Reviewer’s report

Title: Selection Bias and Subject Refusal in a Cluster-Randomized Controlled Trial

Version: 0 Date: 10 Feb 2017

Reviewer: Ann Christina Foldenauer

Reviewer’s report:

The authors presented a well written manuscript on selection bias and subject refusal in the cluster randomized ICARE trial. While methods and limitations are well described, the statistics is not adjusted by the cluster effect leading to biased test results. Therefore, I would like the authors to consider the following points:

1) To analyse the hypotheses, use a cluster-adjusted test for proportions and an adjusted t-test [or use a linear mixed model (eg. fixed group effect, random cluster effect) with linear contrasts. E.g. using SAS via proc mixed]. Methods are available and described in the following book, which is the standard reference for the analysis of cluster randomized trials:'Design and Analysis of Cluster Randomization Trials in Health Research.' by Allan Donner and Neil Klar.

2) It would be of interest, to analyse if one of the clusters showed extreme results due to your remark in the limitations part (p.13 l. 16). I would like the authors to add graphics and/or estimates as supplementary material for the parameters presented in table1 for each cluster distinguishing between control and intervention group. The most interesting differences could then be added in the manuscript.

3) Referring also to (2): I strongly recommend reformulating part of the limitations part, since it seems to imply, that physicians, coordinators and sites have biased the data. Either you know a source of bias and should present it as such (E.g. p.13 l.16: Did a site only enlist hypertensives or not?) or you should formulate the sentences in a way that it remains an open problem (E.g. p.13 l. 18: Did study coordinators state explicitly the use of their knowledge on patient data, or can you simply not exclude the possibility of physicians doing so consciously or unconsciously).

4) The mean should always be presented with standard deviation. E.G: Mean ± SD or as mean (SD). This should be corrected for table 1 after reevaluation.

5) I like that you tried to account your analysis for multiple testing, but it would be more adequate to use adjustment methods which control the family-wise error rate and keep a predefined global significance level, such as Hochberg, Bonferroni, …. You could also re-adjust the p-values with these methods instead of the significance level. E.g. in SAS using proc multtest.

6) Since there are different definitions of selection bias, you could also consider discussing third-order selection bias, developed by Blackwell and Hodges (1957). See ‘Randomization in Clinical Trials: Theory and Practice.’ By W: Rosenberger and J. Lachin.
7) State the statistical software and OS you used in the analysis section.

8) P.9 1.25: The study monitor only choose 20-30% patients for the quality check. Were they chosen at random? If so, add it in the manuscript. This should be discussed as source of bias as well.

**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.

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