Author’s response to reviews

Title: Clinicians Can Independently Predict 30-day Hospital Readmissions as Well as the LACE index

Authors:

William Miller (william.miller@uchospitals.edu)
Kimngan Nguyen (KPNguyen@mednet.ucla.edu)
Sitaram Vangala (SVangala@mednet.ucla.edu)
Erin Dowling (EDowling@mednet.ucla.edu)

Version: 1 Date: 24 Nov 2017

Author’s response to reviews:

Summary of our response:

We broadly agree with the critiques of our reviewer. We have therefore revised our manuscript in accordance with every comment and request of our reviewer.

In our revised manuscript, we have included the confidence intervals that were calculated but not presented in the original manuscript. The confidence intervals for our AUC’s did not cross 0.5, suggesting that our results were indeed better than chance. However, we agree that "better than chance” may not be the same as clinically relevant, and further improvements in predictive tools would be desirable.

While it would be desirable to achieve an AUC of 0.7 with readmission prediction tools, previously reported prediction tools have generally not achieved this in external validation cohorts. Indeed, as we describe in our revised manuscript, the LACE index did not achieve this, but has nevertheless been widely adopted given the lack of simple alternatives and the felt need to predict and prevent readmissions. In the midst of all these efforts, clinician expertise has generally been ignored, perhaps because of the thin body of literature that we describe, which has suggested that clinicians are unable to predict readmissions; or perhaps because clinician expertise is intuitively believed to be the problem, with hospital readmissions being seen as the result of clinician shortcomings instead of challenges at the patient and health system level.

We present a more nuanced view of clinician expertise: clinicians are able to predict readmissions, but they do this incompletely by relying on expertise that focuses on factors such as patient nonadherence. Clinician expertise is therefore significant, but requires supplementation by improved prediction tools.

Itemized list of revisions:
Major Compulsory Revisions

1. Lines 55-59: Could you please specify the range of LACE scores? In some countries, especially in Europe, this index is still not widely known nor utilized.

Response: We have included this in our revised version.

2. Lines 86-88: Using a different scale from that of LACE to assess the providers' predictive ability could have partially biased the results. I think it would have been more sensible to ensure an identical evaluation scale—different survey scales lead to different answers. Maybe the authors opted for a continuous percentage scale in order to facilitate creating ROC curves, but this choice might have "penalized" the LACE index, as it is less detailed (0 to 19/20 vs. 0 to 100 score). These aspects should be at least acknowledged among the limitations of the study.

Response: We have added this limitation in the discussion portion of the manuscript.

3. Lines 110-111: Ideally, to ensure independence between evaluations the authors should have interviewed 377×4 providers (377 residents, 377 nurses, 377 attendings and 377 case managers, one for each patient), but I am well aware that this is impossible. Still, it is worth noticing that the authors have assessed the predictive ability of a restricted number of attendings, residents and, most importantly, case managers. I would not rely on evaluations made by such a little bunch of individuals: six case managers cannot be representative of the category of case managers overall, and results may be heavily unbalanced because of the answers of one or two individuals. Thus, I suggest that the authors delete all analyses related to case managers.

Response: We have deleted all the analyses related to case managers.

4. Lines 123-124: The authors state that they estimated 95% confidence intervals (CIs) for AUCs, but there is no sign of them in the paper. As the sample is not large, I recommend computing exact binomial CIs, rather than asymptotic normal CIs.

Response: We have recalculated the CI's as requested, and we have reported them in the figures, table, and text of our revised manuscript. Thank you for pointing out this important omission.

5. Lines 127-129: Please specify what subgroups you are referring to (poor understanding of disease, poor adherence to therapy, etc.).

Response: We have clarified this in the text.

6. Line 136: Could you please summarize patients' diagnoses and, if any, interventions? This would ease reproducibility and comparability of your results.
Response: We have added a table that describes pertinent patient characteristics (age, sex, primary diagnosis) for the patients admitted to the studied provider teams during the two months of our study period. While we did not collect this specific information on our study cohort, we were able to use internal hospital data that is regularly extracted from our EHR to describe the information from all the patients admitted to the studied services during the entire two months (June-July 2015) during which our study occurred (June 4-July 24, 2015). Unfortunately, information on interventions was not available, and given the broad range of patients treated on these "general" services, it may be difficult to succinctly describe this.

7. Lines 152-154: It is risky to make considerations on power and sample size after analyses are concluded. The authors should have determined the sample size during the study design process. I am prone to think that all tests failed to achieve statistical significance because there was no difference in predictive power between LACE index and providers as well as among providers themselves. Let's consider, for example, the AUCs of LACE index (0.620) and nurses (0.628): the difference is really negligible. Even the difference between LACE and attendants is very low (~0.07). To me, providers and LACE index perform equally.

Response: We broadly agree with this analysis, and have changed the text accordingly.

8. Table 3: Could you please provide the number of patients belonging to each subgroup?

Response: We have provided this information in the relevant table (now table 4).

9. Lines 184-186, 197-202: The authors' statement is not supported by results. Both providers and LACE index exhibited a poor ability in predicting 30-day unplanned readmissions, as the AUCs are always below 0.7 (except for the subgroup of patients with poor understanding of disease). The authors can assert that providers perform better than chance only if 95% CIs of AUCs (not presented in the paper) do not cross 0.5. Results are undoubtedly better than Allaadeen's, but more inferences are needed. Moreover, saying that providers perform "possibly better than the LACE index" is misleading: providers' AUCs are higher than LACE's, but not significantly higher.

Response: We broadly agree. Our CI's, which we have presented in the revised version of our manuscript, do not cross 0.5. It is quite fair to point out that claiming a difference between the AUC's when we cannot prove a statistically significant difference is fraught. We have changed the discussion according to the more rigorous interpretation of the statistical data requested by our reviewer.

10. Lines 228-232: This statement too is not supported by results. How can clinician expertise be paramount, if the LACE index performs as well as (actually, as bad as) providers in assessing the readmission risk of patients? I interpret all the results in a different way: the LACE index performs badly and providers do not seem to do better. In order to make readmission models and tools more useful, efforts are needed to identify variables that may play an important role in predicting this outcome. These variables may include understanding of disease and drug adherence, as also suggested by the authors, as
well as quality of care and organizational structure and processes that were not evaluated in this study. Some authors (Lenzi et al. BMC Health Serv Res 2016, 16:473) also found that low systolic blood pressure levels are highly predictive of 30-day readmission in patients with heart failure.

Response: We again broadly agree with this critique. We have altered the discussion to reflect this more rigorous interpretation of our statistics.

In conclusion, I suggest that the authors rearrange the discussion section in light of these thoughts. As it is, this paper draw a conclusion that is not consistent with the results: 30-day readmissions are still difficult to predict for both providers and published algorithms. Efforts should be done to increase the accuracy and usefulness of readmission prediction tools to help and guide providers adopt tailored strategies for patients at high risk of readmission.

Response: See above. We have revised the discussion of our manuscript in accordance with the request of our reviewer.

Minor Essential Revisions

1. Lines 118-131: I suggest that the authors carry out a sub-analysis whose results might be very useful and informative. I would investigate the association between readmission and the eight risk factors (poor understanding, poor adherence, etc.) using logistic regression models. Given the small sample size, each of the eight regressor could be analyzed separately. Result of this sub-analysis would help understand which factors determine hospital readmissions, guide further research and implement new prediction tools.

Response: We performed this analysis, which is now presented in the text of the results section.

2. Lines 267-287: The study might have another limitation. The "absence" of a specific risk factor, such as poor social support, may reflect two situations: (1) the patient does not have that risk factor or (2) the provider does not know whether the patient has that risk factor. This could have partially biased the results, because answering "No" could mean both "No" and "Unknown".

Response: We have included this limitation in the limitations section of our manuscript.