Reviewer’s report

Title: Resource implications of preparing Individual Participant Data from a clinical trial to share with external researchers

Version: 0 Date: 17 Apr 2017

Reviewer: Marc Buyse

Reviewer’s report:

Sharing of patient-level clinical trial data will, inevitably, become standard practice soon. The cost of, and steps required for, preparing datasets ready to be shared on some platform has not been studied, hence this paper fills a gap. The paper is nicely written.

The biggest limitation of this paper is that the cost estimates described, albeit very detailed, may be substantially different in other environments, whether publicly funded or not. In this respect, it is extremely unfortunate that the only figure coming from industry is from a personal communication stating that "an average of 7 days" is typically needed to prepare data for sharing, without any detailed breakdown between the different tasks involved. It is probably the case that some tasks vary enormously in length, especially (as pointed out by the authors) for legacy trials that were conducted a long time ago, and for which data access is far more time-consuming, when at all possible. While minor differences between the tasks and times required to prepare the data for the SANAD and MENDS trials are here described in exquisite detail, it would have been far more useful to discuss widely different trials for which the tasks and times might have differed more substantially. Along the same lines, this paper describes publicly funded trials carried out in the UK. It would have been nice, if at all possible, to quote figures (if available, even via personal communication) from other clinical trial data centers in Europe and the US.

One interesting aspect of this paper is to provide guidance on the steps required to prepare a "data sharing pack" as well as to anonymize the data for sharing. It would be quite useful, in my view, to add a standard table of contents for the data sharing pack (in addition to the steps outlined in tables 1 and 2). Data sharing is a new paradigm, hence any thoughtful guidance on how to operationalize it is welcome at this stage.

The paper ends with a recommendation to include the costs of data sharing in a trial budget. It is comforting to see that these costs will generally be small compared with the overall costs of a trial, such that costs will not constitute a major obstacle to data sharing in the future. To put things in perspective, could the authors disclose the total costs of SANAD and MENDS?

Level of interest
Please indicate how interesting you found the manuscript:

An article whose findings are important to those with closely related research interests

Quality of written English
Please indicate the quality of language in the manuscript:

Acceptable

**Quality of figures**

All images and figures within the manuscript should be genuine i.e. without evidence of manipulation. No specific feature within an image may be enhanced, obscured, moved, removed, or introduced. If you have concerns about the veracity of the figures you should choose the first option below.

**Statistical review**

Is it essential that this manuscript is seen by an expert statistician? If so, please give your reasons in your report.

**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 am a shareholder and employee of International Drug Development Institute, a company that may indirectly benefit from patient-level data sharing.

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