Reviewer’s report

Title: Consent to Data Linkage in a Large Online Epidemiological Survey of 18-23 Year Old Australian Women in 2012-13

Version: 1 Date: 10 Oct 2019

Reviewer: Reviewer 2

Reviewer's report:

PEER REVIEWER ASSESSMENTS:

OBJECTIVE - Full research articles: is there a clear objective that addresses one or several testable research questions? (Brief or other article types: is there a clear objective?)

No - there are minor issues

DESIGN - Is the current approach (including controls and analysis protocols) appropriate for the objective?

Yes - the approach is appropriate

EXECUTION - Are the experiments and analyses performed with sufficient technical rigor to allow confidence in the results?

No - there are minor issues

STATISTICS - Is the use of statistics in the manuscript appropriate?

Yes - appropriate statistical analyses have been used in the study

INTERPRETATION - Is the current interpretation/discussion of the results reasonable and not overstated?

Yes - the author's interpretation is reasonable
OVERALL MANUSCRIPT POTENTIAL - Has the author addressed your concerns sufficiently for you to now recommend the work as a technically sound contribution? If not, can further revisions be made to make the work technically sound?

Probably - with minor revisions

PEER REVIEWER COMMENTS:

GENERAL COMMENTS: -The authors have been commendably responsive to the reviewer's comments and questions. For example, they have shifted their analytical approach from a parametric regression model based approach to a non-parametric CART (Classification and Regression Tree) approach. Notwithstanding that CART is essentially the same mathematically, there are relaxed assumptions in CART and the output is generally easier to understand for non-specialists. There is a cost for that ease, however. For example, selection criteria to find the best "splitter" variables (e.g., "Importance") are generally less understood than typical regression model diagnostics.

- The authors have collected a nice bibliography of reference material about the issue of "consent to link" different types of data sources. The writing overall has trimmed some less important information and dare I say "simplified" the argument to the good of the paper's general interest. To my mind they have a very simple finding about who consents to link data that they are trying to present in a reasonably robust fashion.

- That said, the very simplicity of their finding begs the question of why this should be of interest. I am still a little fuzzy on how they constructed their CART model, what variable selection criteria were used and so on. That is easily expanded on in their methods section. One major point of confusion for me was not answered until I got to one of the last pages of the mss buried in the Declarations section, to find their explanation of how participants consented into the first part of the study, "...they were not asked for consent to participate because implicit consent is assumed through their completion of the online survey." Ahhh, now I see. I would have liked to see a more careful elaboration of how that phase of the study works because all throughout the mss, whenever they said "consent" I had to wonder about which consent was being referred to. So, these are participants who are already deep into an online survey, finally asked for (extra) permission to link their answers to an administrative database? Finally, I confess additional confusion because isn't a longitudinal epi study going to pretty much insist on several ways of re-establishing contact over time for inclusion in the cohort? And that is what? Where you live, right? Maybe that is obvious to others. Or maybe this is all just part of the epi study recruitment process and these participants are not actually in the study yet until and unless they agree to the data linkage?
REQUESTED REVISIONS:
As I mentioned above, I was a little confused about which level of consent was being discussed. That should be clarified early in the ms. I would like to see some more of the CART details, split and pruning and diagnostics spelled out in the text.

ADDITIONAL REQUESTS/SUGGESTIONS:
As I mentioned above, especially with regard to clarification of consent level which is being modeled, and CART analysis details.

-In addition, I would like to see something like an "Objectives" section in the Abstract which clearly spells out the goals of the paper. This is different from the "Methods" section, which is where the mention of CART should first be introduced, not in the Results section.

-Maybe it is there and I overlooked it, apologies, but I would like to see a statement(s) along the lines that makes it clear if what they are talking about ("consent to data linkage") is actually their shorthand but very "weedy" way of referring to consent to enter the longitudinal cohort.

-I would carefully look at the language throughout but especially in the Conclusions section of the abstract to better qualify some of their assertions. For example, they state "Ensuring that prospective participants [of what, the data linkage part or the longitudinal study enrollment? Is that the same thing?] understand the consent and privacy protocols in place to protect their confidential information [the authors tell us at the end of the paper that they actually DON'T have a consent discussion?] builds confidence in the research process." I would probably take a fresh look at how best to word their remarks throughout about what consents are being talked about both in the field and here in the paper. Otherwise, it was an interesting read and I wish the authors well!

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.

Unable to assess
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 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