Answer to Referee #1

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 review of the paper. We revised the article according to his/her suggestions. Please find below the list of modifications, which are marked in red text in the revised manuscript, and our answers to the referee.

- In the Introduction, we clarified the goal of the manuscript: the presentation of the **penPHcure** package. Moreover, in the last paragraph, we included the structure of the article, which allows the reader to have a clear idea of what to expect from each section:
  - the first part of the section "Methodology" is devoted to the description of the algorithm behind the `penPHcure` function when the argument `pen.type` is set equal to `"none"`;
  - the section "Variable selection" describes the algorithm behind the `penPHcure` function when the argument `pen.type` is set equal to `"SCAD"`.
- The referee noted that the paper of Hendry already provides the R code for generating time-to-event data depending on time-varying covariates. This is true, but only for time-to-event variables on a discrete scale. As we stated in our article, the algorithm implemented in the function `penPHcure.simulate` is similar to the one proposed by Hendry, but it is not the same, in the sense that our algorithm allows the user to generate time-to-events on a continuous scale. Moreover, we provide more details concerning the function $g(t)$ implemented in the function `penPHcure.simulate` (see page 4).
- Concerning the recommendation "*Introduce, describe and document the two main functions of the penPHcure package, similar to the help documentation if necessary*", the addition of a full description of all the arguments of the two functions would increase too much the length of the paper. That is the reason why we intentionally left it out from the article. However, in the section"Application", we already show and explain the usage of the `penPHcure` function to analyze a real dataset. We think that this article should not be seen as the documentation of our package, but a short description/summary of its functionalities.
- At the beginning of the section "Simulation study", we clarified that the "proposed methodology" is the PH cure model and its variable selection technique, which are implemented in the `penPHcure` function.
- Concerning the recommendation "*In Simulation Study, demonstrate how the functionality of the R package implements the Cox PH cure model*", we already added an `.R` file in the Supplementary Material with the code used to run all the simulations, as it is specified at the end of the section "Simulation study". This has been done for a matter of space, otherwise the length of the paper would become too long. Moreover, concerning the recommendation "*In other words, compare the output of the R package's functions to the expected output*", this is something we are already doing when we compute the model errors. We compare the fitted probabilities of being susceptible and the survival probabilities with the true ones.

- In the section "An application to Criminal Recidivism data", we inserted a comment stating that the `cpRossi` data.frame can be found in the `penPHcure` package.
- Concerning the remark "*Documentation states 'install.packages(penPHcure)' to get the package. But proper syntax would be, install.packages("penPHcure").*", many thanks ! We will correct this typo in the next release of the package.