Reviewer's report

Title: Family history of cancer on survival of patients with gastrointestinal cancer in northern Iran, using frailty models

Version: 1 Date: 2 February 2011

Reviewer: Tron Anders Moger

Reviewer's report:

This is a paper which both aims to promote parametric survival models as an alternative to Cox models and introduce frailty models to a wider audience. The main problem with the paper, is that the conclusion that the Cox model fits worse than the parametric models seem to be based on a misunderstanding. One cannot compare log likelihood values (or AIC values which are just functions of the log likelihood values) from parametric models to log partial likelihood values from a Cox model directly to assess the fit of the different models. This argument should be apparent just by using intuition: In the parametric models you use e.g. a Weibull distribution for the baseline hazard. In a Cox model, the baseline hazard would be non-parametric (i.e. more flexible than a parametric distribution). Hence, the Cox model should provide a better fit in all cases except when the baseline hazard is indeed well approximated by a Weibull distribution. In that case, the parametric model may have some interesting additional features compared to the Cox model, which is also the point in the Nardi paper the authors refer to. Another problem is that from the Cox-Snell residual plots, none of the parametric models appear to have a very good fit (there are departures from the straight line for all parametric models). The deviances come at high values of the Cox-Snell residuals, where there might not be much data, making the deviations less serious, but the authors should respond to this. Also, why not show the Cox-Snell residuals for the Cox model (which should also be possible to get)? Perhaps you will then see that the fit of the Cox model is ok. Due to these problems, consider changing the aim of the paper to just introduce frailty models to the journal's readers. In that case, the main content may still be used, as adding a frailty term to the Cox model seem to significantly improve the fit from the partial log likelihood and AIC values (which should be comparable as you then compare to Cox models where the only difference is that you add a frailty to one of them).

Other comments:

In the methods section you should also describe what to look for in the Cox-Snell and deviance residual plots (as you do when explaining the AIC criterion in the methods section), as many readers might not be familiar with these.

It seems from some of the text in the discussion and in Figure 2 that you are using an accelerated failure time modelling of the covariate effects in the parametric models. But this can't be correct? When I read table 3, I read it as having a family history of cancer means a 46% increased risk of dying from the
standard Cox model, or a 85% higher risk from the exponential model without frailty. An accelerated failure time interpretation would be that an individual with family history of cancer lives 85% longer than a person without cancer.

There are not many typos in the manuscript (Affect instead of effect, page 4, bottom and STAT instead of STATA, page 6, bottom), but many sentences throughout are either incomplete or have strange ordering of words (example: Among the parametric models the Loglogistic model with gamma frailty fitted the data was more appropriate, middle of page 7). The authors should carefully check the manuscript for these sentences.

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Not suitable for publication unless extensively edited

**Statistical review:** Yes, and I have assessed the statistics in my report.

**Declaration of competing interests:**

I declare that I have no competing interests.