Reviewer's report

Title: Model-based estimation of measures of association for time-to-event outcomes

Version: 1 Date: 15 March 2014

Reviewer: Paul Blanche

Reviewer's report:

2.1 Major Compulsory Revisions
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

Overall, I found the paper interesting and quite easy to read. In addition, I fully agree with the authors about the fact that:

1. "Sometimes, however, the use of specific measures of association depends also on the statistical methods available for estimation".
2. the pseudo value approach can provide a nice solution to overcome this issue.

However, I only partially agree with the authors when they write: "In this work a simple approach to straightforwardly obtain point and interval estimates of association measures, by using transformation models through suitable link functions, is presented." in the discussion section.

Although I agree about the fact that the approach is simple and straightforward to use, I am questioning on whether the approach is so straightforward to use in a proper way in practice.

2.1.1 About risk of misspecification
------------------------------------

In the section "Model-based estimates of association measures", the authors nicely pointed out the fact that:
- "In principle when a covariate effect is constant using a specific link, it should be time-varying with the other links."
- "Time-dependent effects selection depends therefore on the link transformation used"

Therefore, a short discussion or note of caution about the risk of
model misspecification would probably be welcome. Indeed, a practitioner could be interesting in fitting two different transformation models, in order to obtain two different relevant summary measures. In such a case, he or she must be aware of the fact that, very likely, the two models cannot hold simultaneously (especially if parsimonious models are chosen). My point is not to never recommend the use of two different transformations models that cannot hold simultaneously. However, for a good/valid interpretation of the results, in such a case, one should be quite sure that the model do not "strongly" contradict each other.

2.1.2 About risk of overfitting
--------------------------------

The authors do not discuss the risk of overfitting when using splines, especially when using time-dependent effects and trying to fit the "best model". However, as far as I understood their approach, this may easily happen (especially with a backward selection procedure). Besides, after having fitted many models and chosen the "best one", I am wondering whether a kind of multiple testing issue could affect the interpretation of final p-values and confidence intervals. I would be happy to have the point of view of the authors about this comment.

2.2 Minor Essential Revisions
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

2.2.1 About citation of previous works and discussion on model checking
-----------------------------------------------------------------------

As explained by the author, the proposed approach is of particular interest when the widely used Cox model is misspecified. Conversely, and as previously discussed, the approach proposed by the authors can also lead to misspecified models. Therefore, to my mind, some references to previous works and discussions about model checking when using pseudo-values and transformation models would be most than welcome (see section 4 of Andersen and Perme, SMMR 2010).

2.2.2 About defining and interpreting RD & NNT
Instead of defining and discussing the interpretation of the Risk Difference (RD) and the Number needed to treat (NNT) only in the section "Model-based estimates of association measures", it would have improved the readability and understanding of the paper if the authors would have done it in the introduction section. I especially think about the sentence "Positive and negative values of NNT represents the number of patients needed to be treated for one additional patient to benefit and to be harmed, respectively." that could have fitted better in the introduction section.

In addition, it would be nice to properly define the NNT before displaying an estimator of it at equation (1).

2.2.3 About references to other work and software

Scheike et al. (Lifetime data analysis 2006 & Biometrika, 2008), Cheng et al. (Biometrika, 2005) used Inverse Probability of Censoring Weighting approaches to fit similar transformation models. Besides, it seems to me that some of the approaches are also implemented in the timereg R package. Could the authors comment on that and, if relevant, refer to these related works on their manuscript?

2.2.4 Transformation model equation

As this is an essential point of the manuscript, the equation describing the transformation model, in the text of the "Model-based estimates of association measures" section, should be displayed as the one of the pseudo value (2).

2.2.5 Censoring in the Rotterdam Breast Cancer example

The author wrote "Times to death from other causes were treated as censored." I do not understand this choice. Without additional argumentation, I am not able to understand the point of considering a competing risk issue as a censoring issue. Besides, I believe that
such a confusion between censoring and competing risk issue may sometimes be misleading. Could the author at least give a short explanation for this choice?

2.3 Discretionary Revisions

2.3.1 About the figures

I think that 9 figures is a lot. Some of them are probably not essential for a good understanding of the paper and could probably be summarized by simple and short sentences. In order to improve reader convenience, I thus advice the authors to keep a smaller selection of the most useful figures. This could make the reader more focused on the most important "take home messages".

Besides, I am wondering about the easy reading and understanding of the three dimensional plots, especially if they need to be reduced for the final version of the paper.

2.3.2 About time dependent covariates

As I understood the last paragraph before the "Model-based estimates of association measures" section, the authors claim that the presence of time dependent covariates is generally not an issue. I think a note of caution contrasting what could indeed be done for external covariates whereas not for internal covariates could be of interest for some readers. Excellent discussions about this point are those by Fischer & Lin (Ann. Rev. Public Health, 1999) and by Kalbfleisch & Prentice (2002, Section 8.2.3).

2.3.3 Third example

I am wondering whether the third example, section "Primary Retroperitoneal Soft Tissue Sarcoma" is needed. The "Results" section gives a lot of details and is quite long. Maybe the two previous could be enough. Besides, I do not clearly understand the added value of
this one compared to the two previous. By the way, the introduction of the "Results" section is only about the two first example and only mentions "two examples".

2.4 Minor issues not for publication
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

2.4.1 References
----------------

A lot of "???” are in the references.

2.4.2 Citation style
---------------------

I suggest to replace some awkward citations like "[1] highlights how relative measures [...]" by "Schechtman [1] highlights how relative measures [...]". It would improve the readability of the paper.

**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.