Reviewer’s report

Title: Weight loss as a predictor of cancer and serious disease in primary care: an ISAC approved CPRD protocol for a retrospective cohort study using routinely collected primary care data.

Version: 0 Date: 16 Apr 2017

Reviewer: Katherine Morley

Reviewer's report:

The research questions the authors plan to address are interesting and worthwhile, and I am in principle I strongly supportive of the publication of protocol papers for electronic health record studies. However the manuscript in its current form doesn't provide enough information about certain aspects of the planned work and does not follow the journal Protocol format. Specific comments are listed below:

1. Most of this text appears to have been copied directly from the original ISAC application (see line 49, page 3 for concrete example) and does not follow the journal-specified format for a Protocol. Related to this, under the heading "Data Linkage Required" the authors list requested linkages, but it is unclear whether these have been approved and will form part of the data set. There are many more instances, but I do not list them all here.

2. I note that the authors are using a validated set of readcodes for their cancer endpoint (please insert the appropriate reference for this as it is currently missing). However, it appears that they propose the development of an EHR algorithm for weight loss - what steps do the authors propose to take in order to validate the algorithm developed? If I have understood correctly, the authors also appear to be proposed the development of another algorithm for what they have termed "Serious Disease", but it was not clear what conditions might be included in the endpoint, or how this algorithm might be validated. It was also not clear to me which conditions might be included under "co-morbidity" covariates.

3. The cohort definition was somewhat unclear given that different time windows are proposed for different variables. I appreciate that extracting relevant information from EHR is complex and there will be rationales for the authors decisions, but these are currently hard to follow. A clear statement of inclusion/exclusion, cohort entry, outcome/exit definitions is needed, as well as the time windows used for each set of variables - a diagram may be helpful here?
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 organisation 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