Author’s response to reviews

Title: Concurrent validity of the Fitbit for assessing sedentary behavior and moderate-to-vigorous physical activity

Authors:

Nicklaus Redenius (nicklaus.redenius@ndsu.edu)

Youngwon Kim (Youngwon.Kim@utah.edu)

Wonwoo Byun (won.byun@utah.edu)

Version: 2 Date: 16 Jan 2019

Author’s response to reviews:

Editor Comments:

1. We note the addition of a new author, Youngwon Kim, since the original submission of the manuscript.

In line with COPE guidelines, BioMed Central requires written confirmation from all authors that they agree with any proposed changes in authorship of submitted manuscripts. We would therefore need to ask that all authors complete our change of authorship form, which can be downloaded from http://www.biomedcentral.com/submissions/editorial-policies#authorship.

Please read the instructions in the form carefully, and return the form within 14 days with all signatures (including from any newly proposed/removed co-author(s)) to BMCSeriesEditorial@biomedcentral.com. Please also see our guidelines regarding qualification for authorship (http://www.biomedcentral.com/submissions/editorial-policies#authorship).

We have previously submitted the Change of Authorship Request Form (for the addition of Dr. Youngwon Kim) when we submitted the previous version of the revised manuscript. Also, please note that Dr. Kim has recently relocated from University of Utah to Hong Kong University, thus his affiliation information has been updated in the revision. Please let us know if you need any further information needed.

2. Please provide figure titles/legends under a separate heading of 'Figure Legends' after the References. If Figure titles/legends are within the main text of the manuscript, please move them.
Figure files should contain only the image/graphic, as well as any associated keys/annotations. If titles/legends are present within the figure files, please remove them.

Changes have been made as suggested.

3. Figures should be provided as separate files, and each figure of a manuscript should be submitted as a single file.

All figures are now submitted as separate files with the revised manuscript.

4. Please clarify how the FitBit and Actigraph were acquired for this study, were they donated by the company or purchased? Please provide details in the Methods section. If they were donated by the companies, please declare this in the Competing Interests section, for transparency.

All the GT3X+ and the Fitbit flex monitors used in this study were not donated by the manufacturers, but purchased by the investigative team using our own research funding. We have now clarified this in the Methods section in the revision.

5. Please include the email addresses for all authors on the title page. The corresponding author should still be indicated.

We have indicated all authors’ emails on the title page in the revision.

6. Please remove the response to reviewers' comments letter, as it is no longer needed at this stage. Please do not upload a new letter with this revision.

We have addressed all of the Reviewer 1’s comments in the revised manuscript. We do not include response to reviewers’ comments letter in the resubmission, but just provide here for the editor’s review.

7. At this stage, please upload your manuscript as a single, final, clean version that does not contain any tracked changes, comments, highlights, strikethroughs or text in different colours. All relevant tables/figures/additional files should also be clean versions. Figures (and additional files) should remain uploaded as separate files. We submitted the revised manuscript along with other related files as instructed by the editor.
Reviewer reports:

Dane R Domelen, Ph.D. (Reviewer 1):

- The equivalence testing is unclear to me. Typically in equivalence testing the null hypothesis is that the effect (in this case, difference in mean sedentary time based on Fitbit and GT3X+) is outside of some range, like 30 min/day, and the alternative hypothesis is that it is within that range. But it appears that you are basing it on a percentage of the GT3X+ mean, or 10% of 600.4 which is 60.04. That seems like a large range. Anyway, it would be helpful to give a little more detail on what exactly you are testing, and also how to interpret Figure 2. Does the entire interval for Fitbit have to fall inside interval for GT3X+ to reject the null? Also, I am not sure whether "significantly equivalent" is the right phrase.

As the reviewer inquired, equivalence testing is different than typical tests of differences like ANOVA or tests that provide “evidence of no differences”. However, equivalence testing can provide “Evidence of equivalence” by reversing the typical null hypothesis to specify that the different methods are not equivalent to each other. So rejecting this null hypothesis can lead to a conclusion that the different methods are significantly equivalent to each other.

For this method, we specified a zone of equivalence as +/- 10% of the mean of the GT3X+ values, and then we compared the confidence intervals of the Fitbit with the equivalence zone, and then saw if the CIs are completely included within the equivalence zone.

In the case, we performed equivalence tests separately for SB and MVPA, and as you can see, for MVPA, the Fitbit was equivalent to the GT3X+ on average MVPA since its CIs were completely within the equivalence zone.

We added a sentence “The estimates from the GT3X+ and the Fitbit Flex are equivalent if the CIs of the Fitbit Flex completely fall within the equivalence zone” to better explain this. Also, we removed a word ‘significantly’ throughout in the revised manuscript.

- Pearson correlation does not require normality, so I am not sure you necessarily had to exclude the 2 outliers. But if you wish to exclude the 2 observations, it should be briefly explained in the manuscript.

We excluded the two observations because their values were abnormal due to possible malfunction of activity monitors. It was an oversight, as they should have not been in the analysis for this study. As shown in the revision and acknowledged in the previous response to the reviewers’ comments (for the first revision), an exclusion of those two observations did not cause meaningful changes in the results, but just negligible changes in values reported in tables. Therefore, we feel that it is not necessary to include the explanation of a deletion of two observations that should have been excluded during the initial data screening process.
- Abstract. Why only report the $r = .65$, $\rho = .69$ correlation for MVPA? A range covering all 3 may be more appropriate.

We have updated it as recommended.

- Abstract. Meaning of +/- should be given, e.g. age: $M = 47$ years, $SD = 14$ years.

We made changes as suggested.

- Abstract. Suggest changing "...was estimated using three different cut-points..." to "...were estimated using various cut-points..." to avoid implying that 3 cut-points were also used for SED.

We made changes as suggested.

- P. 4, lines 17-18. Why not state the cut-points here?

We moved the cut-points’ definitions as recommended and updated the corresponding definition for the Freedson Vector Magnitude cut-points (VM3) to the Background section as well.

- P. 5, line 5. This is a very long paragraph; suggest breaking it up.

Thank you for the recommendation. We have created a new paragraph beginning after the updated placement of the cut-points’ definitions (as per the reviewer comment above).

- P. 6, lines 12-14. Again, you mention three cut-points, but really use one for SED and three for MVPA.

We used three different cut-points which are: 1) Freedson, 2) Troiano, and Freedson vector magnitude (VM3). Each set of cut-points defined SED and MVPA as follows: 1) $<100$ counts and $\geq1952$ counts/60-sec, 2) $<100$ counts and $\geq2020$ counts/60-sec, and 3) $\geq2691$ counts/60-sec (for MVPA only). Therefore, the cut-points for SED is the same as $<100$ couts/60-sec. This has explained in the manuscript (Please see the first paragraph in ‘Data reduction’ section).

- P. 6, line 17. Should be convenience sample.

Updated as recommended.

- P. 7. Should be consistent in usage of g vs. grams and hz vs. Hz.

Updated as recommended.

- P. 7, line 18. Should say "...were instructed to simultaneously..." since some non-compliance is expected.

Thank you for the suggestion. We clarified the methods as recommended.
We appreciate this suggestion. The sentence now reads “…which is a standard placement site for PA measurement using wrist-worn accelerometer…”

- P. 8, line 5. Should be e.g. not i.e.

Updated as directed.

- P. 8, line 7. Would not say it was downloaded immediately.

We deleted the word “immediately” and clarified the sentence to read as: Data from activity/sleep logs and activity monitors were retrieved and downloaded respectively at the conclusion of the data collection period.


Thank you for this edit. We deleted the repeat information.

- P. 9, line 16. Isn't there only 1 cut-point for SED?

Yes, the cut-point for SED is the same between cut-points (<100 counts/60-sec). Please also see our response to the reviewer’s comment above.

- P. 9, line 19. I would replace "to determine measurement errors" with "to assess similarity".

Updated as directed.

- P. 9, line 23. Systematic bias usually refers to the mean being wrong, not to two measurement methods having different variances.

Thanks for pointing this out. We have deleted ‘(i.e., systematic bias)’ in the revised manuscript.

- P. 11, line 5. Either say "all P < .01" or maybe omit p-value here since it is given in the table.

We have changed it to “all P < .01”

- P. 11, line 6. Should be “one-way repeated measures ANOVA" to be more precise.

Thank you for this edit. We specified the test as directed.

- For confidence intervals, I would suggest a different notation to avoid confusion between negatives and dashes. In Table 4, I think -74.05, -2.40 would be easier to interpret.

Thank you for the recommendation. We updated all confidence intervals in Table 4 to consistently reflect the new notation.
- P. 11, line 13. I think you need to be more precise with terminology. Pitman's test is for whether the variance of Fitbit and GT3X+ measurements are the same. I don't think it can be interpreted as having anything to do with random bias (do you mean random error?) or systematic bias.

Thank you for the comment. The BA plot is used for quantifying the bias (mean difference) and range of agreement. We have changed the wording as recommended in the revised manuscript.

- P. 12, line 11. Again, it is confusing to talk about 3 cut-points for SED, when it is really just 1 cut-point of 100.

We appreciate the comment. We updated the sentence to clarify that we are simply comparing the SED estimates between devices using the cut-points previously indicated for each in the Background section.

- P. 12, line 17. "Nearly an hour or more" should be reworded.

Thank you for the comment. The phrase has been removed and the sentence now reads “…Fitbit markedly overestimated mean daily MVPA”. The word “markedly” was added maintain the emphasis on the magnitude of difference between the Fitbit and GT3X+ MVPA estimates.

- P. 12, line 22. Should be "strong correlations".

Updated as directed.

- P. 12, line 23. Confusing to say "latter study" when citation hasn't been given yet.

Thank you for the comment. The phrase has been reworded so as to remove confusion: “…with moderate correlations for LPA reported also within the elderly populations”.

- P. 13, line 3. No need for comma after et al.

Comma deleted as directed.

- P. 13, line 20. Not sure what you mean by "...and SED differences were not significantly biased by the total amount of daily SED."

Thank you for pointing this out. We have deleted the words to make the sentence more clear.

- P. 15, line 20. Need citation for "more popular than hip worn models".

Done.

- P. 16, line 1. I think it is too strong of a statement to say that using 3 cut-points makes your results "not confounded by cut-point criteria non-equivalency."

We have revised the sentence to soften the statement.
- P. 16, line 18. Again, not sure about the "significantly equivalent" phrase, and also I think that some readers might view a 37 min/day difference as non-negligible.

Thank you for the comment. The “significantly” was replaced by “statistically” to eliminate confusion about the conclusion.

- Table 1. Footnote still refers to p-value.

Thank you for pointing this out. The footnote has been deleted.

- Table 2. I don't think you need the last footnote, but up to you/journal.

We deleted the footnote in Table 2.