EUROPEAN JOURNAL OF CANCER 47 (2011) 2381–2386

available at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejconline.com

Review

Early stopping rules in oncology: Considerations for clinicians

Som D. Mukherjee [a,*], *John R. Goffin* [a], *Valerie Taylor* [b], *Kelly K. Anderson* [c], *Gregory R. Pond* [a]

[a] *Department of Oncology, McMaster University, Juravinski Cancer Centre, 699 Concession Street, Hamilton, Ontario, Canada L8V5C2*
[b] *Department of Psychiatry and Behavioral Neuroscience, McMaster University, Juravinski Cancer Centre, 699 Concession Street, Hamilton, Ontario, Canada L8V5C2*
[c] *Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Purvis Hall, 1020 Pine Avenue West, Montreal, QC, Canada H3A 1A2*

ARTICLE INFO

Article history:
Available online 16 June 2011

Keywords:
Early stopping rules
Clinical trials
Oncology

ABSTRACT

The number of cancer-related clinical trials has been rapidly increasing over the past decade. Along with this increase, oncology studies stopped early for benefit or harm have also been more common. Clinicians treating cancer patients often are faced with the challenge of having to decide whether or not to incorporate information from these new studies into their daily clinical practice. This review article explains the role of the Data and Safety Monitoring Committee in stopping trials early; provides examples of oncology trials stopped early; and reviews some of the controversies and statistical concepts associated with early stopping rules. In addition, a simple and practical approach to interpreting the findings of trials that are stopped early is provided to assist clinicians in deciding how to incorporate information from these studies into their daily practice.

1. Introduction

The treatment of cancer has undergone considerable advancement over the past 20 to 30 years. A significant proportion of this success is attributable to the numerous clinical trials that have been performed in the field, which have given oncologists the opportunity to apply an evidence-based approach to their patients. Oncologists are becoming increasingly reliant on the results of clinical trials to help guide their day-to-day practice. However, some of these trials may be stopped early, either because of benefit, futility or harm. Trials that are stopped early can potentially present difficulties for health care practitioners downstream, as they must decide how best to apply this new information to their patients. Weighing the scientific and statistical reasons behind early trial termination can be difficult, and a good understanding of these can help oncologists better appraise and interpret this body of evidence.

This article will review examples of trials stopped early in oncology, and will also serve as a helpful guide for clinicians faced with the challenge of interpreting trials stopped early for benefit, futility or harm.

2. The Data and Safety Monitoring Committee (DSMC)

All centres participating in clinical research worldwide are expected to have a system in place to ensure the safety of patients enrolled in studies and to ensure the legitimacy and integrity of data.[1] In 1994, the United States Office of Extramural Research established the Committee on Clinical Trial Monitoring to review the oversight and management practices of

* *Corresponding author:* Tel.: +1 9053879495; fax: +1 9055756326.
E-mail address: som.mukherjee@jcc.hhsc.ca (S.D. Mukherjee).
0959-8049/$ - see front matter
doi:10.1016/j.ejca.2011.05.019

large phase III clinical trials.[1] They recommended that all phase III trials should include an external monitoring body as part of their research project. In this article, we will use the term Data and Safety Monitoring Committee (DSMC) when referring to this independent group of experts, however it should be acknowledged that this committee is often referred to by other names including the Independent Data Monitoring Committee (IDMC), Data and Safety Monitoring Board (DSMB) and the Data Monitoring Committee (DMC).

The primary role of a DSMC is to ensure safe and effective conduct within a clinical study. This committee may recommend stopping a trial early when significant benefits or risks are observed or if the trial is unlikely to be completed successfully. Monitoring can exist on a continuum from careful review by the principal investigator in a small phase I study to an independent DSMC with multiple members for a large phase III clinical trial.[2] A DSMC should continue to review a study after it has been closed to accrual if patients remain in active follow-up, even in cases where treatment has been completed, as patient safety both within and external to the trial can be affected. In addition, decisions regarding early publication, potential cross-over of patients from an inactive or placebo arm to active therapy, or extension of follow-up to determine long-term treatment effects should be reviewed with input from a DSMC.

DSMC's may consider stopping a trial early for a variety of reasons. These include significantly worse outcomes amongst patients receiving the experimental treatment, a large number of unexpected deaths or serious adverse events in the experimental arm, or a statistically and clinically significant benefit favouring the experimental arm. In the latter case, patients receiving the non-experimental arm may be allowed to 'cross-over' and receive the more beneficial treatment. Finally, trials may be stopped for reasons of futility in situations when it is highly unlikely that the trial will accrue the planned number of patients or if the interim analysis shows that it is extremely unlikely that any benefit will be seen if the study is continued. Trials stopped early are more likely to be rapidly published in leading scientific journals, often leading to faster regulatory approval of the experimental treatment.

Unfortunately, the decision to stop a study early also carries with it a number of potential risks. For most clinical trials, statistical guidelines are in place to give guidance to the DSMC as to when they should consider stopping a trial early. These statistical guidelines should not be used as hard-and-fast statistical rules. DSMCs must consider many additional factors when considering whether to recommend early trial closure, including: previous scientific knowledge of the treatment; other newly published studies relating to the treatment in question; the potential future impact of stopping the trial early; and secondary end-points such as ease of treatment or toxicities.

Further, since statistical significance is based on probabilities, there is always a risk of a false-positive result. Due to random variation, an estimated treatment effect in a clinical trial varies around the true treatment effect in the larger population, such that 50% of the time the estimated treatment effect will be greater than the true treatment effect in an unbiased, final analysis. A trial stopped early for benefit is more likely to have a larger estimated treatment effect compared to the true treatment effect.[3] Although it is possible that statistical significance was observed and that the trial was stopped early when there actually was a clinically significant treatment effect, it is also possible that there was no treatment effect and that the trial was stopped early due to a random high in the natural variation in the study results at the point when the data were reviewed. Data from similar trials in the future may yield a more conservative estimate of the treatment effect as 'regression to the truth' takes effect.

There are several examples of trials in the literature whereby a statistically significant improvement in the primary end-point early on in a study, subsequently disappeared at the final analysis. One such example is the National Surgical Adjuvant Breast and Bowel Project (NSABP) C-08 trial that was designed to evaluate the efficacy of adding one year of bevacizumab to modified FOLFOX6 (i.e. infusional/bolus fluorouracil, leucovorin and oxaliplatin) for the adjuvant treatment of patients with stages II to III colon cancer.[4] In this trial, although a statistically significant benefit in DFS was observed after a median follow-up of 15 months (HR 0.61; 95% CI, 0.48 to 0.78; $p < 0.001$), this benefit was not present at the final analysis with a median follow-up of 3 years (HR 1.22; 95% CI 0.98 to 1.52; $p < 0.076$). Similarly, the AVANT trial was a randomised, three arm phase III trial investigating the addition of bevacizumab to either XELOX (capecitabine, and oxaliplatin) or FOLFOX4 versus FOLFOX4 alone as adjuvant treatment for colon cancer.[5] This study showed that patients seemed to derive benefit in the first 12 months of therapy when bevacizumab was being administered (HR's of 0.63 and 0.61 for bevacizumab-containing regimens), but this benefit was lost in subsequent years, similar to the NSABP C-08 trial.

On the other hand, there are also situations where an interim analysis shows no statistically significant benefit in the primary outcome. These studies tend to be terminated early and in many cases, no further information on long-term outcomes is available following study closure. Examples of this include the randomised phase III trial of matrix metalloproteinase inhibitor BMS-275291 in combination with paclitaxel and carboplatin in advanced non-small cell lung cancer[6], as well as the randomised phase III trial of interferon gamma-1b plus standard carboplatin/paclitaxel versus carboplatin/paclitaxel alone for first-line treatment of advanced ovarian and primary peritoneal carcinomas.[7] Trials stopped early due to a lack of benefit at the interim analysis are often associated with increased toxicity or even shorter survival in the intervention arm. These factors often facilitate the decision to stop a trial early.

3. Surrogate end-points and early stopping rules

A surrogate end-point is defined as an alternative end-point (such as a biological marker, physical sign, or precursor event) that can be used as a substitute for a clinically meaningful end-point that measures directly how a patient feels, functions, or survives.[8] Overall survival (OS) is generally considered the most important outcome in most oncology studies

evaluating adjuvant, neoadjuvant and metastatic therapies. The only validated surrogate for OS in oncology clinical trials is the use of disease-free survival (DFS) or progression-free survival (PFS) in colorectal cancer patients treated with fluorouracil-based chemotherapy.[9] Despite this, many trials use other end-points such as tumour response rates (RR's) or time to progression (TTP) as surrogate outcomes for OS. Although employing surrogate end-points can result in smaller sample sizes, reduced overall costs, and a shorter time to trial completion[10], caution must be used in interpreting these studies given previous situations where surrogate measures failed to hold up over time. These examples include the use of arrhythmia suppression as a surrogate for cardiovascular mortality, and the use of bone mineral density as a surrogate for fracture risk.[3] In general, the magnitude of the treatment effect with a surrogate outcome generally does not follow an exact linear relationship with size of the treatment effect for the true clinical outcome.

Even for validated surrogates in cancer trials evaluating survival, a perfect correlation between surrogate outcomes and OS does not exist. Problems relating to an over-inflated treatment effect may be compounded in trials stopped early for benefit based on results that have employed a surrogate end-point. The treatment effect on OS is often smaller than the treatment effect on DFS or other surrogate outcomes given that OS is affected by subsequent treatments received after the trial has stopped.[11,12] In addition, for patients who cross-over to the 'active' treatment regimen following trial closure, subsequently estimating the true OS benefit will not be possible.

4. Case study: MA-17

One example of a clinical trial terminated early is the well-known breast cancer trial MA-17, conducted by the National Cancer Institute of Canada Clinical Trials Group. This study evaluated the potential benefit of an aromatase inhibitor, letrozole, in breast cancer patients.[13] Postmenopausal women with early stage hormone receptor positive breast cancer were randomly assigned to receive either five years of letrozole or placebo, after having received tamoxifen for five years, with follow-up being measured from the time of randomisation. The study was powered to detect an absolute difference of 2.5% in 4-year DFS between the two arms, with a planned enrolment of 4800 women at 2 years of median follow-up.[13] The first interim analysis was planned following 171 events, with an event being defined as a local or metastatic recurrence of the original breast cancer, or a new primary cancer in the contralateral breast. After 5187 women were enrolled, 207 events had occurred with 2.4 years median follow-up.[13] An interim analysis was performed at that time, and a significant improvement in the primary outcome of DFS was noted, with estimated 4-year DFS rates of 93% in the letrozole group and 87% in the placebo group ($p < 0.001$). This translated into an absolute difference of 6%, although there was no improvement in OS at this interim analysis ($p = 0.25$). The DSMC for this trial was faced with the difficult decision of whether to allow the trial to continue, based on the significant improvement in DFS. After much deliberation, the DSMC recommended early study termination. The favourable DFS results were disclosed to all women on the study, and those receiving placebo were given the opportunity to crossover and receive letrozole. The investigators concluded that based on the improvement in DFS, letrozole was a reasonable treatment option for early stage postmenopausal breast cancer patients who had completed 5 years of tamoxifen therapy.[14]

The early termination of this study, based on the improvement in DFS, turned out to be a controversial decision with debate erupting amongst oncologists. It was argued that since there was no statistically significant difference in OS and that the decision to terminate the trial was based primarily on an improvement in DFS, the decision to stop the trial early was premature.[15] Given that patients in the placebo arm were allowed to crossover to the letrozole arm, it became impossible to continue to compare the long-term toxicities or overall survival of letrozole against a placebo group. Furthermore, findings from the trial indicated a higher percentage of osteoporosis, bone fractures, hot flashes and myalgias in the letrozole arm.[15]

In response to the above criticisms, the investigators argued that DFS was an appropriate primary outcome in this study, based on the accumulated experience in adjuvant breast cancer trials to date that suggested that DFS was a good surrogate for OS. The study authors also noted that preventing breast cancer recurrences was a valuable outcome in and of itself. The alternative treatment for patients experiencing a local recurrence is often mastectomy, a potentially traumatic therapy, and it was argued that the study question was answered sooner than expected because the treatment effect was larger than anticipated.[16]

5. Systematic reviews of trials stopped early for benefit

In 2001, the Consolidated Standards of Reporting Trials (CONSORT) Group published recommendations for improving the quality of reports of randomised controlled clinical trials.[17] These guidelines have since been updated in the 2010 CONSORT statement.[18] The CONSORT statement specifically recommends that randomised controlled studies should include a detailed explanation of issues pertaining to sample size determination, interim analyses and stopping rules. At the time of the 2001 CONSORT statement publication, little was known about the extent to which investigators adhered to these recommendations.

To address this issue, Montori and colleagues published the first large systematic review in 2005 evaluating randomised controlled trials (RCT's) stopped early for benefit. In this review of 143 trials, oncology trials were the second largest group of studies, making up 21% of all randomised trials in this review.[3] Of the 143 trials evaluated, 55 were stopped after the first interim analysis, 29 after the second, 19 after three or more analyses, and 40 did not report this information. Interestingly, 48 of 143 studies either did not use or did not specify whether a statistical approach to monitoring the trial was used. Most studies were stopped by the DSMC or principal investigator, and 90 of 143 trials were stopped after the results exceeded a stopping boundary. Only eight of 143 trials re-

ported the following four key methodological elements: planned sample size, the number of interim analyses after which the RCT was stopped, the stopping rules used to make this decision, and adjusted estimates for interim analysis and early stopping.[3] Although this systematic review contained a substantial number of non-oncology trials, the message it sends is highly applicable to oncologists and cancer investigators. There needs to be a higher degree of transparency by oncology trial investigators with respect to the number of interim analyses carried out, and the stopping rules that were applied to the trial should be explicitly stated. Furthermore, a thorough description of the process by which the decision to stop the trial was actually made would be extremely valuable for clinicians attempting to interpret the findings from these studies.

To examine this issue for oncology trials specifically, Trotta and colleagues published a similar review of 25 randomised trials in 2008.[19] They found substantial variation between the trials, particularly in the methods sections: 3 trials performed interim analysis based on a cut-off date, 12 based on a specified number of events, and 9 based on recruitment. They also found that more of the recent studies were being closed early, and the authors questioned whether costs or market considerations might have played a role in influencing premature study closure.

Additionally, Bassler and colleagues published an evaluation of the impact of studies stopped early on the findings of subsequent systematic reviews.[20] A total of 96 published systemic reviews were evaluated, including seven cancer-related reviews. Of the included reviews, 68 of 96 (71%) did not mention that a truncated study was included. Furthermore, the weight of the truncated studies comprised greater than 40% of the total analysis in 16 (34%) of them. This suggests that studies stopped early for benefit make a disproportionate contribution to the pooled effect estimates of systemic reviews.

In contrast, some recent literature suggests that some trials are appropriately stopped early. Korn et al. reviewed 27 studies conducted by the National Cancer Institute (NCI) which were stopped early for benefit.[21] This study found that almost all of the trials being evaluated had sufficient follow-up for conducting an appropriate analysis. In addition, interim results were found to be very similar to the final published results, although the final result tended to have a slightly diminished magnitude. The authors pointed out that all of the NCI trials being evaluated had fairly well designed plans for interim analysis. A separate simulation study of early stopping found that interim results were not overly different from final results when at least 50% of the expected data was available.[22] Together, these studies suggest that early stopping of trials may be reasonable in situations where the trial is well planned and a sufficient number of events have occurred.

6. Early stopping rules and *p*-values

A trial by Lau and colleagues published in *The Lancet* in 1999 evaluated the effects of adjuvant intra-arterial iodine-131-labelled lipiodol for resectable hepatocellular carcinoma in a prospective randomised trial.[23] The original sample size for this study was 120 patients, however, an interim analysis was planned when a total of 30 patients had been followed up for a median of 2 years, with the intention of stopping the trial early if the difference in DFS between the two groups was statistically significant, defined prior to the study as a *p*-value less than 0.03. The interim analysis revealed that the proportion of local hepatocellular cancer recurrences was 3/14 (21%) in the treatment (intra-arterial lipiodol) group and 11/16 (69%) in the control group (no adjuvant treatment) with a *p*-value of 0.01.[18] There was no significant difference in overall survival, although there were three deaths in the treatment group and six deaths in the control group, favouring the treatment group. Therefore, based on these two factors, the investigators decided to stop the trial and the results were made public. However, an additional 13 patients not included in the above analysis were randomised prior to the early termination of the trial, and the investigators decided to postpone the final analysis.[23] The final report, which included these additional 13 patients, was published 18 months after the trial was stopped early, and it revealed that the absolute difference in recurrence rates shrank from 48% to 30% during the interval between study termination and publication.[23] In retrospect, it was realised that this trial should not have been terminated early based on the small number of patients in the trial, and the very low number of deaths.

This trial raises the issue of when it is appropriate to stop a clinical trial based on *p*-values. The most commonly used guidelines are the methods of Pocock, O'Brien-Fleming or Heybittle-Peto, all of which are variations of the Lan-DeMets α-spending method. For each of these methods, a statistical test is performed at each interim analysis. If the treatment effect is large enough to attain statistical significance, the guidelines dictate that the trial be stopped early. In general, a much larger treatment effect and greater level of statistical significance is required in order to meet the requirements for early termination of a trial.

Specifically, the stopping rule utilised by Lau and colleagues is often termed a 'Pocock boundary', which typically requires a *p*-value approaching 0.02.[24] Pocock, who has published numerous articles on early stopping rules in clinical trials, feels that *p*-values approaching 0.05 are simply not persuasive enough for a new treatment to be widely accepted by clinicians.[25] Some statisticians have argued for $p < 0.001$ as a simple and effective statistical boundary for stopping a trial early[24], which would provide 'proof beyond a reasonable doubt', whereas other more complex suggestions of increasing *p*-values have been proposed.[24] In general, it is recommended that statistical boundaries for *p*-values should not be applied too early when few outcome events have been observed.[26] The choice of which stopping boundary to use will vary for any given trial, however, investigators should consider the consequences of this decision. Furthermore, the DSMC should not solely use *p*-values to decide on stopping a trial early. They should consider a number of important variables including: how the results might be interpreted by members of the wider community of clinicians, regulators, and patients; the width of confidence intervals which give a measure of accuracy of the treatment effect; the proportion of the planned and accrued sample size having a defined out-

come at the time of analysis; and previous interim analysis results. Pre-mature termination may result in poor uptake of a potentially effective therapy, while delayed termination may result in excess patients being exposed to an ineffective, and possibly toxic, treatment.

Despite these various statistical approaches, early stopping of clinical trials continues to be a complex and difficult issue for both statisticians and clinicians alike, and no single strategy is likely to prove useful in all trials. As opposed to applying definitive statistical rules in isolation, the decision to stop a trial early should take into consideration all of the relevant issues, and be led primarily by the DSMC.[27,28]

7. Bayesian analysis

As an alternative to traditional analysis methods for RCT's, a Bayesian statistical approach can be applied to monitoring study results and performing interim analyses.[29] Bayesian analyses combine a distribution of prior belief about the treatment effect with new data to produce a posterior belief distribution of the effectiveness of the treatment.[30] This distribution of prior belief involves assigning somewhat subjective probabilities to uncertainties so that they can be analysed and refined with experience over the duration of the study. Essentially, this is equivalent to pooling observed data from a clinical trial with implicit data from a previous distribution.[30] The resulting posterior distribution tends to diminish extremely large treatment effects to produce more conservative estimates, while giving more influence to the data, and less to the subjective prior distribution, as more data become available. With a Bayesian approach, the '*p*-value' is not used as the main means of assessing whether or not to stop a trial. Rather, the DSMC is asked at multiple points during patient accrual whether there is enough cumulative evidence for the study to remain open to recruitment or to close outright.[31] Given that a variety of prior beliefs can be modelled, one can interpret posterior distributions in the light of a variety of different prior beliefs (e.g. from very optimistic to very pessimistic). For example, if conclusions are not consistent across all posterior distributions, a DSMC could interpret this to mean that the data are not sufficiently robust to convince all members of the wider community; therefore, continuation of the trial may be warranted.

8. Interpreting trials stopped early: Important considerations

Studies stopped early for benefit are likely to have a significant impact on day-to-day clinical practice in oncology. Health care practitioners in oncology might ask themselves the following questions when interpreting these trials:

(1) How many interim analyses were conducted prior to the decision to stop the clinical trial early?
(2) Were the statistical methods used to monitor the trial (end-points, *p*-values, and confidence intervals for interim analysis) appropriate?
(3) Was the planned sample size sufficiently large to make conclusions about the study?
(4) Is there sufficient biological plausibility to account for the treatment effect?
(5) Does it seem reasonable that the cancer treatment would result in such a large benefit for the patient population being studied?
(6) Was the decision-making process for early termination transparent and well-described in the paper?
(7) Taking into account all of the literature previously published on this topic, what is the potential impact of this new information on my daily clinical practice?

9. Summary

Oversight by the DSMCs and explicit a priori rules for early termination are necessary components of any oncology clinical trial. Surrogate end-points remain a controversial topic in oncology trials, although future meta-analyses may help to clarify this issue. Ultimately, the decision of whether to stop a trial early is guided by recommendations of the DSMC. Statistical stopping rules serve as a guide to help DSMCs decide whether to continue a trial or make study results available to the public. In reviewing the findings from trials stopped early, evidence-based practitioners will benefit from a careful review of the study design, from identification of whether the stopping rules employed logical assumptions, and from obtaining a clear understanding of what mechanisms led the DSMC to recommend early trial closure. It is hoped that this article will help health care practitioners involved with cancer patients become more aware of the issues relating to oncology trials stopped early, in order to make an educated decision as to whether the results of a particular study should be applied to their clinical practice.

Conflict of interest statement

None declared.

Acknowledgements

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

REFERENCES

1. National Institute of Health Policy for Data and Safety Monitoring. June 1998. Website: http://grants.nih.gov/grants/guide/notice-files/not98-084.html.
2. Wittes J. Forming your phase III trial's data and safety monitoring board: a perspective on safety. *J Investig Med* 2004;**52**(7):453–8.
3. Montori VM, Devereaux PJ, Adhikari NK. Randomized trials stopped early for benefit: a systematic review. *JAMA* 2005;**294**(17):2203–9.
4. Allegra CJ, Yothers G, O'Connell MJ, et al. Phase III trial assessing bevacizumab in stages II and III carcinoma of the

colon: results of the NSABP protocol C-08. *J Clin Oncol* 2011;**29**(1):11–6.
5. De Gramont A, Van Cutsem E, Tabernero J et al. AVANT: Results from a randomized, three arm multinational phase III study to investigate bevacizumab with either XELOX or FOLFOX4 versus FOLFOX4 alone as adjuvant treatment for colon cancer. 2011 Gastrointestinal Cancers Symposium (ASCO GI). Abstract no. 362. *J Clin Oncol 2001*: **29** (suppl 4; abstract 362).
6. Leighl NB, Paz-Ares L, Douillard JY, et al. Randomized phase III study of matrix metalloproteinase inhibitor BMS-275291 in combination with paclitaxel and carboplatin in advanced non-small-cell lung cancer: National Cancer Institute of Canada-Clinical Trials Group Study BR.18. *J Clin Oncol* 2005;**23**(12):2831–9.
7. Alberts DS, Marth C, Alvarez RD, et al. Randomized phase 3 trial of interferon gamma-1b plus standard carboplatin/paclitaxel versus carboplatin/paclitaxel alone for first-line treatment of advanced ovarian and primary peritoneal carcinomas: results from a prospectively designed analysis of progression-free survival. *Gynecol Oncol* 2008;**109**(2):174–81.
8. Temple RJ. In: Nimmo WS, Tucker GT, editors. *A regulatory authority's opinion about surrogate end-points: clinical measurements in drug evaluation*. New York: J. Wiley and Sons; 1995.
9. Sargent D, Wieand H, Haller DG, et al. Disease-free survival versus overall survival as a primary end-point for adjuvant colon cancer studies: individual patient data from 20,898 patients on 18 randomized trials. *J Clin Oncol* 2005;**23**(34):8664–70.
10. Gill S, Sargent D. End-points for adjuvant therapy trials: has the time come to accept disease-free survival as a surrogate end-point for overall survival? *The Oncologist* 2006;**11**:624–9.
11. Meuller 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):878–81.
12. Bassler D, Montori VM, Briel M, Glasziou P, Guyatt G. Early stopping of randomized clinical trials for overt efficacy is problematic. *J Clin Epidemiol* 2008;**61**(3):203–4.
13. Goss P, Ingle JN, Martino S. A randomized trial of letrozole in postmenopausal women after five years of tamoxifen therapy for early stage breast cancer. *N Engl J Med* 2003;**349**(19):1793–802.
14. Goss P, Ingle JN, Martino S. Randomized trial of letrozole following tamoxifen as extended adjuvant therapy in receptor positive breast cancer: updated findings from NCIC CTG MA-17. *J Natl Cancer Inst* 2005;**97**:1262–71.
15. Cannistra S. The ethics of early stopping rules: who is protecting whom? *J Clin Oncol* 2004;**22**(9):1542–5.
16. Pater J, Goss P, Ingle J, Shelley W, Shepherd L. The ethics of early stopping rules. *J Clin Oncol* 2005;**23**(12):2862.
17. Moher D, Schulz KF, Altman D, et al. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *JAMA* 2001;**15**:1987–91.
18. Schulz KF, Altman DG, Moher D. For the CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *Ann Int Med* 2010;**152**(11):726–32.
19. Trotta F, Apolone G, Garattini S, Tafuri G. Stopping a trial early in oncology: for patients or for industry? *Ann Oncol* 2008;**19**(7):1347–53.
20. Bassler D, Ferreira-Gonzalez I, Briel M, et al. Systematic reviewers neglect bias that results from trials stopped early for benefit. *J Clin Epid* 2007;**60**(9):869–73.
21. Korn EL, Freidlin B, Mooney M. Stopping or reporting early for positive results in randomized clinical trials: the National Cancer Institute Cooperative Group experience from 1990 to 2005. *J Clin Oncol* 2009;**27**(10):1712–21.
22. Freidlich B, Korn EL. Stopping clinical trials early for benefit: impact on estimation. *Clin Trials* 2009;**6**(2):119–25.
23. Lau WY, Leung TWT, Ho SKW, et al. Adjuvant intra-arterial iodine-131-labelled lipiodol for resectable hepatocellular carcinoma: a prospective randomized trial. *Lancet* 1999;**353**:797–801.
24. Pocock S. When (not) to stop a clinical trial for benefit. *JAMA* 2005;**294**(17):2228–30.
25. Pocock SJ. When to stop a clinical trial. *BMJ* 1992;**305**:235–40.
26. Pocock S, White I. Trials stopped early: too good to be true? *The Lancet* 1999;**353**:943–4.
27. Armstrong PW, Furberg CD. Clinical trial data and safety monitoring boards. *Circulation* 1995;**91**:901–4.
28. Hill C. Surrogate End-points. *Bulletin du Cancer* 1999;**86**(7–):622–4.
29. Parmar MK, O-Griffiths G, Spegelhalter DJ, et al. Monitoring of large randomized clinical trials: a new approach with Bayesian methods. *The Lancet* 2001;**358**:375–81.
30. Dmitrienko A, Wang MD. Bayesian predictive approach to interim monitoring in clinical trials. *Statistics in Medicine* 2006;**25**:2178–95.
31. Korn E, Yu K, Miller LL. Stopping a Clinical Trial Very Early Because of Toxicity: Summarizing the Evidence. *Control Clin Trials* 1993;**14**:286–95.