Reviewer’s report

Title: The current application of the Royston-Parmar model for prognostic modelling in health research: a scoping review

Version: 0 Date: 31 Aug 2017

Reviewer: Kym Snell

Reviewer's report:

The authors present a 'scoping' review of flexible parametric models for prognostic modelling. This review should be of interest to readers and I'm not aware of any similar reviews having been published. However, I feel that the manuscript still needs considerable work before publication.

I felt the rationale for conducting the scoping review could be better described. The authors state 'the baseline hazard function is estimated smoothly and therefore these models are attractive for prognostic modelling'. Please discuss any potential benefits for prognostic modelling in more detail. Why are they attractive? Please also consider how prediction models are often presented and used, often for a particular time point e.g. mortality within 5 years. Would flexible parametric models still be of benefit?

Although I generally understood what the authors meant, I found the language throughout a bit loose. Please be more specific in the descriptions of statistical methods and concepts. Some examples are given from the background section, but much of the manuscript is written in this way:

* Line 31: "The flexible parametric survival model was developed in 2002". The paper the authors refer to was published in 2002, although the authors also acknowledge an earlier paper in 2001.

* Line 32: Authors state that the R-P model is defined by a RCS used to model the baseline log cumulative hazard function when proportional hazards are of interest. However, these models are broader than this, they can incorporate non-proportional hazards using time-dependent effects and can also model on the proportional odds scale etc.

* Line 68-71: "When regression modelling is used for prognostic models, survival analysis methods are employed because this class of statistical procedures account for the relationship between the predictor(s) and the outcome(s), as well as the time until the outcome(s) occur." This is rather long-winded and not strictly true as logistic regression is often used for outcomes occurring within a short time interval (e.g. if within a few hours, 30 day mortality etc.).
* Lines 79-82: "The baseline hazard function is not estimated in the Cox PH model because the regression parameters are estimated by maximizing a partial likelihood function, thus treating the baseline hazard function as a nuisance parameter." Other way around, baseline hazard function was seen as a nuisance parameter, therefore the partial likelihood function was maximised as a way of removing this "nuisance" parameter.

* Lines 82-84: "Because the baseline hazard function is not estimated, absolute measures of effects can only be predicted at the time points of observed events." Again, I don't think this is quite right. As you go on to discuss, you can use a non-parametric estimate of the baseline hazard function. Also, it is not that you can only predict at the time points of observed events, but rather that the probability is assumed to remain constant until the next event time.

* Line 85: "methods to estimate survival post hoc…” Please be more specific with 'survival' here, survival probabilities or survival function.

* Line 88: please be specific about the 'limitation' you are referring to here.

* Line 94: not 'smoothed', just estimating a function.

A definition of what is meant by 'prognostic model' for the review would be good. For some people, this includes models predicting future onset of disease in healthy individuals, while others might not consider this prognostic. Are you interested in models that predict outcome probabilities for individuals or additionally make inferences at a higher level e.g. net survival for groups of individuals compared to the general population? Models can have very different aims.

Also, do you only consider multivariable models or also include univariable models?

It might help to split the background into two sections. One with the background and rationale for the study and a separate section detailing the flexible parametric models and extensions. Depending on the definition of prognostic models mentioned in the above point, the extensions may/may not be relevant.

I am curious about the choice to exclude prognostic factor studies (associational studies as referred to in the paper) which may still be interested in absolute estimates of hazard or survival, while including studies investigating relative survival, net survival and cure models for which inferences are generally at a population level rather than at the individual level.

Why were studies looking at methodological developments in FP models excluded even if the methods were applied to real data, while studies providing an empirical comparison of methods were included (ref 27, 28)?
If more than one prognostic model was presented in an article, were details summarised for the article rather than for each model?

I am aware of at least one study that used flexible parametric modelling to develop and validate a prognostic model that would be relevant but is not included in this scoping review. See Ensor J, Riley RD, Jowett S, et al. Prediction of risk of recurrence of venous thromboembolism following treatment for a first unprovoked venous thromboembolism: systematic review, prognostic model and clinical decision rule, and economic evaluation. Health technology assessment. 2016; 20: 1-190. This raises concern as to how thorough the search strategy was or how well the inclusion/exclusion criteria were applied.

Lines 215-218: Why were searches in google scholar limited to the first 200 hits and other searches limited to the first 30 search results? Please provide reasoning for why this was thought to be appropriate.

Line 259: Do the authors mean 'abstraction' or rather 'extraction'? 

Line 310 (and earlier): Might be better to refer to d.f. for splines or be more specific about whether you are referring to internal knots or including boundary knots.

Page 13: Flexible parametric survival model specifications section. Number of studies given but not clear out of how many as not always out of 12.

Page 14: When defining calibration and discrimination statistics, please be clearer in your definitions. E.g. discrimination (yates) slopes - separation in what?

Line 352: Say that ref 36 provided full prediction from the prognostic model including the restricted cubic spline function. Having seen the article, they report coefficients for each of the spline terms, however this doesn't make it useable as these correspond to spline terms created in the dataset. Please discuss this more fully.

Line 385: Studies claimed that FP models improved model accuracy. It would be good to know if this was in terms of apparent, internal or external model performance.

Line 390-392: Authors suggest reporting comparisons of performance measures between flexible parametric models. How do you suggest this is done considering it is the external validity of a model that is of most interest?

Line 400: see earlier comment on reporting RCS coefficients. Still need spline terms created in data. This is one of the remaining difficulties with RP models for prediction. It is very difficult to report the full model in a way that predictions can be made over time. Perhaps this should be discussed.
Line 414: Authors mention that more emphasis should be placed on the ability to model time-dependent effects, however these are also difficult to report (as well as interpret). Please discuss this more fully.

Paragraph beginning on line 426, discusses reporting of number and position of knots. However, simply having this information does not make the model any more useable. See earlier comment.

**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.