Reviewer’s report

Title: The EMPCAN Study: Study of The Evolution of The Socio-Economic Position of Workers With Cancer. Study Protocol of a Population-Based Cohort Study.

Version: 1 Date: 22 Nov 2018

Reviewer: Justine Slomian

Reviewer's report:

Ok, the Methods section is now clearer in the abstract but not in the manuscript.

My major concern is that the Methods section is still not well written. I mean that this section does not follow the recommendations in terms of protocol publication. The authors said that they completely reviewed the methods section but nothing has been changed except the added first paragraph.

I think that the first paragraph is helpful. We have a lot of data in the Methods section but it is not always comprehensive because of the lack of structure (headings and subheadings are needed). For example, I asked to add a separate section on the study population with inclusion and exclusion criteria. The author answer was: "We believe that the lines 205-213 describe this inclusion/exclusion criteria + the figure 4". The lines 205-213 are: "The main points under discussion concerned the definition of the cohort (inclusion and exclusion criteria) and the choice of the variables which were the most appropriate regarding the study objectives, i.e. the discussion of the proportionality principle. For the technical and ethical aspects, we received advices and recommendations of lawyers from the governmental platform eHealth. As there was no precedent of such data coupling in Belgium, these different steps of coupling and controlling privacy, required more than two years to finally submit our demand to the national Privacy Commission, which provided its positive opinion in January 2016". I do not understand how these lines are a presentation of the study population. Perhaps the authors meant more precisely about the lines 223-227: "The BCR started the process, with the selection of patients diagnosed from 1st January 2004 to 31st December 2011 included, with: head and neck cancer (C00-C14 and C30-C32); lung cancer (C34); colorectal cancer (C18-C20); breast cancer (C50); corpus uteri (C54); prostate (C61) and testis cancer (C62). Among them, the BCR only retained those aged 16-64 at the date of incidence". Nevertheless, there are not enough details for a protocol. In addition, I cannot read figure 4 because of the picture quality.

Here is an example of the data (corresponding to headings and subheadings) that should be found in a protocol:

* Description of the study
* Study Population
  o Approximate Number of Subjects / sample
  o Inclusion Criteria
  o Exclusion Criteria
* Recruitment
I agree that there are a lot of these data in the actual manuscript but there is a lack of structure that is essential for the presentation of a protocol study.

In addition, some parts of the manuscript should be more detailed as for example:
* Lines 209-210: For the technical and ethical aspects, we received advices and recommendations of lawyers 210 from the governmental platform eHealth. => Which advices and recommendations? It should be presented in the protection of human subjects section.
* Lines 215-216: After this authorization, the process of data extraction, exchange, collection, anonymization, etc. started (Figure 2) and the final files were made available for the researchers in October 2017. => This is the typical information that should be found in the data collection section but more practical details are needed.

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

An article whose findings are important to those with closely related research interests

**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

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