Reviewer’s report

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

Version: 0 Date: 11 May 2017

Reviewer: Theo Georghiou

Reviewer's report:

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:

1. 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?

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)?

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?

2. 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?
3. 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?

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

Results

5. 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?

6. 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.

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

8. 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.

9. 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?

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)?
10. 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.

Discussion

11. 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.

12. 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.

13. 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.

14. 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.

Are the methods appropriate and well described?
If not, please specify what is required in your comments to the authors.

No

Does the work include the necessary controls?
If not, please specify which controls are required in your comments to the authors.

Yes

Are the conclusions drawn adequately supported by the data shown?
If not, please explain in your comments to the authors.

Yes

Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review?
If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

I am able to assess the statistics
Quality of written English
Please indicate the quality of language in the manuscript:

Acceptable

Declaration of competing interests
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I declare that I have no competing interests.

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal