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: 1 Date: 28 Aug 2017

Author’s response to reviews:

Reviewer’s report: The authors are thanked for their study that reviewed the ZB for assessment of HR. This was an interesting topic with some aspects requiring further consideration. Firstly, justification for this study was poor and should be significantly improved to warrant need and publication. Secondly, the Results section was awkward with a review of this needed to clearly delineate the comparisons with a criterion and/or other device. The relevance of some results (and not others) in text was unclear and should be reviewed. Finally, the Discussion was very brief with very little discussion of the reliability/validity of the ZB and other studies that have examined these aspects of the ZB or other similar wearable devices. The authors are encouraged to review and focus the current study and Discussion on these parameters to highlight the unique aspects of the study and need for publication.

Authors’ Response: We have re-written and edited our study Background and provided valuable information to justify the importance of our Systematic Review. We have also re-written our Results section and made the necessary (reviewer-recommended changes). The Discussion section has been edited and comparisons of different studies discussed.

Reviewer’s comment: Lines 36-39…. MCID was not discussed in the manuscript. Relevance?

Authors’ Response: No MCID data was identified or extracted, therefore we have not removed this from our Abstract and Methods sections.

Reviewer’s comment: Line 48…."bronze". Please clarify as not a common term.
Authors’ Response: The term “bronze” is referred to frequently used devices, however not gold standard devices. In this manuscript, it pertains to other commercially available devices. We have removed the term “bronze” as it is not a common term, and replaced it with “other commercially used devices”.

Reviewer’s comment: Line 51…"narrow" Please clarify how this term was defined in the manuscript as one may argue that LOA weren’t narrow.

Authors’ Response: We did not define the term “narrow” in our manuscript. We believe that the term “narrow” can be misleading, therefore we have removed the term “narrow” and include the term “varying” Limits of Agreement (LOA).

Reviewer’s comment: Line 14… Reference 5 - suitable and valid source?? Source was difficult to determine from the listed format.

Authors’ Response: We have edited the format.

Reviewer’s comment: Line 22-44… This section was long and should be condensed to focus the manuscript.

Authors’ Response: We have edited this section and condensed it from 187 words to 91 words.

Reviewer’s comment: Lines 23-27 The aim of the study was written clearly but it was not clear how this aim was arrived at given the Introduction. Greater justification for the review is necessary, other than nobody has done it. Further, why was HR the focus given the capabilities of the harness?? A focus on HR may be of limited application for the harness and therefore a strong justification for HR examination is vital.

Authors’ Response: We have re-written our Background section and have highlighted important points to justify the importance of our Systematic Review. The focus of the review was the HR because it the most frequently reported Zephyr Bioharness variable. Very limited information on measurement properties of respiratory rate and other measures were available.

Reviewer’s comment: Line 34… "physiological responses", Please clarify that it was HR only as the harness can collect other physiological responses (which may be of wider relevance).

Authors’ Response: The harness can collect a wide range of physiological responses. We have stated “physiological responses” under the exclusion criteria – that is studies that report other physiological responses other than heart rate will be excluded. Under the inclusion criteria however, we have stated “heart rate only”. Therefore, we have done exactly what the reviewer has said – so… no changes were made.

Reviewer’s comment: Lines 43-47… SEM, mean differences and CI were extracted. Please clarify how these were used as it was not obvious in the results or interpretation.
Authors’ Response: SEM, mean differences and CI were extracted and have been displayed in Table 3. In the results section, the mean differences have already been mentioned. For the SEM, we have added the SEM range values.

As for the interpretation of these values, for SEM there are no ways to categorise or interpret them except by just reporting the value. This holds true for CI as well (reported in Table – 3). As for the ICC reliability, under the “Data Extraction” section in our methods, we have defined ICC thresholds and interpretations as – ICC < 0.40 indicate poor, 0.40 ≤ ICC <0.75 indicate fair to good and ICC ≥ 0.75 indicate excellent reliability. And used terms “poor” or “fair to good” or excellent in our results section rather than just reporting the value.

Reviewer’s comment: Lines 6-12… MCID and CID were extracted or calculated?? Further, how were these used as there were no results or discussion of these variables in the manuscript?? If not used/discussed, then they were not needed and should be removed.

Authors’ Response: We original planned to report MCID and CID parameters. However, during data extraction, none of the studies had reported these two parameters. Therefore, we have now removed the above-mentioned information concerning the MCID and CID, both from the abstract and methods sections.

Reviewer’s comment: Lines 29-34…. The tool has been used in MS outcomes but clarification was needed regarding its suitability for other variables including the current HR variable. Further, how was "thorough literature review" defined as this could be quite variable between studies and researchers’ interpretations??

Authors’ Response: The tool used in our is a tool used for Quality Appraisal of Clinical Measurement studies (Psychometric properties) and it is independent of kind of variable used. For as long as studies report clinical measurement properties (psychometric properties) including ICC, SEM, Pearsons correlations, agreement properties, or responsiveness, this tool can be used to assess the quality of those studies. Its wide spread applicability regardless of type of variable used (MS or heart rate), is similar to the Cochrane Collaborations Tool for Assessing Risk of Bias. The Tool for Assessing Risk of Bias can be used to determine the risk of bias in Randomized Controlled Trails irrespective of what kind of intervention (MS or Neuro or any…) used.

The terms "thorough literature review" is one of the 12 sub-sections in our Quality Appraisal Tool. Four requirements must be met for a study to be given a maximum score of 2. We did not include detailed explanation of each of these 12 subsections because they are included in the Quality Appraisal Tool user guide, and frankly would take too much space in our manuscript. The "thorough literature review" subsection has four requirements to be fulfilled and given a maximum score of 2:

1. Authors performed a thorough literature review indicating what is currently known, and not known, about the clinical measurement properties of the instrument or tests under the study.
2. Authors presented a critical and unbiased view of what is known about the current measurement properties.

3. Indicated how the current research question fills a gap in the current knowledge base.

4. Established a research question based on the above.

If a study being rated, fulfils all these points, that study will be given a maximum score of 2 on this subsection of the Quality Appraisal Tool.

Reviewer’s comment: Lines 48-53… Please clarify how an article's quality score was calculated as the current results seem to indicate a different methodology than that written here (i.e. total sum divided by best possible score)?? Further, what was the range of scores for each item - 0 to 2??

Authors’ Response: The current results do not indicate a different methodology than that written in the manuscript. In our manuscript, we have not mentioned “total sum divided by best possible score”.

In our manuscript, we have been consistent with how the studies were score. Please see below on how each study was score.

Total score \[\text{sum of subtotals} ÷ 24 \times 100\]. For example, Rawstorn et al. (2015), had a subtotal score of 16. Therefore, 16 was divided by 24 (total items), then multiplied by 100. This gives a percentage score of 66.666 \(\rightarrow\) 67%.

However, if for a specific paper an item is deemed NA (Not Applicable), then, Total score = \[\text{sum of subtotals} ÷ (2 \times \text{number of Applicable items}) \times 100\]. The subsections no. 6, asks for percentage of retention/follow up. This subsection only applies to reliability test-retest studies. For other studies, validity etc…., the sub-section no. 6 was deemed NA. Therefore, Total score = \[\text{sum of subtotals} ÷ (2 \times \text{number of Applicable items}) \times 100\].

Yes… the range of scores for each item was 0 to 2. That is 0, 1 or 2.

Reviewer’s comment: Lines 43… "Table 2" I'd also suggest to refer to Figure 1 that shows the scores for quality of studies.

Authors’ Response: The reference to Figure 1 has been added to the text.

Reviewer’s comment: Lines 43-49… Suggest authors review Table 2 for the most commons flaws as "specific hypotheses" scored poorly and was not stated here in text.

Authors’ Response: The statement on lack of “specific hypotheses” sub-section scored has been added to the text.

Reviewer’s comment: Lines 14… SEM or CV?? Please review and clarify.
Authors’ Response: Some authors reported the SEM and the others CV. In our Table – 3, we have clarified this by placing SEM or CV next to the reported scores.

Reviewer’s comment: Lines 26-39… This was difficult to follow and suggest to combine and simplify. Details of the physical activities was not needed here.

Authors’ Response: We have combined and simplified this section, and removed the physical activity details.

Reviewer’s comment: Lines 41… "device heart rate variable" should be rephrased for clarity.

Authors’ Response: Since this was a systematic review on psychometric properties of Zephyr Bioharness device (used to measure heart rate), We have rephrased "device heart rate variable" to “device”.

Reviewer’s comment: Lines 45-48, & 58…. Authors should clearly indicate here those studies that comparred ZB HR with ECG and those studies that compared HR with a telemetric HR monitor (e.g. study 1 and 18). Was the POLAR HR monitor considered a "criterion" measure?? If so, please explain how it was defined as a criterion measure and the qualifications to be considered a criterion measure. If not, then the Results section should be amended to indicate the criterion devices and the results of such comparisons.

Authors’ Response: The Validity and Agreement results sections have been amended and separated to indicate ZB comparisons against Polar T31 (commercially used device), and against gold standard criterion measure ECG. (as suggested by the reviewer).

Reviewer’s comment: Lines 4-9… Please review this statement and the results as I could not see how 4 were identified. Further, how was narrow LOA defined??

Authors’ Response: We reviewed this statement. Three studies were identified, and we have made the necessary changes. We have removed the term “narrow LOA”, and instead include the actual LOA values.

Reviewer’s comment: Lines 11-15… Please clarify why this result was presented as it was not clear of its relevance.

Authors’ Response: This result has been removed from the text as it was not relevant.

Reviewer’s comment: Lines 16-22… This was the first time that the results were indicated for a comparison between ZB and the POLAR HR monitor. This should be clearly indicated previously in the Results and differences noted between comparisons with an ECG and a POLAR clearly stated.

Authors’ Response: In our Results, under the Validity and Agreement sections, we have now indicated the differences between ZB heart rate comparisons both with the ECG and Polar T31.
Reviewer’s comment: Lines 24-29… Please indicate if this statement was in reference to the ECG or POLAR comparisons or both devices?? As indicated above, this should be clearly indicated and justifiable.

Authors’ Response: This statement was in reference to comparing ZB against ECG and Polar T31. We have now made this clear. We stated that ZB showed better agreement with ECG, because of mean differences of ≤ 3.00 and (-3.10 – 2.42) 95% limits of agreement. However, ZB vs Polar T31, displayed mean differences of ≤ 3.05 with (-79.20 – 79.20) 95% limits of agreement.

There is no specific way to define narrow LOA. However, LOA in pair-wise comparison (for example ZB vs ECG) are used in comparison with other devices (for example ZB vs Polar T31). Since ZB vs ECG displayed (-3.10 – 2.42) 95% limits of agreement, whereas ZB vs Polar T31 displayed (-79.20 – 79.20) 95% limits of agreement, it is more likely that ZB was more agreeable with ECG than with Polar T31.

Reviewer’s comment: Lines 41…. "good to excellent quality" Please clarify how this was defined and shown in the current results.

Authors’ Response: This was defined and shown in our Fig. 1. However, we have now included the same information in our Methods section.

Reviewer’s comment: Lines 51 & 53 "low risk" Clarify how AF patients are low risk. Many clinicians and Exercise Physiologists would consider that AF patients are a high-risk group.

Authors’ Response: We did not state that the AF patients were “low risk”. We have stated that…

“This review suggests that the Zephyr Bioharness device might be useful for prevention or rehabilitation applications where field-based monitoring of heart rate is required in low risk patient populations. The use of the devices in high-risk populations was not studied”.

The psychometric properties of Zephyr Bioharness were established mostly in healthy participants.

We have stated that good to excellent quality evidence exists to support the reliability and validity of the device. And that the device might be used for prevention and rehabilitation applications in low risk patients – based on clinicians’ judgement. Since the device measurement properties were mostly established in healthy participants (good to excellent quality evidence), it is possible that the device can also be used to monitor low risk individuals (for example healthy individuals recovering from and ACL injury or a fracture) heart rates.

Reviewer’s comment: Line 7 The focus here was on use of wearable devices in clinical practice. Please explain why this focus was provided given the focus of the studies in the review (most were non-clinical).
Authors’ Response: it is true that the focus of the studies in the review were mostly non-clinical. However, the use of Zephyr Bioharness to monitoring heart rate can also be extended into clinical settings. We have stated that “It can be extended … when its measurement properties are established (future studies)”

Reviewer’s comment: Lines 33-36 Why were other measurements from the ZB not examined given its widespread use for these measures?? How relevant was the examination of HR alone??

Authors’ Response: The other measurements from ZB were not examined because reliability and validity data were not available/sufficient. Data from other measurements such as respiratory rate, were only available from two studies. We assessed only the quality of evidence concerning the reliability and validity of heart rate using the Zephyr Bioharness device because HR was the most frequently reported variables (HR was reported in 10 studies)

Reviewer’s comment: Table 2 NA?? Please clarify how a scoring of NA was possible. Also, how was each property scored = 0 to 2?? Please provide details.

Authors’ Response: The subsections no. 6, asks for – percentage of retention/follow up? This subsection only applies to reliability test-retest studies. Detailing of how each subsection was score will not be possible in the manuscript because there are 12 subsections, and each subsection has 4 – 5 details points on how a judgement score of 0, 1 or 2 should be considered. We will add the Quality scale as appendix.

Reviewer’s comment: Table 3 SEM or CV?? Please clarify.

Authors’ Response: SEM or CV has been clarified for each study.

Reviewer’s comment: Table 4 and 5 Clarify why some studies that listed different protocols only had one mean difference (e.g. 19). Are both protocols then relevant??

Authors’ Response: the term “baseline” for study 19, was a typo.

For study 18 – only one mean was reported.

For study 16 – only one mean was reported for “Treadmill, Cycle Ergometer and Activities of daily living; sweep and vacuum” protocol.