Author’s response to reviews

Title: The effectiveness and cost-effectiveness of strength and balance Exergames to reduce falls risk for people aged 55 years and older in UK assisted living facilities: A multi-centre, cluster randomised controlled trial

Authors:

Emma Stanmore (emma.stanmore@manchester.ac.uk)
Alexandra Mavroeidi (Alexandra.mavroeidi@strath.ac.uk)
Lex de Jong (lex.dejong@curtin.edu.au)
Dawn Skelton (dawn.skelton@gcu.ac.uk)
Chris Sutton (chris.j.sutton@manchester.ac.uk)
Valerio Benedetto (VBenedetto@uclan.ac.uk)
Luke Munford (luke.munford@manchester.ac.uk)
Wytske Meekes (W.M.A.Meekes@uvt.nl)
Vicky Bell (vicky.bell@manchester.ac.uk)
Chris Todd (chris.todd@manchester.ac.uk)

Version: 2 Date: 16 Dec 2018

Author’s response to reviews:

16th December 2018

Dear Professor Graham,

Thank you for sending us the reviewers’ feedback. We very much appreciate all of the reviewer’s complimentary comments on our manuscript, and have taken on board all of their thoughtful suggestions. Please find attached our updated manuscript, with all of the recommended changes implemented (and highlighted in red). Details of the reviewer’s comments and our respective amendments are listed below.
Manuscript title: The effectiveness and cost-effectiveness of strength and balance Exergames to reduce falls risk for people aged 55 years and older in UK assisted living facilities: A multi-centre, cluster randomised controlled trial

Manuscript reference: BMED-D-18-01363R1

Reviewer #1

The topic of this manuscript (to compare the effectiveness of a strength and balance Exergame programme vs strength and balance home exercise programme leaflet, to improve balance and reduce falls in older people) has a high interest, due to the frequency and morbidity associated with falls in the elderly. The paper is correctly written and presented; abstract and introduction are appropriate, methods are adequately described, and results are correctly shown. Discussion and references are appropriate too. However, there are several limitations which should be corrected and/or discussed prior its publication.

Thank you for your complimentary feedback on the manuscript and for providing these suggestions for improvement. We agree with all of your suggestions, and have taken great care to ensure all of the recommended edits have now been fully implemented within the revised version (as highlighted in red within the article).

1. One of them refers to the age of the participants. Both in the title and in other parts of the manuscript, the authors refer to older people, but the sample includes subjects over 55 years. It is a very low age limit to consider a person as old. It is even debatable that today, in Western societies, a person of 65 years of age can be considered elderly. The authors should delete the word "older" from the title and content of the manuscript.

We fully agree with your comment and recommendation and have deleted the word, ‘older’ from the title and changed this to people aged 55 years and older. We have also deleted, ‘older’ throughout the manuscript; any references to older people that remain, explicitly refer to the background literature on falls or strength and balances exercises for older people, rather than the participants of the study.
2. Another limitation, already included in the discussion, is the no-blinding. It would have been important that at least the blinding of the evaluators had been achieved. This is an important limitation, since it may have introduced an important bias in the interpretation of the results.

We endeavoured to blind all of the assessors/evaluators throughout the study period. The assessors were blinded at baseline but at the 12 week assessment it was not possible for two out of the three assessors to remain blinded due to participants inadvertently revealing the group allocation of their assisted living facility/cluster. We agree that this may have introduced an important bias and have added an additional point to the discussion where this detection bias limitation is discussed that states, ‘and so the results should be interpreted with caution’ (Discussion, lines 516-517, page 21).

3. Finally, assessment of balance is only clinical and with questionnaires. Some instrumental test to assess balance (such us dynamic postuography) would have been useful to quantify the improvement in balance. This should be included as a limitation of the study.

We thank you for this suggestion and agree and have added a line to the limitation section in the discussion that recommends the use of an instrumental measurement of static and dynamic postural control using posturography (Discussion, Lines 542-543, page 21).

Reviewer #2:

This is a well-written manuscript that addresses the effectiveness and cost-effectiveness of an exercise intervention in the elderly population using technology to engage the participants and improve adherence. The technology, termed Exergames, includes real-time motion tracking of the participant using the Microsoft Kinect motion sensor during various games that have been designed to help motivate performance of various exercises that are intended to improve strength and balance. Using a multi-center cluster randomized control study design, the study compares the use the Exergame technology to a control group encouraged to perform a set of home
exercises based on the OTAGO strength and balance program. This type of work is important and relevant to the field.

We are grateful for the reviewer sharing their expertise and insights in the area of Exergames for health, and appreciate the thoughtfulness of their feedback. We thank them for the comprehensive assessment of this article which we believe has strengthened the paper and have taken great care to ensure that the recommended edits have now been implemented within the revised version (as highlighted within the article).

1. There is some concern that changes in the primary outcome (Berg Balance Score, BBS) are overstated and it is not currently clear whether they are clinically meaningful based on within-subject Minimal Detectable Changes reported in the literature of the BBS. The reported between-group change in BBS is based on a small decrease in the control group and a small increase in the intervention group following the training period. This needs to be further discussed in the manuscript.

Thank you for this suggestion, we have added further information and discussion in relation to this point. This includes discussion of the weaknesses in the literature on establishing delta (the magnitude of change on which a trial is powered) for the BBS. We have also noted the recommendations in the recent Delta2 publication (Cook et al., 2018) which state “The target difference for a definitive (e.g. Phase III) trial should be one considered to be important to at least one stakeholder group. The target difference does not necessarily have to be the minimum value that would be considered important if a larger difference is considered to be a realistic possibility or would be necessary to alter practice. Where additional research is needed to inform what would be an important difference, the anchor and opinion-seeking methods are to be favoured. The distribution should not be used. Specifying the target difference based solely on an SES approach should be considered as a last resort, although it may be helpful as a secondary approach.” in our revised discussion. Please see further information in the response to specific comments below, thank you.

2. Furthermore, effect sizes are currently not reported, which is required under CONSORT standards.
Thank you for pointing out that further clarity is needed in signposting the reader to the effect sizes. We have followed the CONSORT standards and checklist and reported the effect sizes in Table 2, column 3 as per the example given in the CONSORT paper (table 6, Moher et al., 2010) and also in the text (e.g. l 355-7 “Using ITT analysis, over 12 weeks the Exergames intervention had a significant positive impact on balance as measured by BBS, relative to Control [6.2 (95% CI 2.4 to 10.0; p=0.003)].” In the footnote to Table 2 we have also added, ‘*Outcome (e.g. BBS) at 12 weeks’ to the footnote text to aid interpretation of the results (Results, Table 2, footnote).

Reference:


Specific Comments:

3. Line 57-61: It is stated that 56 participants were allocated to the intervention and 50 to the control condition and on Line 60 that Intention-to-treat analysis was applied on 49 in the intervention and 43 controls. However, since intention-to-treat analysis means that all subjects allocated to a certain group should be included in the analysis, this statement becomes confusing in the abstract although it is clarified in the Methods section that a sensitivity analysis using multiple imputation was conducted. Please rephrase.

We thank you for this comment and agree that further clarification regarding the implementation of intention-to-treat analysis is needed in the Abstract in addition to the Methods section. We have therefore added the text, ‘using complete-case analysis’ to the Abstract to give further clarification to this statement (Abstract, line 60, page 2)


We have rephrased and broken this sentence up into two sentences to improve readability as per your recommendation (Background, lines 115-118, page 6).
5. Line 134: Please use past tense consistently.
We have amended accordingly (Background, line 151, page 6).

6. Line 134: "suite" should be "a suite".
We have amended accordingly (Background, line 151, page 6).

7. Line 134: Replace "have been" with "were".
We have amended accordingly (Background, line 151, page 6).

8. Line 141: Add hypothesis here or on Line 145.
Thank you for this suggestion. We have added the research hypothesis that a 12 week tailored programme of strength and balance Exergames will improve balance in people aged 55 years and older, in assisted living facilities (Background, line 160-162, page 6).

9. Line 160-161: The statement regarding "sufficient communal space" being >3m2 appears inconsistent with information provided in Image 1 where required space appears to be >8-10m2. Please clarify/revise accordingly.

Thank you for this point which we appreciate you highlighting, we agree that this requires correction. We have amended the text to >10m2 in reference to the communal space required for the participants to safely exercise using the Exergames (Methods, line 183, page 7). This is also now consistent with the information provided in Image 1 that refers to the positioning of the Kinect sensor and participant.

10. Line 167: Please clarify "Acute illness" whether only at time of enrollment or also during the study?
The exclusion criteria, ‘Acute illness’ in Box 1 refers to the enrolment period, however participants who became acutely unwell during the study were withdrawn from the study. The
reasons for withdrawal from the study are reported in the results section, (Results, lines 371-373, page 16) and in the CONSORT flow diagram (Results, Figure 1).

11. Line 167: Was "Mental capacity" assessed using some form of scale and or simply decided by a "trained healthcare professional" and was this only for the purpose of understanding informed consent or understanding the games?

Mental capacity to give informed consent was assessed by a trained healthcare professional (either a research nurse or research physiotherapist) rather than for the purpose of understanding the games. In a previous feasibility study we found that it was difficult to use a screening tool to assess capacity to understand the games because of the individual effects of the cognitive impairment – i.e. we previously found that some people with mild cognitive impairment had more difficulty in processing the game information than those with moderate cognitive impairment. At baseline assessment, the Addenbrookes Cognitive Examination-III (ACE III) was undertaken which gave us a much more comprehensive assessment of the various cognitive domains.

12. Line 167: "Recent fracture or surgery in past 6 months" appears to include the following exclusion criteria "Orthopaedic surgery in the past six months", so duplication not needed.

We fully agree and have omitted, ‘Orthopaedic surgery in the past six months’ (Methods, Box 1, lines 189-190, page 7-8).

13. Line 167: Use of "6" or "six"; please be consistent throughout and follow norms.

Thank you, we have omitted the word, ‘six’ in Box 1 (as per above comment it is a duplication) and in the text we have written numbers up to nine in words, and those over using numerals. There are exceptions such as ratios e.g. 1:1 and in the tables where numerals are used throughout as per the journal norm.
14. Line 167: Was any "Wheelchair use" an exclusion or subjects who were completely wheelchair dependent?

A dependence on wheelchair use was the exclusion criteria and we have amended the text to add clarity (Methods, Box 1, lines 189-190, page 7-8).

15. Line 167: How were "Peripheral neuropathy or other uncontrolled….." assessed?

The trained researchers (physiotherapists) assessed the participants and exclusion criteria were checked and verified using general practitioner records. We have added further description to aid clarity in Methods, lines 198-199, page 8.

16. Line 167: "Current use of gaming technology to exercise" was an exclusion criterion. Were subjects involved in regular forms of exercises allowed to participate?

Yes, participants involved in regular forms of exercise were allowed to participate. Levels of physical activity were measured using the Physical Activity Scale for the Elderly (PASE) tool.

17. Line 175: Spell out "GP".

We have made this amendment in Methods, Line 198, page 8.

18. Line 190: This section would benefit from a simple table to illustrate and compare the two interventions and make it easier for the reader to contrast what the two groups experienced.

A simple table (Box 2) has been added to enable the reader to easily compare the interventions for the two groups (Methods, line 215, page 9).
19. Line 206-214: This section appears to be background information and not current Methods.

We agree and have moved this point to the background section (Background, lines 136-149, pages 5-6).

20. Line 235: "is" should be "was".

Amendment made as recommended (Methods, line 262, page 11).

21. Line 236-246: The section starting with "In an earlier feasibility study....." partly duplicates previous information in Methods. Some of this information would be more relevant in the Background section. Please revise and/or remove as relevant.

We agree and have moved to the background section and removed any duplication of this information (Background, lines 155-157, page 6).

22. Line 256: Clarify ".....were posted monthly". Were calendars mailed to the investigators?

That is correct, to give further clarification in the text we have added, ‘to the researchers’ in the section that explains about the reporting of falls using the calendars (Methods, line 283, page 12).

23. Line 261: Please clarify why sample size estimation was based on the Minimal Detectable Change of 8 points (according to reference #68, Conradsson et al. 2007) when results interpretation and discussion refers to a different study (#80, which appears to be the incorrect article by Donoghue, see comments below) that reports an MDC of 5 in the BBS?
We apologise for the lack of clarity and incorrect reference which we have addressed in our revised manuscript. The sample size estimation was based on the Minimal Detectable Change (MDC) as per reference 68 (Conradsson et al., 2007). This was the information identified by the team and then statistical advisor and the sample size was based on this. However, as recognised by the recent DELTA2 publication (Cook et al., 2018), the use of a value of MDC (from a single publication) is now not recommended for powering a trial. We have therefore considered a wider variety of publications on the minimal important change (MIC), including the publication by Donoghue and Stokes (2009), and its potential magnitude in this population and have now clarified our interpretation in the Discussion (Discussion, lines 451-458, page 18-19).

24. Line 276: It is not clearly stated what statistical procedure was used to compare pre- and post-data. Also, what statistical software was used?

Pre- and post-data (within groups) were only compared using the descriptive statistics detailed in the manuscript. Pre- and post-data (between groups) were compared using linear mixed effects modelling, as clarified in the amended Methods section. All analyses were performed using Stata 14; we have added this to the manuscript (Methods, lines 312-317, page 13-14).

25. Line 281-283: Sentence appears incomplete and/or ends abruptly. End of parenthesis beginning on line 281 is missing. Please clarify/revise.

Thank you for identifying these errors. We have corrected the typographical error (‘;’ instead of ‘’) and inserted the missing word (‘data’) at the end of the sentence (Methods, line 310-312, page 13).

26. Line 334-336: Where reported differences statistically significant?
While it is an interesting observation that the Exergames group improved by a similar amount as the Control group deteriorated on the BBS over the 12 weeks’ follow-up, this is not directly related to our trial’s objectives. As recommended by Bland and Altman (2011; BMJ 2011;342:d561), we have not performed tests of change from baseline in each of the two groups. Given the age of participants (92% aged over 65 years, with a median age of 78 years), we would expect a small deterioration in BBS, on average, without an active intervention. Hence, of interest is the difference in BBS (or, perhaps, the difference in change in BBS, although it is worth noting that ANCOVA of BBS will give the same estimate of effectiveness as ANCOVA of change in BBS when the outcome is adjusted for the baseline BBS) between groups as this demonstrates the effect of an intervention, such as our EXERGAMES, relative to a suitable comparator.

27. Line 343: Quantify "Several".

As we did not formally collect data on the number of participants who wanted to continue the Exergames after the 12 week period we have omitted this information as, ‘several’ is too vague to be included in the results section (Results, lines 376-377, page 16).

Results

28. Throughout Results: Please report effect sizes as described under the 2010 CONSORT standards. See 17a. Outcomes and estimation.


As detailed earlier in our response, effect sizes, as described under the CONSORT 2010 standards are reported throughout the Results section (Results, lines 382-383, 391-392, 396-397), including in Table 2 (Results, Table 2).

29. Line 348: It would be informative to report how many subjects in the two respective groups improved their BBS by at least the MDC (# of "responders", 8 or 5 for this population; see comments above and below).
In line with our response to the reviewer’s previous comment, we would prefer not to add information about change in BBS within groups in this manner. This is, firstly, given the lack of clarity about the MDC for BBS and, secondly, given that the trial’s objectives related to differences (of differences in change) between groups, rather than change within groups and we do not believe that additional informative and objective evidence would be provided by adding these statistics to the manuscript. Within this trial we did not attempt to distinguish between potential subgroups of participants and did not collect data that would enable us to do so. It may be an interesting investigation to perform in future research to investigate who and how EXERGAMES affects different subgroups of people. However, we can inform the reviewer that 5/42 (12%) of the Control arm participants and 15/49 (31%) of the Intervention arm participants had an increase of 5 or more on the BBS over the 12 weeks of the trial, compared with the 2.5% expected if no participant in either group experienced a true improvement.

30. Line 367: Were falls reported pre-intervention?

Yes, a 1 year history of falls (none, single or multiple falls) was self-reported by the participants and this is presented in Table 1 as per the ProFaNE consensus on the reporting of falls (Lamb et al., 2005)(Results, Table 1).

31. Line 379: The last sentence, "This would represent….", reflects interpretation of the data and should be moved to Discussion.

Thank you for drawing our attention to this. This sentence is deleted as a similar sentence is already provided in the discussion (Discussion, line 412-414, page 17).

32. Line 392-398: It would be helpful to the reader to indicate the various costs and associated probabilities mentioned in this section directly in the figure. Simply add lines pointing to the curve.
Thank you for this suggestion. These have now been added to Figure 3, where we included dashed lines at the NICE thresholds of £20,000 and £30,000 per QALY. The text on line 392 (“…higher cost than controls in 95% (9490 out of 10,000) of replications.”) relates to the number of ‘dots’ above the x-axis (Results, Figure 3).

33. Line 407-408: This sentence represents interpretation of the data and should be moved to Discussion.

We agree and have deleted the interpretation of the data from the results section (Results, lines 440-441, page 18). A similar sentence is already provided in the discussion (Discussion, lines 492-3, page 20).

Discussion

34. Line 418-423: The difference in change between the groups reported in this trial (6.2 points "adjusted difference") was due to a small decrease in BBS in the control group and a small BBS increase in the intervention group. Please clarify (in Methods) how this "adjustment" was made. The information provided below Table 2 is insufficient.

We have clarified the description of the analysis methods as to how the adjustment was made, making specific reference to the focus of the analysis (group indicator) and the variables for which adjustment was made (assisted living facility unit (random effect); baseline measures of the corresponding outcome variable (fixed effect); location indicator (Manchester; Glasgow; fixed effect)). As the method used is a standard approach, and probably the most common approach, to analysing data from cluster-randomised trials – see e.g. Eldridge and Kerry (2012; Cluster Randomised Trials in Health Services Research; John Wiley) we have not included a specific reference to this method (Methods, lines 313-317, page 13-14)

35. Comparing the reported "adjusted" between-group difference of 6.2 with the within-subject MDC of 5 (or 8) required to be 95% confident that a true change in functional balance as measured with BBS has occurred as a result of an intervention, appears to overstate the results of the study. The mean (presumably "raw"?) BBS values reported in
Table 2 indicate a similar magnitude mean increase (2.6 points) and decrease (3.0 points) for the intervention and control groups, respectively.

Consequently, on average the MDC of 5 (or 8) points increase in BBS was not achieved. How many subjects in the two groups achieved a BBS change of 5 or more? Please address this concern in the discussion.

We have revised the discussion to provide a broader interpretation of the effect size against relevant values from the literature. We hope that this revision to the discussion and our previous responses satisfactorily addresses this point (Discussion, lines 451-467, page 18-19).

36. Line 419: The reference #80, Donoghue et al. 2017 is incorrect for the statement made regarding "….clinically important change .....". It should be Donoghue & Stokes 2009.

Thank you for identifying this error. We have amended this reference (Reference list, line 866, page 30).

37. Line 422: Please confirm that reference #78 is correct for the statement made.

We confirm that the reference is correct, this NICE guidance provides economic evidence and guidance on evaluating how well a treatment works in relation to how much it costs - does it represent value for money (Reference list, line 859, page 30).

38. Line 432: Clarify "traditional costs"; is it for the Otago program, regular physical therapy?

Traditional costs refer to the Otago exercise programme, we have added this to the sentence to aid clarity (Discussion, lines 482-3, page 19).
39. Table 1: One subject in the Exergame group was "Registered blind". Please explain how the subject was able to participate.

Although this participant was certified as, ‘Registered blind’ and therefore had severe visual impairment we found that she could make out the game ‘objects’ and follow the rules of the Exergames due to the bright colours and contrasts, the audio prompts given throughout the games and the support of the supervising therapist.

40. Appendix 1: The appendix should illustrate, name and describe the various exercises but not indicate that exercises has a specific effect unless there is a reference that supports the statement. The current study as presented in this manuscript was not designed to answer questions regarding effects of any specific exercises. Therefore, statements such as e.g. "Can improve balance and reduce the risk of falls" for the Sit to Stand exercise etc. should either be taken out, revised or be associated with a reference that has shown the statement is valid. For example, it would be ok to state "This exercise targets strength and balance during sit to stand activities", but not suggest any improvement unless such evidence is available. The wording "Can improve…." in this context is hypothetical and should be removed. The following exercise descriptions currently contain such statements that need to be revised:

- Sit to Stand
- Squats
- Hip Frontal Flexion
- Elbow Flexion
- Full Body Turn
- General - Shoulder
- Hip Abduction
- Knee Flexion
- Shoulder Abduction
Thank you for this very helpful recommendation. In the appendix, we have revised all of the descriptions that refer to the exercises improving balance, reducing the risk of falls or other in, ‘Sit to Stand’, ‘Squats’, ‘Hip Frontal Flexion’, ‘Elbow Flexion’, ‘Full Body Turn’, ‘General – Shoulder’, ‘Hip Abduction’, ‘Knee Flexion’ and, ‘Shoulder Abduction’ and instead used your suggested text, ‘This exercise targets….’ (Appendix 1, Table 1, middle column, pages 1-3).

Thank you for your feedback and we hope to hear from you soon.

Yours sincerely,

Dr Emma Stanmore (on behalf of all authors)