Reviewer’s report

Title: Using the Distance between Sets of Hierarchical Taxonomic Clinical Concepts to Measure Patient Similarity

Version: 2 Date: 29 Sep 2018

Reviewer: Chonghui Guo

Reviewer's report:

The authors have addressed all my comments, and the revision has been greatly improved.

I have some slight suggestions that the authors might consider to revise it, pg2 lines 15-16 in "Abstract" section, "patient similarity measurements" should be "patient similarity measure or measures"? pg3 lines 24-25 in "Background" section, electronic medical records ("EMR") should be "EMRs", pg18 lines 35-36, SS #5 lack of "Formula". Some contents should be placed in the manuscript, not in the appendix, for example, "Supplemental Table 1: numbers of patients of four pre-defined subpopulations" in Supplementary Material Data Interpretation and Subpopulation Criteria. docx. Besides, the format of references needs to be unified, for example, case sensitive of the titles, the abbreviation of the journals, pages, and so on.

So the authors need to recheck the language, format, structure and content of the paper, and make the paper more readable.

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 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