should bear in mind that approximately 25% of new drugs eventually prove to have serious adverse effects that were unexpected when they entered the market (19).

In summary, understanding the biases associated with stopping randomized trials early, adopting conservative stopping rules, and considering the potential harms to subsequent patients of overestimating treatment effects will reduce the ethical problems resulting from stopping trials early.

From Mayo Clinic College of Medicine, Rochester, Minnesota; University Children's Hospital, Tuebingen, Germany; and McMaster University, Hamilton, Ontario, Canada.

Potential Financial Conflicts of Interest: *Consultancies:* B.A. Koenig (Affymetrix Corporation).

Requests for Single Reprints: Paul S. Mueller, MD, MPH, Program in Professionalism and Bioethics, Department of Medicine, Mayo Clinic College of Medicine, 200 First Street SW, Rochester, MN 55905; e-mail, mueller.pauls@mayo.edu.

Current author addresses are available at www.annals.org.

References

1. **Montori VM, Devereaux PJ, Adhikari NK, Burns KE, Eggert CH, Briel M, et al.** Randomized trials stopped early for benefit: a systematic review. JAMA. 2005;294:2203-9. [PMID: 16264162]
2. **Emanuel EJ, Wendler D, Grady C.** What makes clinical research ethical? JAMA. 2000;283:2701-11. [PMID: 10819955]
3. **Pocock SJ.** When (not) to stop a clinical trial for benefit [Editorial]. JAMA. 2005;294:2228-30. [PMID: 16264167]
4. **Fish WH, Cohen M, Franzek D, Williams JM, Lemons JA.** Effect of intramuscular vitamin E on mortality and intracranial hemorrhage in neonates of 1000 grams or less. Pediatrics. 1990;85:578-84. [PMID: 2179850]
5. **Brion LP, Bell EF, Raghuveer TS.** Vitamin E supplementation for prevention of morbidity and mortality in preterm infants. Cochrane Database Syst Rev. 2003:CD003665. [PMID: 12917978]
6. **Buchanan D, Miller FG.** Principles of early stopping of randomized trials for efficacy: a critique of equipoise and an alternative nonexploitation ethical framework. Kennedy Inst Ethics J. 2005;15:161-78. [PMID: 16149206]
7. **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. [PMID: 12559643]
8. **Poldermans D, Boersma E, Bax JJ, Thomson IR, van de Ven LL, Blankensteijn JD, et al.** The effect of bisoprolol on perioperative mortality and myocardial infarction in high-risk patients undergoing vascular surgery. Dutch Echocardiographic Cardiac Risk Evaluation Applying Stress Echocardiography Study Group. N Engl J Med. 1999;341:1789-94. [PMID: 10588963]
9. **Eagle KA, Berger PB, Calkins H, Chaitman BR, Ewy GA, Fleischmann KE, et al.** ACC/AHA guideline update for perioperative cardiovascular evaluation for noncardiac surgery—executive summary a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Committee to Update the 1996 Guidelines on Perioperative Cardiovascular Evaluation for Noncardiac Surgery). Circulation. 2002;105:1257-67. [PMID: 11889023]
10. **Brady AR, Gibbs JS, Greenhalgh RM, Powell JT, Sydes MR; POBBLE Trial Investigators.** Perioperative beta-blockade (POBBLE) for patients undergoing infrarenal vascular surgery: results of a randomized double-blind controlled trial. J Vasc Surg. 2005;41:602-9. [PMID: 15874923]
11. **Juul AB, Wetterslev J, Gluud C, Kofoed-Enevoldsen A, Jensen G, Callesen T, et al.** Effect of perioperative beta blockade in patients with diabetes undergoing major non-cardiac surgery: randomised placebo controlled, blinded multicentre trial. BMJ. 2006;332:1482. [PMID: 16793810]
12. **POISE Trial Investigators, Devereaux PJ, Yang H, Guyatt GH, Leslie K, Villar JC, et al.** Rationale, design, and organization of the PeriOperative ISchemic Evaluation (POISE) trial: a randomized controlled trial of metoprolol versus placebo in patients undergoing noncardiac surgery. Am Heart J. 2006;152:223-30. [PMID: 16875901]
13. **Freedman B.** Equipoise and the ethics of clinical research. N Engl J Med. 1987;317:141-5. [PMID: 3600702]
14. **Antman EM, Lau J, Kupelnick B, Mosteller F, Chalmers TC.** A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts. Treatments for myocardial infarction. JAMA. 1992;268:240-8. [PMID: 1535110]
15. **Knowler WC, Barrett-Connor E, Fowler SE, Hamman RF, Lachin JM, Walker EA, et al.** Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. N Engl J Med. 2002;346:393-403. [PMID: 11832527]
16. **Pan XR, Li GW, Hu YH, Wang JX, Yang WY, An ZX, et al.** Effects of diet and exercise in preventing NIDDM in people with impaired glucose tolerance. The Da Qing IGT and Diabetes Study. Diabetes Care. 1997;20:537-44. [PMID: 9096977]
17. **Tuomilehto J, Lindström J, Eriksson JG, Valle TT, Hämäläinen H, Ilanne-Parikka P, et al.** Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance. N Engl J Med. 2001;344:1343-50. [PMID: 11333990]
18. **Slutsky AS, Lavery JV.** Data safety and monitoring boards. N Engl J Med. 2004;350:1143-7. [PMID: 15014189]
19. **Lasser KE, Allen PD, Woolhandler SJ, Himmelstein DU, Wolfe SM, Bor DH.** Timing of new black box warnings and withdrawals for prescription medications. JAMA. 2002;287:2215-20. [PMID: 11980521]

Current Author Addresses: Drs. Mueller, Montori, and Koenig: Mayo Clinic College of Medicine, 200 First Street SW, Rochester, MN 55905.
Dr. Bassler: University Children's Hospital, Department of Neonatology, Calwerstrasse 7, 72076 Tuebingen, Germany.
Dr. Guyatt: McMaster University, 1200 Main Street West, Room 2C12, Hamilton, L8N 3Z5 Ontario, Canada.