Reviewer’s report

Title: Automated Chart Review Utilizing Natural Language Processing Algorithm for Asthma Predictive Index

Version: 1 Date: 02 Dec 2017

Reviewer: Pietro Pirina

Reviewer’s report:

This manuscript submitted by Kaur H et al, entitled "Automated Chart Review Utilizing Natural Language Processing Algorithm for Asthma Predictive Index" is well done and well written and the study design is very interesting. If the validity and applicability of this clinical data mining method will be confirmed in a larger population samples, its usefulness on asthma epidemiological studies and on public health interventions will certainly be relevant.

However, after the above considerations, some limitations of the study should be underlined:

1) Given the complexity of the methodological approach and the specificity of the used terminology, I generally suggest that less abbreviations are used in the text. This will make the reading of the manuscript more pleasant to a greater number of readers.

2) Effective extrapolation and reliability of collected data is essentially related to the care by which the electronic database is managed. Missing data and incorrect interpretations of the clinical data can completely alter the final clinical information. This limitation should be stressed in the discussion section.

3) On the heterogeneity of the different electronic databases (electronic health record) that are different in the different local health settings and on the applicability (portability) of the NLP algorithm in these contexts, the authors have already expressed the same as several limitations of the paper.

4) Small and "selected" population sample. Can we reproduce the same sensitivity, specificity, positive predictive value and negative predictive value in a larger population sample and, above all, in a different geographical area?

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.

Unable to assess
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.

Not relevant to this manuscript

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