Author’s response to reviews

Title: Reporting performance of prognostic models in cancer: a review

Authors:

Susan Mallett (susan.mallett@cs.m.ox.ac.uk)
Patrick Royston (pr@ctu.mrc.ac.uk)
Rachel Waters (Rachel.Waters@icr.ac.uk)
Susan Dutton (Susan.Dutton@cs.m.ox.ac.uk)
Douglas Altman (doug.altman@cs.m.ox.ac.uk)

Version: 2 Date: 24 February 2010

Author's response to reviews: see over
Response to Editor and reviewers' comments sent in email on 17th Dec 2009

Two companion articles: MS: 1301625644319197
- Reporting methods in studies developing prognostic models in cancer: a review.
  Susan Mallett, Patrick Royston, Susan Dutton, Rachel Waters and Douglas Altman
- Reporting performance of prognostic models in cancer: a review
  Susan Mallett, Patrick Royston, Rachel Waters, Susan Dutton and Douglas Altman

18th Jan 2010

Dear Editor,

We appreciate the time and attention to detail of these reviews, by both yourself and the reviewers. We appreciate the reviewers' comments to improve our manuscript and have taken these on board wherever possible.

We have address the Editor and reviewers' comments below. We have included our responses to reviewer's comments in blue.

Thank you for careful and thoughtful reading of our manuscript and helping us to improve it.

Best regards,

Susan Mallett
Corresponding author
Response to Editor’s comments

Please note that one of the referees raised the possibility that the two manuscripts could be merged. While we are happy to keep the two as individual sister manuscripts and would not insist on them being merged, we would urge you to shorten the methodological details in one of the manuscripts (given that the search strategy and data set is the same) and instead cite and refer to the related manuscript for such details, in the interest of clarify and brevity. The same applies with other sections where significant overlap may occur between the two manuscripts. We have shortened the second article by replacing the duplicated methods and results sections from the second article with a citation to the first article. Thank you.

Response to Reviewers’ comments

1. REVIEWER: PANAYIOTIS KYZAS

Article 1: Reporting methods in studies developing prognostic models in cancer: a review

Major Compulsory Revisions
1. Abstract, Conclusion: “Many published prognostic models have been developed using poor methodological choices that…” The message from this sentence is clear from the first sentence of the conclusion section; it doesn’t have to be repeated.

This was deleted and combined to the sentence above
"The use of poor methods compromises the reliability of prognostic models developed to provide objective probability estimates to complement clinical intuition of the physician and guidelines."

2. Methods/Literature Search: “…planned hand search of 10 high impact cancer journals in 2005…” How was the journal population selected? I assume this was based on the impact factor (i.e. the 10 with the highest IF). Is that the case? If yes, which year’s IF was used? Please clarify.

These 10 journals were chosen based on the higher impact factors (range impact factor in 2005 3.7 to 15.2) of Web of Science "oncology" section journals. These were the highest impact factors for journals found to publish higher numbers of prognostic articles. Actually these journals did not have sufficient numbers of articles, so we developed a search string to search all journals.

3. Methods/Inclusion Criteria: “…published in specialist cancer journals” (This comment applies to both articles). In theory, well designed and better quality articles may get published in major general journals with wider audience (i.e. NEJM, JAMA, Lancet, e.t.c.) instead of specialty journals. There is no easy way to assess this issue – I suggest selecting randomly one of these journals and using the search string trying to identify studies on cancer prognostic models. The identified articles (if any) should be compared with the 47 eligible articles.

Thank you for this comment. Actually post submission of this article to BMC we found that our search string was not restricted to specialist oncology journals as we had thought. This is
because the search term "neoplasms[subject term]" restricts the search to specialist journals only in the PubMed journal listing. Unfortunately in the main PubMed database, this search term does not work in the same way, but generated an error and by default completed an all text search by "neoplasm". This had the result that it restricted the search to articles using the MeSH term "neoplasm", a disease indexing term for cancer.

Inclusion of this search term reduced the number of articles identified by our search string from 2976 to 2076. As the search term did not work as intended, we took the precaution of screening all abstracts of the 900 articles excluded and found that none were eligible for inclusion. This is not surprising as our review was of prognostic models in cancer, and so articles should have been indexed under "neoplasms". The restriction to the MeSH term "neoplasms" was beneficial to the search and we have altered the search string to be explicit and more robust, by searching for "neoplasms[MeSH Terms]" in the appendix and the text in the methods and results sections.

We also attempted the search of several general medical journals as suggested by the reviewer, but found JAMA, NEJM and Lancet had no eligible articles. We also note that the review was not after all restricted to specialist oncology journals as we had previously stated.

4. Methods: Please see my comment on the next article regarding data extraction duplication.

Thank you. We've removed the duplicated sections in the methods and results from the second article, and referenced the first paper.

5. Figure 1 is the same in both articles. It doesn't have to be present in both. Also, at the end of the figure there is a statement about “duplicate data extraction on all studies”. Is this correct? Please clarify.

We've deleted Figure 1 from the second article. We've removed the "duplicate data extraction" - you are correct, this was not done for all data items.

Minor Essential Revisions

Thank you to the reviewer for these details that help to improve our article.

1. Methods/Page 5: Please correct the typo: “abstracts,16 full” to “abstracts, 16 full”
   Done

2. Results/Page 8: Please correct “metastatis” to “metastasis”
   Done

3. Page 16: Please replace “a couple” with “two”
   Done

4. Please add n=47 in Table 2.
   Done

5. In the Discussion section, the authors often refer to numbers and percentages from the results. Although this is obviously done to emphasize the findings of the study, I believe it would be better to be kept to a minimum.

Thank you. We've reduced this.
6. Discussion/Page 18: Please replace “has highlighted” with “highlights”.
  Done

Discretionary Revisions
1. Introduction: “…by a review”. The current paper is a systematic review; and this would be important to emphasize.
  Done. Also used phrase "systematic review" in second article p4 " We assessed, by a systematic review, the methods used in 47 articles on prognostic models"

2. Is there a reason to split the studies according to cancer site? (Results, page 8). Have the authors performed any analyses depending on the site? I believe they haven’t – and I don’t think is necessary. One may argue about the way they have splitted the studies (i.e. pancreas with UGI, bone mets in a separate category); I therefore suggest to omit this categorization.
  The listing by cancer site was reported to provide a description of the cancer sites in our sample for the reader.

3. Please see my 1st comment on the Discretionary Revisions section for the next article. The current article reads well they way Results are presented; I wouldn’t, therefore, suggest changing it.
  Thank you

4. Results/Page 12: “…unclear whether all deaths or only cancer deaths were included in overall survival” If a study doesn’t clarify which deaths were included in overall survival, would be unreasonable to conclude that all deaths have been included? Does this make it unclear? Please explain.
  We feel it is unreasonable to conclude all deaths are included in overall survival, as in this study there were at least 10 studies where overall survival included cancer deaths only (Table 2). This has also been found in other studies.
  In our definition of a clearly defined outcome we have included 28 studies (Table 2). These are 17 studies using OS (7 explicitly death any cause and 10 cancer death only) and 11 studies using DFS (5 DFS including death and 6 DFS not including death). The 14 studies from OS where the type of death is unclear were counted as unclear, plus the 10 DFS studies where it was not clear if death was included.
  We have clarified this in the article on p 12
  " In 40% of articles the outcomes were not clearly defined; in 30% (14) of studies it was unclear whether all deaths or only cancer deaths were included in overall survival and in 10% (5) whether death was included in DFS outcomes."

Article 2: Reporting performance of prognostic models in cancer: a review

Major Compulsory Revisions
1. Abstract, Background/Methods: Although the aim of the study is very clear in the full text, this is not the case in the Abstract. I would suggest the authors to rephrase so that the aim of the current study becomes clearer.
  Thank you. We've moved the aims from the Methods section of the abstract to the Background section, so it's clearer.

2. Methods: Both articles share the same methodology for the Literature search
and the Inclusion criteria. I strongly recommend to the authors to describe in detail these aspects in one of the articles (i.e. Reporting methods in studies developing prognostic models in cancer: a review) and then refer to this on the companion paper. I appreciate that an attempt was made to keep the Methods section short in this article, but there are many identical sentences in both articles that don’t read well when someone assess both articles together. The same should stand for the first paragraph for “Validity assessment and Data Extraction”, up to “current literature”.

Thank you. We’ve removed the duplicated sections in the methods and results from the second article, and referenced the first paper.

3. Page 7, “Validity assessment and Data Extraction”: Is there a specific reason that data extraction was not duplicated for all items? Please state.

Thank you, I’ve added our reason for this to the text. "For 8 items examination of disagreements led to refinement of data items and re-assessment by a single reviewer (SM) due to study resource and timeline limitations."

The main items were extracted in duplicate, although the detail was often extracted by a single reviewer. For example in Table 2, the number of models where components of the final model were used to create the PI was assessed in duplicate (as all, some, or none of final model). This then sparked the question of what was going on if the final model was not being used, or only some of it was used. A single reviewer (SM) then separated this into categories Components of final model used to create PI

- Same variables and coefficients
- Same variables but not same coefficients
- Neither same variables nor coefficients
- Method unclear

4. Page 7, Results: “A detailed description of the characteristics of these studies is reported in the accompanying paper” Again, a reference should be enough – the description that follows is not very brief and I fell is not necessary.

Thank you - we’ve taken these out and referenced the first paper.

Minor Essential Revisions
1. Abstract/Results: Please delete “)“ after studies.
   Done

Discretionary Revisions
1. In the Results section, the items examined (Reporting of model, Reproducibility of model development, Development of prognostic index, Development of risk groups, Discrimination and calibration, Usability of model, Model validation) are presented with an introduction paragraph and the methodology used by the authors to assess these items, followed by the results. Although this is acceptable in a systematic review, I would prefer to see a paragraph in the Methods section incorporating the background and the methodology to assess these items, and only the results in the Results section. I leave this change to the authors’ discretion, because the Results section reads quite easily. The authors should also be congratulated for the quality of the tables
and the appropriate references to these, which enhance the understanding of the findings.
Thank you. We've left this, but appreciate the comment.

2. Page 17, Discussion: Please replace “This research” with “This study”.
   Done

3. Page 18, Discussion: “…consensus that some methods are not advisable”. I agree with the authors, but this statement should be properly referenced.
   Thank you. We have included a couple of references.

4. I don’t feel that adding titles (i.e. Reporting methods to develop risk groups, Lack of reporting of model performance, e.t.c) in the Discussion section aids to the understanding of the manuscript. The paragraphs are clearly separated and the message that the authors want to pass is very sound. I would suggest deleting these titles from this section.
   Thank you. We are happy to work with editor preference on this.

5. Discussion, page 20: “external data…only 11% studies”. This is a very important finding of this study. It is mentioned in the Abstract, but not with the percentage. It would be better if this % appears in the Abstract as well.
   We have combined this with a comment from the other reviewer and included the percentage of models where an external validation was reported either in the original paper or in a subsequent publication.

Extra text in abstract
" Including reports of validation with external data from publications up to four years following model development, external validation was attempted for only 21% (10) of models."

6. I would expect a paragraph listing the limitations of the current study, at the end of the Discussion section.
   We have added a phrase to signpost our paragraph (p17 in discussion) on limitations of the model. We have also added to this.
   "Though this research relates to prognostic models in cancer... A further limitation of our study was that only 47 articles were reviewed, however we judged little further value would be obtained from review of a larger number of articles."

Level of interest: An exceptional article
Quality of written English: Acceptable
Statistical review: No, the manuscript does not need to be seen by a statistician.
Declaration of competing interests: 'I declare that I have no competing interests'

2. REVIEWER: LUCA CINDOLO

Reviewer's report
Title: Reporting performance of prognostic models in cancer: a review
Version: 1 Date: 27 November 2009
Reviewer: luca cindolo
Reviewer's report:
Discretionary Revisions
- how many studies have been validated on external cohorts after the initial publication?

Thank you for this comment. We have searched for reports of external validation of these models in subsequent publications using a citation search.

Extra text in abstract
"Including reports of validation with external data from publications up to four years following model development, external validation was attempted for only 21% (10) of models."

Extra text in methods
To identify articles, published subsequent to the original articles on model development, that might include external validations of the 47 prognostic models, we completed a citation search on each of the 47 original articles (30th Dec 2009). We used the Web of Science® citation search (Citation databases: Science Citation Index expanded, Social Sciences Citation Index, Arts & Humanities Citation Index and Conference Proceedings Citation Index - Science 1990 onwards) through the ISI Web of KnowledgeSM (Thomson Reuters 2009). Titles and abstracts for the cited references were screened and where appropriate full articles (SM).

Extra text in results
In addition to the five articles in our review that included external validations, we also searched for subsequent publications that included external validations for the 47 prognostic models, using a citation search in December 2009. For eight prognostic models [17,19,28,33-35,45,46] subsequent articles have been published that used external patient data and reported completion of a model validation [78-91]. For three models, a model evaluation was reported in one subsequent article per model [28,34,35] whereas in five models, evaluation was reported by more than one article [17,19,33,45,46]. The same authors as had developed the prognostic model had published reports of evaluation for two models [28,35], whereas different authors reported evaluation for six models [17,19,33,34,45,46]. Overall, in the same or subsequent publications, 21% (10 of the 47) of models were reported as evaluated using external datasets, although the quality of evaluations was often poor and uninformative.

Extra text in discussion
Reported external evaluation of eight models was found in subsequent publications. In total only 21% (10 models) were reported as externally validated in either the original articles or in the subsequent four years.

- I know that every prognostic model (nomogram, score, tables, etc) is not self explanatory. A short legenda as "instruction for the use" is desirable. Probably the authors could add in the analysis how many articles contain this hel for the readers and finally for the patients.

Thank you for this suggestion.
I have added into Table 3 in the section on model usability from the article

- 6 articles (Mariani 401, Mariani 221, Stephenson, Sorbellini, Hoang, Kim) gave instructions likely to be suitable for physicians, either as a specific section or a clearly explained example of how to calculate a patient's prognostic score and link it to survival. Three of these articles included the same author (Kattan in Mariani 221, Stephenson,
4 of these articles used nomograms to calculate prognostic scores (Mariani 402, Stephenson, Sorbellini, Kim)
- 2 articles also supplied example text for the physicians to explain to patients the interpretation of their scores (Stephenson, Sorbellini)

We've added the text in the results

"However, instructions likely to be suitable for physicians on how to use the prognostic model, either as specific instructions or a worked example, were included in only six articles (refs listed). In two articles (refs listed) example text for the physicians to explain to patients the interpretation of their scores, was also included."

Addition to table 3

<table>
<thead>
<tr>
<th></th>
<th>Articles with risk groups (n=36)</th>
<th>Articles with PI but no risk groups (n=6***)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model usability from article$$</td>
<td>Prognostic score or risk group can be assigned</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>Survival presented for risk group and/or prognostic score</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>Instructions for use suitable for physicians included</td>
<td>3</td>
</tr>
</tbody>
</table>

**Level of interest:** An article of importance in its field

**Quality of written English:** Acceptable

**Statistical review:** No, the manuscript does not need to be seen by a statistician.

**Declaration of competing interests:**
'I declare that I have no competing interests'