Reviewer's report

Title: Estimating and modelling cure in population-based cancer studies within the framework of flexible parametric survival models

Version: 1 Date: 18 March 2011

Reviewer: Laurent Remontet

Reviewer's report:

Major Compulsory Revisions

1) This article is interesting; however, in its present form, it seems a bit long, especially section Results (five pages!). In fact, this section includes paragraphs that would better belong to sections Methods or Discussion. A shortening of this section and a reallocation of some paragraphs would greatly improve its legibility.

2) I do not think it is appropriate to evaluate the sensitivity to the location of the knots by comparing predicted survival with Ederer 2 estimates. Indeed, different cure models can have different cured fraction estimates but, at the same time, have similar predicted survivals (due to different estimates for the non-cured population). Furthermore, as « the cured proportion is estimated from the cumulative excess hazard at the last knot », the position of the last knot is necessarily crucial (?). A last knot set very early during follow-up will generate bias.

The authors should give the value of the last knot in years (and not percentiles) for each studied setting (left panel of Figure 1) so one could see how much these values differ from each other. They should give also the range of the cure fraction according to each combination age group * calendar period. Besides, the authors may, for illustration, give the estimates of the cured fraction with last knots set at 6, 7, 8, 9 and 10 years.

3) The authors seem to have solved the above-mentioned problem by setting the last knot at 12 years. The patients being censored 10 years after diagnosis, what is the difference between this model (restricted FPM with cure in Table 3) and the standard FPM? In Table 3, it seems that there is no difference between the two models between 0 and 10 years after diagnosis (period of data availability). This is a key point that should be clarified.

4) One important problem in cure models is the strong negative correlation between the parameters of cured people and those of non-cured ones. The authors have added flexibility to the model (by avoiding a parametric distribution) but they have also added a strong constraint on the cumulative hazard (slope zero after the last knot). The consequences of these two conditions on the correlation between parameters should be discussed.

5) The modelling uses the log of the cumulated rate. It would be interesting to
show some graphs relative to this rate and not to survival because, with survival, the differences are difficult to see (for example, in Figure 2, the differences between the curves relative to « age group 80+ »)

Minor Essential Revisions

6) According to the notation in equation (9), it seems that when K=5, the knots are kmin, k1, k2, k3, and kmax. However, this does not seem to be the case in equation (12): what is k4 (when K=5 and j=1)?

7) What are the natures of the parameters in Table 3? log hazard ratios? Else?

8) The references to absent sections should be deleted (for example, « section 2.4 » page 10, « section 3.3 » page 11)

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

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’