Author’s response to reviews

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

Authors:
Harsheen Kaur (hakaur@unm.edu)
Sunghwan Sohn (sohn.sunghwan@mayo.edu)
Chung-il Wi (wi.chung@mayo.edu)
Euijung Ryu (ryu.euijung@mayo.edu)
Miguel Park (park.miguel@mayo.edu)
Kay Bachman (bachman.kay@mayo.edu)
Hirohito Kita (kita.hirohito@mayo.edu)
Ivana Croghan (croghan.ivana@mayo.edu)
Jose Castro-Rodriguez (jacastro17@hotmail.com)
Gretchen Voge (Gretchen.Voge@childrensmn.org)
Hongfang Liu (liu.hongfang@mayo.edu)
Young Juhn (juhn.young@mayo.edu)

Version: 2 Date: 20 Dec 2017

Author’s response to reviews:

POINT-BY-POINT RESPONSE TO REVIEWERS’ COMMENTS:

We are submitting our point-by-point response to reviewers’ comments and have revised our manuscript according to these responses. We have carefully considered each of the comments offered by the reviewers. Specific changes are marked and highlighted using track change in the manuscript, and we are submitting both revised manuscript and the point-by-point to the review.

Reviewer 1 (Pietro Pirina)
1. 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.

Response: We appreciate reviewer’s comments.

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

Response: We have revised the manuscript.

3. 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.

Response: Given our study as a retrospective study, we acknowledge the reviewer’s concern and included it as a limitation in the Discussion section. We also included our earlier work based on a prospective cohort study showing a close correlation between medical events (mild acute illnesses) in EMR and those captured by prospective follow up[1-2].

“Intrinsic limitation of EHR reliability collected from the sources may not represent complete history of patient’s conditions (eg, parents’ asthma collected from family history section and patient provided information) and thus may affect the final asthma ascertainment. However, our earlier work based on a prospective cohort study showed a close correlation between medical events such as common mild acute illnesses in children captured by EMR and those by a prospective follow up.”

4. 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.

Response: We appreciate the reviewer’s thoughtful comments. To address the reviewer’s concerns, we assessed whether NLP algorithms for asthma can be implementable in a different study setting in Sioux Falls, South Dakota (ie, different clinical practice, different study population, different documentation style and different EMR system) and we reported the results
The results showed external validity (portability) of NLP algorithms for asthma. We included the summary of this response in the Discussion section.

5. 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?

Response: We addressed the reviewer’s comments above. The intention of our work on the development of NLP algorithms for asthma was to make it available to other study settings so that large scale studies can be conducted with a consistent method of determining asthma. [3-4]. At present, while we have not replicated the similar work to NLP-PAC, we are planning to replicate it in different local health setting in the future.

Reviewer 2 (Davor Plavec, M.D., M.Sc., Ph.D.)

1. This is an interesting study showing the possibility to use computer based algorithms to competently assess large scale data for epidemiological studies or epidemiological follow up. The methodology is adequately described an used and all limitations were adequately described suggesting needed follow up studies to improve the instrument used.

Response: We appreciate reviewer’s comments.

References


