Author’s response to reviews

Title: Feasibility and Assessment of Outcome Measures for Yoga as Self-Care For Minorities with Arthritis: A Pilot Study

Authors:

Kimberly Middleton (middletonk@cc.nih.gov)
Michael Ward (wardm1@mail.nih.gov)
Steffany Haaz Moonaz (smoonaz@muih.edu)
Miriam Magana Lopez (miriam.magana.lopez@gmail.com)
Gladys Tataw-Ayuketah (Tatawayuketahg@mail.nih.gov)
Li Yang (yangl8@cc.nih.gov)
Ana Acevedo (atace@aol.com)
Zavera Brandon (zavera.brandon@nih.gov)
Gwenyth Wallen (gwallen@cc.nih.gov)

Version: 1 Date: 11 Jan 2018

Author’s response to reviews:

-Reviewer #1

This an interesting feasibility study of offering yoga to ethnically diverse groups with arthritis. In my opinion the results hardly suggest feasibility of the intervention in the targeted population but this could be discussed with reference to other intervention participation rates.

1) Since most people will only read the Abstract, this should state that the subjects came from Washington DC. This was added to the abstract.

2) The crucial numbers (Lines 179-180) are 128 screened, 59 eligible 30 recruited over 4 years, 18 dropped out leaving 12 subjects (or 11?) just under 10% of the original screening population available for analysis. A perfect example of Lasagna’s law! These numbers should be in the Abstract should be part of the decision about feasibility. Recruitment (51%), participation (60%), and adherence (67%) rates are offered in the abstract. The number of patients screened was impacted by converting from referral only to using a list of potential participants. This is discussed in detail in the manuscript.
3) The numbers in Tables 2 and 3 are not consistent. In Fig 2 if 30 were enrolled and 16 withdrew that leaves 14. Also if 10 withdrew prior to the intervention that leaves 20 initiated. An extra 2 have been lost. In Fig3 we are left with only 11! The numbers in Figure 2 were corrected. Twelve participants completed the study and physical evaluation, 1 did not complete the final questionnaire, and therefore there are data for 11 participants for those measures.

4) Fig 4 is poor- no scales on the graphs. Scales were added to the graphs.

5) P11/12 I would omit all the statistical analyses. Feasibility studies are too small to test whether the intervention works, Presentation could just give means before and after. The fact that a very highly selected group showed improvement is no surprise, and given a lack of a control group, regression to the mean cannot be discounted. What is needed is predictors of who drops out. Those who enrolled were compared to those who refused using cross-tab chi-square. However, as indicated by the reviewer, a multivariate analysis of predictors of drop out would not be feasible due to the sample size.

6) Give numbers of subjects as well as percentages for CASI data collection and Exit interview Minor Clarification was added to the text. All participants (n=30) used the CASI instrument during the original enrollment process. Exit interviews (n=16) were completed in the yoga studio following the last class. Additionally, an attempt was made to complete exit interviews with those who withdrew from the study after attending the first class.

1) L373 is 'satisficing' a word? Yes, as described in the referenced article* and used within the field of survey methodology, ‘satisficing’ refers to a respondent providing a satisfactory answer instead of optimally answering a survey question that would require substantial cognitive effort.


2) Line 389-90 No confidence intervals were given. This statement was slightly edited; however, confidence intervals were added to the text per the 2nd reviewer’s request.

3) Table 2 omit SD in last 6 rows (except perhaps for BMI), SD is meaningless for such skewed variables. Variables in table 2 were evaluated only three (length of time in US, arthritis duration, and BMI) were found to potentially be skewed. However due to the sample size, median and interquartile range (IQR) are now listed as an alternative to mean (SD).

4) Table 3 I would omit the last column with the non parametric results, but include the median and mean change. Table 3 was updated. Mean change was added to the last column for clarification. However the non-parametric Wilcoxon signed-rank test was used to evaluate statistical change as indicated in the table footnote.

-Reviewer #2:
This is a pilot study with some interesting findings. Most of my comments are around the structure and of the paper and presentation of results. I think this would improve the readability of the text and the ability of readers to interpret, and possibly use, the information in it.

1. In the methods section of the abstract a second objective is mentioned but the first objective is not explicitly articulated. Could this be done? This statement was added to the abstract.

2. In the results section of the abstract, proportions should be reported with confidence intervals to give readers an idea of the precision with which these estimates are estimated. The percentages offered in the results section of the abstract are descriptives of the sample participants; these are not estimates therefore measures of precision are not included.

3. I expect that the study was underpowered for many if not all of the hypotheses tested and would therefore prefer not to see any hypothesis tests in this paper. There are a lot of comparisons made so it would not be surprising to see some significant results even if null hypotheses were true. In addition with no control group, changes pre and post intervention could be due to other factors eg secular trends or regression to the mean, so I would definitely not put the results of such tests in the abstract even if they have been done. The second objective of this study was to evaluate the appropriateness of the selected measures. Since there are no gold standard global measures for this intervention with this population, the non-parametric (Wilcoxon) test was used to investigate change from baseline to final in the selected measures.

4. Lines 93/94 mention the feasibility of adapting an intervention, but later mentions of feasibility suggest that it was not only feasibility of the intervention adaptation that the investigators were interested (eg line 112) in but also feasibility of running the trial eg recruitment rates. I think there needs to be some rethinking of the overall aim(s), and then objectives within those aims and that structure needs to run through the whole of the paper. At the moment the feasibility criteria section doesn't really match any of the three stated objectives. The language for the primary objective was expanded to clarify, “the feasibility of adapting a previously studied arthritis-based yoga intervention for a bilingual ethnic minority patient group.”

5. Lines 164-166 I think it's fine to present correlations but this text ties them to "effect sizes" without any recognition of the natural history of the condition, regression to the mean etc. I don't think these should be presented as effect sizes. It is relevant to highlight that large effect sizes were observed with some measures, adding credibility for their continued use in future yoga studies with similar populations.

6. Line 167 What happened when missing data >10%? Missing values >10% were not present in the dataset, this comment has been deleted from the text.

7. Line 177 18 months seems a long time to wait to acquire a list of eligible patients. Is this what would happen in practice? Can there be some more comment on how pragmatic this
is as a strategy/whether this would happen in the main trial? Additional details were added to the text in an attempt to clarify. The initial recruitment strategy was clinician referral only. The first participant was enrolled 10/2012, and the first class started 3/2013, the 2nd 5/2013. After about 18 months referrals slowed and a new strategy (calling from a list) was employed.

8. Line 204 Present mean and confidence interval. The current presentation is not usual or helpful for this type of study. Confidence intervals were added to the text.

9. Line 212-226 Any statistical tests will be underpowered. I would advise against presenting the results of any statistical tests here. Estimates of effect and confidence intervals would be much better. I could not understand lines 217 and 218, particularly in terms of brackets and the words "sometimes" and "often". This header for this section was changed to Evaluation of Outcome Measures. It describes the changes seen by using the selected variables. The terms "sometimes" and "often" are variable labels for the HPLP measure. Full scale labels [1=never, 2=sometimes, 3=often, 4=routinely] were added to the text for clarification.

10. Line 232 Change this in line with comment 5. See response for comment 5.

11. Lines 239-244 Again present estimates and confidence intervals rather than beta and t, and rewrite this section. As per comment 3, it would not be unusual to find some significant results even if the null hypothesis was true, and there is a need to take into account natural history and regression to the mean. All exploratory predictive models were removed. We are only evaluating change from baseline to final assessment. Compared to results from the non-parametric (Wilcoxon) test which only used data from 12-pairs, the advantage of running the model is to use all the data from two time points.

12. The discussion looks rather long and parts of it are a mixture of results and methods, for example lines 278 to 291, and lines 340-344. Suggest substantially shortening the discussion and putting some parts into the results section. This section was reduced.

13. Lines 304-315 In this section it is unclear which study is being talked about at times. Needs to be clear whether it is the present study or the past study. Reference to the previous ‘Yoga for Arthritis’ study was removed from this section.

14. Line 320 This is not a sentence. This statement was corrected in the text.

15. Line 350-361 In view of my previous comments I think this section contains over-interpretation of the results. This section was revised and reduced.

16. Line 388 In this section on generalisability, authors should make it clear whether they are drawing out generalisability to a possible future trial or to other studies in the same area. There may be implications for both. This section was revised to clarify generalizability implications.
17. In keeping with previous comments I don't think the different sorts of lines on figure 1 are justified. The differentiating lines and the corresponding key were removed from figure 1.