Author’s response to reviews

Title: Predicting 7-day, 30-day and 60-day all-cause unplanned readmission: A case study of a Sydney hospital

Authors:

Yashar Maali (yashar.maali@mq.edu.au)
Oscar Perez-Concha (oscar.perez.concha@gmail.com)
Enrico Coiera (enrico.coiera@mq.edu.au)
David Roffe (David.Roffe@svha.org.au)
Richard Day (r.day@unsw.edu.au)
Blanca Gallego (blanca.gallegoluxan@mq.edu.au)

Version: 1 Date: 08 Sep 2017

Author’s response to reviews:

Dear Editor and Reviewers,

Thank you very much for your detailed comments and suggestions. We would like to extend our appreciation for the time devoted to this matter.

Please find our point-by-point response to the reviewers’ comments in the lines below.

Kind regards,

Oscar Perez-Concha on behalf of all the authors.
Editor Comments:

Thank you very much for your submission to BMC Medical Informatics and Decision Making. We would now be grateful if you could revise your manuscript according to the reviewer reports below. Please provide a point-by-point response and indicate exactly where changes to the text have been made. Please also provide a version of the manuscript with all revisions indicated, either through the track changes function or through text highlights.

In addition to the revisions requested by the reviewers, we would be grateful if you could make the following editorial revisions:

Comment 1: In the ethics approval and consent to participate section of the declarations please provide a statement detailing the consent sought from the participants. If the need for consent was waived by the ethics committee due to the retrospective nature of the analysis then please provide details of this.

Response 1: The consent was waived by the ethics committee. I have indicated this in the respective section.

Comment 2: In this section please also provide the Ethics committee’s reference number if available.

Response 2: The reference number HREC/13/CIPHS/29 has now been included in the manuscript.

Comment 3: Please include the tables in the main manuscript file rather than uploading them as separate files and place them at the end of the main text after the references.

Response 3: We have now included the tables in the main manuscript after the reference section.

Comment 4: Please provide a description of the supplementary material at the end of the main text (after the references) and include the following:

- File name (e.g. Additional file 1)
Response 4: We have added these details. In addition, we have created a new table (Table S1) with the parameters of the gradient tree boosting algorithm.

Reviewer reports:

Theo Georghiou (Reviewer 1): This study reports on the development of models to predict unplanned readmissions over a small number of short term periods. The cohort used was a group of people discharged from one Sydney hospital, over a 4.5 year period.

The study is an interesting one with clear merits, but I have a number of questions about the presentation of the work. These range from the very minor, to some more fundamental points of confusion that need further explanation. (Hence my selection of 'No' in response to the question of whether the methods appropriate and well described. The methods are appropriate, but the description needs some clarification).

Methods:

Comment 5: Lines 161 to 167: Readmissions were defined as the first admission to any hospital in NSW, with subsequent admissions 'ignored for the purpose of this study'. Later in line 199 the authors describe building the models 'ignoring planned readmissions'. These comments led me to be fundamentally confused about the design and interpretation of the study. Were unplanned admissions included if they were preceded by a planned admission?

Response 5:

Let us start by the definition of index admission: an index admission was defined as the first admission by the patient during the study period (lines 159 and 160). This index admission can be either an unplanned (via emergency department) or planned admission.

A readmission is defined as the first admission to any hospital in NSW from the index admission (line 183). This readmission can be unplanned or planned. Any subsequent readmission by the same patient is ignored for the purpose of this study. Readmissions were further stratified as
occurring within 7 days, 30 days or 60 days from discharge from the index admission (as displayed in Figure 1). This has now been clarified in the manuscript.

For the purpose of prediction, we train and test predictive algorithms considering pairs of (index admission, unplanned readmission). That is, we ignore planned readmissions.

Comment 6: If planned admissions disallowed future unplanned admissions - then some of the differences between the three models will be explained by the relationship of the predictors with planned admissions (of which there are a larger number)?

Response 6: As explained above, the index admission can be any kind of admission, either planned or unplanned.

Comment 7: Or - if this interpretation is wrong and if planned admissions after discharge were ignored entirely (as suggested by the line 199) - why are they referred to at all in the paper?

Response 7: This is a very good question. Thank you for bringing it up. As you well pointed out, the predictive model only deals with unplanned readmissions. Nonetheless, we wanted to show how different unplanned readmissions are from planned readmissions. Planned readmissions have a clear weekly pattern. In addition, planned readmissions most likely occur in the same hospital of the index admission.

We thought it was important to highlight these differences since many of the predictive algorithms do not distinguish between planned and unplanned readmission.

Comment 8: Figure 1 is slightly ambiguous. I had to work out that eg the 8.4% on the 60 days level box was a proportion of 62,255, and not of 51,768. Would the figure not be less ambiguous if the "X not readmitted within 7/30 days" boxes were not connected to the 30 days/60 days vertical lines?

Response 8: Thank you for the suggestion. We have changed the figure as suggested.

Comment 9: The variables used as possible predictors are generally well considered, although there was little justification for the way the authors used some of the predictors. For example, two of the most important predictors of unplanned admissions - age, and prior admissions - could possibly have been better used. The age categories used were very broad and not very discriminatory for older age groups (65-79 and 80+ only), and it doesn't seem that number of
prior admissions was used - just cumulative LOS, and days since last admission. Why were these decisions taken?

Response 9: We agree with the reviewer in that there is more than one way of pre-processing variables for predictive modelling. Discretization of continuous variables is common in the generation of scores. In this study (line 194), categorical groups were chosen based on a combination of knowledge and their distribution in the dataset. In particular, age was discretised into 7, groups (all with enough sample size): Under 25, 26-45, 46-65, 66-85, and over 85. Number of previous hospital admissions was strongly correlated with cumulative LOS across these admissions. The latter was chosen as the better proxy for acute care utilisation. This has now been commented in the discussion section. The effect of discretising continuous variables was not explored in this study and may have influenced the prediction performance of the scores.

Comment 10: (For information) I can't comment directly on the appropriateness, or implementation, of gradient tree boosting machine learning and the subsequent feature selection.

Response 10: Gradient tree boosting is a popular machine learning technique. It is computationally more efficient than random forests and tends to provide more accurate results.

Results

Comment 11: Line 240-242. Interpretation of all numbers is unclear (related to comment 1). For eg the 8.4% in line 240/figure 1 - is this all patients who returned to hospital via ED within 60 days, or all patients who returned to hospital via ED within 60 day with no preceding planned admission?

Response 11: We have now updated Figure 1 to make it easier to read. It represents a summary of patients discharged and readmitted over three overlapping periods: 7-days, 30-days and 60-days postdischarge. Percentages in the box to the right of the 7 days arrow (3.6%, 4.4%, 0.1%) represent percentages of the number of live discharges above (62,255). That is, 3.6% of 62,255 live discharges have an unplanned readmission within 7 days of discharge. Similarly, percentages in the box to the right of the 30 days arrow (6.6%, 9.9%, 0.4%) represent percentages of the number of live discharges above (62,255). That is, 6.6% of 62,255 live discharges have an unplanned readmission within 30 days of discharge. Finally, the percentages in the box to the right of the 60 days arrow (8.4%, 13.7%, 0.9%) represent percentages of the number of live discharges above (62,255). In this way, 8.4% of 62,255 live discharges have an unplanned readmission within 60 days of discharge. As mentioned before, the 62,255 live discharges correspond to the index admissions, which can be planned or unplanned.
Comment 12: Figure 2. Can the authors clarify - is this all admissions, or first admissions? If the latter, then using 'rate of ... readmissions' as the description is not quite right.

Response 12: This has now been clarified in the Figure caption.

Comment 13: Figure 2. Is there any reason for distinguishing between other area and same area readmissions - is this information used elsewhere?

Response 13: Yes, hospital readmissions are often reported at different levels: hospital level, area of health services level, the State level and the national level. It has often been the case for a hospital to measure hospital readmission only as returns to their own hospital. This significantly underestimates the problem of hospital readmission and hinders the hospital’s understanding of what happens to their patients after discharge. Looking at the area health service provides additional information on the significance of missing information outside the area health service.

We have added a reference to Figure 2 in the Discussion session that was missing.

Now it reads like this (lines 387-393):

“In this study, we found that a significant number of unplanned readmissions took place in hospitals different from the hospital of the index admission (see right panel of Figure 2). This was confirmed in the latest report on readmissions in New South Wales [3]. It reflects the importance of maintaining medical record systems that are patient (as opposed to visit) centric, and can follow the patient across institutions. It also has implications for the implementation of financial penalties for unplanned returns to hospital.”

Comment 14: Table 1.1 - PPVs, sensitivity and specificity. In line 228 it's stated 'Thresholds for these measures were chosen as those that optimized the sum of sensitivity and specificity in the training sets'. Can there be different thresholds for each model? Comment 11 has some more to say about presentation of these measures.

Response 14: Thresholds are estimated independently for each model. This information has now been added to the manuscript.

Comment 15: Line 263-264. '.can be found in Table S6' - should this be Table 1.2 (which is otherwise unmentioned)? S6 looks like it contains info about all scores (which is useful). Can the authors clarify what observed and expected refer to - is observed when applied to the 20% validation set?
Response 15: Both tables should have been mentioned. Table 1.2 is a simplified version of Table S7 (old Table S6). We have added the reference to Table 1.2 in the text (line 280): “Observed and expected rates for selected scores can be found in Table 1.2 and Table S7 in the Supplementary material”.

Comment 16: In S6 - the sum of number of admissions is 11905, 11210 and 10,704 for the 7, 30 and 60 days models respectively. Why is there a difference between these numbers? Does this represent the 20% validation set (20% of 62,255 is 12,451)?

Response 16: These numbers represents 20% of 62,255 after removing the unplanned readmissions. In this way, 10,704 is 20% of (62,255-8560=53,695).

Comment 17: Where a person has had a readmission within 7 days - presumably the person is also considered to have had a readmission within 30, and within 60 days as well? I think this is the case - but comments in line 283-292 (and discussion line 358) about the '8 to 30 day' and '31 to 60 day' readmission groups gave me some small doubts about this interpretation. Can you confirm that the 30 and 60 day readmissions groups do include all people in the 7 day readmissions group? And that the use of '8 to 30' day group is a shorthand for 'those in the 30 day group, and not in the 7 day group'? It might be helpful to have some brief clarification about this in the methods/results.

Response 17: Yes, this is it, when a person has had a readmission within 7 days, that person is considered towards the 30 and 60 days readmission. We explored the characteristics of patient features within 8-30 days and 31-60 days a posteriori, to explore the change of the distribution of important features over different time periods. We have now clarified this issue throughout the manuscript.

Discussion

Comment 18: Line 313 to 318. It's stated that the 'score was built to optimise the combined sum of sensitivity and specificity'. In the methods (line 228) it's only suggested that the thresholds chosen for presentation of the model performance (ie in table 1.1) were chosen such that the sum of sensitivity and specificity was optimised.

Can the authors clarify whether the sum of sensitivity and specificity was used in a superficial/presentational way - or whether it was used to fundamentally select features of the model?
The Parr-30 model is mentioned as having been built to optimise PPV - but I'm not sure this is the case - it's just that the PPV and sensitivity were presented at threshold of 0.5 (after the practice of the Parr model's predecessors).

The authors of this current study could have presented the results of their models at equivalent thresholds to those of other models (including Parr-30) to have compared more directly between them.

For example, scores above 0.5 in Parr-30 represented only 1.1% of the cohort (table 2 of ref 23). From S6 - scores of 25 and above (30 day model) represent 1.2% of the group. 48 readmissions occurred in this group, representing a PPV of 35% and a sensitivity of 6%.

High PPVs are more likely for groups of those at highest predicted risk (analysis of data in S6 bears this out). Practitioners interested in using such models (who have some intervention aiming to prevent future readmissions) are most likely to want to target a relatively small number of individuals at highest risk (and the authors effectively acknowledge this saying that high PPVs 'could contain the costs of readmission strategies'). In practice, I feel that reporting the PPVs and sensitivity for the top few percent of those at high risk might have been more useful in the results (table 1.1). When looking at such models I'm rarely interested in knowing that there are a very large number of true negatives for those at lowest predicted risk.

Response 18: Thank you for your comment. The choice of cut-off point to report sensitivity and specificity did not alter the loss function of the algorithms. They were chosen as the cut-off points to maximise the sum of the sensitivity and specificity after training. This has now made more clear in the manuscript. Our scores roughly go from 0 to 35 points and the chosen cut-off points were 12 for RETURN7 and RETURN30 and 11 for RETURN60. This information has now been included in Table 1.1. Comparisons with other scores have been left for future work. However, values at additional cut-off points have now been included in the supplementary material (new Table S8). We have also modified the discussion section, taking into account the reviewer’s comments.

Comment 19: Lines 353 to 363. This relates to my first comment above. The interpretation of differences between the three time periods depends on how planned admissions affect the authors' use of subsequent unplanned admissions. If planned admissions disallow subsequent unplanned admissions from the model then some of the changes in predictors found could be due to changes in planned admissions over time. Perhaps this competing risk needs to be addressed in the discussion.

Response 19: Please refer to our response in Comment 5/ Response 5.
Comment 20: The discussion could perhaps be improved with addition of some further reflection on the main motivation for this work. Is it the 7, 30, 60 days comparison that's the key reason for the reader to be interested? Or perhaps the comparison between the gradient tree boosting models and the corresponding logistic regression models? Or to produce a model specifically designed for this single hospital? Who are the intended users of this information?

The background section does suggest that the focus in on the 7, 30 and 60 day comparison - but I don't feel that the focus of the paper overall really highlights this.

Response 20: Thank you for your comment. The aim of this study was to investigate the factors associated with unplanned readmission in a Sydney hospital. We started by measuring the number of unplanned readmissions per day to the same hospital, as well as to other hospitals within the State (Figure 1). We then developed and compared validated readmission risk scores using routinely collected hospital data to predict 7-day, 30-day and 60-day all-cause unplanned readmission. This has now been clarified in the manuscript. Also note that the title has been changed following this comment as well as comment 22.

Comment 21: The authors acknowledge that the sample size is relatively modest especially for the 7 day model (line 381) and that more data could improve the models. I'd like to see in the discussion a reflection of the fact that the cohort were all discharged from a single hospital over a relatively long period of time. There will be particular factors at play in this hospital that will mean that the models found might not necessarily be generalizable to others.

Response 21: We have now acknowledged this in the Limitations section.

Hadi Kharrazi (Reviewer 2): Thanks for submitting your article to BMC MIDM. The manuscript focuses on predicting 7, 30, and 60 day all-cause unplanned hospital readmissions in Australia. I was pleased reading your manuscript. The manuscript is well written, conforms to most items recommended by the TRIPOD statement, presents a novel methodology and provides some interesting results by comparing the variables contributing to the different readmissions (7, 30 and 60 days); however, there are some comments to be addressed:

Comment 22: *Title: following TRIPOD recommendations, the title could be more specific (i.e. "Developing and validating a predictive model for 7-day, 30-day and 60-day all-cause unplanned readmission to Australian hospitals")

Response 22: Following this recommendation, we have changed the title to: "Predicting 7-day, 30-day and 60-day all-cause unplanned readmission: A case study of a Sydney hospital"
Comment 23: * Line 63: "where" instead of "were"
Response 23: This has been corrected. Thank you.

Comment 24: * Line 68: "It indicates that shorter-term…" - please clarify "it" with the "different predictors of 7-day versus 30 and 60-days"
Response 24: The “it” refers to “the study”; this has now been changed in the manuscript.

Comment 25: * Line 65: "associated with" instead of "associated to"
Response 25: This has been corrected. Thank you.

Comment 26: * Line 108-111: What was the method for finding the past evidence? Literature review, systematic review, or expert feedback?
Response 26: A literature review was carried out before the commencement of the study.

Comment 27: * Line 116-120: requires a citation
Response 27: A citation has been added.

Comment 28: * Line 169: It is not clear how the list of 88 variables was determined? It doesn't look like that a univariate analysis with the variables (which is part of TRIPOD recommendations) has been performed - please describe the reason.
Response 28: The preliminary list of 88 variables correspond to the variables available in the electronic health record and commonly used in predictive algorithms. This has now been clarified in the manuscript. For example, medication administration was not available at the time of the study and it is not often found in predictive scores, although, if available, it could have been used as a proxy for patient comorbidity and health status. In this study, multivariate variable selection is carried out using the XGBoost algorithm instead of performing univariate analysis. Multivariate selection using machine learning algorithms has been proven superior to the univariate technique in high-dimensional, highly-correlated data.
Comment 29: * Line 189: "Pre-processing" describe how lab value thresholds are determined? Only expert opinion? If yes, explain why didn't you use a regression tree to determine the most efficient thresholds?

Response 29: We used the lab value thresholds provided by the pathology electronic system. This is the most straightforward simple way to categorise lab results. As with other continuous variables, we could have chosen to retain the continuous variables without discretisation and this may have influenced the prediction performance of the scores. This has now been acknowledged in the limitations section.

Comment 30: * Line 189 + Appendix: Not clear which laboratory results were used (i.e. first ordered tests upon admission, last ordered tests before discharge, any abnormal result over admission)

Response 30: As part of the original 88 preliminary variables, we used the two last common pathology results available before discharge. This has now been made clear in the manuscript. After variable selection, only selected last pathology findings had predictive power (as reflected in Table 2).

Comment 31: * Line 193: The project uses a derivation set and a validation set; however, does not use a revalidation set (external to their underlying data)… please mention this and potential generalizability issues in the limitation of the manuscript.

Response 31: No external data was used as a validation set in this study. This has now been included in the limitations section.

Comment 32: * Line 199-205: should not repeat in the appendix anymore

Response 32: We have removed the double mention of the appendix.

Comment 33: * Line 199/320: stylistic, but inconsistent capitalization/italicizing of gradient tree boosting

Response 33: We have removed the capital letters.
Comment 34: * Line 217: t-test was used to compare distribution of the selected features - please cite a paper explaining why t-test is the best method to do this

Response 34: The t-test assesses whether the means of two groups (of features in this case) are statistically different from each other. This has been rephrased for clarity.

Comment 35: * Line 268: "good predictor" - please provide numerical value

Response 35: The ORs for RETURN 7, 30 and 60 have been added to the manuscript.

Comment 36: * Line 275: "two socio-economic" … please mention them

Response 36: Marital status and payment status. They have been added to the manuscript.


Response 37: Thank you. The citation has been added to the manuscript.

Comment 38: * Line 427: "Availability of data" please explain if other researchers/analysts can access the data (and if not, why [e.g., protected health information])

Response 38: A brief explanation has been added to this section

Comment 39: * Line 491: Kansagara's review is outdated. Please double check and add/cite the new systematic review: https://www.ncbi.nlm.nih.gov/pubmed/27354072

Response 39: Thank you. We have added this citation.

Comment 40: * Line 326-329: "Outcome variables are very unequally represented" - the paragraph doesn't relate to the previous paragraph

Response 40: We clearly separted this paragraph to the previous one.
Comment 41: * Line 338: explain "Medicare-Holder" public patients in the Australian context?
Response 41: We have added a description of Medicare in Australia.

Comment 42: * Line 351: explain the "weekend effect"
Response 42: We have added an explanation.

Comment 43: * Table 1.2 - this table can be moved to appendix?
Response 43: The appendix contains table S7 that is a more detailed version of Table 1.2.

Comment 44: * Table 2: why "normal last lipase" has a positive effect?
Response 44: We believe this is a reflection on how having a lipase test at all (vs not having the test) is associated with higher probability of readmission. In this test, the percentage of abnormal lipase findings is very small. This has been added to the discussion section.

Comment 45: * Table 2: "Overseas visitor" and "Referred by other practitioner" have a negative effect - please explain (could be simply because the data had an incomplete set of measures such as not showing the readmissions?)
Response 45: We can only speculate why there is an association between these variables and lower probability of readmission. One possible explanation of the effect of ‘Overseas visitor’ could simply be a reflection of lack of follow-up, which is why it appears to be important for the prediction of 60-day readmission and not for shorter follow-up times. We have added further discussion on these potential associations and have included a discussion on causality vs association as part of future work.

Comment 46: * Table S6: indicate issues with using the higher scores given the low number of admissions (less than 10) in those ranges
Response 46: Limitations due to small sample size have been acknowledged in the Limitations and Future Work section.