Reviewer’s report

Title: Use Case Driven Evaluation of Open Databases for Pediatric Cancer Research

Version: 0 Date: 31 Aug 2018

Reviewer: Joseph Romano

Reviewer's report:

Jeanquartier et al present a study in which they assessed the quality and utility of existing pediatric cancer data resources via 5 use cases.

* Major comments

- Substantial edits need to be made to fix grammar and sentence structure.

- The Results section needs to be restructured. The goal of this paper is to take 5 (?) data sources and evaluate their ability to satisfy 5 use cases (summarized in Table 2). However, the Results section consists of a bulleted list with 7 items, where each item discusses zero (?) or more of the use cases with respect to zero or more of the 5 pediatric cancer data sources being considered. Furthermore, these use cases are not referred to consistently---in Table 2, it is done in a way that could easily be extended to the main text of the manuscript (e.g., "UC1", "UC2", etc...). Additionally, a bulleted list is unsuitable for describing results in long narrative form like this (subsections for each data source with dedicated paragraphs for each of the 5 use cases would be much more suitable). As it currently stands, I can't tell which of the use cases are being answered for each of the data sources / tools.

- I'm having some trouble with your conclusion, in which you state the data resources on pediatric cancer are lacking (compared to non-pediatric cancer resources). However, you only systematically assessed the pediatric data sources, so what quantitative basis are you using for stating that pediatric cancer resources are limited in comparison? The conclusion makes logical sense, but it needs to be demonstrated using data.

* Minor comments

- Some pediatric cancer initiatives are not discussed or mentioned in the paper. One that comes to mind is the Alex's Lemonade Stand Foundation and their associated projects.
- How did you arrive at the 5 use cases described in the Methods? They vary greatly from one another---are they meant to describe stereotypical questions posed by different types of stakeholders involved in cancer research and treatment, or something else? In other words, what specifically should the success of these use cases tell you about the data resources?

- Bold highlighting for pediatric cancer resources is used inconsistently in the Results section, making it hard to scan for these data sources.

- Use Case 5 (API availability) isn't really a use case. It's a feature.

* Overall opinion

The authors present a look at an important and rapidly growing area within cancer research and biomedical data science---namely, pediatric cancers. While the authors clearly describe their intent and explain the motivation behind their work, the paper does not fully convince me that their conclusions are justified. I suggest that the authors either relax their interpretation of the results, or bolster their methods to resolve the issues described above.

**Level of interest**
Please indicate how interesting you found the manuscript:

An article of importance in its field

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

Not suitable for publication unless extensively edited

**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

Do you want to get recognition for reviewing this manuscript? Add a record of this review to Publons to track and showcase your reviewing expertise across the world’s journals. Signing up is quick, easy and free!

No