However, the notable findings in the study by Talcott et al[1] and the earlier report by Zietman et al[2] were that (1) at 5 years, a significantly higher proportion of men remained free of biochemical failure in the higher-dose radiation group (80.4%) vs the conventional-dose treatment group (61.4%) and that (2) at a median of 9.4 years following treatment, there were no significant differences in key functional parameters between conventional-dose and high-dose treatment, such as urinary obstruction, urinary incontinence, bowel problems, and decreased sexual function. Furthermore, at a median follow-up time of 8.9 years, long-term cancer control was still significantly higher in the high-dose vs conventional-dose treatment groups.[3]

We suspect that these findings about symptoms probably were not necessarily only "because eventually the patients will adapt and stop complaining" (as Vikram suggests), but instead may represent important and meaningful outcomes for patients. Similarly, Talcott et al also found that patients who had received conventional doses were significantly more concerned about control of their cancer than patients who had received high doses, and also tended to express more regret.

The quality-of-life consequences for men treated for prostate cancer and the psychological toll experienced by men concerned with disease recurrence should not be underestimated. The results from the study by Talcott et al add to the evidence that physicians can use when discussing treatment options for men with early prostate cancer by showing that high-dose radiation treatment is not associated with worse long-term adverse effects than conventional-dose treatment.

Boris Pasche, MD, PhD
boris.pasche@ccc.uab.edu
Division of Hematology/Oncology
University of Alabama at Birmingham

Robert A. McNutt, MD
Department of Medicine
Rush University Medical Center
Chicago, Illinois

Phil B. Fontanarosa, MD, MBA
Executive Deputy Editor, *JAMA*
Chicago

Financial Disclosures: None reported.

1. Talcott JA, Rossi C, Shipley WU, et al. Patient-reported long-term outcomes after conventional and high-dose combined proton and photon radiation for early prostate cancer. *JAMA*. 2010;303(11):1046-1053.
2. Zietman AL, DeSilvio ML, Slater JD, et al. Comparison of conventional-dose vs high-dose conformal radiation therapy in clinically localized adenocarcinoma of the prostate: a randomized controlled trial. *JAMA*. 2005;294(10):1233-1239.
3. Zietman AL, Bae K, Slater JD, et al. Randomized trial comparing conventional-dose with high-dose conformal radiation therapy in early-stage adenocarcinoma of the prostate: long-term results from Proton Radiation Oncology Group/American College of Radiology 95-09. *J Clin Oncol*. 2010;28(7):1106-1111.

Bias and Trials Stopped Early for Benefit

To the Editor: We believe that the analysis reported by Dr Bassler and colleagues[1] incorporated an important scientific and logical error that led to invalid conclusions. The authors separated trials into 2 groups: one that stopped early because of apparent efficacy and another that did not stop early. The main conclusion was that trials that stopped early for efficacy yielded biased estimates for the treatment effect, and the authors presented analyses suggesting that the relative risk (RR) in studies that stopped early was 0.71 relative to those that did not stop early.

This analysis was analogous to a comparative clinical trial with 2 groups: group A consisting of trials that stopped early and group B consisting of trials not stopping early. The authors drew conclusions as though the membership in groups A and B is a random one. However, this is not true. The membership of each trial in group A or group B is based on the data within each trial. Trials that stop early for efficacy tend to have conservative stopping boundaries[2] to control the type I error of the trial. For a trial to hit an early stopping boundary, the experimental treatment must be performing very well compared with the control: the observed effect size must be large.

To illustrate the issue, consider a clinical trial in which analysis is as follows: participants found to be performing better are retrospectively placed in the experimental group and participants found not to be performing well are retrospectively placed in the control group; a statistically significant difference in outcome is found when the groups are compared. It is clear that posttreatment selection of participants, based on their outcomes, would be responsible for any observed difference. This is logically equivalent to the analysis reported by Bassler et al. The bias created by the outcome-based assignment of trials to groups created the difference in RR observed between the 2 groups.

Early termination of clinical trials, for either clear efficacy or harm, is a cornerstone of efficient and ethical trial design. It does not lead to substantive bias in the estimation of treatment effects.

Scott M. Berry, PhD
scott@berryconsultants.com
Berry Consultants
College Station, Texas

Bradley P. Carlin, PhD
School of Public Health
University of Minnesota
Minneapolis

Jason Connor, PhD
Berry Consultants

Financial Disclosures: Dr Berry reported being president and part owner and Dr Connor reported being an employee of Berry Consultants, a company that specializes in designing adaptive trials for pharmaceutical and medical device companies.

1. Bassler D, Briel M, Montori VM, et al; STOPIT-2 Study Group. Stopping randomized trials early for benefit and estimation of treatment effects: systematic review and meta-regression analysis. *JAMA*. 2010;303(12):1180-1187.
2. O'Brien PC, Fleming TR. A multiple testing procedure for clinical trials. *Biometrics*. 1979;35(3):549-556.

To the Editor: Dr Bassler and colleagues[1] claimed that trials stopped for efficacy are prone to major bias. Unfortunately, their conclusions are based on faulty mathematical reasoning.[2,3] We believe that they misinterpreted their data and provided incorrect guidance about ethical trial design.

The authors' claim of a large bias is spurious because they have confused "bias" with the observation that results from trials addressing the same question will differ because of random variation. Consider an unbiased group of trials that go to completion. Trials with the largest effects will have similarly large interim effects, potentially leading to truncation.[3] So comparing the truncated trials to the nontruncated trials is similar to comparing completed trials with large effects with those with lower effects. The difference the authors observed was both predictable and uninformative.

Estimates from trials stopped early for efficacy have negligible bias.[4] Such estimates, on average, are therefore correct, not subject to empirical disproof. Also, bias is a property of study procedures; it is not logically applicable to a subset of results, eg, from truncated trials. The only way to know if a particular result is too high (or low)—whether or not from a truncated trial—is if the likely magnitude of effect is known from other studies.

The authors' legitimate concern is about implausibly large effects, which can be seen in any kind of trial—small, large, or truncated. The real value of this exercise is that it quantifies the range of effects seen in the published literature, which helps define "plausible." These effects tend to be of null to moderate size, making any larger estimate likely to be an overestimate. The proper correctives are statistical approaches that appropriately shrink the estimated effect size.[2] This is only necessary when dealing with a single trial. If there are several trials, averaging them performs that shrinkage function. The authors' contention that such meta-analytic averages are meaningfully biased by truncated trials is incorrect.[5] It is the omission or discounting of such trials that would produce substantive bias.

Implementation of adaptive designs and use of stopping rules are critical for efficiently and ethically developing therapies for patients.

Steven Goodman, MD, MHS, PhD
sgoodman@jhmi.edu
Department of Oncology
Johns Hopkins School of Medicine
Baltimore, Maryland

Donald Berry, PhD
Division of Quantitative Sciences
M. D. Anderson Cancer Center
Houston, Texas

Janet Wittes, PhD
Statistics Collaborative
Washington, DC

Financial Disclosures: Dr Berry reported being part owner of Berry Consultants, which advocates for and develops Bayesian adaptive designs for pharmaceutical and medical device companies. In addition, a substantial number of clinical trials designed and carried out at M. D. Anderson Cancer Center under Dr Berry's direction take a Bayesian adaptive perspective. Dr Wittes reported being the president of Statistics Collaborative, an employee-owned company of which she is a majority stakeholder. The company consults for many pharmaceutical, biotechnology, and device companies, as well as for government agencies and not-for-profit organizations; it also serves as the independent statistical organization reporting to data monitoring committees, performing interim analysis of ongoing randomized clinical trials. No other disclosures were reported.

1. Bassler D, Briel M, Montori VM, et al; STOPIT-2 Study Group. Stopping randomized trials early for benefit and estimation of treatment effects: systematic review and meta-regression analysis. *JAMA*. 2010;303(12):1180-1187.
2. Goodman SN. Stopping at nothing? some dilemmas of data monitoring in clinical trials. *Ann Intern Med*. 2007;146(12):882-887.
3. Freidlin B, Korn EL. Stopping clinical trials early for benefit: impact on estimation. *Clin Trials*. 2009;6(2):119-125.
4. Emerson SS, Fleming TR. Parameter estimation following sequential hypothesis testing. *Biometrika*. 1990;77:875-892.
5. Goodman SN. Systematic reviews are not biased by results from trials stopped early for benefit. *J Clin Epidemiol*. 2008;61(1):95-96.

To the Editor: Dr Bassler and colleagues[1] suggested that stopping rules for clinical trials may not satisfy an ethical obligation to future patients because they can substantially overestimate the treatment effect. Stopping a trial and releasing the information early allows current and future patients to benefit from new therapies as soon as possible. Furthermore, it allows the new treatments to be tested in combination therapies and in earlier stages of disease when their benefits may even be greater. In life-threatening diseases such as cancer, these benefits outweigh not knowing precisely how efficacious the new treatment is. The number of events required for convincing evidence of benefit that is sufficient for early stopping should be prespecified in the protocol and derived with consideration of the disease and treatment being tested.

To assess the clinical utility of early-stopped trial results, we examined NCI/CTEP-sponsored early-stopped cancer treatment trials included in the review by Bassler et al. RTOG-85-01,[2] testing the addition of radiotherapy to chemotherapy for locally advanced esophageal cancer, stopped randomizing patients after 90 evaluable participants had been accrued. This is a very small number of patients, but the effects were large, with 30% (95% confidence interval [CI], 19%-43%) vs 0% (95% CI, 0%-6%) alive at 3 years and 26% (95% CI, 15%-40%) vs 0% (95% CI, 0%-6%) alive at 5 years. The results appear to be consistent with the nonstopped trials (Figure 2, question 14, in the review by Bassler et al).

NSABP-B-31/N9831,[3] testing the addition of trastuzumab to adjuvant chemotherapy for surgically removed *HER2*-positive breast cancer, reported early after 394 events with a disease-free survival (DFS) hazard ratio (HR) of 0.48 (95% CI, 0.39-0.59), and an overall survival (OS) HR of 0.67 (95% CI, 0.48-0.93). These results were consistent with another early-stopped trial of adjuvant trastuzumab (as demonstrated by the pooled early-stopped results, question 33[1]), but not with the reported RR of 0.80 based on a meta-analysis of 3 nonstopped trials (question 33[1]). However, of the trials included in the meta-analysis, only one[4] was in the adjuvant setting (DFS HR, 0.42; 95% CI, 0.21-0.83). The

other 2 trials were for metastatic breast cancer, making their inclusion inappropriate for quantitative pooling of treatment effects.

NCCTG-N9741,[5] testing oxaliplatin for metastatic colorectal cancer, stopped early with 305 patients with DFS benefits (HR, 0.55; 95% CI, 0.43-0.70) and OS benefits (HR, 0.76; 95% CI, 0.60-0.97). The results appear to be consistent with the nonstopped trial (question 10[1]).

These results suggest that in oncology, even with small sample sizes, the precision and accuracy of the data from these stopped cancer trials have been sufficient to improve cancer treatments in a timely manner.

Edward L. Korn, PhD
korne@ctep.nci.nih.gov
Boris Freidlin, PhD
Biometric Research Branch
National Cancer Institute
Bethesda, Maryland

Margaret Mooney, MD
Cancer Therapy Evaluation Program (CTEP)
National Cancer Institute

Financial Disclosures: None reported.

1. Bassler D, Briel M, Montori VM, et al; STOPIT-2 Study Group. Stopping randomized trials early for benefit and estimation of treatment effects: systematic review and meta-regression analysis. *JAMA*. 2010;303(12):1180-1187.
2. Cooper JS, Guo MD, Herskovic A, et al; Radiation Therapy Oncology Group. Chemoradiotherapy of locally advanced esophageal cancer: long-term follow-up of a prospective randomized trial (RTOG 85-01). *JAMA*. 1999;281(17):1623-1627.
3. Romond EH, Perez EA, Bryant J, et al. Trastuzumab plus adjuvant chemotherapy for operable *HER2*-positive breast cancer. *N Engl J Med*. 2005;353(16):1673-1684.
4. Joensuu H, Kellokumpu-Lehtinen PL, Bono P, et al; FinHer Study Investigators. Adjuvant docetaxel or vinorelbine with or without trastuzumab for breast cancer. *N Engl J Med*. 2006;354(8):809-820.
5. Goldberg RM, Sargent DJ, Morton RF, et al. Randomized controlled trial of reduced-dose bolus fluorouracil plus leucovorin and irinotecan or infused fluorouracil plus leucovorin and oxaliplatin in patients with previously untreated metastatic colorectal cancer: a North American Intergroup Trial. *J Clin Oncol*. 2006;24(21):3347-3353.

To the Editor: The apparent purpose of the analysis in the review by Dr Bassler and colleagues[1] was to argue that trials terminated early will overestimate treatment benefit, thereby misleading physicians and patients. It is true that a trial terminated early for benefit will tend to overestimate true effect; this happens because there always is variability in estimation of true effect, and when assessing data over time, evidence of extreme benefit is more likely obtained at times when the data provide a random overestimate of truth. However, statistical methods are available to adjust for this random-high bias,[2-5] although they are not applied as often as they should be.

For settings in which the need for long-term efficacy and safety data are very important to making treatment decisions, and it is highly unlikely that the continuation of the trial would induce unacceptable increased risk of irreversible morbidity or mortality to trial participants, then having either no or highly conservative monitoring boundaries would be appropriate. But in circumstances in which a clearly superior treatment in terms of preserving life or essential life functions might be identified, there would be serious ethical concerns about continuing to randomize participants for the purpose of getting a somewhat more precise or longer-term estimate of the extent of its superiority.

If the authors had chosen to focus their arguments on the need to adjust treatment effect estimates in studies terminated early, we would have no disagreement. They seem, however, to be warning against early trial termination. This is a much more complex issue on which the problem of modest upward bias of the effect estimate, readily remediable by existing methodology, should have little bearing.

Susan S. Ellenberg, PhD
sellenbe@mail.med.upenn.edu
University of Pennsylvania School of Medicine
Philadelphia

David L. DeMets, PhD
University of Wisconsin School of Medicine and Public Health
Madison

Thomas R. Fleming, PhD
University of Washington School of Public Health
Seattle

Financial Disclosures: Drs Ellenberg, DeMets, and Fleming reported receiving compensation for consulting and for serving on data monitoring committees for many pharmaceutical, biotechnology, and device companies and for government agencies and receiving honoraria or royalties for books or book chapters.

1. Bassler D, Briel M, Montori VM, et al; STOPIT-2 Study Group. Stopping randomized trials early for benefit and estimation of treatment effects: systematic review and meta-regression analysis. *JAMA*. 2010;303(12):1180-1187.
2. Hughes MD, Pocock SJ. Stopping rules and estimation problems in clinical trials. *Stat Med*. 1988;7(12):1231-1242.
3. Emerson SS, Fleming TR. Parameter estimation following sequential hypothesis testing. *Biometrika*. 1990;77:875-892.
4. Pinheiro JC, DeMets DL. Estimating and reducing bias in group sequential designs with Gaussian independent structure. *Biometrika*. 1997;84:831-843.
5. Fan XF, DeMets DL, Lan KK. Conditional bias of point estimates following a group sequential test. *J Biopharm Stat*. 2004;14(2):505-530.

In Reply: Authors of all 4 letters acknowledge that truncated randomized controlled trials (tRCTs) overestimate treatment effects; controversy remains as to by how much and what should be done about it. A pivotal issue is what real-world data should be used to quantify overestimates. Ideally, trials with fixed sample size and no provision for early stopping would be available. Such trials are rare and typically not identifiable through reading their methods. Feasible alternatives are therefore all trials not stopped early (our primary comparison, with which Dr Berry and colleagues and Dr Goodman and colleagues disagree, describing it as a logical error) or all trials including tRCTs (our secondary comparison).

Goodman et al rely on simulation rather than empirical evidence for inferences on effects from stopped-early trials.[1,2] However, real-world data do not conform to idealized simulations. Goodman has pointed out that tRCTs should in theory contribute little weight to a pooled estimate.[2] We have shown that tRCTs often contribute far more than one would expect in theory.[3] Possible explanations include publica-

tion bias, a "freezing effect" on subsequent trials by the publication of an apparent success, and the undisciplined application of stopping rules. Whatever the explanation, real-world data should drive inferences and actions.

Neither our primary or secondary comparator is perfect. However, either is superior to ignoring the real world and relying exclusively on simulations. To the (unknown) extent that nontruncated RCTs had stopping rules that they did not apply, they will yield small underestimates of treatment effect and result in small overestimates of the conditional bias associated with truncation. To the extent that publication bias and a freezing effect are at play, a pooled estimate including tRCTs will yield spuriously large treatment effects and underestimate the exaggerated treatment effects of tRCTs. These considerations led us to choose our primary comparison.

Goodman has previously endorsed our secondary comparison: tRCTs should be compared with meta-analyses that include the tRCTs.[2] He and his coauthors now state that this comparison should perform the necessary shrinkage function. In our secondary comparison, the tRCTs still overestimated treatment effects (RR, 0.85).

Dr Korn and colleagues point out that 3 cancer tRCTs in our data did not substantially overestimate treatment effects. They imply that there is something special about cancer trials that protect them from the general phenomenon. The ratio of RRs for all hematology-oncology trials in our review (0.75; 95% CI, 0.65-0.87; $P < .001$) is similar to the overall result.

Whether the best estimate of the ratio of RRs of tRCTs to best estimates of effect comes from our primary (0.71) or secondary (0.85) analysis—or something in between—tRCTs on average substantially overestimate treatment effects. Goodman et al and Dr Ellenberg and colleagues believe the solution to this problem is applying a shrinkage function to tRCTs. Until authors of tRCTs focus on shrunken rather than apparent estimates, other solutions to the problem are required.

Ellenberg et al and Korn et al highlight the difficult ethical issues affecting the decision to stop a trial early. A large proportion of the trials stopped early for benefit followed accepted statistical procedures,[4] some of which are advocated as means to run less expensive trials that may yield impressive results and faster regulatory approval.[5] We have previously discussed the ethical implications of our findings.[6] When it matters to have accurate estimates of benefit (eg, when other treatments exist, or when the new treatment has important harms and costs) and when truncation can lead to large overestimates of treatment effects (when the number of events is not large—our data suggest fewer than 500), stopping early is ethically problematic.

Gordon H. Guyatt, MD, MSc
Department of Clinical Epidemiology and Biostatistics
McMaster University
Hamilton, Ontario, Canada

Dirk Bassler, MD, MSc
Department of Neonatology
University Children's Hospital Tuebingen
Tuebingen, Germany

Victor M. Montori, MD, MSc
montori.victor@mayo.edu
Knowledge and Encounter Research Unit
Mayo Clinic
Rochester, Minnesota
on behalf of the STOPIT-2 Study Group

Financial Disclosures: None reported.

1. Emerson SS, Fleming TR. Parameter estimation following sequential hypothesis testing. *Biometrika*. 1990;77:875-892.
2. Goodman SN. Systematic reviews are not biased by results from trials stopped early for benefit. *J Clin Epidemiol*. 2008;61(1):95-96.
3. Bassler D, Ferreira-Gonzalez I, Briel M, et al. Systematic reviewers neglect bias that results from trials stopped early for benefit. *J Clin Epidemiol*. 2007;60(9):869-873.
4. Montori VM, Devereaux PJ, Adhikari NK, et al. Randomized trials stopped early for benefit: a systematic review. *JAMA*. 2005;294(17):2203-2209.
5. Whitehead J. Stopping clinical trials by design. *Nat Rev Drug Discov*. 2004;3(11):973-977.
6. Mueller PS, Montori VM, Bassler D, Koenig BA, Guyatt GH. Ethical issues in stopping randomized trials early because of apparent benefit. *Ann Intern Med*. 2007;146(12):878-881.

RESEARCH LETTER

Validity and Reliability of the Schamroth Sign for the Diagnosis of Clubbing

To the Editor: Digital clubbing is characterized by the enlargement of the terminal segments of the fingers or toes, resulting from proliferation of connective tissue between the nail matrix and the distal phalanx[1] and may be a sign of respiratory and other diseases.[2] Confirming clubbing requires using instruments to determine the nail bed angles or the phalangeal depth ratio (PDR) and is not performed routinely. In 1976, Schamroth[3] reported a clinical sign associated with clubbing: obliteration in clubbed fingers of the diamond-shaped window normally produced when the dorsal surfaces of the corresponding finger of each hand are opposed (FIGURE). The reproducibility and accuracy of this sign have not been formally tested.[4,5]

Methods. This cross-sectional study was conducted in July through September 2009 in a tertiary care hospital in northwest Spain. Participants were all consecutive patients who were admitted or treated in outpatient services during these months with disease categories associated with clubbing (infectious, pulmonary, cardiac, digestive, metabolic, and malignant diseases). After obtaining oral consent, clinicians first independently evaluated the presence of Schamroth sign on the third fingers (or, if not possible, the fourth fingers) of both hands. They subsequently calculated the PDR[4,6] by measuring the distal interphalangeal depth (IPD) and the nail bed diameter (distal phalangeal depth [DPD]) with a precision caliper (Figure). Observers were not blinded to the results of the Schamroth test when assessing PDR.