Author’s response to reviews

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

Authors:

Ryan Ng (ry.ng@mail.utoronto.ca)
Kathy Kornas (kathy.kornas@utoronto.ca)
Rinku Sutradhar (rinku.sutradhar@ices.on.ca)
Walter Wodchis (walter.wodchis@utoronto.ca)
Laura Rosella (laura.rosella@utoronto.ca)

Version: 2 Date: 17 Jan 2018

Author’s response to reviews:

Manuscript ID: DAPR-D-17-00017R1

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

Dear Dr. Debray,

We would like to thank the reviewers for their thoughtful feedback throughout the review process, as it has been critical in further developing and improving our manuscript. We have provided a response to each of the additional comments from the reviewers, and have indicated corresponding changes in the manuscript using yellow highlighted text.

On behalf of my co-authors, I would like to thank you for considering our manuscript.

Sincerely,

Ryan Ng
Point-by-point response to reviewers:

Reviewer reports:

Reviewer #1: The authors have addressed most of my previous comments. I have only a few minor comments, mostly for clarification or further information:

Line 72: I find it rather confusing to say 'time-specific predictions for any time point'. Perhaps refer to it as a function over time.

  • Thanks. The statement has been re-written to describe predictions as a function time (pg 3, line 73).

Line 193: Not sure what is meant by 'reporting mathematically as differences in cubes via basis functions'. Can you provide a reference or further information here please.

  • Yes, this was unclear. The statement has been rewritten to reference an equation mentioned earlier in the paper (equation 4), which replaces the basis function (pg 8, line 193). All equations in the manuscript have been numbered.

Line 69 & 102: Please provide references for statements such as 'most commonly used method'. E.g. for survival analysis being the most common regression model - logistic regression also commonly used. Fine to say 'a common approach'.

  • Yes, that is a valid point. We have changed the wording to say that survival analysis is a common approach to prognostic modelling (pg 3, line 73) and the Weibull distribution is a common-used distribution for parametric survival models (pg 5, line 102).

Line 152: Please define what the basis functions are or say which part of the equation they refer to.

  • We agree that there was some ambiguity around basis functions. The statement has been rewritten to reference an equation mentioned earlier in the paper (equation 4) (pg 7, line 152).
Reviewer #3: The new version has much improved and only some less critical points remain.

1) Ref 41 has P Royston as a co-author. Just state whether any of the other 12 papers (not all have the full author list) has any of the developers (Lambert, Royston, Parmar, p 11, l252) as a co-author. Could be done in Table 1 but a sentence in the text could be sufficient.
   • Thanks for the suggestion. A statement about papers with developers as the co-author has been added (pg 14, line 322).

2) L 393. I am confused about this sentence. Estimates change when you allow non-proportionality. Please re-phrase.
   • We have rewritten the sentence to clarify that some studies included a second restricted cubic spline to model the interaction (pg 16, line 393).

3) Your 'reporting recommendations' concentrate on the spline part (baseline) whereas the regression part and other issues are completely ignored. I agree that other parts are well handled in the TRIPOD and REMARK guidelines (l 470, refs 3, 45) but you need to better 'combine' these general guidelines with your proposal to report the spline part. Some subheadings could help.
   • Thank you for the suggestion. We agree that subheadings are required. We have restructured the discussion to have headings describing ‘Model building considerations for the baseline cumulative hazard function’ and ‘Reporting considerations for the baseline cumulative hazard function’ (see pages 19-23). These two subheadings focus the discussion specifically to the baseline cumulative hazard function. We wholeheartedly agree that other reporting recommendations are important, but due to space considerations we decided to provide references in which the reader can go to for general model building strategies and reporting guidelines at the end of each subsection:
     o “Lastly, general model building strategies for prediction models (e.g. missing data, variable selection, validation) should also be considered when modeling with the Royston-Parmar model[47].” – Regression Modeling Strategies textbook by Frank Harrell (page 21, line 522)
     o “The suggestions provided for reporting the baseline cumulative hazard function here should be used in conjunction with general reporting guidelines (e.g., TRIPOD, REMARK) for prognostic models so that all aspects of model development and/or validation are transparent[3][48].” – TRIPOD and REMARK (page 22, line 542)
4) Your reporting part is also mixed with proposals for analysis. Is there any evidence for your proposals? Guidance would be most helpful but it needs to be based on empirical evidence. Please restructure the long discussion, consider subheading and whether deletion of some parts improves the paper.

- We agree with this suggestion, and have addressed this in the previous comment #3.

5) I welcome the extension of Table 2. However, reporting guidelines clearly express the importance of providing information about the number of variables investigated, not only the number of variables in the final model. Please add this information.

- Thank you for the suggestion. We have added an additional column in Table 2 indicating the number of candidate variables in the prediction model (page 32, Table 2).