Reviewer's report

Title: Development, Implementation, and Prospective Validation of a Model to Predict 60-day End-of-Life in Hospitalized Adults upon Admission at Three Sites

Version: 0 Date: 02 Jul 2020

Reviewer's report:

This paper presented an interesting study of developing prediction model of 60-day mortality in hospitalized adults using random forest. The overall study design is well-planned, and more importantly, the application itself is clinical meaningful and can be implemented in the future. However, there are a few issues need to be addressed before considering publication:

1. In Methods section, the author stated that "patients with a lack of data were omitted", is there any particular reason for the exclusion? The amount of patients excluded (19.6%) was fairly large, and may affect the model performance. The authors may consider investigating the missing values in more details, for example, whether they are missing at random or not? Does the missing values provide any clinical insights? Maybe they represent some missed opportunities in predicting and preventing death? The authors could also try imputation techniques on these data.

2. The author mentioned the random forest incorporated 9614 variables in the abstract, but did not mention how these variables were selected in the Methods section, and only a few were presented in Table 1 (I was also wondering how they authors choose which predictors/comorbidities can be listed in the table?) In addition, with that many variables, it is destined to introduce bias, have the authors considered filter or preselect a set of variables?

3. The authors only used random forest as their modelling approach, have the authors tried other algorithms? For example decision tree. A common practice would be test different models and identify the best one based on both performance and interpretability.

4. The authors described a different performance when applying the model to Brooklyn population, I would love to read more discussion on this as to why and what may cause the observed difference, this may provide important insights on the generalizability and representativeness of the study results.

5. The authors may also consider explain how and why the model works in clinical practice, for example, the authors could try some explainable techniques and models in addition to random forest. An explainable component would improve the feasibility of the application as many clinicians would like to know how the model works before using it.

Other minor issues:

1. Table 1 has too many comparisons, the p-values reported should be adjusted or corrected accordingly.
2. The authors may want to work on the some language corrections, for example, the first sentence in methods section of abstract (A prognostic study of ......within 60 days of admission) is hard to read.

Publons Reviewer Recognition. Springer Nature can send verification of this review directly to Publons (a subsidiary of Clarivate Analytics). If you would like to take advantage of this service, please click on the “Yes” option below. Your name, email address, title of the reviewed manuscript, name of the journal, and date of your review submission (the “Review Data”) will then be transmitted to Publons upon publication of the manuscript. If you have already registered at Publons, they will notify you of the receipt of this review and update your profile as per your settings and their policy. If you are not registered with Publons, you will receive an email from them asking you to register in order for them to be able to recognize your review on your new profile page. Publons may use the Review Data to generate derivative metadata for the benefit of Publons and you as a reviewer, carefully considering the sensitivity of such information. For example, Publons may verify your record as a reviewer by updating your profile published on its webservice if you have registered for such service or help editors to identify candidate reviewers. Please find the details of processing in Publons’ privacy policy https://publons.com/about/terms

No

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

Yes

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:
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

Reviewer Publication Consent. 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) if this manuscript is accepted for publication. Any comments that I do not wish to be included in the published report have been included as confidential comments to the editor, which will not be published. If you are not happy for us to publish this report, please contact the editorial office before completing the review. If you wish, you can include your name in your published report. Please note you must decide whether to include your name at the start of the process and confirm this decision whenever prompted. If you change your decision later, we will not be able to publish your name.

I agree to the terms of the CC-BY 4.0 license; please do not publish my name with my report. (default)