Author’s response to reviews

Title: A randomised trial of adaptive pacing therapy, cognitive behaviour therapy, graded exercise, and specialist medical care for chronic fatigue syndrome (PACE): Statistical analysis plan

Authors:

Rebecca Walwyn (R.E.A.Walwyn@leeds.ac.uk)
Laura Potts (lauramrpotts@googlemail.com)
Paul McCrone (paul.mccrone@kcl.ac.uk)
Anthony L Johnson (tony.johnson@mrc-bsu.cam.ac.uk)
Julia C DeCesare (Julia.Decesare@mundipharma-rd.eu)
Hannah Baber (Hannah.Baber@icr.ac.uk)
Kimberley Goldsmith (kimberley.goldsmith@kcl.ac.uk)
Michael C Sharpe (michael.sharpe@psych.ox.ac.uk)
Trudie Chalder (trudie.chalder@kcl.ac.uk)
Peter D White (p.d.white@qmul.ac.uk)

Version: 2 Date: 5 August 2013

Author’s response to reviews: see over
Dear Sir/Madam,
Many thanks for the reviewers’ comments and for your invitation to respond to them. Enclosed is a revised manuscript entitled “A randomised trial of adaptive pacing therapy, cognitive behaviour therapy, graded exercise, and specialist medical care for chronic fatigue syndrome (PACE): Statistical analysis plan”. Changes are tracked and below are point-by-point responses to each of the concerns raised.

Reviewer: Charlie Goldsmith
The authors of this SAP should consider having a section that defines all short forms used in this document.
I have added a glossary to the statistical analysis plan

The references cited are not in the Trials suggested format.
I have updated the format of the references

1. Page 2, paragraph 6. Include the date of registration as well as the date the first patient was randomized using a common defined date format. The date formats in the manuscript are not currently consistent.
I have done this. The date format adopted throughout is dd/mm/yyyy.

2. Page 5, paragraph 6. Specify the date format. See page 13, paragraph 2 where a different format is used.
The date format adopted throughout is dd/mm/yyyy.

3. Page 6, paragraph 1, lines 5 and 6. Delete [to be] as it is now published.
Deleted

4. Page 6 and 36. Consider putting all the References in the same location.
We have considered this but feel that it retains the self-contained nature of the SAP if its references are separate from those in the introduction. To make this clear we have started the references at [1] again in the SAP itself and ensured all references that appear in the SAP are in the reference list at the end of it.

5. Page 6 and 36 ff. Trials likes to publish the first 30 authors before using [et al], so many of these Rs need to be revised. A random sample of the Rs listed here have been checked for citation accuracy. Trials does not use [p] as part of the citation.
I have included all authors for any references that were abridged to et al. I have amended the references to be in Trials format.

6. Page 6, Reference 1. Insert (9044) after [349].
I have done this.

I have done this.

8. Page 6, Reference 4, line 2. Insert (9668) after [373].
I have done this.
I have done this.

10. P 6, R 10, I 1. There are 4 more authors and on I 3, insert (6) after [11].
I have done this.

I have done this.

12. P 6, R 17. Insert (9780) after [377].
I have done this.

I have done this.

I have done this.

15. P 7, R 20. Insert (9780) after [377].
I have done this.

16. P 7, R 21, I 2. Insert (9780) after [377].
I have done this.

17. P 7, R 22. Insert (9780) after [377].
I have done this.

18. P 7, R 23. Insert (9780) after [377].
I have done this.

19. P 7, R 24. Insert more authors and insert (9780) after [377].
I have done this.

I have done this.

21. P 8, S(ection) 2.3, Flow diagram. There is an error on the right and no page number.
I have corrected this.

22. P 8, S 2.3, I 3. Spell out [incl.]. Also P 9, S 10.1, I 2. Notice I 3 where it is spelled out. Also P 13, heading Basic design.
I have done this.

23. P 10, Primary Objectives. The Economic Analysis (See P 11) contained the comparison of CBT vs GET, so why not here as well even though you have too little power? The time at which these are to be done should be stated. A comparison between CBT and GET was not considered of interest for effectiveness as they were expected to be equal in this regard. Such a comparison was however of interest on cost effectiveness as there were possible differences here. I have included the time at which the objectives are to be done.

24. P 11, p 1, I 4. Why include [physical] when it was not on the previous page?
I have deleted “physical” here.
I have replaced “significant” by “substantial” here.

26. P 12, p 2, SF-36. Which version? Also all outcome measures should have Rs and the size of the MCID (minimum clinically important difference) should be listed for each, even though the sample size has not been computed for each. Also each scale should have its measurement properties defined and how to interpret each scale.
I have included the version of the SF-36 as version 2. I have also included references for all the outcome measures where these are available (see section 3.2) and indications of how to interpret each scale (see Section 4.2). We defined the MCID for the primary outcomes only in response to a request by the Trial Steering Committee. While I can see that it would be desirable to define this for all the outcomes, it would be post-hoc and potentially influenced by the trial results. As such, we have not provided the MCID for other outcomes. The measurement properties are contained in the references provided so we have not reproduced them in the SAP.

27. P 12, p 3, l(tem) 1 b. [CGI] is not defined until 6.
Amended

28. P 12, p 4, l 12. [CDC] is not defined.
Defined

29. P 12, last p. The selection should be stated as this leads to improper reporting unless the papers are specified. How will the selection be made?
The selection was based on space constraints depending upon the target journal and this in turn depended on the results. If we had a null result we would have more space than if we had a complex result that needed more explanation. I have added a sentence to clarify this.

It is not published. I have put a note to say that it is available from the website, along with all other manuals.

31. P 14, p 4, l 1. Provide a R to the protocol if it is published. On l 3, this would be a good place to see the MCID sizes.
The protocol that was published was an abridged version. However this section was reproduced here so I have added the reference. See point 26 regarding the MCID sizes.

32. P 15, p 3, l 2. Provide a R for ITT. There are many versions of this.
This has been added.

33. P 18, NB. Provide R where this list can be obtained, with an email address or some contact information.
This has been added.

34. P 18, S 3.5, p 2, l 4. Was some form of imputation considered? See the R to reporting of studies and missing data in NEJM 2013-Oct-04 for the Little RJ et al paper and the editorial as to how things seem to be going for publications. Include Rs for your strategy.
Yes we did consider multiple imputation where the proportion of missing data was not minimal (see Section 7.4, including references). Thank you for the reference. We have not updated the SAP to include it because it post-dates the SAP and final analysis.

36. P 18, S 3.5, p 4, l 1. How do you decide [large] or [differ]?
Good question. We were going to make a subjective judgement with reference to the data. Not ideal, but it would change what we actually did to specify “large” and “differ” now. I have added a sentence to clarify that it will be made by a consensus judgement of the authors.

37. P 19, S 4.2, p 1, l 1. Will you use percentages as well as proportions?
I have replaced proportions by percentages.

38. P 20, S 5.3, p 1, l a and b. Suggest replacing [Less] by [Fewer].
I have replaced less by fewer.

39. P 20, S 5.3. Consider adding a measure of the randomization integrity.
I have added a sentence under Section 8.5 to the effect that eyeball comparisons of the baseline characteristics across arms will be carried out as a measure of the randomisation integrity.

40. P 21, S 5.5. Consider adding the CONSORT R to the flow chart and reporting.
The section referring to the CONSORT flow chart is section 8.3 where a reference is given.

41. P 23, p 2. Is there a R to this technique?
I have inserted a reference for the multiple membership technique.

42. P 23, S 6.4, p 2, l 3 and 9. Suggest changing [is] to [are] since data is a plural word.
Also P 23, S 6.4, p 4, l 1
According to the Oxford English Dictionary data is singular. However I am aware that both forms are used generally and in this paper. As such I have changed “is” to “are” in line with the recommendations of the Chair of the Trial Steering Committee as these are instances she and I missed.

I have done this.

44. P 23, S 6.4, p 3, l 3. Why 100 when most Rs suggest 5 or 10 will be adequate. Do you have a R as to why?
I have included a reference. This states that the proportion of missing data should be used to guide the number of imputations. So 20% missing data should lead to 20 imputations. We were being possibly unnecessarily conservative but that was the plan.

45. P 24, p 2, last l. Consider measuring the impact of missingness on the conclusions as per the Little R in 34.
As we were expecting the proportion of missing data to be minimal, we felt this was unnecessary. If we had needed to follow the strategy for substantial missing data, we had specified the “best” approach to handling it and would have reported that. For our own sakes we would have included a sensitivity analysis however. I have not updated the SAP because this decision may be viewed as post-hoc.

46. P 24, S 6.5, l 1 and 2. Consider using 6 since the economic analysis will consider 6, not just 5. Also P 29, S 9.3, p a, l 1. Also P 35, p 1
The considerations about multiplicity related to the evidence that would have been used to make the principal conclusions about the results (i.e. the 5 effectiveness comparisons). The cost effectiveness data was considered separately so new considerations would apply here. These weren’t specified in the SAP.

47. P 24, S 6.5, l 1. Replace [P] by [alpha].
I have done this.
48. P 25, S 7.1, p 2, l 2. Which definition of these two measures will you use: IQR= difference between quartiles 1 and 3; or the two quartiles separated with a dash? Similarly with range? Also P 27 does not make it consistent? See P 28 S 9.1 range use is not statistical.
I have changed this to reflect our intention to give the two quartiles separated with a dash, and the minimum and maximum likewise.

We did not intend to use a test of normality. Instead we planned an examination of the histograms.

50. P 29, S 9.3, p 2. Is the a R for the MCID?
We had two primary outcomes, neither of which had well validated MCID defined. In specifying 0.5 SD as a clinically useful difference we were guided by (Sloan, Cella, & Hays 2005) and (Cohen 1988) but in reality we chose the values via discussion within the team. As such, there is no reference.


51. P 33, last p, l 2. Suggest replacing [two fold increase] by [doubling].
I have done this.

52. P 34, S 10.3, p 6. Are the utilities contained in some R?
I have added a reference here.

I have done this.

54. P 35, p 2. Provide Rs for all software with version numbers that is planned for use. Also P 36, S 13
I have added a sentence to state that the most up-to-date version available will be used. We planned to use Stata but other software was not ruled out.

I have added a reference here.

I have done this.

57. P 36 ff. These Rs are not standardized for trials reporting.
I have updated the references for Trials

58. P 38, R 58. Is this DG Altman?
Yes. I have updated the reference

Reviewer: Jinhui Ma

Major Compulsory Revisions:
1. My major concern with this kind of articles is about their length. Their values may fade away when readers find it is difficult to catch the essential message quickly, and consequently lose their interests of reading them in further detail. Apparently there are no strict guidelines on how to report a statistical analysis plan. But based on my experience as a biostatistician for years, I found when variables (primary outcomes, secondary outcomes, explanatory variables, adverse events, etc.) were summarized in a table with variable name, how it is measured, when it is measured and type of the variables (continuous or categorical) etc., both statisticians and clinicians found it was easier to get the key information quickly and clearly. Same thing for the statistical analysis, a table can be built with several columns indicating variable name to be analyzed, statistical method used for analyzing this variable, confounding variables to be adjusted for in the analysis, methods used for model assumption check, methods used for goodness of fit test, etc. Another advantage of providing these tables is that it may be helpful on shortening the main text since the tables speak for itself. The SAP is not intended to be read in the same way as a normal paper where an essential message is given. It is there for the reasons outlined in the introduction. It has a contents page with page numbers that allows the reader to navigate it and quickly find the information they are looking for. We don't agree that a table of the statistical analysis would be helpful.

2. My another general suggestion is that a statistical analysis plan for a trial/study, if intent to publish, should be beneficial to other researchers in their future studies, rather than just simply reporting the statistical analysis plan for the present trial/study.

We agree that the SAP should be beneficial to other researchers in their future studies but believe that it is in its current form. In addition to statistical strategies for handling a variety of issues anticipated in the data, it gives researchers a template and example for their SAPs. I have received positive feedback from colleagues developing their own SAPs.

3. In ABSTRACT on page 2, I am expecting some summary of the results, such as what are reported etc., rather than simply saying “The SAP is given in full as approved by the Trial Steering Committee”.

I have updated the results in the abstract in line with this comment.

4. In ABSTRACT on page 2, The ‘Conclusion’ section should state what you can conclude from the results (for example the advantage of publishing the SAP etc), rather than “Online journals offer the greatest potential for publishing….”, which cannot be supported by the results presented in this manuscript.

I have updated the conclusion in the abstract in line with this comment.

5. In Section 2.1 Secondary Objectives on page 11, it is not clear what the first objective really means by “As for the primary objectives with outcome as the participants’ self-rated clinical global impression change rating”.

I have clarified what was meant by this.

6. In Section 6.4 Method for Handling Dropouts and Missing Data at the first and second line on page 24, the authors stated “There is specific guidance for missing baseline scale data and this will be followed [48]”. This specific guidance should be summarized here rather than just providing a reference.

I have included a sentence summarising this guidance.

7. This manuscript provided a statistical analysis plan for the principal papers of the PACE trial. Since a statistical analysis section must be provided in each principal paper when publishing it, the present manuscript should emphasize that the present manuscript is not just the sum of all the statistical analysis sections of the principle papers. Why is publishing this present manuscript important or necessary?
I have emphasised this in the abstract. The SAP was written and signed off before any of the statistical analysis sections were written for the principal papers and none are simple abstractions from the SAP. It is important to publish the present manuscript because it reports the plan in full as it was developed. The statistical analysis sections report what was actually done. These may differ in the same way as the main report may differ from the protocol. It ensures complete transparency, which is important considering the public interest in this trial.

**Minor Essential Revisions**

1. In Section 2.3 Trial Design regarding “sample size calculations taken from the protocol”, I did not see how the sample size calculation matched with the primary statistical analysis method (mixed-effects linear regressions including participant as a random intercept and investigating adding a random slope on time, as you presented on page 29). How is the clustering effect taken into account?

   Clustering was not taken into account in the sample size, only in the analysis. This reflected the change in practice over the course of the trial. At the point when the sample size was calculated there was no consensus in the medical statistics community about allowing for clustering in individually randomised trials. By the time the analysis was undertaken it was part of the CONSORT statement for non-pharmacological treatment trials. Even so, there was no recommended method for handling clustering in a trial with the multiple membership data structure this trial had. As such this trial is an exemplar and this SAP of interest.

2. Following my previous question, the sample size calculation was based on the primary outcome/analysis. I am thinking whether it is helpful to provide what statistical power can be achieved for analyzing the secondary outcomes and economic outcomes based on this sample size.

   The sample size calculation reported was used to design the trial. We did not carry out power calculations for the secondary and economic outcomes and doing so now would be post-hoc and of limited value as a result. The confidence intervals reported in the principal papers provide estimates of the precision of the results.

3. On page 39, the title of Table 1 is too short to provide enough information to understand this table. For example, what does the ‘X’ mean when presented in the column “Discontinuation of Follow-up” and row “SF-36PF”? In addition, many abbreviations were used in this table, and detail information regarding what they stand for should be listed at end of the table.

   I have added a longer title to Table 1. The abbreviations are explained in a glossary and other clarification is included in a footnote.

**Discretionary Revisions**

1. Why do the figures in this manuscript use different arrow symbols? Personally I believe the arrow symbol used in Figure 1 makes the figure look better.

   We do not agree and have left the figures as they are.

Yours sincerely,

Dr Rebecca Walwyn on behalf of the co-authors

Principal Statistician