Reviewer's report

Title: f-treeGC: A questionnaire-based family tree-creation software for genetic counseling and genome cohort studies

Version: 1 Date: 17 Apr 2017

Reviewer: Lori Ann Orlando

Reviewer's report:

This manuscript describes the development of f-treeGC, a questionnaire-based pedigree-creation software program built in Adobe Air.

I applaud the authors for undertaking this effort as there is a great need for tools to help facilitate family health history data collection, particularly for genomic studies. There are a number of nice features- most specifically the fact that it is compliant with the AHIC criteria for good quality family histories.

The authors however make significant claims about the accuracy and the facile nature of the tool without providing any data to support their statement. The manuscript guidelines for software recommend comparison to existing tools and I would strongly urge that this the tool be evaluated in a way that allows comparison of the features to existing tools. Even without a direct comparison some metrics regarding the tool could be provided. who has used the tool? where has it been implemented? how many patient's histories have been collected with it? How much data was collected? how long did it take (mean and range) to collect the data in clinical setting? Who is entering the data in the clinical setting? How does it compare to the data available in the chart? And so on... At the very lease present the data on validation of the tools accuracy.

There are a also some considerations that need to be addressed in order to determine if the tool will indeed be useful for genomic studies - outside of its ability to be used. Three that are most pressing are:

1) if you are limited to only 4 medical conditions per person how can this be enough data to adequately link to genomic information for analyses? In most countries the average number of medical problems (except for the very young) exceed 4. Is this different for your country? If so please provide data to support that. Even in this case it would limit the usefulness of the tool to your country.

2) how can the data be analyzed in relationship to genomic data if it is not codified in some way. It seems from the description that the problems are entered as free text. That creates problems with analysis as chronic kidney disease could be entered in any number of different
manner making it practically impossible to evaluate across individuals what conditions are present and compare to genome findings across populations. There are any number of data standards that could be used to help resolve this issue or you could use drop down lists to prevent items from being coded differently. Other solutions exist but I think if you want to say that this tool will facilitate large scale genome studies this has to be addressed.

3) Address the question of patient lack of knowledge regarding their family history. If the data is collected at the point of care without any guidance to patients about talking with relatives and what type of information they should ask about- the amount and accuracy of the family history is limited. This should be addressed - by further developing the tool or by listing it as a limitation as it is a huge barrier to the quality of family history data.

Are the methods appropriate and well described?
If not, please specify what is required in your comments to the authors.

No

Does the work include the necessary controls?
If not, please specify which controls are required in your comments to the authors.

No

Are the conclusions drawn adequately supported by the data shown?
If not, please explain in your comments to the authors.

No

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.

Not relevant to this manuscript

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.

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/](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