Reviewer’s report

Title: Representation of people with comorbidity and multimorbidity in clinical trials of novel drug therapies; an individual-level participant data analysis

Version: 0 Date: 29 Jun 2019

Reviewer: Christiane Muth

Reviewer's report:

Manuscript number BMED-D-19-00712
Title Representation of people with comorbidity in clinical trials of novel drug therapies; an individual-level participant data analysis.

Summary:
This paper presents - to my knowledge - the first comparison of comorbidity counts and patterns between randomized controlled drug trials and a primary care database, which is based on an IPD meta-analytic approach. The authors have to be congratulated, as their investigation is related to one of THE key issues in modern health care (whether or not results from clinical trials are applicable to real world patients) and - at the same time - their methodological approach is a novelty and posed an enormous challenge on the investigators.

Major strengths: The methodology is overall well described, it contains information about the source of data, participants, outcomes, and statistical analysis methods. The manuscript is well written in a concise manner, the tables and figures presented are well selected from a large amount of results and the complexity of the study is explained in a digestible way to the readers, which may not always be familiar with the methods applied.

Comments
1. The authors provided a large (and with >100 pages somewhat bulky) supplement with further details about methods and results. I suggest partitioning it into smaller sections, which can be picked up easier (and more specifically) by interested readers.

2. The following information is inconsistent between manuscript, appendix and PROSPERO registry: the manuscript tells about two data repositories of clinical trial data while appendix and registry list three. Although, the exclusion of all trials from the third registry was reported at the supplement on p. 11, I suggest a clarification in the manuscript.

3. The search and selection process of trials as well as of conditions and medications is difficult to understand. What I understood from the appendix is that it was a forward-backward procedure where some of the difficulties with repository data could not be foreseen and procedures had to be adapted (e.g. about the way data are 'held in separate data "safe havens"' (supplement p. 51); about 'redacting of diagnosis' in trial data etc.). Although the authors do not claim to have drawn a systematic sample of
studies, these difficulties should be more recognizable in methods section, as most readers may expect a widely standardized and straightforward methodology from IPD-MA's. Furthermore, the authors presented a novel methodological approach including repository data and primary care data. In particular, I would highly suggest the presentation of a flowchart (or a tabulation) of the search and selection process and its discussion.

4. I missed a more differentiated discussion of certain comorbidities. The authors provided rich results on 22 distinct comorbidities but rather focused their discussion on disease counts. While the latter is a typical measure in epidemiologic studies on multimorbidity, clinicians are more interested in comorbid conditions or patterns of these conditions itself. Where certain (distinct) comorbidities more frequently excluded / underrepresented in clinical trials than others (not explained by low prevalences)?

5. The authors pointed out that not all sponsors share data at all or of all of their studies. Is there any information whether non-sharing of data might have a similar impact on meta-analytic results like a publication bias? If there is no evidence available, what would the authors suggest about further research regarding this potential source of bias?

6. I missed the following information in the manuscript:
   a. Methods, searches: What was the start and end date of searches in the registries / repositories?
   b. Methods, inclusion / exclusion criteria: was a restriction to industrial funded studies a priori intended?, What was the definition of 'new drugs'?, Was the cut-off 'trials conducted after 1990' applied and if so, why was this year chosen?, How did the authors define 'chronic conditions' (did they develop an own list or did they use an existing list)? The authors restricted conditions to 'medical' conditions in the manuscript and to 'medical and urological conditions' in the supplement - perhaps, it would be helpful to avoid these rather broad categories but report the number and names of the included conditions specifically.
   c. Methods, community sample: why was the time span between 1st January 2011 and 1st January 2012 chosen?
   d. Methods, statistics: what was analyzed to compare included trials with other trials not included for several reasons (see lines 171-174) and what were the 'other trials'? This is also important to understand Tables S21 to S24 and Figures S3 and S4 (supplement).
   e. Results, Table S17 (supplement): I suggest presenting this table with the main text (or at least as a separate appendix) and to order and group the trials in accordance with the presentation of the trials in Figure 3.
   f. Results, Figure 3: I suggest to present trial identifiers (identical to Table S17) in Figure 3. The ratios of mean count between community patients and trials vary widely between trials of the same indication. Clinicians may want to identify the evidence, which may be more applicable to their patient with multiple conditions (and can look up trials with higher comorbidity counts presented in Figure 3).

Limitations of my review:
I am not an expert in the statistics applied in this study and do not feel confident to appraise the appropriateness of statistical analyses therefore.

Kind regards,

Yours sincerely, Christiane Muth
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.

Not applicable

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 recommend additional statistical review

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