Author’s response to reviews

Title: Handgrip Force Steadiness in young and older adults: A Reproducibility Study

Authors:
Andreas Blomkvist (andreas.wahl@gmail.com)
Fredrik Eika (eikafredrik@gmail.com)
Eling D. de Bruin (eling.debruin@hest.ethz.ch)
Stig Andersen (lasa@rn.dk)
Martin Jorgensen (mgj@rn.dk)

Version: 2 Date: 01 Feb 2018

Author’s response to reviews:

Response letter

Manuscript: BMSD-D-17-01189R1 entitled “Handgrip Force Steadiness in young and older adults using the Nintendo Wii Balance Board: A Reproducibility Study”

Dear Charlotte Beaudart,

We greatly appreciate the thorough review of our manuscript. We are confident that the manuscript has been improved by the adjustments based on the reviewer’s comments. In addition to these adjustments some additional changes are marked in the revised manuscript. These can be found at:

• Page 2, line 11 and 13: corrected typo
• Page 2, line 15: specificsed which group was tested with reproducibility analysis
• Page 5 line 8 and 21: corrected typo
Below please find the reviewer’s comments followed by our point-by-point response and further adjustments.

Response to reviewer #1 (Charlotte Beaudart):

“This is a good and well-written manuscript.”

Thank you for the kind and encouraging comment.

“A validation against a gold standard (e.g. Jamar dynamometer) has not been done in this study. Why?”

We agree that validation against a gold standard is warranted as the next research step within a comprehensive approach of determining the psychometric properties of the measurement approach. However, to our knowledge a gold standard is lacking for measuring force steadiness. The Jamar dynamometer measures maximal voluntary contraction and is not validated for force steadiness.

“Did the authors ensure that no particular physical changes occurred between the two assessments? If yes, how did they proceed?”

We aimed for a pragmatic protocol to portray daily practice. Regarding the testing environment, in order to simulate a simple community assessment scenario, we repeated the test one week later
in the same room under similar conditions assuming that within this time frame no dramatic changes in force steadiness would take place.

“Only healthy populations participated to this study which could be interpreted as a limit.”

We agree, and the next step could be to investigate a more frail population. However, we chose to include healthy participants in this study for two reasons. Healthy individuals is a reasonable starting population as an initial test of a novel method, as they make up a reference population useful for later comparisons with frail populations. More importantly, they are a more homogenous sample. If the reproducibility results are good, then the method is more likely to be also useful for a more heterogeneous sample (e.g. patients with neuromuscular disorders).

“Figure 1 should be changed with a cleared picture.”

We agree with the reviewer that the figure was not optimal for the purpose of our manuscript. In the revised version, we have added an alternative figure (Alt_Figure 1.jpg), which we believe is of higher quality.

“In the protocol, did the evaluators encourage participants (to get the best measurements)? Verbal encouragements have been showed to strongly impacted the results of strength assessments. If yes, how did the authors control results for inter-rater reliability?”

This is a fine point by the reviewer. We did not verbally encourage the participants. To the best of our knowledge, no studies have investigated the influence of this on force steadiness measures. We tried to minimize the unspecific influence by assessors. The testing procedures
was done by the participant working with the Nintendo Wii Balance Board and the personal computer. Practically speaking, this amounts to only one rater (the software).

Response to reviewer #2 (Médéa Locquet):

“This manuscript brings original elements, previously very little developed in the scientific literature. The article is well written, very complete and the statistical analyzes are appropriate and well thought out.”

We greatly appreciate the kind comments.

“Abstract – line 9: typing error -&gt; edit «measuring» in «measuring»”

Thank you. This has been corrected (page 2, line 6 in the revised manuscript).

“Background - line 57: to measure muscle steadiness I guess? Clarify that.”

We would like to clarify any unclear statements. However, we are unsure where in the manuscript the reviewer is referring. Line 57 in the manuscript is in the results section of the abstract, and the Background section has less than 57 lines. We will correct any unclear statement as soon as the statement is clear to us.

“Methods: line 44: AWB - FE -&gt; define in full letters a first time.”
We agree with the reviewer, and the manuscript has been edited accordingly (page 6, line 19-20).

“Results - Table 1 and text line 1: Why not compare groups statistically by adjusting for age, to see if the difference observed are maintained?”

As the reviewer points out, table 1 refers to the two groups (young and old) we statistically tested for age differences in force steadiness. These results are presented in Table 4. The purpose of the test was to test for age differences, hence, we did not adjust for age.

“In general: as the authors have pointed out, a major limitation is the sex difference in steadiness. Indeed, muscle variables are markedly different between men and women. So if the distribution is not fifty-fifty (which is the case), the differences and results observed in muscle steadiness between the groups could simply reflect the difference in gender distribution. Why not do the analyzes men and women separately?”

This is a relevant observation by the reviewer. As found in some studies, there is likely a sex difference in force steadiness, with males being steadier than females, possibly specific to the upper limb. We know of no studies that have investigated this difference for handgrip steadiness specifically. Still, the reported sex difference in steadiness is of a lower magnitude than the reported age difference in steadiness. In our study, the possible sex difference would augment the age difference observed, since there is a slightly higher proportion of men in the young group. Splitting the two groups into four by sex would yield a small sample size and, hence, high uncertainty in the estimates. A post hoc analysis using males and results from CVM measurements confirmed this. Thus, age difference were present for all MVC values (as reported in table 4), but none were statistically significant with wide confidence intervals. We are willing to add these results, but we are hesitant when weighted against the added value to our paper.
Still, we agree with the point raised by the reviewer, and we have added it to the discussions section as follows:

“It should be noted that some studies have found small sex differences in force steadiness\textsuperscript{46,47} with males being more steady than females. Thus, the small difference in sex distribution between age groups in our data may have augmented the age-difference (table 1). Our sample size was too small to provide estimate of the age-difference for male and female separately according to a post hoc analysis.” (page 10 and 11, line 22-25 and 1-2, respectively)

Response to reviewer #3 (Olivier Bruyère):

“The advertisement for the Nintendo Wii should be removed from the title.”

We have changed the title to “Handgrip Force Steadiness in young and older adults: A Reproducibility Study” (page 1, line 2). For clarification, none of the authors have any connection to Nintendo, financial or otherwise.

“In the abstract, relative reproducibility and absolute reproducibility must be defined.”

This is well observed by the reviewer. We used the terms “relative” and “absolute” reproducibility in the abstract while “reliability” and “agreement” were used in the manuscript. We have aligned the abstract with the rest of the manuscript. In addition, we have added a definition of reproducibility in the abstract (page 2, line 9-10). Further clarification of these terms are found in the method section. Thank you for pointing this out.
“The assessment of the differences between young adults and older adults is not the objective of this study and must be removed from the manuscript.”

The reviewer is correct in that the secondary objective of this study was not stated clearly in the abstract. The comparison between young and old adults was a secondary objective, as stated at the end of the introduction paragraph. We have clarified the objectives of the study in the abstract and introduction (page 2, line 6-8, and page 4, line 23).

“One of the major issues in this manuscript is related to the validity of the Nintendo Wii Balance Board to assess hand grip force steadiness. As a matter of fact, it would be very interesting to assess the validity of the assessment using a defined force steadiness to be sure that what is assessed by the apparatus is true.”

We agree with the reviewer, and these studies have been done previously as validity aspects belong to a comprehensive assessment of psychometric measurement properties. In this phase of the assessment we, however, focused on the reliability of the approach. The validity of the apparatus, both as a measurement device for accurately assessing force and the assessment of maximal handgrip strength has been previously confirmed , .

“The second issue in this manuscript is related to the use of the 5, 10, and 25% of MVC. As a matter of fact, since other studies have also assessed the 50% of MVC, the rational is not very clear. The question is whether it would be clinically relevant to assess the 5 or 10% of MVC.”

This is a good and well taken point. We chose 5 and 10% of MVC for several reasons. Firstly, force steadiness at these lower force levels have shown to vary the most with respect to age and sexe.e.g. , . For a given method, the more heterogeneous a sample is with respect to the measure
under investigation, the more reliable (i.e. higher intraclass correlation coefficient) the method becomes since it is less likely that the individuals change rank with repeated tests. If our method was unreliable for lower force levels, it is unlikely that it will be reliable in higher force levels. However, if the method was unreliable at higher force levels, it would simply be unknown if it was reliable for lower force levels. This effect is confirmed in our study where the reliability is higher for lower force levels. Secondly, it is at the lower force levels studies have found the largest fluctuations in force production (i.e lower force steadiness) among individuals with neurological disorders. Interestingly, force steadiness is best correlated with functional tests at lower force levels, and sometimes only at the lower levels e.g. . Hence, it is at the lower force levels we can first hope to develop a clinically useful method to assess functional limitations with the greatest accuracy. Thirdly, higher percentage of MVC requires more force to maintain target force, and for durations around 20 seconds or more (which is likely necessary to accurately assess force steadiness), we increasingly risk testing muscle strength endurance rather than the accuracy of force production. This was important for us to avoid because previous research on sustained handgrip contraction has found this measure to be unreliable and not recommended for clinical use . As a final point, 50% of MVC is also higher than what is normally needed for normal everyday activities, such as when holding small items. We thus settled with 5, 10 and 25% based on these considerations and have added information on our rationale to the manuscript.

“More information must be given regarding the SRD.”

We agree with the reviewer that more information on SRD could be provided. We have expanded the relevant paragraph to read: “Beckerman et al.50, introduced the SRD to indicate whether a subsequent measurement represents a real change and not random variation. SRD represents the smallest change in measurement, which can be attributed to a real change in the parameter assessed. A SRD of 15% means that the subsequent force steadiness measurement must exceed a 15% difference to indicate a true change in force steadiness.” (page 12, line 2-6)
“The reviewer does not fully understand the added value of the assessment of the area between the target force line and the force produced.”

A priori the area between the target force line and force production could be an interesting measure of the total force deviation from target force as it is independent of the standard deviation and the mean force produced. It is an approximated measure of the absolute deviation between target force line and the force produced. We did not know, whether this measure would prove useful. As the results indicate, this measure was unreliable.

“The reliability of the visio-motor feedback must be discussed.”

We have added the following specification at the relevant paragraph: “Specifically, old adults, as opposed to young adults, seem to have significantly lower steadiness with visual feedback compared to no visual feedback at lower forces52.” (page 13, line 2-3)

“The authors concluded that the WPD method is a relatively cheap, widely available, pragmatic and portable method. However, the reviewer is wondering whether the software used in this study is freely and widely available to assess the hand grip force steadiness.”

This is a good point by the reviewer. The FysioMeter software is not for free but is low-cost compared to similar type of laboratory equipment. The sentence the reviewer is referring to has been rephrased to accommodate this point: “The WBB is a relatively cheap, widely available and portable device, and the FysioMeter software is accessible to support the method.” (page 13, line 9-10)
“At last, the reviewer is not sure that this method could be relevant in a clinical setting because of the limited reproducibility of the method but believes that it could be very useful in epidemiological studies (i.e. at the population level). Could the authors comment on that?”

We support the point on the relevance for epidemiological studies and agree that extra work is needed to assess the utility for clinical settings. With the reliability results observed in our study (indicated by the intra-class correlation coefficient), the method is useful for comparisons between groups of individuals, especially for larger epidemiological samples. However, with the systematic difference in mean between sessions and the lower agreement results (especially for 10% and 25% of MVC) compared to the reliability results, the method is less suitable for a clinical setting. Still, the systematic difference, particularly for the left hand, is an interesting finding that could be a significant training effect after a single test. If improving force steadiness at low force levels is an effective way to increase manual dexterity and hand muscle performance, as clearly indicated by the study by Kornatz KW et al., the method could serve as an intervention tool and an intra-individual index for hand muscle performance. Still, this is for future studies to confirm.

References


