Reviewer's report

Title: Relationship Between Multimorbidity, Demographic Factors and Mortality: Findings from the UK Biobank Cohort

Version: 0 Date: 05 Oct 2018

Reviewer: Dexter Canoy

Reviewer's report:

This study investigates the prospective relation between multimorbidity and mortality. Using data from the UK Biobank involving 0.5 million men and women, the investigators identified the number of long-term conditions (from a list of 43 conditions) the participants had at baseline. Over a median follow-up of 7 years, the study accrued over 14,000 deaths. the investigators reported that those with 4 or more long-term conditions had a higher risk of all-cause, vascular and cancer mortality than those with less numbers of long-term conditions, suggesting that multimorbidity is associated with increased mortality risk. Many reports have looked at multimorbidity in relation to mortality, but this study covers the 'middle age' range whereas others mostly included the older age range. This study also includes large numbers of events. However, the study does not go beyond what many have examined in terms of what predicts an adverse outcome - the number of long-term-conditions, a simple parameter to use but does not shed light on the underlying determinants or aetiology (hence, the clinical impact is difficult to characterise). This approach limits examination of any interactions between diseases, or the timing/temporal sequence of the accumulation of the conditions which may have an impact on the outcome.

It would be useful to know the distribution of broad types of conditions (e.g. cardiometabolic, cancer, etc) across the number of long-term condition categories). Some conditions (e.g. cancer, cardiovascular conditions) are likely to have other coexisting conditions (therefore classified in categories with higher number of long-term conditions) and strongly associated with mortality particularly in the short term (and impact likely to persist despite various adjustments). Considering that both cancer and vascular conditions are prevalent in the UK Biobank population and associated with other conditions, these conditions are likely disproportionately represented in those with 4 or more long-term conditions and therefore this group would be strongly associated with subsequent mortality. Ref 31 alone shows a very strong impact of the combined effects of three cardiometabolic conditions - suggesting perhaps that the types of coexisting conditions, rather than just the number, are more important and informative clinically?

The authors refer to the multimorbidity report of the Academy of Medical Sciences (ref 1). I suggest that the authors adhere to the guidance and research framework set out by this report to allow meaningful summary, integration and interpretation of findings and reduce fragmentation as a result of lack of consistency in the various aspects of conducting research into
multimorbidity. For example, I suggest that the definition of multimorbidity for this study is aligned to that suggested by the report.

Although the authors cited ref 31, they did not discuss this study which, in my view, is relevant. Although the conditions included in the study were limited, they used a similar resource (UK Biobank). I suggest that a comparison (of similarities/differences) is made and demonstrate why this current investigation adds new insight about multimorbidity.

In survival/time-to-event analyses, it should be useful to provide the context of the follow-up. Dates of beginning and end of follow-up would be useful. What was the underlying time variable?

In relation to the comment on defining follow-up periods, it seems to me that the underlying time variable is follow-up duration, and the author adjusted for 'age' in the model. Since we are dealing with mortality, age-specific death rates increase dramatically in older ages, and simply adjusting for age may not be sufficient to control for age differences across the categories of numbers of long-term conditions. I think it is more valid to model Cox regression using age as the underlying time variable, so that event rates are compared at the age of the event (after all, when age-specific mortality rates are typically reported such as in censuses, these rates are calculated using age at the time of death, not the baseline age). [Consider those who entered the study at age 37 years, and followed for 7 years (which is the median duration of follow-up): they would just be 44 years old and the mortality rate at this age would be low. Yet for those aged 65 years at baseline and followed up for 3 years, at age 68 years, the mortality rate for this age group is going to be higher.]

The difficulty in investigations into the prognosis of multimorbidity is how to adjust for confounding factors appropriately. Presumably, given the classical definition of what confounding is, different long-term conditions (in relation to the different outcomes) would have different confounding factors. Thus, in this type of multimorbidity research, this source of residual confounding is a problem. How did the authors minimise the impact of residual confounding in the analysis?

In the sensitivity analysis, the authors adjusted for 'other long-term conditions' when estimating risks associated with the number of long-term conditions. I do not understand how could this work. Because the exposure variable simply counts the number of long-term conditions regardless of the type of condition, adjusting for other types of condition could mean that what one adjusted for could be a condition included in any of the exposure category and, therefore, an overadjustment.
In another sensitivity analysis, the investigators defined multimorbidity using data from Hospital Episode statistics, rather than relying on self-reported conditions. However, there was no detail about how multimorbidity was defined using hospitalisation data (many of the 43 conditions considered are probably not the conditions that may be captured from hospitalisation data). It is important to provide more details in terms of the method used. As the date of exposure would have occurred after baseline, how was the duration of exposure quantified?

The survival plot seems to be unadjusted for age. I'm not really sure how useful an information is showing survival plot in relation to the exposure, when the exposure of interest is strongly correlated with age.

Please clarify the Townsend score which seems different in the text (page 8, line 168) when compared to the values reported in table 1.

In presenting results, the numbers of events for each category of multimorbidity should ideally be reported (not just the total number of the outcome).

In reporting the absolute event rates, was this age-adjusted?

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.

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.

Statement on potential review bias
Please complete a statement on potential review bias, considering the following questions:

1. Did you co-author any publication with an author of this manuscript in the last 5 years?

2. Are you currently or recently affiliated at the same institution as an author of this manuscript?

If you can answer no to all of the above, write 'I declare that I did not publish with these authors in the last 5 years and also meet the affiliation criteria". If your reply is yes to any, please give details below.

I declare that I did not publish with these authors in the last 5 years and also meet the affiliation criteria.
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