Reviewer's report

Title: Predictors of working beyond retirement in older workers with and without a chronic disease - Results from data linkage of Dutch questionnaire and registry data

Version: 0 Date: 02 Jun 2017

Reviewer: Matthias Bethge

Reviewer's report:

The draft describes a study on predictors of working beyond retirement in older workers with and without a chronic disease. Data come from a large cohort study which uses both administrative data and survey data. The topic is in the scope of the journal. However, I have major concerns with the statistical analyses in the paper as they are not appropriate to answer the main research question. There are also some other reporting issues. Nevertheless, I am confident that these issues can be resolved when revising the draft.

1. The introduction is concise and well written (maybe the authors can avoid the repeated use of "To illustrate,...").

2. The methods section is clear when describing the data, the study population and measurements.

3. The study aimed at predicting working beyond retirement in persons with and without chronic disease. The authors tried to explore if predictors are different and finally concluded this from their analyses. However, they did stratified analyses. I strongly advise to calculate one model with chronic disease as a moderate or to calculate one model with segment-specific estimates (and then to test if the estimates are different). This will clarify if predictors really differ.

4. Procedure of model building is not fair for comparing patients with and without chronic diseases. The sample of the patients without chronic diseases is smaller. Effects need to be larger in order to be included in the final model.

5. Odds ratios do not express a probability.

6. Model fit is usually described by a pseudo R2 in logistic regression. Please identify which measure you used. A pseudo R2 is not a measure of explained variance (please check the whole draft).
7. As you use the area under the curve for describing the accuracy of the prediction model you may provide thresholds for good accuracy and explain the measure.

8. Please identify the software package.

9. Page 15, discuss: Furthermore, his study? this study?

Table 1: Perceived health: Very good for a quarter of the persons with chronic diseases. In patients without chronic diseases probably nearly all persons stated good health? The note of the table states that the number cannot be shown as the number is below the privacy-based norm if Statistics Netherlands. Would it be better to use a different approach for categorizing health in order to avoid that are "all" in one category? Also for model building?

Table 1-2: "Good health" instead of "(very) good health" as you categorized excellent, very good and good in one category.

**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.

Yes

**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.

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