Reviewer’s report

Title: Early warnings and repayment plans: novel trial management methods for monitoring and managing data return rates in a multi-centre phase III randomised controlled trial with paper case report forms

Version: 0 Date: 04 Nov 2018

Reviewer: Doris Lanz

Reviewer’s report:

The authors make a very good case for the need a) for more evidence-based methods for improving data returns, and b) for making these improvements sustainable across the life of a study; all too often we approach data returns as binary problem, or in the worst case as emergency exercises ahead of data cut-off.

The approach is presented and discussed very clearly, and its applications and limitations are laid out very thoroughly; the paper also eloquently accounts for the less quantifiable aspects of managing such processes. The paper merits consideration by a wider trial management audience. My suggestions below are largely minor and should be viewed as suggestions to clarify and potentially enrich some of the findings.

The "repayment plan" approach strikes me as a key element of the presented method. Perhaps it would merit a little more detail, in either the Methods or Results section, e.g. by illustrating this with a practical example.

The Results section is somewhat sparse compared to other sections; I would be interested here in the structure of data reminders and general communications. For instance, you say that 10 sites were identified with diminishing DRR rates, and that they were contacted initially saying only that they hadn't submitted forms recently. Were all these sites contacted again at every subsequent extract / CTU meeting, or only once a change in trend had emerged? Could you give some (perhaps more descriptive, qualitative) detail on how these communications progressed? Given that I would expect you to be in communication with the remaining 25 sites in some form, this would help to isolate the nature of the "intervention" a bit more clearly. It might also serve to illustrate how sustainable the intervention was during its limited time of application, and to help the reader assess how much effort is involved in producing a DRR increase per site.

While this may have been outside the remit of the project, but I wonder if any effect on data quality (not just completeness) was seen, e.g. in terms of number of data queries required. My expectation would be that as sites that are lagging behind are not required to submit all overdue data immediately, this results in less of a rush to submit data and might contribute to relative better data quality; are the authors able to comment?
In the discussion, the authors mention the increased efforts required for communicating with sites e.g. through telephone. I wonder how these efforts compare with the more traditional, static model of data collection reminders?

Conclusions: I fully agree that the method presented merits further exploration. Perhaps the approach would even lend itself to a potential SWAT (study within a trial), randomising between the different data reminder methods?

Minor comments:

The abstract would benefit from a little more context: word count permitting, I would suggest adding e.g. that the particular trial entailed a very long-term follow-up, and that the 10 sites identified for DRR intervention were among 35 total.

Escalation policy: Are you able to quantify what would have constituted "consistently" (line 178) or "prolonged" (line 179) to warrant an escalation?

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

An article of importance in its field

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 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

Were you mentored through this peer review?

No