Reviewer's report

Title: Are big data analytics helpful in caring for multimorbid patients in general practice? - A scoping review

Version: 0 Date: 10 Dec 2018

Reviewer: Joseph Lee

Reviewer's report:

Thank you for this interesting read. I think these are important results.

As this is a review with a 'negative' result the search and exclusion criteria are important. A little more detail would allow the reader to judge if it is negative because of these factors.

Specific comments: These are mostly optional language tweaks, offered in the spirit of trying to be helpful, but is not an exhaustive list of English corrections…

Abstract:

Line 26: 'in high velocity' should be changed, perhaps to 'generated at high velocity'

Background:

Line 59: 'population is accountable for' - to 'population accounts for'
Line 62: delete 'most' or even the whole line
Line 76: this is very useful - could you define the 3 Vs in the abstract?
Line 82: delete 'yet'

Methods: Search strategy - looks reasonable to me, I wonder if more terms for co-morbidity would have increased the numbers of included studies. Could you give exact dates for the searches? (was there a limit)

Was there a registered protocol?

Could you provide a PRISMA scr checklist?
How did you define big data analytics for inclusion? (given this is a difficult definition and important for your decisions)

Line 131: Arguably, the flowchart is a result?

Results:

Lines 146 to 148: Could you clarify if this means none of the grey literature was included in the final six? or did all six come from the grey literature, but only after further discussion?

It would be nice to have a supplementary table with reasons for exclusion for full texts - covidence can generate this.

Discussion:

Line 198: 'proof' isn't quite right, ? 'apply'

Strengths and limitations:

Are there lots of studies using electronic health records that don't self-identify as 'big data' and therefore were missed, or were these excluded for not meeting your criteria? E.g. is a cohort study of lots of electronic records included for being big or excluded for a lack of 'velocity'? Or perhaps for using classical statistics?

Either way there is a limitation that I think should be identified.

Table 1: Col 5 is headed 'No. of used datasets' but some of the lines have large numbers that are presumably the number of records or participants?

Supplmentary material:

Could be slightly more complete, could you provide the rest of the strategy here as well? i.e. the dates and databases searched, just to make it self contained.
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 am able to assess the statistics

Quality of written English
Please indicate the quality of language in the manuscript:

Needs some language corrections before being published

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 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.