Reviewer's report

Title: Parametric Frailty Models for Clustered Data with Arbitrary Censoring: Application to Effect of Male Circumcision on HPV Clearance

Version: 1 Date: 24 February 2010

Reviewer: Andreas Wienke

Reviewer's report:

Frailty models are becoming more and more popular in survival analysis during the last two decades, especially in multivariate survival analysis. Unfortunately, estimation procedures for more complex models are not supported in standard statistical software which limited the applicability of these models. Especially in the log-normal frailty model the unconditional likelihood can not be written down in an explicit form. Consequently, parameter estimation in this model requires approximation methods for the integrals in the likelihood. The authors of the present paper suggest adaptive Gaussian quadrature methods in a parametric shared frailty model to circumvent the above mentioned problem in the case of clustered event times with arbitrary censoring. The paper is written in a very clear and concise style. I would like to make a few minor comments:

page 5 formula (1) and later: x should be xij
page 5 and later: expression 1 should be expression (1)
page 6 line 6: I do not understand the proportional hazards property of the Gompertz distribution. You can use any distribution on the nonnegative real line as baseline distribution h0 in formula (1), resulting in a parametric PH model. This is also connected to your statement on page 7, lines 9-12.
page 7, line -10 and later: exp # should be exp(#)
page 8 line 9 including fixed effects # and the variance of the random effect
page 8 last line (formula) I miss a statement that the observations are reordered in the sense that the first observations are the left censored ones, then the interval censored observations and finally the right censored observations follow.

For the discussion of other frailty distributions on page 16 I miss the reference
A likelihood reformulation method in non-normal random effects models.
Statistics in Medicine 27, 3105 - 3124

Furthermore, it would be useful to expand the discussion on page 16 by including more detailed information about available software for frailty models with special emphasis to left and interval censored event times. One example would be
Henschel, V., Engel, J., Hölzel, D., Mansmann, U. (2009)
A semiparametric Bayesian proportional hazards model for interval censored data with frailty effects. BMC Medical Research Methodology 9, article 9

For general discussion I miss the reference
The use of Gaussian quadrature for estimation in frailty proportional hazard models.
Statistics in Medicine 27, 2665-2683

**Level of interest:** An article of importance in its field

**Quality of written English:** Acceptable

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

**Declaration of competing interests:**

I declare that I have no competing interests