Reviewer’s report

Title: A visual interactive analytic tool for filtering and summarizing large health data sets coded with hierarchical terminologies (VIADS)

Version: 0 Date: 30 Nov 2018

Reviewer: Hallvard Laerum

Reviewer’s report:

Thank you for the article, that describes an interesting tool for analysis of hierarchical data. I agree completely that analyzing large hierarchical data sets is difficult for most users, and that effective visualization tools are needed to identify patterns and dependencies in the data. I understand that the article is about the tool itself, so I will limit my response to the description as such.

1. The article states that "To improve comprehension, new tools are needed to extract the overviews of such data sets" and that "the functionality of VIADS cannot be achieved by any single tool. Is this entirely accurate? The tool uses quite generic algorithms that seems to apply to any hierarchy. I suspect you might find very similar functionality in generic analysis packages, even as a library in R or a module in Mathlab, SAS or other comprehensive analysis packages? You do compare VIADS to two software tools, but they are quite specialized, and do not have the same capabilities as VIADS, as you demonstrate quite nicely. I miss a similar comparison to generic analysis packages (where the hierarchical terminology would be uploaded with the data) as well.

2. The VIADS seems to aim at being a tool for the clinician or manager of clinical departments. Still, the user has to extract and aggregate the data before uploading it to the web tool. This makes the tool less useful - it would be much easier to have it integrated in the EMR or population health management system, extracting the data directly from the medical records. Also, do the big EMR vendors like Cerner, Intersystems or Epic have the ability to traverse, filter and aggregate hierarchical data already? If so, the article should state that. If not, the authors would have demonstrated the usefulness of the tool even better, by describing that weakness of the big EMR suites.

3. The system uses several different libraries and programming languages to achieve its goal. I think the discussion might mention the challenge of version control of the modules in further development or local implementation.

4. In the description of the system, it would have been nice to see a few examples or scenarios, i.e. insights that the user might achieve by analyzing hierarchically, not just using the separate codes.

Otherwise, the article is nicely structured, the language is clear and the figures and tables makes it easy for the reader to understand how the system works. I like that.
**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:

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 organisation 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