Reviewer’s report

Title: Combining information from a clinical data warehouse and a pharmaceutical database to generate a framework to detect comorbidities in electronic health records

Version: 0 Date: 13 Apr 2017

Reviewer: Jane Carrington

Reviewer's report:

Thank you for the opportunity to review this paper. Authors sought to increase the rate of co-morbidity detection for more accurate coding. I agree this is a problem that influences reimbursement and can influence continuation of patient care. Authors took the approach of linking medications and notes to detect the presence of co-morbidities. Authors started with ENT dataset. The annotation process using two medical experts. These roles should be better defined for credibility of the annotation. Were nurses included in this process? 98.2% agreement is very high for this process and warrants some explanation. The classifier development stage was not specifically identified in this paper. What was that process? The limitations of this project as similar for those using the EHR in algorithm development and I would broaden the comment on page 15 line 299 to healthcare, not medical field. Paragraph 297 pg 15 begins a discussion of limitations that includes lack of NLP, machine learning, and rule based approached towards algorithm development. I question the comment that although NLP yield good results, they often target specific diseases rather than co-morbidities. I disagree with this statement and suggest that co-morbidities are diseases and NLP could be very useful here. I understand this is a very simple straightforward algorithm. A full description of how it is going to be used next would be useful, what are next steps to move this to practice for coding? Is there an interface or autopopulation of data fields? These should be addressed. Thank you.

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 recommend additional statistical review

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.

We have a provisional patent through the University of Arizona for an algorithm and interface.

However, our algorithm seeks to fill a different need in healthcare.

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