Author's response to reviews

Title: Treatment Use in Prognostic Model Research: a Systematic Review of Cardiovascular Prognostic Studies

Authors:

Romin Pajouheshnia (r.pajouheshnia@umcutrecht.nl)
Johanna Damen (J.A.A.Damen@umcutrecht.nl)
Rolf Groenwold (R.H.H.Groenwold@umcutrecht.nl)
Karel Moons (K.G.M.Moons@umcutrecht.nl)
Linda Peelen (L.M.Peelen@umcutrecht.nl)

Version: 1 Date: 01 Sep 2017

Author’s response to reviews:

Dear Dr Snell,

We would like to thank you for your consideration and evaluation of our manuscript, “Treatment Use in Prognostic Model Research: a Systematic Review”. We would like to thank the reviewers for the comments and suggestions provided. The feedback was thorough and informative, and has provided the basis for necessary improvements to the original manuscript. We have addressed all comments and suggestions provided by the reviewers. Changes have been made using “track changes” as requested, and we submit a revised manuscript with revised figures (where necessary) alongside this letter.

We hope that the outlined adjustments to the manuscript are sufficient in clarifying your concerns, and we would again like to thank you for your suggested revisions.

With kind regards,

On behalf of all authors,

Romin Pajouheshnia MSc

Reviewer 1
1. Abstract: Research - The first two statements feel speculative given that they appear before the results of the systematic review. I would therefore suggest that the authors move these to the end of this section.

- As these two sentences are not results from the systematic review itself (instead, the first statement is background information and the second is the result of our own discussion/logical reasoning on this topic), we have decided to remove both sentences from the abstract.

2. Abstract: Conclusions - The statement "is often ignored" does not truly reflect the findings in my opinion. For example 70% of studies did mention treatment use. Therefore, the authors should modify the conclusions section of the abstract to more accurately reflect the results.

- We have amended the phrasing of the first sentence of the conclusion to better represent our findings:

“The use of treatments has been partly considered by the majority of CVD prognostic model studies. Detailed accounts including, for example, information on treatment drop-in were rare”. The term “treatment drop-in” is now mentioned here and earlier in the abstract, upon request from Reviewer 2.

3. Methods: 2.1, Line 66 - I believe there is a word missing between "as well" and "direct targeted".

- We have corrected this omission in the text:

“as well as direct targeted”.

4. Methods, 2.1.3, Line 139 - Again, I think that there is a word missing between "lower than it would" and "been, had they..."

- We have corrected this omission in the text:

“lower than it would have been, had they”.

5. Methods, 2.2.2 - Please make it clear whether these methods of data extraction relate to the original review or to this current systematic review. I believe they relate to the current review based on the authors involved in the extraction but it needs to be clarified within the text.

- We have clarified this in text (lines 240-241):

“…a list of key items (Additional file 2) for extraction was derived for the current review by one author (RP) and updated after group consideration…”. 
6. Table 1 - The footnote suggests that the percentages are not reported, however the final two rows of the table include percentages. Therefore, the authors should be explicit as to how they have estimated these percentages.

- We have clarified the meaning of the values in the table by amending the footnote to Table 1, as follows:

“† Values represent as follows: median, (lower quartile, upper quartile); percentage of studies that did not report this information.”

7. Results, 3.2.1-3.2.3 - The authors should define what they mean by each of these study types.

- The definitions of the three study types are now given in the methods section (lines 228-232):

“…reporting the development (derivation of a new model) or external validation (evaluation of an existing model in a new population) of a prognostic model, and “incremental value studies”, in which the additional value of a certain predictor or (bio)marker was assessed on top of either an existing risk score or a model consisting of a core set of conventional predictors (e.g. age, sex, smoking, systolic blood pressure, cholesterol, diabetes).”

In addition, we summarize these definitions in each respective section of the results (sections 3.2.1-3.2.3).

8. Results, 3.2.3 - What do the authors mean by “incident surgical procedures”?

- We have clarified this in text (lines 298-299): “…incident surgical procedures (cardiovascular surgeries occurring after the study baseline)…”.

9. Table 2 - It would be more mathematically pleasing if the subcategory percentages added up to 100. Therefore, it would be helpful to add a fourth category of "other" to assist with this.

- We suggest that Table 2 remain unchanged. The addition of an extra column to represent articles that reported multiple study types (e.g. development and validation) would result in a column with a very heterogeneous mix of articles, which could be difficult to interpret. In terms of the columns themselves adding to 100%, this should not be the case as individual studies could report multiple treatments.

10. Table 3 - What do the authors mean by the starred footnote? It is not clear to me.

- Articles that reported evaluating incremental value (IV) of a predictor can be divided into two categories: 1) those that assessed IV over a set of individual core predictors, requiring the selection and fitting of a model, only for this purpose, and 2) those that assessed IV over an
existing prognostic model (e.g. by using the linear intercept). As the latter kind of IV assessment does not involve the fitting of a model (or at least the selection of predictor variables), we decided it did not make sense to include these articles within the cell counts for the section of Table 3 “Treatment modelled as a predictor”. Thus these values in the IV column represent only those studies that assessed IV over a set of predictors that the authors of those studies selected themselves (and thus has the opportunity to include variables related to treatment use).

We clarify this in the footnote of Table 3:

“* Only studies that assessed incremental value over a core set of individual predictors (n = 81) and thus had the opportunity to include treatment variables within the core set of predictors; studies that assessed incremental value over an existing prognostic model or risk score did not derive a new prediction model and are not included in the calculation.”.

11. Discussion, page 16, first paragraph - The final two sentence (lines 299 and 300) are unclear to me. Do the authors mean people who are not on treatment? The clarity of this section should be improved.

- We have amended this section in order to provide greater clarity (lines 369-376):

“Finally, in some studies treatments may not have been considered by the authors to be relevant to the prognostic question being addressed. One article did not model treatment effects on the grounds that “The prediction of initial CHD [coronary heart disease] events in a free-living population not on medication is emphasized” (16) i.e. the model was designed for use in individuals who are not already on treatment. However, as already discussed, this rationale does not take into account treatment drop-in that may have occurred during the follow-up period of the study.”

12. Discussion, page 16, second paragraph - Within lines 302-303 the authors discuss this review as the first overview of how treatment information has been reported. However, the authors have already referenced a review (13) which considered external model validation studies. Therefore, the authors should clarify what makes their review unique. I presume it is because they have considered development studies as well as external validation ones?

- Related work has indeed been previously published. The review by Collins et al (#13) focussed on general methodological considerations in validation studies. The review discusses the relevance of “case-mix” differences between development and validation studies, which is related to the issue we highlight in our review article. However, we focus specifically on one aspect of “case-mix” and develop this issue to a much greater extent than the publication by Collins et al, or a previously published review of CVD prediction models by Liew et el (#6). Furthermore, this article is the first to consider the issues caused by ignoring treatment use across development, validation and incremental value studies, building upon and providing a much more complete account of the CVD field than the review by Liew et al (#6).
As suggested, we provide greater explanation of the novelty of the review as follows (lines 379-383):

“While other studies have broadly addressed related methodological issues (13), or have focussed on a single aspect of CVD modelling, such as model development (6), we provide comprehensive coverage of CVD prediction model studies and support this with a conceptual framework describing when and how treatments can affect a prognostic study.”

13. Discussion, page 16, final line - Unjustified statements suggesting which clinical field might be most advanced should be avoided in my opinion.

- We have amended this statement, as recommended:

“As the CVD domain is a highly active field in prognostic model research…”

We suggest that this statement is more justifiable than the previous statement.

14. Figure 4 - The authors should talk about the Figure rather than just including it. This would help to clarify each point within it. For example, what is "sufficient information of treatment use"? Why is inverse probability weighting the most appropriate recommendation for guided treatments?

- We now provide a paragraph of text to further explain Figure 4 (lines 427-441). This should make the recommendations clearer.

15. Conclusion - I think that the authors should be more specific. As per the abstract, 70% of studies did mention treatment use so statements suggesting that treatment use within prognostic models has yet to be widely recognised are not strictly true. Therefore, the authors should be specific about the current lack of appropriate reporting and conscious inclusion or exclusion of treatment within models.

- We agree that the conclusion would benefit from being more specific. We have amended it as follows (lines 445-453):

“…Our review shows that while the importance of treatments for prognostic prediction has been recognized in many studies, reporting rarely covers all relevant treatments, and changes in treatment have hardly been acknowledged. Furthermore, we found no clear consensus within the published literature over how treatments should be considered in the analyses of prognostic studies…”

16. Figure 3 - This would benefit from an improved legend. Reported and not reported relate to treatment I presume, but it would be helpful to more carefully define the two groups.
- Figure 3 now has a more complete legend. In addition we provide further clarification of the interpretation of this figure in the legend, as follows:

“Reporting of treatment in CVD prognostic modelling studies over time. Articles were classified as having reported information on treatment if the use of at least one potentially risk-lowering treatment in the study was reported, or if the effect of a treatment on the study findings was discussed…”.

Reviewer 2

1. As the authors themselves acknowledge, the findings are based on a systematic review that was conducted in 2013 - and is hence now 4 years out of date. I do find this a little concerning, especially given - as the authors report in Section 3.2 - that the handling of treatment has improved over time. This raises the possibility that reporting of treatment is 2017 is now far better. I am not suggesting that the authors repeat the systematic review but they should consider checking recent papers - perhaps in the Discussion section where they raise this limitation (QRISK3 being an obvious example).

- We agree that a more elaborate discussion of key models/studies from 2013-2017 will add greatly to the article. Following the recommendation, we comment on QRISK-3, along with two other important developments in CVD prediction: the ACC/AHA Pooled Cohort equations and the Globorisk tool. This discussion can be found in lines (387-395):

“Three important developments in the past four years include the ACC/AHA Pooled Cohort equations (16), the Globorisk CVD assessment tool (19) and the Qrisk-3 calculator (20), each developed as tools for the prediction of CVD in the general population. Among these three currently implemented CVD risk estimators, there is no clear consensus over how treatments should be taken into account in prognostic models for CVD: treatment use at baseline is modelled differently in each of the prognostic models, and none of the studies accounted for the effects of treatment drop-in. Thus, questions have been raised regarding the validity of these models and their respective validation studies (9, 21), and treatment use remains an issue at present.”

2. The title needs to be clear that the review is restricted to CVD.

- We have adapted the title as follows:

“Treatment Use in Prognostic Model Research: a Systematic Review of Cardiovascular Prognostic Studies”.

3. Example 1 (line 108) feels out of place, given the CVD focus of the rest of the paper.
We understand the reviewer’s view on the first example we give in section 2.1.3. While there is a distinct CVD theme in the article, our primary message is with respect to the impact of treatment use in prognostic model research in general. Thus we would ideally prefer to keep an example from outside of the CVD research domain, to support the idea that this is a widespread phenomenon.

In order to slightly improve the flow of the text, we now include an additional clause when introducing the scenarios:

“We illustrate the distinction between different types of treatment with two hypothetical examples, from two different clinical domains.”

If a second example from the cardiovascular domain is still greatly preferred, we can provide one, or elaborate on example 2.

4. Readers may find the paragraph starting on line 33 a little unclear. However, the paragraph starting on line 75 is much clearer about the issue the authors wish to raise, about the difficulty of estimating, and the consequences of it being hard to estimate, treatment naïve risk when using prognostic models to guide treatment initiation. I ask the authors to consider (at their discretion) restructuring such that a paragraph more like the one on 175 features in the Introduction.

- We agree with this suggestion, and have restructured the introduction, such that paragraphs two and three are combined and follow the structure of paragraph two of the Methods section more closely.

5. Treatment initiation after baseline ascertainment but before outcome has been termed 'treatment drop-in' in the literature (e.g. Liew S, Doust J, Glasziou P. Cardiovascular risk scores do not account for the effect of treatment: a review. Heart 2011.). Can the authors use this term? Also note that the guidelines given by the author in Fig 4 do not offer any guidance on handling treatment drop-in.

- We agree that the term “treatment drop-in” is a useful term and now use this term throughout the article (for example, line 144, figures 1, 4, and in the abstract).

At this stage we hesitate to make specific recommendations for handling time-varying treatment use as this is a topic of ongoing research, though the methods we recommend may help to address more simplistic scenarios of treatment drop-in, as supported by simulations in references (8 and 10). For more complex scenarios, existing methods need further evaluation before recommendations can be given. Instead, we emphasize a need for better reporting of information on treatment drop-in, to enhance the interpretation of the findings of prognostic studies.

6. Please check the abbreviations list: I didn't see PCE anywhere in the paper.
- We have now corrected the abbreviations list and have removed PCE accordingly.