Answer to Referee #2

Alessandro Beretta
alessandro.beretta@uliege.be

Cédric Heuchenne
c.heuchenne@uliege.be

January 2021

First of all, we would like to thank the Referee for his/her careful review of the paper. We did our best to revise the article according to his/her valuable suggestions, for which we are grateful. Please find below our answers to the comments raised by the Referee; all modifications in the manuscript have been made using red text.

It would be useful to define both the incidence component $L_1(\mathbf{b})$ and latency component $L_1(\beta, h_0)$ after Equation 1 instead of the M step of the E-M algorithm. Similarly, it would be helpful to directly state the relationship between the incidence component and the time-fixed covariates in Equation 1 as well.

At the end of page 2, we added the definitions of L_1 and L_2, as well as their relationships with the covariates $\mathbf{x}_i$ and $\mathbf{z}_i'(t)$, respectively.

It is unclear what is meant by R_j is the risk set at $t_{(j)}^-$. Further explanation is necessary.

We added the definition of the risk set R_j at time $t_{(j)}^-$, see page 3.

In addition to specifying the tuning parameters and the information criteria, the authors may want to mention how the starting values are selected for each penalty type.

At the beginning of page 4, we now mention that, by default, the initial values for $\mathbf{b}$ and $\boldsymbol{\beta}$ are vectors with all elements equal to zero and that the user can provide his own starting values using the argument `SV`.

It is unclear if the percentages for cure in Table 1 are with respect to all the data or given the censored data.

At the beginning of paragraph 2 of page 5 and in the caption of Table 1, we specified that the percentages of censoring and cure are both computed with respect to the whole sample size.

The authors mentioned fitting a standard PH cure model with all covariates (FULL), but did not provide results in Table 2. Moreover, the results in Table 2 can be condensed into a figure faceted by censorship and cure percentage to summarize relationships and highlight trends. Furthermore, the paper should use the mean of the model errors and the mean of the relative model errors for consistency as well as taking into account large model errors.

For a matter of space, in the previous version of the article, we did not include the absolute errors of the FULL model, we only kept the relative errors of the other models with respect to the FULL model. In this new (revised) version of the paper, we provide in Figure 1 and 2 the errors of the incidence and latency components, respectively, which summarize the results of all

applied methods, including the errors for the FULL model. Moreover, in Figure 3, we provide the number of correct/incorrect zeros identified by the SCAD method (both incidence and latency components). Finally, as suggested by the referee, we used mean quantities instead of the median. For a matter of space, we kept the mean of the relative errors in the supplementary material (see column 6 and 10 of the table contained in the file *'beretta- heuchenne-suppl.pdf'*).

The authors can expand the interpretation of Table 2 and Table 3 with respect to censorship and cure as well as between incidence and latency.

We rewrote the paragraph containing the interpretation of the results in the Section "Simulation study" and we added an additional comment describing the behaviour of the model in situations of high censoring and low cure rate.

Lastly, 100 resamples may not be sufficient for the comparison of the coverage probabilities of the estimated 95% confidence intervals for the ORACLE model using the basic and percentile bootstrap methods, namely for large sample sizes or estimating latency components $\beta_1, ..., \beta_4$.

We increased the number of resamples from 100 to 500, we ran again all the simulations and we inserted the new results in Table 1.

Alternative approaches in R were proposed and limitations were discussed. Specifically, the authors were unaware of other R packages for estimation of PH mixture cure models with time-varying covariates where the user is able to perform variable selection. Additionally, the authors listed available R packages In the context of cure models for right censored data. A point of contrast may be to compare spduration package versus the method proposed in the paper since both approaches are cure models with time-varying covariates.

In the Introduction, we rewrote the end of the second to last paragraph, which now clearly states the differences between the `spduration` and the `penPHcure` packages.