Reviewer's report

Title: Case Identification of Mental Health and Related Problems in Children and Young People using the New Zealand Integrated Data Infrastructure

Version: 0 Date: 19 Aug 2019

Reviewer: Hugo Leroux

Reviewer's report:

This paper describes research into case identification method for research into the mental health of children and young people in New Zealand. This is a very well-written paper that is substantiated with really good evidence.

My main concern is that the authors present the data sets used and mention the "algorithm" used to link and reason over the data. However, the algorithm is rather implicit in nature and is not presented clearly to the reader. For example, on page 16, lines 25-32, the authors state: "The algorithm identified 84,000 individuals …. This algorithm is not designed to estimate prevalence …". This may well be true (and I don't dispute this fact), however, the content of the paper provides no guidance as to how this is achieved.

Furthermore, on page 17, line 20, the authors state: "… used a careful and transparent process to assign codes …". It would be useful to the reader if they would indeed explain the process to make it more transparent.

Page 6, line 43: "… according to Statistics New Zealand protocols …". Please state the protocols or include a reference (I'm assuming that it's "Statistics New Zealand 2017c")

Page 16, line 46: please insert a comma after "compose" and after "far" - i.e. "compose, by far, most of …"

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 recommend additional statistical review

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