Author’s response to reviews

Title: Psychometric Properties of the Zephyr Bioharness Device: A Systematic Review.

Authors:

Goris Nazari (gnazari@uwo.ca)
Pavlos Pavlos Bobos (bobos@uwo.ca)
Joy MacDermid (macderm@uwo.ca)
Kathryn Sinden (sinden@lakeheadu.ca)
Julie Richardson (richard@mcmaster.ca)
Ada Tang (tang@mcmaster.ca)

Version: 2 Date: 05 Dec 2017

Author’s response to reviews:

Reviewer’s comment: 2 48-52 "Agreement error of…” Please amend as only some studies showed this result with others providing poorer agreement.

Authors’ Response: Please see our discussion section on how and why we have stated this. Here’s what we have stated…

“In regards to ZB vs ECG comparisons, Flanagan et al. (2014) study rated at “Very good”, reported (-2.84 – 2.42) 95% LoA. Similarly, Dolezal et al. (2014) and Smith et al. (2014) studies rated at “Good”, reported (-0.21 – 0.14) and (-3.01 – 0.70) 95% LoA between ZB vs ECG respectively. It is important to note that there are no thresholds to help categorized 95% LoA into excellent or poor, however, narrower 95% LoA between ZB vs ECG, is suggestive of better agreement and possible interchangeable use. On the contrary, three studies, Kim et al. (2013), Rawstorn et al. (2015) and Gatti et al. (2014) reported somewhat wider 95% LoA in pair-wise device comparisons between ZB vs ECG. However, these studies had lower methodological scores (22,25,26). Therefore, studies with higher methodological quality scores that assessed ZB vs ECG agreements, displayed narrower 95% LoA than studies with lower methodological scores”.

IMPORTANT… if higher quality studies reported good agreement between ZB vs ECG, and lower quality studies reported poor agreement between ZB vs ECG, then, one must assume that the poor agreement measures reported between the two devices were influenced by how the study was conducted (study quality). Also note … Johnstone et al. (2012) and Johnstone et al. (2012a), although rated to have a high methodological quality, the studies did not assess ZB agreement measures against a criterion gold standard measures (e.g. ECG).
The reason why we stated “agreement error of ≤ 3.00 beats per minute”, was because we calculated the average score for Gatti et al. (2014) study. However, we understand that it can cause some confusion, therefore, we have now considered, and reported the range of point estimates based on all the included studies, i.e.

“Zephyr Bioharness agreement error ranged from -4.81 (under-estimation) to 3.00 (over-estimation) beats per minute, with varying 95% limits of agreement, when compared with gold standard measures”.

Reviewer’s comment: 4 33-36 The link between wearable devices and HR monitoring was non-existent and should be developed further.

Authors’ Response: The section was re-written to better link the wearable devices to HR monitoring.

Reviewer’s comment: 5 9-24 The focus of this section was on validity with only a minor statement about reliability. Justification for the current work needs further work.

Authors’ Response: Edited. Statement about reliability has been to highlight the importance reliability measures, and ultimately the importance of measurement properties.

Reviewer’s comment: 5 21 & 26 & 43 "clinical" What was the relevance of this as the current study did not examine clinical population studies?? Justification for the current work needs further work.

Authors’ Response: We mentioned clinical measurement properties, referring to the psychometric properties (reliability, validity, responsiveness and agreement parameters), and NOT referring to a specific clinical (disease/pathology) population. The measurement properties of Zephyr Bioharness have not yet been assessed within a clinical (disease/pathology) population. We understand that this may cause confusions as it has, therefore, we have now removed the term “clinical”, and have only include “measurement properties”.

Reviewer’s comment: 10 29-34 3 devices were reported here as being within good limits but another 3 studies showed quit different results between ZB and ECG. Why were the good studies reported on and others not?? Clarify the 'selective' reporting.

Authors’ Response: This is because if studies reported completely opposite results (as seen here), one must consider the quality of the studies – therefore, the reason for conducting this systematic review. And, the most important point here is … if higher quality studies reported good agreement between ZB vs ECG, and lower quality studies reported poor agreement between ZB vs ECG, then, one must assume that the poor agreement measures reported between the two devices were influenced by how the study was conducted (study quality). This is a standard
practice; if two studies demonstrate opposite results, one must pay attention to the quality of the studies (or risk of bias in RCTs). This is what we have done. This does not constitute 'selective reporting'. Had we not included the studies with poor qualities in our Table – 1, or assessed their quality in Table – 2, then… Yes, that might have implied selective reporting.

Reviewer’s comment: 10 55 "clinical" Why was this referred to as clinical given the populations and measure that was examined??

Authors’ Response: The word “clinical” has been removed.

Reviewer’s comment: 11 16-22 1 sentence does not make a paragraph in my opinion. Given the focus of the review, expansion of the reliability section was needed - what have other studies shown??

Authors’ Response: Four studies were identified, and all four studies demonstrated high reliability parameters. There were very little to almost no variations in results. We did not further expand on this because then we would have to re-state the same information that we stated in our results section. However, we have edited and re-written the paragraph.

Reviewer’s comment: 11 51- This section was very repetitive and should be condensed.

Authors’ Response: Edited.

Reviewer’s comment: 12 16 "formulaic method" Not clear what this meant and suggest to rephrase. I expect that the authors were indicating that no thresholds or limits exist to help catagorise LOA.

Authors’ Response: “formulaic method” has been rephrased, as suggested.

Reviewer’s comment: 12 26- Suggest to expand why 50% of the studies demonstrated narrow LOA and 50% showed very large LOA. This is very important to minimise selective and/or bias reporting.

Authors’ Response: Edited. However, there were no selective/reporting bias. We have included all the studies and reported the data in the Tables.

Reviewer’s comment: 12 28-56 This section should be condensed to focus the Discussion in line with the review aims.

Authors’ Response: Condensed.
Reviewer’s comment: 13 26-32 Please review and amend in light of the above comments as only some studies identified good agreement-validity with ECG.

Tables 3 and 4 Please check footnotes/legend for errors ("Mead")

Authors’ Response: Edited.