Reviewer’s report

Title: Sex differences in dementia: On the potentially mediating effects of educational attainment and experiences of psychological distress

Version: 1 Date: 02 May 2020

Reviewer: Christopher R Beam

Reviewer's report:

The authors "seek to shed further light on the complex and multifaceted relationship between dementia and social inequity by exploring the extent to which the effect of female sex on dementia is mediated by differences in levels of educational attainment and experiences of general psychological distress." This is an important question, which has not received as much as attention was one would assume prima facie. As a reviewer of this resubmission, my aim was to improve the impact of the study. All critiques are offered in the spirit of strengthening the authors presentation of their study.

In the opening paragraphs, it is true that women have greater lifetime risk of dementia, including in Sweden (see Beam et al., 2018). The authors, however, cite that dementia risk is greater in women than men at age 65, but this might not be the case even in Sweden. Beam et al. (2018) found that age 65 rates of dementia risk in a Swedish sample of men and women were nearly identical. The authors are encouraged to acknowledge these differences in Sweden.

Page 1, line 48, the sentence starting with "Today" introduces a new line of thought about effects on educational disparities, rather than causes, and so should begin a new paragraph.

The authors posit that depression is a plausible risk factor of dementia, a view not widely supported including in the Study of Dementia in Swedish Twins (Brommelhoff et al., 2009). While the authors acknowledge that reverse causality is a plausible rival hypothesis, they are strongly urged to consider depression as a prodromal syndrome of dementia.

Karlsson et al., 2009 and Karlsson et al., 2010 (p. 3, line 24-26) are not indexed properly, that is, they are in APA Style. This occurs elsewhere in the manuscript (e.g., p.4). The authors should review the manuscript for style used by BMC Psychiatry.

It is unclear whether excluding the 23 persons with MMSE scores less than 24 reduces the possibility of reverse causality. The listwise deletion of these cases may introduce unnecessary bias in the study (namely MNAR missing data mechanism), as these persons probably are not a random sample of the population sample. What proportion of these persons were eventually diagnosed with dementia? And does the rate significantly differ from the rest of the sample? If the rate reflects the base rate of diagnosis in Sweden, it seems reasonable to include them in the analytic sample.
How did the authors handle the APOE E4 e2/e4 allele configuration? Please justify the decision to include/exclude these cases.

Given how loneliness was coded, the authors should characterize loneliness as chronic loneliness. For example, persons who report feeling lonely for less than 5 years are, nevertheless, still lonely at time of data collection.

What proportion of data were missing? Can the authors supply arguments (i.e., univariate t-tests; Enders, 2010) for whether the missing data mechanism is MAR or MNAR? That is, what are the correlates of missingness in the model. And were these correlates included in the imputation model? What was the imputation procedure used (e.g., data augmentation or multiple imputation with chained equations)? Data augmentation assumes that all variables have the same distribution, but this seems an unwise decision given the mixture of nominal and ordered categorical responses included in the data table. How many iterations were required for the missing data algorithm to converge? Finally, what was the rationale for using only 5 imputed data sets? The general consensus is to use as many imputed data sets as the mean percentage of missing data in the data table. Increasing the number of imputed data sets will improve precision of the standard errors and may encourage confidence that hypothesis tests were unbiased.

TLI and CFA both are incremental fit indexes. Because they are highly correlated, only one should be reported (typically the TLI is more conservative and preferred).

The modification indices seems to suggest that a one-factor model may be too simplistic, as including residual correlations suggests a multifactor solution. The authors might present an exploratory factor analysis to identify whether a single factor solution is adequate. One possibility is that the latent variable is not really one "psychological distress" variable, but rather two variables that may index social connection and depressive symptomatology.

Age and cohort appear to be conflated. How was cohort taken into account? It also seems strange to include age as a mediator? Might it be worth testing whether age is a moderator of the effect of education on dementia rather than mediator?

Please report unstandardized regression parameters from probit analyses. In analyses with binary dependent variables and binary covariates, it is easier to interpret unstandardized results, which do not depend on the variances of the covariates.

In the tables, please report point and interval estimates rather than point estimates and p-values. Probability values do not allow for conclusions to be drawn about the precision of the point estimates.

The authors suggest that lack of power may have caused effects of distress on dementia may have vanished when time to diagnosis of dementia was increased. What was the overall power to detect hypothesized mediating effects given the sample size?

The discussion section mainly summarizes the findings presented in the results. Could the authors provide implications of their finding that age and psychological distress mediate effects of female
gender on dementia risk whereas education does not? Further, what are the implications that education mediates effects of female gender on psychological distress, which then has a direct correlation with dementia risk?

Thank you for the opportunity to review your work. I look forward to seeing it published in the future.

**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.

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:

Needs some language corrections before being published

**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.

1. No
2. No
3. No
4. No
5. No

Reviewer Publication Consent. 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) if this manuscript is accepted for publication. Any comments that I do not wish to be included in the published report have been included as confidential comments to the editor, which will not be published. If you are not happy for us to publish this report, please contact the editorial office before completing the review. If you wish, you can include your name in your published report. Please note you must decide whether to include your name at the start of the process and confirm this decision whenever prompted. If you change your decision later, we will not be able to publish your name.

I agree to the terms of the CC-BY 4.0 license; please publish my name with my report.