Reviewer's report

Title: Disease-specific survival for limited-stage small-cell lung cancer affected by statistical method of assessment

Version: Date: 27 May 2006

Reviewer: Esa Läärä

Reviewer's report:

General

Analysis of medical survival data by other approaches than the ubiquitous Kaplan-Meier (KM) estimator and the proportional hazards (PH) model introduced by Cox (1972) is refreshing. Sir David Cox himself has expressed his concern about overuse of the PH model, which he prefers not to be called as "Cox' model", and has urged survival analysts to consider more seriously suitable parametric alternatives. In light of this the present manuscript has promising aims.

The statistical analysis and presentation has several shortcomings.

-----------------------------------------------------------------------------------------------

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

1. Given that the focus of the paper is in comparison of different statistical methods, it is strange that the article starts with a lengthy discussion about the TNM classification and its alternatives. I suggest that the two first paragraphs of section "Background" are totally removed.

2. Given the focus of the paper, the description of the different statistical approaches in section "Methods" is inadequate. The less familiar log-normal models would merit a more explicit specification maybe supported by key mathematical formulas in an appendix. It would be very important here to know in particular, how the "cure-rate" parameter in Boag's model was treated; was it indeed assumed to be > 0 and hence estimated? For consistency, I would expect that the log-normal model incorporating the prognostic factors included a cure-rate component, too, but is this so? Moreover, a more detailed account of the methods of fitting the various models and estimation of their parameters at the end of this section is essential. A phrase "All analyses were performed with ... package" as such is equally informative as "The manuscript was written with Word".

3. The prognostic factors were assessed "in step-wise forward model building", but the authors do not provide any principle or criteria, according to which this model building proceeded. The use of automatic variable selection procedures based on arbitrary "significance" levels is widely condemned by statisticians because of their biased performance. A sounder approach to variable selection in a prognostic model would be based on information criteria like AIC, BIC, etc. Nevertheless, it is not difficult to believe that the four chosen factors are relevant in this context.

4. Section "Results", paragraph 1: "... reasonably supported the use of log-normal survival analysis..." is a misinterpretation of the quoted P=0.42. A more adequate comment would be "... were also reasonably consistent with the log-normal model..."

It is clear that the data set was not large enough to discriminate, which of these models would be a better description of the "truth".

5. Section "Results", paragraph 1, and Tables 2 to 6: The results presented in Tables 3 to 6 do not provide any new information upon Tables 2 but are largely redundant, because all of these "major subgroups" constitute 80 to 94 percent of the total patient cohort. Hence, tables 3 to 6 deserve to be removed from the manuscript.

6. Section "Results", Tables 2 and 7: The comparison of the survival proportions by different methods is only based on point estimates, but no assessment of the precision of them is given. I would expect that most of the observed differences would well be within a reasonable random error margin. Hence, I would
like to see adequate assessment of random error.

7. Figures 1 and 2: These could be combined into one figure in which solid lines could represent the estimated survival curves from the PH model and dashed lines those from the log-normal model. Only then the reader would be able to compare the survival curves predicted by the two different models. The four different risk groups do not require any plotting symbols, which are chart-junk given that the groups are so well separated from each other. The four groups are easily indicated by locating their identifiers A to D close to the right ends of the pertaining curves.

8. "Discussion": When the error margins of the observed differences in the survival proportions were not quantitatively assessed at all (see item 6. above), the interpretation (para 2) of the more extreme contrasts remains fuzzy.

9. "Discussion", Para 3: "Where the model-type is supported by the data..." A statistical model is rarely "supported" by data; we mostly encounter a situation that our favourite model is at best concordant with the data within a reasonable error margin, but so are typically many other alternative models, even very different ones. "The smooth log-normal survival analysis is advantageous over the Cox plot" is a strange statement, which probably means "The smooth log-normal is advantageous over the semiparametric proportional hazards model", but which even then is questionable given the problem in its premise. The discontinuity of the PH survival estimates is presented as a problem, which it is not. An experienced statistician would have a low trust on survival estimates at the right tails of the curves anyway, whatever model was used in their estimation: parametric or non-parametric.

10. "Discussion": The last three pages comparing the empirical results of this study to others is largely irrelevant and should be removed, if the focus is on comparison of the survival analysis methods. Even if relevant, the quotation of P-values associated with some factors in different studies is completely void of useful information.

-----------------------------------------------------------------------------------------------
Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

-----------------------------------------------------------------------------------------------
Discretionary Revisions (which the author can choose to ignore)

What next?: Unable to decide on acceptance or rejection until the authors have responded to the major compulsory revisions

Level of interest: An article of limited interest

Quality of written English: Acceptable

Statistical review: Yes

Declaration of competing interests:

I declare that I have no competing interests.