Reviewer’s report

Title: Harnessing repeated measurements of predictor variables for clinical risk prediction: A review of existing methods

Version: 0 Date: 11 Mar 2020

Reviewer: Nicole Erler

Reviewer's report:

The manuscript "Harnessing repeated measurements of predictor variables for clinical risk prediction: A review of existing methods" is a useful addition to the literature. I also very much liked the precise definition of terms in the Methods section and section 3.2 (including Figure 2).

Below are some comments that I would like the authors to address:

# Comments
## Background
p.4, line 8/9:
The authors only refer to binary or time-to-event outcomes (throughout the manuscript). Why are continuous outcomes not included in this review?

## Methods
p.7 line 11:
The authors mention that they extracted the computer software used from the screened articles. Could the authors provide some information about how often the software used was reported at all, and what type of software is being used in the field of CPMs?
I realize that, for an applied researcher, searching for a solution to his/her prediction problem, this may not be of primary interest, but I think in the context of a general review of the methodology this type of information could have a place.
If there was a lack of reporting of the software used, I think this should also be mentioned on page 27, lines 14-19.

## Results
p.9 Figure 1:
In the "Screening" section, it says that "method[s]" with no application were excluded. Why?
The title of the manuscript states that this is "A review of existing methods". Should that not include purely statistical publications?
Why was "patient-reported outcome framework" an exclusion criterion?
For completeness, could the authors provide a list of the included articles in an additional supplementary file (for example, a .csv)? This is the data they worked with, after all.

p.10 line 17:
It is not entirely clear to me what A1 "to better specify the predictor-outcome" relationship" exactly means. I can see how "to better capture the ..." would be an aim (for example, when using a more flexible approach that does impose fewer restrictions and could, therefore, give a more accurate representation of the actual relationship). Is this what the authors mean?

p.11 line 1:
I think the definition of A2 should not use the "true value". I don't think it is possible to get the true value in most settings. I suggest using, for instance, to infer or predict (or estimate?) the value of a covariate at a pre-specified time.

p.12 line 1:
The statement that methods F4 to F7 require repeated measurements at the time of prediction is not entirely clear to me. (1) does this mean that a value measured at exactly that timepoint has to be available? Or does it mean that we can only do prediction for subjects for whom some (at least one?) of the repeated measurements is already available?
(2) I'm not sure if this is correct (at least in my interpretation of what is a requirement):
With a two-stage approach or JM, for instance, the value at the exact time of prediction could be estimated from the model imposed on repeated measurements that are available at other time points for the same patient.
If a patient has no values at all at the time of prediction, technically, it should still be possible to impute a value based on the distribution estimated by the fitted model conditional on other observed covariates. (Maybe not in all estimation frameworks, but the authors make a general statement here.)
I expect that in most cases in practice such an imputed value will not be very meaningful/reliable, and maybe the available software is not capable of handling this. However, since there is a way to handle cases with no information, in theory, I would not consider it a requirement of the method.
From my personal Bayesian missing data point of view, allowing for cases without any longitudinal information would even be a natural thing.

p.12 and following:
I think it would be helpful for readers if the authors would state within the first few sentences of each of the subsections about the different modeling frameworks explicitly if the approach can handle both types of outcomes or only survival/binary.

p.13 lines 4-6:
How does "those still at risk" go together with "those who have ...experienced an event at that specific time", since the latter are no longer under risk? Please re-phrase.
The authors write that applying a baseline CPM to follow-up data would lead to over-estimated risk predictions. If the risk (of having an event) is overestimated, this means that the survival probability is underestimated. The given reference, however, states the opposite: "...time-fixed models may tend to overestimate survival if they are applied to follow-up data."

It might be worth mentioning that the term "joint model" generally refers to any models that are estimated jointly. The term "joint model" is often used as an abbreviation of "joint models for longitudinal and survival data", and probably most joint models in the prediction context are of that type. However, there are also other joint models, such as, for example, a joint model for a binary outcome and repeatedly measured covariates. In my experience, when first encountering a new area of methodology, it can be difficult to distinguish whether a term is the name of one particular type of model or refers to a group of models. It might be useful for readers if this context is provided more clearly.

I find this section a bit confusing. It is my understanding that the issue is that, for a new patient, the random effects, which are used in the linear predictor of the survival sub-model, are unknown. To deal with this, random effects can be sampled from the posterior distribution that depends on the new patient's longitudinal measurements up until the time of prediction, the fact that the patient is still alive at this time, and the parameter estimates obtained from the fitted joined model. Using Monte Carlo simulation allows us to "integrate out" the unknown random effects and thereby take into account the added uncertainty in the estimate of the survival probability for the new patient. Is this what the authors mean? Maybe they can find a way to re-phrase.

It is not entirely clear to me how this approach works. I can follow the description of the model in lines 6-10. In the Bayesian framework, we could then obtain the posterior distribution for the unknown binary outcome for a new patient, given his/her repeatedly measured covariate information. Something similar should, of course, be possible using likelihood theory. I do not understand, however, what the authors mean in lines 10-12. Could the authors please clarify?

Moreover, is this latent class approach related to the joint latent class model detailed in the supplementary materials? It might be useful to include a mention of this connection (or the difference).
Since many of the points I raise here could be just matters of opinion, a detailed point-by-point reply to each of the points is not necessary here.

### Abstract
At first read, I was wondering if "prediction of future events" in the Methods paragraph referred to time-to-event outcomes only or also included binary "events". The authors could maybe clarify this here for those readers who are specifically looking for a solution to only one of the two settings.

### Background
p.3, line 10/11:
Maybe add "available", to make clear that it is any information that is available for a patient at the time of prediction, not necessarily information measured at that time.

p.4 line 11:
I assume "in this field" refers to "clinical prediction modeling", but on the first read, I was a bit confused about whether it referred to the "particular clinical application" mentioned in the previous line. I would find a more specific wording like "in the field of CPMs" helpful.

p.4 line 14:
The authors write "Our main objective" which makes me wonder what the other objectives were.

### Methods
p.5 line 1:
Maybe a rather philosophical question, but should it be "potentially predictive" instead of "predictive"?

### Results
p.9 Figure 1:
The second box in the "Screening" section states that "method[s] with no application" and "no clinical application" were excluded. Are these different things?

p.11 line 13:
Is the use of F1 to F7 necessary? Throughout most of the article, the acronyms of the approaches are used, so maybe this additional notation can be avoided.

p.14 line 2:
I would say that the issues of LOCF are "usually" (not probably) dependent on the window size but also on how stable the variable is over time.

p.17 line 18:
"The simplest ... methods ... are aggregated data and ...". I do not think aggregated data is a method. It would have to be something like "the use of aggregated data" or "to aggregate data".

p.18 line 22:
I suggest using "may or may not" instead of "might" since I assume the authors are listing the possibilities (and that wording would then seem more natural to me).

p.19 line 11-14:
I do not understand why linear models, quadratic growth curves, and cubic splines are grouped, but fractional polynomials are mentioned separately.
In my understanding, the issue is that the "default" assumption of a linear trend over time (= linear model?) often does not fit the data.
Alternatives to model the trajectories more flexible are quadratic growth curves or data-driven techniques such as splines or fractional polynomials.

p.20 line 7:
I would add the word "typically" (... sub-models typically involve shared ...) since "joint models" is such a general term, and they might be connected in other ways, for instance, by using one outcome in the linear predictor for another one.

## Discussion
p.27 line 9:
Why (only) epidemiologists? Is this not relevant to (applied) researchers in general?

## Appendix 2: Joint modeling variations
In the section on shared random effects joint models, the authors state that "Rizopoulos (2014) proposed a Bayesian Moving Average algorithm". Should this not be "... Bayesian Model Averaging algorithm"?

Typo: it should be "JMbayes" (small b)

In the section on joint frailty model for recurrent events, it should be the "frailtyPenal" function (capitalization).

**Level of interest**
Please indicate how interesting you found the manuscript:

An article of importance in its field

**Quality of written English**
Please indicate the quality of language in the manuscript:

Acceptable
Declaration of competing interests
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I declare that I have no competing interests

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal