Reviewer’s report

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

Version: 0 Date: 30 Aug 2018

Reviewer: Reviewer 2

Reviewer's report:

PEER REVIEWER COMMENTS: To view the full report from the academic peer reviewer, please see the attached file.

REVIEWER COMMENTS FROM REPORT: In this paper the authors aim to compare the concurrent validity of the Fitbit Flex to the ActiGraph GT3X+ in 67 participants with a wide spectrum of ages (20 - 70 yrs) and BMIs (18.6 - 40) during free-living over a 7 day period. Using three different (and validated) cut-points for the GT3X+ when worn at the waist, minutes spent performing MVPA and sedentary behaviour per day are extracted from the device and compared to outputs from the Fitbit Flex, which was worn at the wrist. The Fitbit estimated significantly higher MVPA compared to the GT3X+ but was comparable with regards to sedentary time. While of some scientific merit (to help advance the public health agenda regarding the objective measurement of physical activity dimensions using wearable devices) this paper has a number of issues that require addressing.

REQUESTED REVISIONS:

Major points

As the devices are worn at different anatomical locations and utilise different on-board processing steps what explains the potential differences between the two devices? Is this study a fair comparison? I understand consumer-based devices are more of a black box with proprietary algorithms etc. but is it possible to provide the dynamic range of the Fitbit Flex and sampling frequency so the reader can make a more informed comparison between the devices?

Why did the researchers choose the non-dominant wrist to wear the Fitbit Flex device? Can you please provide rationale or support for this decision? Similar to my point above, dominance is not standardised between devices as the Actigraph is worn on the dominant hip. It seems like this comparison is between apples and oranges.
Take a closer look at your references throughout. For example, in the data reduction section of the manuscript you present reference years rather than number, and in some instances both.

The fact you see no differences in mean bias between devices is merely an artefact of under and over estimations cancelling each other out. To truly understand the differences between the two devices the authors should consider calculating mean absolute error (Please see example: http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0171720). You could also consider adding equivalence testing also reported in this paper.

As neither of these devices are considered a representative criterion (gold-standard) isn't the discussion regarding under or over predicting somewhat arbitrary? For example, without a criterion measurement, if the Fitbit Flex over predicts MVPA compared to the GT3X+ this could also be that the GT3X+ under-predicts relative to the Fitbit Flex with no indication of which device is more accurate. From a public health perspective, what is an acceptable level of agreement between different devices if we are to make recommendations regarding optimal levels of MVPA for various health outcomes? A difference of 60 - 79 min/day translates to a 420 - 553 min/week differences that would have significant implications for observing whether someone is meeting the MVPA guidelines of 150 min/week. The authors should comment on these issues.

Given how quickly the physical activity monitoring field is progressing with technological developments and analytical improvements consumer-based wearable technology can become obsolete very quickly. The Flex has already been replaced by the next generation Flex 2. Therefore are the findings from this study now irrelevant given researchers will be attracted to the latest technology or generations of a wearable device. What is the specification differences between the Flex original and Flex 2 and are your findings translatable? I'm afraid we do not currently know this, which drastically limits the generalizability of your findings. This is a major limitation of this study.

ADDITIONAL REQUESTS/SUGGESTIONS:

Minor points

Abstract:

Line 14: You've abbreviated sedentary behavior (SED) but not PA, please correct.

Line 26: Should be written as 'time spent performing sedentary (SED) and moderate-to-vigorous PA (MVPA) was estimated…'

Line 34: Pearson correlations do not evaluate agreement, merely associations. Please revise this sentence.
Line 41: I'm not sure why you are presenting a range for MVPA (MD = -59 - 77 min/day) without specifying so. Also you state MVPA is significantly higher yet the values are proceeded with a minus symbol. Please also state the P value for significance.

Background:

Line 9: I think it would be useful to specify the health outcomes, please provide examples for the reader.

Line 21: What do you mean by feasibility in this context? Please explain, are you referring to wearability, cost etc?

Line 29 -30: consistency with abbreviations at the first time of use; IC (indirect calorimetry?), MET (metabolic equivalents)

Line 38: do you mean minutes of 'moderate to' vigorous PA which is the most common and clinically relevant category captured by wearable devices.

Line 46: insert 'intensity' after PA. Given the preceding sentence you are discussing the intensities of PA here specifically, not for example PAL or PAEE. Clarity around the terminology you use is very important and seemingly jumps around throughout the manuscript. The end of this sentence doesn't make sense. What do you mean by 'duration of activities used to calibrate the equation'. What equation are you referring to here (do you mean cut-point) and what is the calibration step? Perhaps you should also discuss that these cut-points are not validated for specific population groups (over-weight obese men and women or children) for example:

https://www.jsams.org/article/S1440-2440(11)00475-0/pdf

https://symbiosisonlinepublishing.com/exercise-sports-orthopedics/exercise-sports-orthopedics02.pdf

Pg 3. Line 12. Perhaps define the popularity of these consumer-based wearable devices. For context, how much of the general population own one of these devices? Moreover, what do you mean by acceptability? Do you mean that these devices are commonly worn in everyday life? Providing evidence and a reference for my first question would help support this assertion.

Pg 3. Line 17. Models? It is advisable to be more specific when referring to wearable devices used in each study rather than simply stating the manufacturer name. What Fitbit models are you referring to exactly; i.e. Fitbit charge 1,2,3, Versa, Ionic??? Please define throughout the manuscript. It is also pertinent to address whether any of these utilise multi-sensor signals (i.e. HR) as this would explain an improved estimate of EE.
Pg 3. Line 29. What sort of activities does the Fitbit under and over-predict for specifically (i.e. cycling, household chores etc). This paper demonstrates that the Fitbit Charge HR over-predicts walking and loaded walking (which presumably would be the mainstay of EE in free-living for the general population) and under-predicts for cycling: http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0171720

Methods

Pg 5. Line 38. Can you please provide more information pertaining to the definition of non-wear time? It is frustrating as a reader to be directed to another source.

Pg 6. Line 11. Did you check the normality of your data before proceeding with a parametric analysis (Pearson correlation coefficients)?

Pg 6. Line 29. Re-write as significant overall ANOVA 'effects' were followed by…

Pg 6. Line 34. BA also identifies random bias (LoA) as well as systematic. Consider also explaining that this approach is identifying the agreement between devices.

Results

Pg 6. Line 43. Consider changing subjects to participants throughout the manuscript: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1115535/

Pg 6. Line 53. I would not say that 14.9 hour/day demonstrates high compliance when you've asked participants to wear the devices during all waking and sleep hours (Pg 4. Line 20). Is it that these 14.9 hours represents those that you've removed as 'sleep'? Perhaps you should make this clearer in this sentence.

Discussion

Pg 7. Line 32: change to 'suggesting that the Fitbit Flex may overestimate MVPA in active individuals compared to the GT3X+'.

Pg 8. Paragraph from Line 8. I admire you trying to explain your results relative to the wider literature and identifying that sex might play a part (and your sample is mostly made up of females). However, a clearer description of how this is relevant needs to be provided. Is it that males are more active than females? Is this important because the range of values in a comparison can strengthen the magnitude of associations (i.e. larger range in males, 5 - 500 min/day vs. 3 - 200 min/day in females may explain the differences across studies)?
Author contributions: please define.

Are the methods appropriate and well described?
If not, please specify what is required in your comments to the authors.

No

Does the work include the necessary controls?
If not, please specify which controls are required in your comments to the authors.

No

Are the conclusions drawn adequately supported by the data shown?
If not, please explain in your comments to the authors.

No

Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review?
If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

Not relevant to this manuscript

Quality of written English
Please indicate the quality of language in the manuscript:

Acceptable

Declaration of competing interests
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?
6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

This reviewer has been recruited by a partner organization, Research Square. Reviewers with declared or apparent competing interests are not utilized for these reviews. This reviewer has agreed to publication of their comments online under a Creative Commons Attribution License attributed to Research Square and was paid a small honorarium for completing the review within a specified timeframe. Honoraria for reviews such as this are paid regardless of the reviewer recommendation.

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal.