



Journal club critique

Trials stopped early for benefit? Not so fast!

Alan C. Heffner¹, Eric B. Milbrandt², and Ramesh Venkataraman²

¹ Clinical Fellow, Department of Critical Care Medicine, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, USA

² Assistant Professor, Department of Critical Care Medicine, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, USA

Published online: 22 February 2007

This article is online at <http://ccforum.com/content/11/1/305>

© 2007 BioMed Central Ltd

Critical Care 2007, 11: 305 (DOI 10.1186/cc5676)

Expanded Abstract

Citation

Montori VM, Devereaux PJ, Adhikari NK, Burns KE, Eggert CH, Briel M, Lacchetti C, Leung TW, Darling E, Bryant DM, Bucher HC, Schunemann HJ, Meade MO, Cook DJ, Erwin PJ, Sood A, Sood R, Lo B, Thompson CA, Zhou Q, Mills E, Guyatt GH: Randomized trials stopped early for benefit: a systematic review. *JAMA* 2005, 294:2203-2209 [1].

Background

Randomized clinical trials (RCTs) that stop earlier than planned because of apparent benefit often receive great attention and affect clinical practice. Their prevalence, the magnitude and plausibility of their treatment effects, and the extent to which they report information about how investigators decided to stop early are, however, unknown.

Methods

Objective: To evaluate the epidemiology and reporting quality of RCTs involving interventions stopped early for benefit.

Design: Systematic review up to November 2004 of MEDLINE, EMBASE, Current Contents, and full-text journal content databases to identify RCTs stopped early for benefit.

Study selection: Randomized clinical trials of any intervention reported as having stopped early because of results favoring the intervention. There were no exclusion criteria.

Data extraction: Twelve reviewers working independently and in duplicate abstracted data on content area and type of intervention tested, reporting of funding, type of end point driving study termination, treatment effect, length of follow-up, estimated sample size and total sample studied, role of a data and safety monitoring board in stopping the study, number of interim analyses planned and conducted, and existence and type of monitoring methods, statistical

boundaries, and adjustment procedures for interim analyses and early stopping.

Data synthesis: Of 143 RCTs stopped early for benefit, the majority (92) were published in 5 high-impact medical journals. Typically, these were industry-funded drug trials in cardiology, cancer, and human immunodeficiency virus/AIDS. The proportion of all RCTs published in high-impact journals that were stopped early for benefit increased from 0.5% in 1990-1994 to 1.2% in 2000-2004 ($P < .001$ for trend). On average, RCTs recruited 63% (SD, 25%) of the planned sample and stopped after a median of 13 (interquartile range [IQR], 3-25) months of follow-up, 1 interim analysis, and when a median of 66 (IQR, 23-195) patients had experienced the end point driving study termination (event). The median risk ratio among truncated RCTs was 0.53 (IQR, 0.28-0.66). One hundred thirty-five (94%) of the 143 RCTs did not report at least 1 of the following: the planned sample size ($n = 28$), the interim analysis after which the trial was stopped ($n = 45$), whether a stopping rule informed the decision ($n = 48$), or an adjusted analysis accounting for interim monitoring and truncation ($n = 129$). Trials with fewer events yielded greater treatment effects (odds ratio, 28; 95% confidence interval, 11-73).

Conclusion

RCTs stopped early for benefit are becoming more common, often fail to adequately report relevant information about the decision to stop early, and show implausibly large treatment effects, particularly when the number of events is small. These findings suggest clinicians should view the results of such trials with skepticism.

Commentary

Randomized controlled trials (RCTs) stopped early for benefit are increasingly common and frequently earn publication in high impact journals. Such trials typically report impressive treatment effects and generate considerable enthusiasm. Yet, there may be reason to be cautiously skeptical when a trial is stopped early for benefit. In the current study, Montori and coworkers [1] systematically reviewed RCTs of any intervention reported as having been stopped early because of results favoring the intervention. Their review included trials with results published in MEDLINE, EMBASE, Current Contents, and full-text journal content databases up to November 2004. The authors identified 143 RCTs stopped early for benefit, the majority of which were industry-funded drug trials in cardiology, cancer, and AIDS. They noted that the number of trials stopped early has increased significantly over the past 15 years and that 94% of these failed to report at least one of several key details about trial design or the decision to stop, as required by the CONSORT guidelines [2,3]. Montori and coworkers also found a strong inverse relationship between the reported event rate and estimated treatment effect, meaning that small trials with few events were likely to report large treatment effects. The median risk ratio (RR) of the truncated trials was an implausible 0.53. Comparatively, not a single study with more than 195 outcome events generated a RR < 0.50. Because the decision to stop is typically driven by highly significant p-value thresholds, trials stopped early because of apparent benefit frequently show large treatment effects. Ioannidis recently highlighted a startlingly similar association between small sample size and likelihood of a trial's findings being challenged and refuted over time [4].

A timely example of a trial that could have been stopped early but was not is OPTIMIST (Optimized Phase 3 Tifacogin In Multicenter International Sepsis Trial) [5]. At a planned interim analysis, tifacogin appeared to provide a clinically and statistically significant survival benefit compared to placebo (29.1% vs 38.9%, $P=0.006$). However, the survival difference was not large enough to activate the predefined stopping rule and the study continued. Interestingly, the benefit of tifacogin vanished as the study reached completion (34.2% vs 33.9%, $P=0.88$) (Figure). Despite thorough investigation of drug formulation, study procedures, and clerical data, the best explanation for the early inclination of benefit is the play of chance [6]. In other words, early on in the trial, the tifacogin group was on a "random high" [7,8].

The points raised by Montori and coworkers are notable for the discipline of critical care medicine; seven percent of reviewed articles were generated by our specialty. Four trials that were stopped early (perioperative beta-blockade [9], low tidal volume ventilation [10], recombinant human activated protein C [11], and tight glucose control [12]) have altered the landscape of critical care practice within the past 10 years. These trials have generated more than their fair share of controversy, which might have been minimized if they had been continued to completion.

Montori's group draws attention to legitimate issues and engenders additional questions. Most importantly, should

therapeutic trials ever be halted for benefit? The accompanying editorial [13] suggests that trials should only be stopped early for benefit when there is a highly significant p-value (e.g., $p<0.001$) and after sufficient outcome events have been observed. This more stringent approach may result in randomizing subsequent patients to a potentially inferior treatment, an option that seems at odds with safety ethics. Viewed more broadly, however, avoiding premature conclusions that could be both costly and harmful to patients ultimately compounds both clinical and societal value.

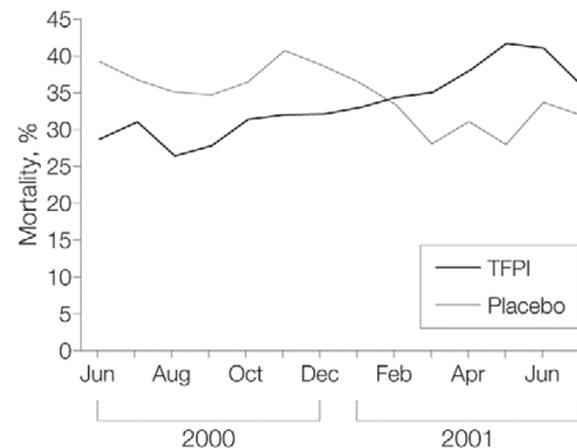


Figure: Three-month moving average for mortality, TFPI (tifacogin) vs. placebo. From Abraham, E. et al. JAMA. 2003;290:238-247. Used with permission.

Recommendation

Montori and coworkers have highlighted the potential problems associated with premature cessation of clinical trials. RCTs should only be stopped early when there is overwhelming evidence of benefit and after sufficient events have occurred. Just as clinical trial registry is now a requirement for publication, journals should also require adequate reporting about trial design and the decision to stop.

Competing interests

The authors declare no competing interests.

References

1. Montori VM, Devereaux PJ, Adhikari NK, Burns KE, Eggert CH, Briel M, Lacchetti C, Leung TW, Darling E, Bryant DM, Bucher HC, Schunemann HJ, Meade MO, Cook DJ, Erwin PJ, Sood A, Sood R, Lo B, Thompson CA, Zhou Q, Mills E, Guyatt GH: **Randomized trials stopped early for benefit: a systematic review.** JAMA 2005, **294**:2203-2209.
2. Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, Pitkin R, Rennie D, Schulz KF, Simel D, Stroup DF: **Improving the quality of reporting of randomized controlled trials. The CONSORT statement.** JAMA 1996, **276**:637-639.
3. Moher D, Schulz KF, Altman D: **The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials.** JAMA 2001, **285**:1987-1991.
4. Ioannidis JP: **Contradicted and initially stronger effects in highly cited clinical research.** JAMA 2005, **294**:218-228.

5. Abraham E, Reinhart K, Opal S, Demeyer I, Doig C, Rodriguez AL, Beale R, Svoboda P, Laterre PF, Simon S, Light B, Spapen H, Stone J, Seibert A, Peckelsen C, De Deyne C, Postier R, Pettila V, Artigas A, Percell SR, Shu V, Zwengelstein C, Tobias J, Poole L, Stolzenbach JC, Creasey AA: **Efficacy and safety of tifacogin (recombinant tissue factor pathway inhibitor) in severe sepsis: a randomized controlled trial.** *JAMA* 2003, **290**:238-247.
6. Angus DC, Crowther MA: **Unraveling severe sepsis: why did OPTIMIST fail and what's next?** *JAMA* 2003, **290**:256-258.
7. Pocock S, White I: **Trials stopped early: too good to be true?** *Lancet* 1999, **353**:943-944.
8. Schulz KF, Grimes DA: **Multiplicity in randomised trials II: subgroup and interim analyses.** *Lancet* 2005, **365**:1657-1661.
9. Poldermans D, Boersma E, Bax JJ, Thomson IR, van de Ven LL, Blankensteijn JD, Baars HF, Yo TI, Trocino G, Vigna C, Roelandt JR, van Urk H: **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**:1789-1794.
10. **Ventilation with lower tidal volumes as compared with traditional tidal volumes for acute lung injury and the acute respiratory distress syndrome. The Acute Respiratory Distress Syndrome Network.** *N Engl J Med* 2000, **342**:1301-1308.
11. Bernard GR, Vincent JL, Laterre PF, LaRosa SP, Dhainaut JF, Lopez-Rodriguez A, Steingrub JS, Garber GE, Helterbrand JD, Ely EW, Fisher CJ, Jr.: **Efficacy and safety of recombinant human activated protein C for severe sepsis.** *N Engl J Med* 2001, **344**:699-709.
12. Van den Berghe G, Wouters P, Weekers F, Verwaest C, Bruyincx F, Schetz M, Vlasselaers D, Ferdinande P, Lauwers P, Bouillon R: **Intensive insulin therapy in the critically ill patients.** *N Engl J Med* 2001, **345**:1359-1367.
13. Pocock SJ: **When (not) to stop a clinical trial for benefit.** *JAMA* 2005, **294**:2228-2230.