

High-Dose Chemotherapy With Hematopoietic Rescue as Primary Treatment for Metastatic Breast Cancer: A Randomized Trial.

Bezwoda WR, Seymour L and Dansey RD.

J Clin Oncology, 13 (10) : 2483-2489, Oct 1995

STUDY DESIGN AND CONDUCT

Study Aim

The aim of the study stated in the methods section were to determine whether high-dose chemotherapy given to patients with metastatic breast cancer but without prior chemotherapy exposure could result in a complete response (CR) rate of greater than or equal to 50%, and to compare the CR rates of the high-dose chemotherapy arm with a conventional dose treatment regimen.

The aim stated in the abstract, the introduction and the discussion is to compare, in terms of CR rate, duration of response and duration of survival, the results of high-dose CNV with a conventional-dose treatment schedule of comparable drug composition, as first line treatment of metastatic breast cancer. There is an inference that the study is able to determine whether there is a dose-response relationship for the schedule by comparing the responses of the high and low-dose arms.

These aims were to be answered by undertaking a randomized trial with two arms: one high-dose cyclophosphamide, mitoxantrone and etoposide (HD-CNV), the other conventional-dose cyclophosphamide, mitoxantrone and vincristine (CNV). In the introduction, the authors state that the trial is ongoing, however it is not clear whether the authors are presenting an interim analysis or whether the trial closed after the last patient was entered in February 1993. This needs to be spelt out, and it would be essential for the authors to publish the trial in full if the end results show no difference between the two arms.

The study design would be able to compare results of the two arms and to determine the CR rate of the high-dose arm, however the study is not able to assess a dose-response relationship for a particular combination of chemotherapy since two different combinations were used. One combination contained etoposide while the other contained vincristine. The use of CR rates as the main outcome variable is suboptimal, and will be discussed later. The choice of CNV as the conventional arm will also be further discussed.

Sample size and power

It is stated that a sample size of between 40 and 50 patients was required (for the "first phase" of the study) to detect, with 80% power, an increase of the proportion of patients with CR of at least 30% over that which could be expected to result from conventional-dose treatment regimens. The authors do not state what CR rate could be expected from conventional regimens, although in the discussion they report previous experience of CR rates with CNV of between 6 and 20%. It is also possible, based on comments made in the "methods" section, that the authors wish to detect a difference of 50% in the high-dose arm versus 20% in the conventional arm. The expected CR rate should have been stated when discussing power calculations in order for the reader to check them. The method of power calculations (eg whether one or two sided tests were used) should have been discussed. In all, 45 patients were randomized to each arm, which is a small sample size and thus further reason for more detail regarding power calculations.

The authors need to delineate what is meant by the "first phase" of the study: once again it is unclear whether this study was set up as it is presented, or whether these patients are only part of a larger study. If part of a larger study, it should be stated how often analyses were to be prospectively performed, and what stopping rules were being employed.

The patients were accrued over 26 months which is an acceptable timeframe, with the last patient being entered in February 1993.

Patient selection

Source

The source of the subjects is not described. We are not given any information regarding referral base, number of centres involved or refusal rates of eligible patients. The possibility of selection bias thus arises. This should be addressed by the authors.

Patient eligibility

Aged 50 years of age or less. Although this is a valid entry criteria, menopausal status may have been more appropriate for patient selection when studying breast cancer patients.

Histologically or cytologically confirmed metastatic breast cancer: there is no mention of independent review of specimens which we must presume did not occur. There is also no definition of metastatic disease: the most recent TNM update defines

axillary nodes as regional metastases and the supraclavicular node as distant metastatic disease. Patients with soft tissue sites are included and the question arises as to whether this is in fact local recurrence alone, and not metastatic. Patients with loco-regional recurrence would be expected to do better than patients with distant metastatic disease, and thus this may be a confounding variable. The authors should better define this, as well as stating what quality control features were instituted.

No previous chemotherapy for metastatic disease was permitted, although presumably prior hormonal therapy was allowed (a potential confounder). However the definition does not exclude adjuvant chemotherapy which may also be a confounding variable. The authors should provide some figures regarding this.

Normal renal and hepatic function tests, however a precise definition of normal is not described. Normal cardiac function is well defined. The exclusion of patients with abnormal liver function tests may exclude poor prognosis patients with liver metastases, and restrict extrapolation of results.

ECOG performance status of greater than or equal to 2 defines a good prognostic group.

Patient stratification and randomization

Patients were not stratified for known prognostic variables; stratification may be important since it avoids gross imbalances between the groups for known prognostic variables and helps prevent retrospective data dredging. Randomization was performed by a random-number, closed-envelope technique: it is uncertain whether this is an indication of blindness. There is no indication of the time from randomization to treatment commencement.

Treatment description

There are several points to be made about the treatments described:

- It may be argued that in attempting to obtain two arms of the study with similar drug combinations, the conventional arm does not contain standard treatment. One combination which is commonly used in metastatic breast cancer is CAF (cyclophosphamide, adriamycin and 5-FU). This argument is further supported by the very poor results obtained in the "conventional" arm. These results in fact are even poorer than the authors own experience with the same regime, which cannot be accounted for except for other confounding factors: the authors need to comment on the reasons that this may be the case. Several letters and an editorial were written after the Bezwoda trial was published (5,11,12,15), some making the above points also.
- The two arms of the study do not contain the same drug combinations, which would invalidate any pure dose-response comparison.
- Mitoxantrone doses vary within the high dose arm depending on previous radiotherapy. In the authors' own dose-escalation pilot study, patients who received more than 35 mg/m² of mitoxantrone appear to have done the best in terms of complete response rate. However the HD-CNV arm has patients who are on 35 mg/m², a further confounder.
- Patients who responded to treatment were placed on tamoxifen, which also may influence outcomes to an unknown extent. The authors should comment on this fact in the discussion.
- Patients with progressive disease were taken off the trial and included in analysis of response rates. However there is no indication of what salvage treatment was given to these patients or those who relapsed after responding. This is important as it may directly influence outcomes analysed.
- In the high dose CNV arm 3 different types of haematologic support were used: autologous bone marrow, peripheral blood stem cells (PBSC) and PBSC with granulocyte colony-stimulating factor (G-CSF). It appears when reading the "discussion", that at the start of the trial, 9 patients received autologous bone marrow; however due to the longer than expected haematologic recovery and subsequent delay in proceeding to the second cycle of HD-CNV, patients were changed to PBSC. These 9 patients were the only HD-CNV subjects who had only one cycle of HD-CNV (of the two planned), and the only two non-responders in the high-dose arm were in this group also (stated in the dose-intensity results). Therefore the change in treatment support which was initiated after the trial began, probably had a direct effect on outcomes to be analysed, and is another confounder. The best available support measures should have been stipulated in the protocol, and not changed during the running of the trial.
- The priming dose of cyclophosphamide given to patients receiving haematologic support may also influence outcomes measured, although it does seem to have been included in the dose-intensity calculations.
- There were no strict guidelines in the protocol for use of other support measures such as blood products and anti-emetics. This may influence toxicity reporting to an unknown extent.

Quality control

There are no indications of quality control measures used in this trial. Trial design, conduct and analysis should all be subject to some form of quality control. For example, measures should have been instituted to ensure that entry criteria (such as pathological diagnosis), and assessment of response (eg by imaging) were subject to some form of review. Collection of data, accuracy of instruments and random checks for errors are all measures that may help improve the quality of the data (8).

Outcome measures

The main endpoint chosen is CR rate, with several other secondary endpoints analysed. These secondary outcomes have not been consistently stated in the abstract, introduction, results and conclusions; the authors should report the prospectively determined endpoints to be analysed consistently.

The choice of CR rate as the main endpoint in a trial investigating the management of metastatic breast cancer is not optimal. This is a group of patients with a relatively short life expectancy, who are often most interested in quality of remaining life. More useful and clinically relevant measures of patient welfare are survival and quality of life. Response rates are also more subjective than survival, and depend on individual physician perceptions and intensity of follow-up. Response duration is also subject to adequate follow up assessments, and is not a direct measure of patient welfare. Treatment costs might also be considered in this group of patients, but are not even alluded to.

Dose intensity is a difficult endpoint to quantify, and the authors use a formula to attempt to produce a dose-intensity index. This formula is not referenced or its derivation explained. Although not stated, it is presumed that a mean dose-intensity is calculated from each treatment arm before calculating the dose-intensity index. The authors should address these concerns. The measurement and comparison of dose-intensities obtained in the two treatment arms of this trial appears flawed for many reasons:

The two arms contain different drug combinations: the comparison of the two common drugs in isolation from the combination is not optimal, since drugs in the combination interact in known and unknown ways.

The use of a formula cannot take into account all of the known and unknown confounding variables, which interact in unquantifiable ways. Some of these confounders include: previous adjuvant chemotherapy, differing doses of mitoxantrone within the HD-CNV arm, differing cycle numbers, varying time periods for treatment, priming cyclophosphamide, tamoxifen for responders, salvage treatments used and differing support measures.

Some of these points were also made by Kennedy (11) in his editorial.

Toxicity endpoints are important, particularly when we are dealing with palliation in metastatic disease. Unfortunately toxicity outcome measures have not been delineated as secondary endpoints, although some toxicity data is provided. Most of the confounding variables affecting dose-intensity calculations may also interact with toxicity measures: for example different time periods between chemotherapy cycles will affect the side effect profile.

Survival outcomes seem to have been relegated to secondary outcome measures at. Only median survival times and a survival curve were reported for each arm: disease free survival and prospectively determined survival assessment (eg 1 year survival) may add valuable information, although short follow-up is again a problem.

The assessment of multiple factors on outcome variables should also have been prospectively stated in the trial protocol. The few factors analysed seem to have been done so retrospectively, and those chosen are suboptimal.

Treatment completion

Not all patients completed their intended treatment schedule:

- 9 of 45 patients only had 1 cycle of HD-CNV, as previously discussed.
- For the patients who did complete 2 cycles of HD-CNV, the mean time to complete therapy was 11 weeks (8 weeks was anticipated in the protocol).
- 21 of 45 patients in the conventional CNV arm received < 6 cycles of therapy: all of these patients were withdrawn due to progressive disease.
- 2 of 45 HD-CNV patients were withdrawn due to disease progression.
- There is no mention of how many patients (if any) required dose reduction: this needs to be stated.
- There were 3 toxic deaths with no other comment regarding whether intended treatment was completed.

Side effects

Table 8 states that toxicity resulted in 3 deaths, however the text contradicts this and states that no treatment related deaths occurred. Bezwoda (2) acknowledges that the text statement is an error in a subsequent reply. Besides the 9 HD-CNV patients who had delayed haematologic recovery and did not have their second course of treatment, there is no indication of whether toxicity resulted in dose reduction or withdrawal. Toxicity data is not complete, with no breakdown of data by grade. The presentation of toxicity data should be altered to include the above points in order to allow the reader to better determine the comparability of the two treatment arms.

Follow up

Median follow up time is short at 72 weeks, and it is not clear whether the trial protocol stated the total duration of the trial and patient follow-up. There is no indication of whether patients were lost to follow-up. These points need to be clarified. The criteria for patient and disease assessment is not well documented, particularly in terms of follow-up. Regularity and

duration of follow-up, and investigations required after chemotherapy had ended are not described. Thus patients might relapse after a response and not be detected if there is inadequate follow-up, introducing potential bias. This is an area that must be standardized in the protocol, particularly since response (which is more subject to bias than say, survival) is the main endpoint. The authors should clearly state intended follow-up procedures.

ANALYSIS AND RESULTS

Patient characteristics

There is no attempt to use stratified randomization for major prognostic variables, which helps prevent gross imbalances and prevents retrospective data dredging, as discussed in Paper 1.

Some of the patients' baseline characteristics are listed in Table 3. It is interesting to note that there is an uneven distribution of oestrogen receptor status among the two groups. This may have been prevented by stratification. There are a high proportion of unknown results however, which is cause for concern. The fact that this baseline data is incomplete leads us to question whether baseline characteristics were considered prospectively, or determined retrospectively, whether entry criteria were strictly adhered to, what quality control measures were utilized, and whether there was any prospective intention to determine the effect of prognostic factors on outcome variables. The authors should address these points.

The baseline characteristics do not include other variables such as time to develop metastatic disease, and number and type of metastatic sites. The latter two are later analysed, presumably retrospectively: this is not optimal, as subset analyses should be stated in the trial protocol. There is also no indication of the number of patients with regional lymph node involvement or local recurrence alone: presumably this would be a better prognostic group. The distribution of these variables should be delineated.

Statistical analysis

One of the trial's major deficiencies is the failure to describe or reference any of the statistical analyses undertaken. This makes questionable all of the results and subsequent conclusions. For those outcomes tested, hypothesis testing is the main statistical test used: the authors give no indication of the cut-off level for statistical significance (eg 0.05), nor whether one or two-sided P values are used (two-sided tests would be most appropriate). The statistical tests utilised should be better described and referenced. The use of confidence intervals is preferable to quoting P values alone, and confidence intervals should be included in the results.

When analysing response rates, the authors state that there is a significantly higher complete and overall response rate for the HD-CNV arm compared to the CNV arm, however only one P value is given ($P < .01$): both analyses should be given. Confidence intervals were given only for CR rate for HD-CNV (51%; 95% confidence interval, 36% to 65%) and conventional dose CNV (4%; 95% confidence interval, 2% to 7%). There was no confidence interval given for the comparison between the two arms: as previously stated, 95% confidence intervals should be given for all relevant results. The CR rates were only analysed by number of metastatic sites. The number of metastatic sites for comparison (2 or more v 3 or more) may have been arbitrarily decided in retrospect (ie as part of a data dredging exercise): the intent to analyse this should have been stated in the trial protocol to avoid bias. If the data relating to metastatic sites was not specifically addressed in the protocol, there may also have been missing data (as with receptor status), further invalidating such analyses.

The dose-intensity endpoints were presumably calculated by determining the mean dose-intensities for the HD-CNV and CNV arms, and using these to calculate a dose-intensity index. The authors subsequently state that "patients on HD-CNV actually received 3.8 times more cyclophosphamide and 2.2 times more mitoxantrone as patients on conventional-dose CNV". However this observation is not backed up by any statistical analysis. The calculations use a formula that is not referenced or explained, with no discussion as to its validity. Although some readers may argue that it is of interest to be shown the figures of dose-intensity index calculated and actually received, I do not believe that any conclusions can be drawn from a comparison between the two arms as discussed in the "outcome measures" section. If the authors do wish to proceed with such a comparison, the statistical analysis should be presented.

The treatment toxicity results are commenced with the statement: "There were no treatment-related deaths on the study". As previously discussed, this is in direct conflict with the data shown in Table 8 which clearly lists 3 treatment related deaths (of which 2 occur in the conventional-dose CNV arm). This requires correction. Despite the obvious importance of treatment toxicity, there was no statistical analysis of any potential differences between the two arms. The only apparent analysis is a statement declaring "a significantly more rapid rate of hematologic recovery (median, 12 days; range, 9 to 17) for patients who received peripheral-blood stem-cell plus growth factor rescue as compared with bone marrow rescue or peripheral-blood stem cells without growth factor". However, once again no P value is given to back up this statement. In addition, this is an analysis performed retrospectively with small patient numbers and no confidence intervals, and thus no conclusions can safely be drawn from it. There is no mention of recovery rates for the conventional arm. Toxicity endpoints need to be better assessed in the results section by the authors.

The toxicity data in Table 8 are of interest, however could have been further broken down into actual grade. There is no indication of which patients (suffering toxicities such as neutropaenic fever, haemorrhage or death) received G-CSF support, nor what toxicity patients actually died from: this information would have been of considerable interest. The data in Table 8 indicate a substantially higher toxicity in the HD-CNV arm. Also these high-dose patients had a 9 day mean hospital stay,

with no comment about that of the low-dose arm. This data is of interest when considering the use of a new treatment schedule in palliation of cancer, and should be quoted.

Response durations were quoted as medians, and a figure provided. The difference between the two treatment groups was stated to be significant, with no P value or confidence intervals given. This level of information is not adequate.

Median survival times are quoted (45 weeks for CNV v 90 weeks for HD-CNV), and a survival curve is shown, however there is no other survival data given, and the median follow-up time is only 72 weeks. Once again the difference between the two arms (in terms of median survival) is quoted as being significant, yet there is no formal statistical analysis, no P value given and no confidence intervals used: the authors must give the analyses. Although survival curves display useful information, they should not be compared by visual impression: tests such as the logrank test should be utilised. Median survival times should be derived from the Kaplan-Meier curve (Altman, 1991, p.386): the apparent median survival times calculated from the displayed curve (figure 2) seem to differ from those quoted in the text (for example, the median survival time for the HD-CNV group is closer to 100 weeks when using the curve, although it is a little difficult to estimate since the axes are not adequately scaled). In addition, when determining survival, there is no indication as to whether those patients withdrawn from the trial were included in results. These points require clarification. One of the best ways of analysing survival is using the logrank test and hazard ratios (Altman, 1991, p.385). The authors should draw attention to the short follow-up time of only 72 weeks when discussing survival (and other results).

Survival is analysed for complete versus partial responders, an analysis that is flawed since the groups are defined by a factor not known at the start of treatment (Altman, 1991, p.387). No other patient subsets are analysed. In general the analysis of the impact of factors on outcomes measured has not been adequately considered (for example Cox regression was not performed in the survival analysis).

Presentation

The main concern with this paper's graphical presentation was the survival curve. The curve's axes are not calibrated and the numbers at risk are not displayed. This makes interpretation difficult, and should be rectified.

TRIAL CONCLUSIONS

The following conclusions are reached in the discussion:

Dose-intensities achieved in the HD-CNV arm for cyclophosphamide and mitoxantrone were 2.2 and 3.8 fold that achieved in the conventional arm. One reason that this was less than that originally calculated, was the longer time to haematologic recovery (and subsequent delay in initiating the second high dose cycle) that occurred in those who received bone marrow grafts.

Time to haematologic recovery was substantially reduced by a change to peripheral blood stem cell rescue, especially when combined with G-CSF. The more rapid haematologic recovery obtained would probably not make much difference in treatment outcome.

Patient selection had a substantial influence on outcome, as demonstrated by a difference in CR rate depending on disease bulk (patients with 2 or more metastatic sites had a significantly higher CR rate than those with 3 or more metastatic sites).

There was a substantial and significant difference in overall remission and CR rates between the HD-CNV and CNV treatment arms: the CR rate was increased 11 fold in the HD-CNV arm. There was also a significant increase in duration of response and survival of approximately 2 fold in the HD-CNV arm. Durations of response and survival were not increased in direct proportion to the increase in CR rate. This may be due to CR rates being overestimated, with substantial patients having residual, undetected disease.

The clinical efficacy of the HD-CNV drug combination may be determined by the individual drug dose-intensities, not the drug combination.

High-dose chemotherapy seems to be a suitable treatment approach for selected younger patients with breast cancer.

Several problems apparent in this trial cast some doubt as to the conclusions reached by the authors:

It is uncertain whether this trial is part of a larger trial, and if so, whether it was prospectively intended to analyse the data at the present time. If the trial is still being conducted, the intended accrual numbers and time of final analysis should be stated. If the trial was intended to stand alone, the reference to the "first phase" of the trial should be explained.

Although the trial purports to have sufficient patient numbers, there are only 90 patients in the trial, with a smaller number actually receiving the intended treatment (it is unclear exactly how many since no dose reduction figures are given). Power calculations are not referenced, and the CR rate expected from the conventional arm is not quoted: these should have been addressed in the methods section of the paper. Power calculations are only based on a single outcome variable (the main endpoint), in this case CR rates. Since this is a small trial, with some uncertainty with regards to power calculations, confidence intervals should have been used in the results section, but were not.

The short follow-up time is not commented on.

The use of CR rate as the main outcome measure is questionable, as previously stated. The authors comment on the fact that the increased CR rates in the HD-CNV of "11 fold" compared to the CNV arm did not seem to directly flow on to an 11 fold increase in duration of response and survival (only "2 fold"). This is hardly surprising, and is a very good reason not to use response rates as the main outcome variable. In this group of patients, survival, quality of life and cost of treatments are probably more important outcomes. Response rates are also more subjective than survival endpoints.

Dose-intensity endpoints are probably not valid for the many reasons previously discussed. Due to this and the many confounding factors (not least being responders receiving tamoxifen), the conclusion that the clinical efficacy of a combination may be determined by the individual drug dose-intensities and not the combination itself, cannot be sustained.

Toxicity endpoints have not been well considered, with no analysis of the difference between the two arms. Toxicities listed in Table 8 appear to be much worse in the HD-CNV arm, and this is reflected in the mean hospitalization time of 9 days. These endpoints should have been analysed secondary outcome measures which should be considered when putting forward the HD-CNV treatment schedule as a promising schedule.

The statistical analysis is suboptimal in many respects, the most important being:

- There has been no referencing of statistical procedures used.
- Some statements that differences between the two arms are significant have not been backed up by any statistical analysis.
- For the main statistical analyses hypothesis testing alone has been used (with only one mention of confidence intervals).
- There has been scant attention paid to the effect of other variables on outcome measures. The conclusion that a higher CR rate depends on number of metastatic sites is flawed as this appears to be a retrospective analysis with an arbitrary number of sites chosen.

Very poor results have been obtained in the "conventional-dose" arm (even compared to the authors' own past experience), leading us to wonder why this has occurred. Some of the possible explanations include:

- CNV may not be the best available treatment (CAF may be better)
- Randomization procedures may not have been adequate
- There may be an imbalance in the treatment groups (eg uneven distribution of patients with loco-regional disease alone).

The HD-CNV arm has three different haematologic supports used, and furthermore these appear to have been changed during the trial due to unexpected length of haematologic recovery. This change of patient management mid-trial is suboptimal, and has probably affected outcomes for the HD-CNV arm: the 9 patients who had bone marrow support appear to be the only HD-CNV patients who received only 1 cycle of chemotherapy, and the only HD-CNV non-responders were in this group. Thus the authors' comment that more rapid haematologic recovery would not make a substantial difference in treatment outcome seems to be in conflict with the trial's own results.

RECOMMENDATIONS

This trial explores a very important area of patient management, and purports to be the first randomized trial in the area studied. The trial does however contain several important problems which do not allow us to draw any firm conclusions: at best the trial provides us with some interesting results that need to be further studied with better designed and analysed trials with larger patient numbers.

As a trial reviewer, I would accept this trial for publication after several suggestions for amendment or review. These would include:

- A definitive statement describing whether the trial was designed to stand alone as is, or was part of a larger trial. If it was part of a larger trial, details should be provided regarding prospective intention to analyse data, intended trial accrual and duration and expected time to final analysis. The reference to "first phase" requires explanation.
- A clear statement of prospectively determined outcome variables and methods of assessment, including interactions of other variables on endpoints. Toxicity outcomes to be analysed should also be stated.
- Power calculations should be more precisely described.
- Degree of blindness should be stated.
- Source of patients should be stated.
- Intended follow-up procedures and time should be stated.
- The formula for dose-intensity calculations should be discussed and referenced.
- The number of patients in each arm actually receiving intended treatment without dose-reduction should be

described.

- Prognostic information should be given for loco-regional disease alone, time to metastatic disease and use of prior adjuvant chemotherapy.
- Reasons for incomplete data should be elucidated, and the authors should discuss what quality control measures were instituted in the trial.
- The effect of excluding patients with abnormal LFTs (ie with potential metastatic disease) should be discussed, perhaps providing the number of such patients excluded, and limitations for extrapolation of results.
- The effect of using 3 different types of haematologic support should be discussed in relation to their effect on outcomes, as well as whether provision was made in the trial protocol for changing such supports. The authors' comment that more rapid haematologic recovery would not make a substantial difference in treatment outcome seems to be in conflict with the trial's own results, as previously discussed. The authors should review or justify this conclusion.
- Statistical analyses should be referenced. All analyses performed should be presented. Confidence intervals should be given for all relevant results.
- All retrospectively determined analyses should be stated as such.
- Indicate whether those patients withdrawn from the trial were included in results (ie state whether intention-to-treat analyses were used). The use of logrank test and hazard ratios would be encouraged when analysing survival. Overall survival should be quoted. In the discussion of survival results, the authors should comment on the short follow-up time and the poor results for the conventional arm. Survival analysis for complete versus partial responders should be omitted. Numbers of patients at risk should be displayed on the survival curve.
- The authors should consider better analysis and presentation of toxicity data.
- The error relating to patient deaths should be corrected.
- The effect of variables on outcome measures should be analysed (eg using regression analysis).
- If the authors wish to conclude that the clinical efficacy of a combination may be determined by the individual drug dose-intensities and not the combination itself, then they should as a minimum discuss the potential confounders inherent in the study (for example responders receiving tamoxifen).
- In the conclusion, the poor results obtained in the conventional arm should be further discussed, along with the possible implications for results.
- The discussion of the results should include an assessment of the confidence intervals obtained if these affect the interpretation of the results.
- The authors should consider resubmitting their results after longer follow-up in order to determine if there are any prolonged survival benefits.

Addendum:

Unbeknown to this candidate and all other oncologists at the time, this article was assembled from fabricated evidence. It is not politically correct to mention this before a trial is discredited, but one should maintain at the back of one's mind the maxim "If the deal appears too good to be true, it probably is!" AAM 27/12/2007