Reviewer’s report

Title: Is pathology necessary to predict mortality among men with prostate-cancer?

Version: 2 Date: 20 June 2014

Reviewer: Sarah Lewis

Reviewer’s report:

Major Compulsory Revisions

The author must respond to these before a decision on publication can be reached. For example, additional necessary experiments or controls, statistical mistakes, errors in interpretation.

1. Aside from the smaller (largely grammatical) revisions below, there needs to be some better metric/evidence/argument as to why, you state in your Conclusion (Lines 283-285) ‘The current study demonstrates that the additional cost associated with the use of detailed clinical data may not be justified for an outcome of all-cause mortality, but is vital to study prostate-cancer specific mortality.’ This statement builds off of part of the ‘Results and Discussion’ section (Lines 214-223) that states that a 14.8% change in risk category is ‘modest’ while you seemed to consider 28% a more substantial difference. What is your evidence/thinking behind this conclusion (aside from intuition)? Is there a meaningful difference between these numbers aside from one being larger than another?

Additionally, why, as you state in your Conclusion, is studying PC-specific mortality ‘vital’ and worthy of the associated costs but for all-cause mortality it ‘may not be justified’? Much stronger evidence/rationale is needed both for the magnitude of the differences you found in risk-change between all-cause and prostate-specific mortality as well as your cost threshold/justification.

Without further justification/clarification of your Results, Discussion, and Conclusion in this area, I cannot yet determine whether to ‘Accept’ or ‘Reject’ this work. If it can be sufficiently clarified, I would recommend this paper be Accepted as the issue of using purely administrative vs. clinical data sources in cancer research is a timely and increasingly important one facing many researchers, and evidence as to the pros and cons of using each type data source would be of interest to many researchers.

- Minor Essential Revisions

The author can be trusted to make these. For example, missing labels on figures, the wrong use of a term, spelling mistakes.

Line 71: Change ‘outcome’ to ‘outcomes’
Lines 73-74: Take out ‘traditionally’ OR keep and change include to past tense ‘included’

Line 81: change to ‘subject’s’

84: add ‘data’ after pathology

84-86: Revise to: ‘Capturing information from pathology data is labor intensive and expensive. Therefore, if the addition of these pathology clinical variables to a predictive model with variables attained solely from administrative data does not enhance model performance, their inclusion should be avoided.’

Methods Section: (Lines 115-125) A CONSORT Diagram or including number of patients excluded from analysis by each step of inclusion criteria (ie Line 117 – how many total patients identified over age 66? Lines 119-120 – how many patients excluded for PC prior to diabetes? Etc…do this for each exclusion criteria to see how much/percentage of total population sample included/excluded through all of your exclusion criteria)

136-137: What does ‘documented and considered for the analysis mean’? Included in the analysis? Be more specific of the impact of this statement.

137: remove ‘-’ after biopsy

156-157: Add semicolon: We chose these categories as we believe they represent clinically meaningful categories;’ similar categories were used by others to asses all-cause mortality among prostate cancer patients.

176: replace ‘Grade’ with ‘Gleason grade’ to remain consistent throughout paper – check to ensure you are always using the same terms

189: add ‘patients’ after ‘(14.5%)’

190: replace ‘of’ with ‘patients’

195: replace ‘Grade’ with ‘Gleason grade’ to remain consistent throughout paper

229: Delete ‘,’ after ‘Although’

233: Change to $21,000. Explain further how you came up with this dollar amount (ie Add in how much you are assuming each extraction costs)

236: put reference in superscript to match the others

241: put references in superscript to match the others

249: change maybe to ‘may be’

250: put references in superscript to match the others

259: Remove ‘,’ after Although
259-260: This is not a convincing argument: ‘since we aim to assess the utility of adding clinical variables...’ as to why the intermediate category is ‘so large’ by your own admission – explain more clearly why this is the case – why wouldn’t it matter in this case that the intermediate category is so large?

266-267: add an ‘and’ between ‘diagnosis’ and ‘most’

267: comorbidities is one word

268: add comma after diabetes, no apostrophe after etc.

270: comorbidities is one word

273: change ‘co-variants’ to ‘covariates’

274: ‘in order’

- Discretionary Revisions

These are recommendations for improvement which the author can choose to ignore. For example clarifications, data that would be useful but not essential.

Line 80: contradicts Line 72 of ‘Most cancers are fast growing’ – would recommend removing this portion of Line 72 and amending first few Lines of 71-74: ‘Administrative databases are often used to create models to predict clinical outcomes, including survival. Models using detailed oncologic information to predict survival among patients with cancer are important due to cancer’s often significant impact on overall survival.’

Line 82: remove ‘a cohort’

Level of interest: An article of importance in its field

Quality of written English: Needs some language corrections before being published

Statistical review: Yes, but I do not feel adequately qualified to assess the statistics.

Declaration of competing interests:
‘I declare that I have no competing interests.’