Author's response to reviews

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

Authors:

David Margel (sdmargel@gmail.com)
David Urbach (david.urbach@uhn.ca)
Lorraine L Lipscombe (lorraine.lipscombe@wchospital.ca)
Chaim M Bell (cbell@mtsain.on.ca)
Girish Kulkarni (Girish.Kulkarni@uhn.ca)
Jack Baniel (jbanien@clalit.org.il)
Neil Fleshner (neil.fleshner@uhn.ca)
Peter C Austin (peter.austin@ices.on.ca)

Version: 3
Date: 21 August 2014

Author's response to reviews: see over
To: Arlene Pura  
Journal Editorial Office  
BioMed Central

Dear Mrs. Pura:

RE: Manuscript Number 8531300971300935, "Is pathology necessary to predict mortality among men with prostate-cancer?"

Thank you for allowing us the opportunity to revise the above mentioned manuscript. We have found the reviewers comments very helpful and consequently made the necessary changes to our manuscript. We believe the manuscript has improved significantly.

Below, please find a detailed point-by-point reply to the Reviewers’ comments. We hope you find the revised version suitable for publication in the BMC Urology.

We look forward to hearing from you soon.

Sincerely,
David Margel
E mail: sdmargel@gmail.com
Reviewer 1

1. Comment: more detail needed on variables included in administrative databases.

Response: We thank the reviewer and now added more detail on the variables included in our models.

Changes made: Methods Page 6 lines 10-14.

2. Comment: While the study is very well conducted and the appropriate metrics are applied (discrimination and calibration). HOWEVER there is a more efficient way to derive a variable from the administrative database...one could abstract a random sample of medical records and use that abstraction to infer pathology information by developing a claims only algorithm. So pathology data could be defined from the information derived in the claims (ie via an algorithm validated against the medical record). I'd like the authors to discuss this option (see Seeger et al, pharmacoepidemiology and drug safety 2006; 15: 784–792) and have explain why or why not it is applicable here.

Response: The reviewers point is well taken indeed in many cases an algorithm from medical claims can be used to efficiently abstract data instead of chart review. However this rarely applies to pathology. There is no unified clinical pathway that may identify Grade or volume of tumor. For example two patients with a similar Grade and stage may receive different treatments, and vice versa patients with different pathology may receive similar treatment. We have added this point to the discussion.


Reviewer 2

1. Comment: 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.

Response: We thank the reviewer for this very important point. As far as we know there is no valid mathematical or statistical method to compare changes in two different NRI analyses. We contacted Michael Pencina (who is the person who developed the NRI method) regarding this issue. He has replied that a recent paper (in press in Statist Med) now shows that NRI for events is approximately equal to change in sensitivity at the higher threshold plus change in sensitivity at the lower threshold. Similarly, non-event NRI is a sum of specificities. With this, one can ask how much change in sensitivity/specificity is sufficient to call something meaningful. This therefore depends on the field and research question. If for instance 5% change might be considered reasonable in the field of prostate cancer, I’d say 10% in sensitivity and 10% for specificity for a total of 20% may be a threshold to consider. We have softened our conclusions and addressed this important point in-depth in the discussion.


2. Comment: 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.

Response: All minor revisions have been made as per reviewer's suggestions.

Changes made:
Page 3 lines 2-9: Changed 'Administrative databases are often used to create models to predict clinical outcomes, in particular survival. Most cancers are fast growing, and once diagnosed have an enormous impact on survival. Therefore, commonly, models to predict survival among these subjects include detailed oncologic information.' However, earlier cancer diagnoses and advances in treatment have been associated with reduced cancer mortality, such that in 2003 there were an estimated 10 million cancer survivors in the United States. Consequently, patients are living longer after a diagnosis of cancer to the point where existing comorbidities may have a substantial impact on their overall survival.' to 'Administrative databases are often used to create models to predict clinical outcomes, in particular survival. Due to the enormous impact of cancer on overall survival, predictive models for survival among cancer patients frequently include detailed oncologic information.' However, earlier cancer diagnoses and advances in treatment have been associated with reduced cancer mortality, such that in 2003 there were an estimated 10 million cancer survivors in the United States. Consequently, patients are living longer after a diagnosis of cancer to the point
where existing comorbidities may have a substantial impact on their overall survival.'

Page 3 line 12: changed 'subjects' to ‘subject’s’

Page 3 line 13: Removed ‘a cohort of’

Page 3 line 15: added ‘data’ after pathology

Page 3 lines 15-17: Replaced "Capturing information from pathology data is labor intensive and expensive. Therefore, if the addition of these clinical variables to a predictive model, with variables attained solely from administrative data, does not enhance its performance it should be avoided." by: ‘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.’

Page 4 line 23- page 5 line 6: included number of patients excluded from analysis by each step of inclusion criteria

Page 5 line 21: Replaced ‘documented and considered for the analysis' by ' documented and included in the analysis'

Page 5 line 22: Removed ‘.-’ after biopsy

Page 6 line 21: Replaced comma by semi-colon

Page 8 line 23: added 'subjects' after ‘(14.5%)’

Page 8 line 23: replaced ‘of’ with ‘patients’

Page 11 line 5: Added "(that costs approximately $25 an hour)"

Page 11 line 6: Changed 21,000$ to $21,000.

Page 11 line 21: Changed maybe to ‘may be’

Page 12 line 6: Deleted ‘,’ after ‘Although’

Page 12 line 6-10: Added additional explanation for the large range of the intermediate category- "Although the intermediate risk category range is rather large (10-50% risk), these patients are often grouped together and treatment decisions are made at the extremes. Since we aim to assess the utility of adding clinical variables to a prediction model, we believe that these categories are sufficient."
Page 12 line 14: Replaced comma by 'and' between ‘diagnosis’ and ‘most’

Page 12 line 15: Added comma after diabetes, and removed apostrophe after etc.

Page 12 line 20: Replaced ‘co-variants’ by ‘covariates’

Page 12 lines 20-21: Replaced 'in-order' by 'in order'

Throughout: replaced ‘Grade’ with ‘Gleason grade’

Throughout: Changed references character to superscript

Throughout: Replaced 'co-morbidities' and 'co morbidity' by 'comorbidities'