Author's response to reviews

Title: An Audit of Sample Sizes for Pilot and Feasibility Studies being undertaken in the United Kingdom registered on the United Kingdom Clinical Research Network database

Authors:

Sophie A.M Billingham (sbillingham1@sheffield.ac.uk)
Amy L Whitehead (a.whitehead@sheffield.ac.uk)
Steven A Julious (S.A.Julious@sheffield.ac.uk)

Version: 2 Date: 22 March 2013

Author's response to reviews: see over
Reviewer's report

Title: An Audit of Sample Sizes for Pilot Studies being Undertaken in the United Kingdom

Version: 1 Date: 26 November 2012

Reviewer: Lawrence Mbuagbaw

Reviewer's report:

- Major Compulsory Revisions

1. The title is misleading hence my second point. Audit may not be the most appropriate term and according to the authors this paper is about pilot and feasibility studies. The title/abstract should have some component of study design.

   COMPLETED – we have kept in the word audit but have reworded the title to include feasibility studies and the fact the data is from a database

2. The research question is also unclear. The objectives in the abstract differ from those in the main manuscript.

   COMPLETED – abstract re-written

3. It is unclear why there is a distinction between pilot and feasibility studies, as the use of the term may be interchangeable or the authors may not adhere to the definitions reported in the article.

   COMPLETED – more clearly expressed in paper

4. The format of the paper doesn’t follow BMC recommendations. There are tables in the text. The subtitles and title page should be modified.
Complete – tables are now at the end

5. The search in not described in sufficient detail to be replicated and data extraction is not done in duplicate. The search terms should have been more specific. This would have prevented them from getting all the irrelevant hits.

COMPLETED – search described more thoroughly - could not search pilot OR feasibility as not possible to use Boolean operators in the UKCRN search. Used pilot and feasibility study to maintain consistency with previous work.

6. Statistical tests should be used to show if the sample sizes are different in groups.

RESPONSE – we did do statistical tests – the P-values ranged between 0.2 and 0.4 but we have not included. This is not because they are not statistically significant but because the study was not intended as one formally test hypotheses

7. The conclusions seem to be a repetition of the results. How can these results help other researchers designing pilot studies?

COMPLETED – Conclusions edited

8. No limitations are stated for this research.

COMPLETED – entered into discussion

9. Are these results relevant outside the UK?

RESPONSE – this of course is in the eye of the reader. There is no reason why not under certain provisos

10. There are no recommendations.

COMPLETED – the intention was never to make recommendations per se but to capture the reality of what sample sizes people use for pilot and feasibility studies. We do make one brief recommendation now in the conclusions
- Minor Essential Revisions

1. There are many spelling and grammatical errors throughout the manuscript.

For example in the abstract:

a. Pilot studies were found to have a smaller sample size per arm (median = 30) then (THAN) feasibility studies (median = 38.)

COMPLETED

b. Public and Industry funded pilot trials (TRIALS) appear to be of an equal size (both medians = 30.)

COMPLETED

2. UKCRN and ISRCTN should be written in full in the abstract.

COMPLETED

3. The URL for the websites searched should be provided [UK Clinical Research Network (UKCRN)].

COMPLETED

4. Normal and binary should be replaced with continuous and dichotomous; likewise survival should be replaced with time-to-event.

COMPLETED

5. Differentiate between sample size (what was achieved) and target sample size (what they hoped to achieve)

COMPLETED

6. The tables are titled mean sample size, yet means and medians are reported. The authors should specify why they chose to report both means and medians
(table 3). Average is also used in some tables (table 4).

**COMPLETED** – table titles changed, only medians presented

7. “Observational trials” are a misnomer. Trials are typically not observational.

**COMPLETED** – observational here is non-randomised studies which we exclude

8. Data are not presented consistently in the article. It should be n (%) or median (range) or mean (Standard deviation).

**COMPLETED**

- Discretionary Revisions

1. The data in the appendix can be presented better, such that the tables are smaller and there are less zeros.

**COMPLETED** – deleted tables in appendix.
Most pilot and feasibility studies did not justify their target sample size. However, it is important to provide rationales for the minimum number of patients needed to meet the objectives of those studies even though they are not the main studies. In this paper, the authors reviewed the target sample sizes of on-going and eligible pilot/feasibility studies in UK. They aimed to assess the sample sizes by the type of endpoints, funders, interventions, etc. The findings are valuable to researchers in understanding this issue.

- Major Compulsory Revisions

1. The design is appropriate but the methodology can be improved. The methods were insufficiently described.

a. The authors relied on a single trial registry in UK to search for eligible studies. I wonder what other efforts the authors had made to identify the studies that were not registered but potentially eligible for inclusion? Could that potentially change the conclusion?

COMPLETED - LISTED AS A LIMITATION

b. The authors only used two key words in the search: ‘pilot’ or ‘feasibility’. Would that miss some pilot/feasibility studies that were framed as, e.g. ‘experimental’, ‘exploratory’, ‘test’, ‘try-out’, or ‘preliminary’ studies?

COMPLETED - LISTED AS A LIMITATION – however, the search was conducted in this way to maintain consistency with previous work carried out

c. It was unclear if more than one reviewer conducted the search independently. If so, what was the agreement on retrieving the studies among reviewers? How was any disagreement resolved?

COMPLETED - LISTED AS LIMITATION – only one reviewer conducted the search
d. The authors did not provide a statistical plan for their analysis. For example, what statistical test would they use for testing if there was any difference between the type of funders/endpoints?

**COMPLETED -** The analysis was undertaken by summary statistics. Medians and ranges were estimated overall for the different types of trial and then broken down by factors such as type of endpoint, trial population being investigated and whether the trial was public or industry funded. Pg 9, section 2.3 Analysis Plan

2. Although this is not a systematic review, the authors need to describe their search in a more systematic way. For example, how many searches did they conduct? What issues did they come across? How did they screen the studies initially?

**COMPLETED –** more detail added

3. The authors tended to separate pilot from feasibility studies. Most current studies would use those two words interchangeably, and even for some other meaning. The authors need to provide some details as how they made sure the selected studies were consistent with their definitions of pilot/feasibility used in this paper.

**COMPLETED -** Arain found that there are some methodological differences between the two so we are looking at whether there is a difference in the sample sizes used between the two. The entry criteria were used to try to filter out trials using pilot and feasibility for other reasons – else the data were analysed regarding to the definition used by the investigator. Page 6 first paragraph background section.

4. I am not clear why the authors chose to exclude the trials that were using healthy volunteers in the search. That needs some clarification.

**RESPONSE –** ‘Trials being conducted on healthy volunteers were not included as these are not usually efficacy studies’ page 8 paragraph beneath eligibility criteria

5. The authors solely relied on descriptive statistics to draw the conclusions, e.g. a difference between publically funded pilot studies with a binary endpoint and those with a continuous endpoint. It would be more credible if the authors conducted some statistical testing to make the inference and report in p value.
RESPONSE – we did do statistical tests – the P-values ranged between 0.2 and 0.4 but we have not included. This is not because they are not statistically significant but because the study was not intended as one formally test hypotheses

6. The authors would need to provide deeper discussion as how the findings could help and guide future researchers to plan the sample size to meet the objectives of their pilot/feasibility studies. What is the impact of the finding in designing pilot/feasibility studies?

COMPLETED – we have put some discussion about how trials need a sample size justification and our results can help with this

7. The authors did not address the limitations of the study, for example, the population, the methods, the generalizability of the results.

COMPLETED – in discussion section

- Minor Essential Revisions

1. Although the authors acknowledged previous work upon which they were building, they did not explain why they were interested in the impact of the factors: the type of funders, study endpoints, etc. on the sample size of pilot/feasibility studies. It would be good if the authors can provide some rationales as why they chose those factors not the others.

COMPLETED – we mainly collected these data for descriptive reasons. The main factor is type of endpoint – we expect Normal endpoints to require less subjects – and we have a reference for this and it is mentioned on the discussion. We also have a reference to funder influencing

2. The title “An Audit of Sample Sizes for Pilot Studies being Undertaken in the United Kingdom” is a bit too broad, given the scope of the review and the factors being assessed. The authors may consider reframing the title to convey what has been actually targeted and found.

COMPLETED – we have added “registered on the United Kingdom Clinical Research Network database”
3. The authors need to report the n for each mean calculated in all tables. For example, in tables 2, 3 and 4, that information was missing.

COMPLETED

4. The paper needs revision on some spelling and grammatical errors, and citation for the software.

COMPLETED
This paper sought to assess the target sample sizes of pilot and feasibility studies currently running in within the United Kingdom. The significance of such a report is not stated and the recommendations from the analyses are not supported by their data.

Major concerns:

1. The conclusion of the article is completely not supported by the data provided. The data provided are purely a descriptive look at the range of sample sizes of ongoing pilot trials. There was no critical review of the sample sizes to see if they were appropriate for the aims of the proposed trial. Just because pilot trials are being conducted with sample sizes of 30 does not mean that is the best sample size for all pilot trials. How is one to know these trials have enough or too few participants for their goals?

   RESPONSE – given the data source we restricted as to what we can review but our data source complements other reviews as ours is was looking at a snap shot in time while other reviews were looking back at trials designed some time back

   Sample sizes for pilot studies often do not have a formal justification and the hope was by simply capturing what sample sizes are being used this could assist researchers when designing trials

2. Not clear what this analysis contributes to the literature. Would have been better to critically review the trials to see if the sample size aligned with their objectives.

   RESPONSE – many pilot trials do not have formal objectives. We believe the work will assist applied researchers when designing future such trials.

3. Not clear why the authors looked for differences by funding status and by primary endpoint. No research questions or hypotheses were framed around these variables.

   COMPLETED – research questions outlined in background section
4. Seems important to know if the primary endpoints were efficacy endpoints or feasibility outcomes. The paper could be made richer with a description of the types of endpoints beyond normal or binary. Same applies to the description of the interventions, particularly the health technology interventions.

LISTED AS A POSSIBLE EXTENSION

5. The authors acknowledge the skewness of the sample size data but still present the means. Only medians should be presented.

COMPLETED

6. Conclusion: This paper does not assist researchers in sample size justifications for pilot investigations. Just because trials are practicing 30 patients per trial does not mean 30 patients per trial provides these pilot trials with the answers (preliminary data) they are seeking.

RESPONSE – Obviously we are not saying 30 is the answer for a pilot sample size but we do believe that knowing this is useful when framing a sample size justification

Minor concerns:

7. The title of the manuscript should state “Pilot Trials” instead of “Pilot Studies” which would imply any type of study design, including purely observational.

COMPLETED

8. Multiple grammatical and syntax problems throughout the manuscript. Many contractions used throughout article such as “won’t” – not ideal for a scientific manuscript. Possessives used incorrectly (funders’ website vs funders’ websites). The word ‘data’ is plural.

COMPLETED

9. The abstract is written with abbreviations/acronyms such as UKCRN that will not be known to the reader. These should be removed and substituted with more general language (ex. Trial database)
EXPANDED ABBREVIATIONS

10. The 7th paragraph of the background begins with a sentence that is not clear. In the same paragraph the citations listed do not have a citation.

COMPLETED

11. More description needed on UKCRN. Who must register, how to access, who can access.

COMPLETED - The United Kingdom Clinical Research Network (UKCRN) database, [http://public.ukcrn.org.uk/search/] was used to identify pilot and feasibility trials currently ongoing in the UK. The UKCRN study portfolio is a database of clinical research studies being undertaken in the UK. The database comprises of the National Institute for Health Research (NIHR) portfolio in England, and the corresponding portfolios of Northern Ireland, Scotland and Wales. The studies benefit from the support given by the clinical research network however, it is not compulsory for researchers to register with the UKCRN. The database is accessible by anyone online through the URL listed above. Pg 7, methods section paragraph 1

12. Only trials that were randomized were included. Data components included presence or absence of a control. If no control was in the study, what were the participants randomized to (two interventions)?

COMPLETED – mistake rectified and analysis re-run

13. What information was sought from the internet search and clinicaltrials.gov

COMPLETED - when information was missing. Pg 8 data extraction section first paragraph

14. What is meant by ‘very limited information’ for 62 trials – did the PI and funders respond?

COMPLETED – ‘After conducting all of these searches 62 (75%) of the trials did not have complete information and so, in these cases, the principal investigator or funder(s) were contacted by email for the relevant study protocol, in all cases responses were received.’ –pg 8, methods section 2.2 data extraction, fourth paragraph
15. Results: ‘no data available’ in results means the sample size was not available?

COMPLETED – ‘Studies where no data were available means that, although the trial was registered, no information regarding the trial was listed. In these cases (n = 5) the trial coordinators were contacted however none of these replied and the trials were assessed as ineligible.’ –pg 9, results section, paragraph 1.

16. Table 1: MSK not defined

COMPLETED

17. 3.2 sample size: author states “all trials found…..provided a target sample size” but this conflicts with statement that some of the trials had missing data.

COMPLETED – the missing data was not about the sample size, this has been more clearly defined in the paper.

18. Most important part of this paper is that none of the trials explained justification for their target sampled.

RESPONSE – the UKRCN database does not give sample size justifications

19. Tables should be condensed to one table. Each table has so little information. Sample sizes (number of studies) are needed for each row of the tables.

COMPLETED – there are now less tables

20. No definition of public or industry.

COMPLETED – defined in background section

21. Declarative statements are made about differences or lack of them without any statistical testing.

COMPLETED
22. No justification for fully stratifying results by funder and endpoint.

COMPLETED – added in research questions and justifications into background section

23. Discussion: statement that target sample size for pilot study is often likely to be adapted during the study or main study is not cited? From where do the authors draw their conclusion. Same paragraph makes statement about what information participants should receive – this seems completely out of line with the rest of the discussion.

COMPLETED – we have not added a reference but toned down to say “and has a great degree of uncertainty”

24. Why is it significant that funding source has been found to drive sample size in a previous study?

COMPLETED – we are just looking at pilot studies where as the other paper looked at many types of studies (not just pilots). We have highlighted this.

25. Strong statement made about “30% less patients required for a study with a continuous outcome” -- what is meant by required? How do you know these studies are the “right” size?

COMPLETED – the reference to this work is given. For the same power a binary endpoint will require less than a continuous.

26. This statement is completely unsupported by the analyses “It is also apparent that when preliminary studies are published, it appears to be only those which have found ‘significant’ results.”

COMPLETED – Offending sentence removed – we were paraphrasing Arain
Reviewer: Julius Sim

Major Compulsory Revisions

Clarify how the data in this paper, which describe current practice rather than optimum practice, can assist researchers in justifying sample sizes for pilot studies (which is claimed on page 3 in the abstract and on page 15).

Be more explicit as to the argument on page 5 (lines 11-13) regarding appropriate degrees of freedom for a sensitivity analysis. Readers may otherwise not grasp the point without recourse to the paper referenced.

COMPLETED – ‘Julious [9] highlights that a sample size that is too small will have an imprecisely estimated variance, which could impact on the design of a future study.’

Clarify the point about ‘without the results being skewed’ on page 7, line 7. What sort of skewing might occur, and how?

COMPLETED – we have took out this sentence and only now quote medians (with ranges and IQR)

On page 11, it is said that the median per arm of 31 is consistent with Browne’s recommendation. Presumably Browne recommended a minimum sample size, and it is not clear how a median sample size across studies is ‘consistent’ with this recommendation.

COMPLETED – deleted comment

Throughout the paper, the emphasis is on median values, but the appendix reports only means. Is this intended?

COMPLETED – appendix tables removed

Minor Essential Revisions
Page 5, line 5: change ‘unnecessarily be being’ to ‘be unnecessarily’.

COMPLETED

Page 5, line 7 from end: insert ‘that’ after ‘poorly reported and’.

COMPLETED

Page 5, line 2 from bottom: sentence is grammatically incomplete.

COMPLETED

Page 6, line 2: reference numbers missing for Arain et al and Lancaster et al.

COMPLETED

Page 9, line 6: omit ‘in’ before ‘mental health’.

COMPLETED

Page 11, line 9: change ‘then’ to ‘than’.

COMPLETED

Page 11, line 12: change ‘mean’ to ‘average’, as both means and medians are reported. Do the same in the body of Table 3.

COMPLETED

Page 11, line 13: change ‘trails’ to ‘trials’.

COMPLETED
Page 11, line 2 from end: insert semi-colon before ‘however’.
COMPLETED

Page 12, line 1: change ‘regards’ to ‘regard’.
COMPLETED

Page 15, line 4: change ‘less’ to ‘fewer’.
COMPLETED

References: check consistency of capitalization in titles of journal articles.
COMPLETED

Change ‘publically’ to ‘publicly’ throughout.
COMPLETED

Sometimes both an n value and a percentage are given, sometimes just the percentage – suggest inserting the relevant n value in the latter case.
COMPLETED

Discretionary Revisions

Page 2 (Results): is it necessary to have initial capitals for ‘Normal’ and ‘Binary’?

COMPLETED – changed to continuous and dichotomous

It certainly seems inappropriate for ‘binary’, and an initial capital for ‘normal’ is probably only required to avoid ambiguity, which is unlikely to occur in this paper.

SEE PREVIOUS COMMENT
Change ‘then’ to ‘than’.

COMPLETED

Page 5, line 3 from end: suggest ‘calculation’ rather than ‘estimate’. The sample size is not really a parameter, and thus not strictly estimated.

COMPLETED

Page 11: if there is only one industry-funded feasibility study, is it worth reporting a mean or median at all?

RESPONSE for completeness we have kept in

Throughout the word ‘data’ is treated as singular, but it is probably preferable to treat it as plural – changing ‘was’ to ‘were’ etc as appropriate throughout.

COMPLETED