Reviewer’s report

Title: Six underlying health conditions strongly influence mortality based on pneumonia severity in an ageing population of Japan: A prospective cohort study

Version: 0 Date: 15 Jan 2018

Reviewer: Henri Van Werkhoven

Reviewer's report:

The authors determined which of a list of 23 candidate predictors (termed underlying health conditions) independently predicted 30-day mortality in patients hospitalized for community-acquired pneumonia (CAP). They found 6 conditions to be associated and compared discrimination of a model consisting of these factors to that of the frequently used CURB65 and PSI scores. Their model performed better than both of these scores.

The accurate prediction of mortality in CAP patients in specific population is an important field of study. However, the current study has some risks of bias and the manuscript would benefit from some additional results to be reported. Therefore, the following comments should be addressed:

General:

- With 140 endpoints, the authors evaluated 23 candidate predictors. Optimism due to overfitted model is likely and was not analysed or corrected for. The authors should complement their findings by an internal validation with bootstrapping to determine the likely performance of their model in an external setting (see e.g. Steyerberg et al. J Clin Epidemiol. 2001 Aug;54(8):774-81. PMID 11470385)

- For the same reason, comparing the AUC of externally derived model (for which there is no optimism) to AUC of internally derived model, is biased towards finding better performance of the derived model. This should be acknowledged.

Abstract:

- "The 30 day-mortality was 7.9%." Please also show the number of patients died in the abstract.
- It is confusing that first the independent candidate underlying health conditions are selected based on a model with CURB included, and next the AUC is determined from a model of only the independent underlying health conditions (I presume without CURB). Why this rationale?

- Added predictive value of underlying health conditions is claimed (abstract conclusion) and seems reasonable given the data, but this has not been tested or quantified. It would be valuable to show the added predictive value of a model including CURB65 or PSI score and the 6 underlying health conditions.

- Performance against the PSI score was also tested but these results are not provided in the abstract. Please include this relevant piece of information.

Introduction:

- Line 130: "For the current analysis, we used data collected between September 2011 and March 2013." Please inform why data from April to August 2013 were ignored, it doesn't make sense at first sight. (Or maybe the data can be included.)

Methods:

- Line 140: "If a patient developed the disease and was referred to our hospital more than 48 hours following admission to another hospital, the patient was classified as having hospital-acquired pneumonia" Any time window after hospital discharge the patient was still considered to have hospital-acquired pneumonia?

- Line 165: "Those conditions are not only theoretically but also practically untreatable or unmodifiable." I do not agree that all of these factors are unmodifiable. Many of the comorbidities can be modified positively by proper medical treatment. (Not so much their presence, but their severity.)

- Line 190: "PSI scoring system ... is not practically useful in some clinical environments, such as a busy emergency room or primary care clinic." It is widely used, including in the emergency room, I don't think it is very impractical and might do better than CURB65 score.

- Line 201: In the final model, only variables with p-value < 0.05 were selected. Please indicate this here clearly, as the wording of the methods suggests that variables with p-value < 0.20 were retained in the model. (e.g. age 65-74 seems to be significant at the p<0.2 level judging from the confidence interval.) Please notice that several methodological papers
suggest that p<0.05 is probably too stringent for prediction model development as it results in overfitting of the model to the data.

- Any missing data of subjects that were lost to follow-up? How did the authors deal with that?

Results:

- Line 246: "27.0% had a CURB score of two or more." Is CURB without age meant here? (Hard to compare with literature since most use CURB65.) If not, why did you use two or more as the break point? Line 179 says that CURB65 score >= 3 is used to define severe CAP.

- Line 249: "Because of their small numbers, home infusion therapy (n=8) and home wound care (n=13) were combined into a "home care" variable." I would suggest to replace this to the methods section.

- Line 248-254: "Mortality rates were higher ... highest CURB score died." Most of what is mentioned here is not informative (most is already known from the literature, and not clear what is the mortality for each category). Also what is relevant is mentioned in the section "Risk factors for 30-day mortality". I would suggest to delete this paragraph.

- Table 2: "Foreign bodies interfering with swallowing" has a crude and adjusted OR, which is surprising because it is not significant in the crude analysis and far from significant in the adjusted. "Bed-ridden state" on the other hand only has crude OR which is significant. Please revisit table and analysis.

- Line 257-268: This section has overlap with table 2, so could be substantially shortened.

- Line 288-291: "The mortality rate of severe pneumonia ... high risk patients (risk ratio 1.53, 95% CI 0.94-2.50, p=0.1)." This repeats the preceding lines and can be removed.

Discussion:

- The AUC of the CURB65 score is substantially lower (and of the PSI score is also a bit lower) than in most external validations (roughly around 0.8, e.g. Kwok et al. Biomed Res Int. 2013;2013:504136.). Can the authors elaborate in the discussion about why these models perform less in their population? Does it have to do with the different age or comorbidity distribution?
I do not completely agree that CURB-65 does not adequately predict mortality in the patients with >=2 conditions. The data suggest an interaction on a relative scale (RR 7.2 vs. 1.5) but on an absolute scale the effects are quite similar (4.8% vs. 6.1%). So there is no sign of interaction on an absolute scale. Therefore, the CURB65 score may still be informative in the patients with high number of conditions. Please provide a more balanced discussion of the results.

Line 312: "However, if the number of underlying health conditions is two or more, CURB65 poorly predicts mortality since treatment cannot substantially modify these underlying health conditions." It is not clear to me where the authors have demonstrated that CURB65 doesn't adequately predict mortality in this subgroup. Please provide calibration plots for the subgroups to determine whether prediction is inadequate. It may be that the RR is only 1.5 but still the risk could be adequately predicted. Also, the authors should assess whether the AUC of the CURB65 score is different in the two groups. (Probably not as the absolute risk difference is the same and this drives the AUC.)

Also, if prediction is indeed poor, whether this is because "treatment cannot substantially modify these underlying health conditions" cannot be derived from this data. To me it is questionable whether this would be the reason for the poor prediction. A more likely reason would be that the CURB65 score does not well generalize to populations with multiple comorbidities.

Line 314: "Thus, in the latter patients group, clinicians should spend special effort on pneumonia prevention, such as vaccination." This does not logically follow from the results presented. The decision to target preventative interventions to a specific subgroup should be informed by cost-effectiveness studies of that intervention, which depends on subgroup specific incidences of the disease, vaccine effectiveness, costs of an episode, attributable health outcome of an episode, etc. It cannot be based on the observation that comorbidities, rather than pneumonia severity, determine the prognosis for patients with pneumonia.

Line 322: "The results of the present study provide evidence supporting the rationale behind such DNR decision-making." I don't see the reasoning how the presented data support the rationale behind DNR decision making. Please explain or remove DNR section. (It doesn't add value to the discussion in my opinion.)

Line 325: Of course it is not calendar age that drives (in a causal sense) the mortality risk, however, for prediction purposes it doesn't really matter whether age is an independent predictor or confounded by (uncaptured) age-related conditions. Same for gender in line 336.
- Line 353: "However, PSI predict mortality based on calculating a single score, consisting of both treatable and untreatable parameters combined." This is not a problem as long as the aim is to just to predict the risk of an unfavorable outcome.

- Line 364: "Last, the outcome of this analysis was only 30-day mortality because we could not follow all survivors after hospital discharge." This sentence suggests that patients discharged alive before day 30 were not assessed for survival after discharge up to day 30. Is this the case? If so, please mention this in the methods section and discuss the limitation. If not, please rephrase this sentence.

- Line 370: "CURB65 score was not useful for mortality prediction when the patient had multiple underlying health condition." As argued above I do not (yet) agree with this statement. Check of discrimination and calibration in the subgroups and adjustment for optimism by internal validation should be performed before making this conclusion.

**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.

No

**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