Author's response to reviews

Title: Chronic Headache Education and Self-management Study (CHESS) - a mixed method feasibility study to inform the design of a randomised controlled trial

Authors:

Kimberley White (Kimberley.White@warwick.ac.uk)

Rachel Potter (r.potter@warwick.ac.uk)

Shilpa Patel (Shilpa.Patel@warwick.ac.uk)

Vivien Nichols (V.P.Nichols@warwick.ac.uk)

Kirstie Haywood (K.L.Haywood@warwick.ac.uk)

Siew Wan Hee (S.W.Hee@warwick.ac.uk)

Dipesh Mistry (D.Mistry@warwick.ac.uk)

Dawn Carnes (d.carnes@qmul.ac.uk)

Stephanie Taylor (s.j.c.taylor@qmul.ac.uk)

Martin Underwood (M.Underwood@warwick.ac.uk)

Manjit Matharu (manjit.matharu@nhs.net)

Version: 1 Date: 06 Nov 2018

Author's response to reviews:

Technical comment: Authors’ Contribution - Please represent authors' names using their full initials, not their full name, in the Authors’ Contributions section. For example, the initials of John Smith would be ‘JS’. If there are any duplicated initials, please differentiate them to make it clear that the initials refer to separate authors, for example, by adding their middle name initial.

Our response: Amended

Please check the reviewers’ comments carefully, they raise a number of interesting points to improve the manuscript. Feasibility studies can be considered essential before conducting an RCT but are often not done for a number of reasons (e.g. resources), please consider slightly broadening the introduction/discussion on this point to make your study an example for our readers with an interest in medical research methododology; what is known in the literature
about the feasibility and design of feasibility studies, what components should be in there (you chose four, which ones not and why), added value? As one of the reviewers suggests, your motivation for not having done a pilot trial could be an interesting point of discussion as well, is that always necessary, pros/cons? Better remove it from the methods, where you should just describe what you have done.

Our response: The ethics statement now only in declaration section and the date of the study has been moved to the methods section. Thank you, we have included a section on the aim and scope of a programme grant and included more detail for the methods used for the four components of the study.

Reviewer 1 Overall, this is a much-needed area of study and intervention development, and I applaud the research team for taking on this work. The feasibility assessment was comprehensive in terms of examining 4 key areas needed to support a large trial, and is a strength of the work. Another strength of the work is the inclusion of a lay advisory group.

Despite this, there are a few fundamental concerns that warrant re-examining the purpose of the manuscript, and what it can contribute to the literature that is new. First, while the comprehensiveness of the feasibility domains is an overall strength of the research, explaining all of these domains in a single manuscript limits the detail that directly relates to the methods used (which is the focus of this journal). In an initial read, it seems the manuscript tries to do and cover too much territory. This also causes problems for the amount of background information that can be provided -- where there is only a cursory or superficial attention given to each of the 4 areas of interest.

Our response: We appreciate that including all four key elements of the feasibility study raises challenges. The purpose of the paper is to provide the reader with an overview of the complexity of the work involved prior to embarking on a full trial and we hope that by adding more detail to each of the four elements we have be able to strike a better balance now.

Second, and probably a consequence of the first issue raised here, some of the methods described have been published in detail elsewhere, while others have not. Specifically, the classification interview methods seem to have been published in detail elsewhere. However, assessing the feasibility of the PROMs using cognitive interviewing is not well-described in terms of cognitive interviewing, and the use of the SmartPhone App as a method are only briefly mentioned. There is a whole literature and skill level required to conduct cognitive interviews, and it isn't clear whether the accepted principles of (see, for instance, NCI's website on this) and training needed to conduct the cognitive interviews was carried out according to these standards. Similarly, there is an entire literature around daily/weekly journaling of symptoms (and the accuracy thereof), relative to newer methods that might incorporate an Ecological Momentary Assessment approach, for example.

Our response: We have tried to make it clearer to the reader that two of the elements are described in more detail elsewhere, and have added more detail to the remaining two elements.
We have added references to support our justification for testing the feasibility of using a smartphone app and provided more information in the methods section. And included - Structured cognitive interviews were also conducted to explore the acceptability and relevance of the measures. Informed by good practice guidance, the interviews explored how responder’s made judgements when completing the PROMs, including aspects such as question comprehension, recall and ease of completion (27, 28). The cognitive interviews and their analysis was carried out by an experienced qualitative team with expertise in this area.

Third, there are at least two areas in the manuscript where the authors go out of their way to explain why a full pilot study was not conducted. This seems unnecessary, as the rationale for completing all of the aspects of feasibility that are needed is sound and justifiable. It is not clear whether a pilot was not planned or conducted at all but seems to be the case, as the there is reference to a trial that had just recently begun in the conclusion. It seems defensible to not have conducted a full pilot at the same time as the feasibility study, but that is different from suggesting a pilot is not needed at all. Be clear on this point.

Our response: Thank you we have removed the duplication and make the issue clearer

Finally -- unrelated to the comprehensiveness of what was covered . . . the authors seem to sweep over or ignore what was found by participants around time constraints and giving feedback that an additional half day might be too burdensome, and then moving forward with adding an entire day of the intervention based on the nurse delivery feedback for testing in the RCT. Why did the intervention delivery recommendation supercede the participant feedback? The rationale is not clear and yet this is a critical component of the delivery.

Our response: We have rewritten this section to demonstrate more clearly how the qualitative interviews with participants and facilitators were used to inform the final content and structure of the intervention that is now being evaluated in the RCT.

Given these points, it seems reasonable to think more about what new (and sufficiently in-depth) is presented that could be useful for researchers in the way of methods. It might be better to focus on a more detailed methods paper describing one of these 4 areas (i.e., the PROMs and cognitive interviewing) as the focal point of the paper, rather than trying to present a very broad view of all feasibility components.

Or response: addressed above

Reviewer 2 Thank you for the opportunity to review an interesting study on a worthy topic. Overall, I felt that the study rationale and the reasons behind some methodological options could be clearer and that a more in-depth description of the procedures and results could be provided. Please see a detailed description below.
Abstract. In Table 1 you state that there were 131 participants, but in the abstract you refer 130? It seems to me that neither the results nor the conclusion of your abstract conveys the difficulties that you had, in particular, during recruitment. When I first read your abstract, it seemed that everything run smoothly but, for example, you ended up with a recruitment rate inferior to what was initially expected.

Our response: This was an error, now amended, the number of participants was 131

Thank you, we have made sure this is now reflected in the abstract, as you point out this was an important learning point from the study.

Background. I found that this section does not adequately contextualizes the need for your study. You need to provide support for each of the 4 objectives. For example, you fail to strongly advocate for the need to design a new intervention (how does it differ from previously used interventions), or about the need to test the headache classification (is this a new classification? Has it been used previously?....), is there any evidence suggesting that commonly used PROMs may not be the most adequate? You need to provide a clear rationale for your study.

Our response: We have added more support for each of the four objectives to make it clearer for the reader and a clearer rationale for the feasibility study itself

Lines 96 to 99 of page 5 should be removed from the introduction and placed in the methods with a context, i.e., you need to state how did you use the information collected from reviewing the literature. For example, literature on headache classification tools was used to identify specific tools that meet a set of criteria…This is very clear in your study diagram. Was this a systematic literature search?

- lines 100 and 101 should be in the methods section

Our response: We have removed the paragraph about the reviews from the introduction and added detail about each review in the relevant section.

Methods. It would be easier to follow if you use headings that directly link each of your aims to the methods used. There is also some repetition. See lines 124 to 126, you fail to state anything on the recruitment procedures, rather you repeat the aims.

How were inclusion and exclusion criteria ascertained?

Line 172 - which instruments were used at baseline? What was assessed?

175-176 - How did you ask for feedback? Did you interview GPs? Use questionnaires?...

You need a reference for the method you used to calculate sample size.
You state the sample size was revised during the study. Based on what?

Lines 192-193 - you state you needed 6 to 10 practices, but on lines 131-134 you state you used 14 (5+additional 9). These different statements have (apparently) different justifications. You need to combine them and state the number of practices and respective justification in one place only. I would suggest under the sample size sub-heading.

The heading "Feasibility of the headache classification interview" should not be under the heading "intervention".

Why did you chose to conduct a two-day intervention (as opposed to several sessions of shorter duration)?

Line 277 - Did you use any criteria to select those participants that tested the app?

Please include a section on data analysis (quantitative and qualitative data).

It is unclear to me whether the 85 participants contacted to receive the intervention were part of the 131 who were eligible and sent consent. If I exclude those that failed to respond and withdrew, I get 131 -44=87 and not 85.

Were participants that received the intervention assessed at baseline and post-intervention?

Were the interviews structured, semi-structured…not structured at all? If they were structured/semi-structured, can you provide a brief description of the interview questions?

Our response: The headings in the methods section and results section now relate to the four key elements of the study.

We have described the recruitment process in more detail and removed any repetition

CHQLQ, HIT-6, SF-12 and EQ-5D-5L used at baseline (amended in manuscript)

Reference added for sample size calculation

From the consensus meeting the classification changed from three sets of binary levels to 2 sets. Thus the final analyses of the study changed from 3 pairwise comparisons to 2 pairwise comparisons. This therefore changed the multiplicity adjustment hence giving a revised sample size of 153. This explanation has now been provided in the text.

Feedback from GPs was gained via an emailed short structured questionnaire

We have removed the number of practices from the methods section and explained in the sample size how many we anticipated needing (6-10) and include in the results the number needed (14) and discuss the implications.
The members of the intervention design meeting agreed that two days were needed to cover all the areas considered important to include in the intervention and thought it would be more practical for participants and facilitators to attend two days rather than several short sessions.

A small number of participants were approached to test the app that had agreed to take part in the study at the time the app was ready for testing (which was near to the end of the recruitment period).

The participants in the intervention completed questionnaires and the smartphone pre and post intervention, but the purpose of the feasibility wasn’t to test the effectiveness of the intervention but to test our data collection methods and the acceptability of the intervention.

Results. You are very descriptive in your results section and do not give data that support your statements. See these examples:

Lines 334- 343 - it is unclear whether feedback was consensual or not. Can you tell how many patients found the time commitment too great; did lay facilitators fail the intervention sessions? How many and how many times? And the nurse facilitators?

You state that "Measurement data quality, reliability and validity for the headache-specific and generic measures was acceptable,..." please do provide data that supports this statement.

Our response: We acknowledge that the results are largely descriptive as this is a feasibility study, and have amended the manuscript to make the data support the statements better. References have been added to support the statement: Measurement data quality, reliability and validity for the headache-specific and generic measures was reached at acceptable standards (30, 31), supporting application of the measures with groups of patients with chronic headache. The comparative analysis will be reported in full in a paper currently under development.