Author’s response to reviews

Title: Validity and bias on the online Active Australia Survey: Activity level and participant factors associated with self-report bias

Authors:

Rachel Curtis (Rachel.Curtis@unisa.edu.au)
Timothy Olds (Timothy.Olds@unisa.edu.au)
Ronald Plotnikoff (ron.plotnikoff@newcastle.edu.au)
Corneel Vandelanotte (c.vandelanotte@cqu.edu.au)
Sarah Edney (sarah.edney@mymail.unisa.edu.au)
Jillian Ryan (Jillian.Ryan@unisa.edu.au)
Carol Maher (Carol.Maher@unisa.edu.au)

Version: 1 Date: 13 Nov 2019

Author’s response to reviews:

Validity and bias on the online Active Australia Survey: Activity level and participant factors associated with self-report bias (BMRM-D-19-00398 Revised)

Tamara Hughes, PhD
Editor, BMC Medical Research Methodology

Dear Dr Hughes,

Thank you for the opportunity to revise and resubmit our manuscript, now titled “Validity and bias on the online Active Australia Survey: Activity level and participant factors associated with self-report bias” (BMRM-D-19-00398). We have revised the manuscript according to the Reviewers’ recommendations and respond to each point below. We would like to thank the Reviewers for their insightful and thorough comments, and believe our changes have significantly improved the manuscript. Changes to the manuscript are highlighted in yellow. We look forward to hearing the outcome of the next stage of review.

Sincerely,

Rachel Curtis
Reviewer 1 Volkert Siersma

1. My first comment is on the generic use of validity. What is investigated is the association between AAS and a direct accelerometer measure of activity. Hence, the primary aim of the analysis (aim 1) is the investigation of criterion validity: for AAS to be any good, there has to be an association with the accelerometer measure, else we will not believe the AAS. Thereafter, the authors proceed by imagining that the accelerometer is some gold standard, while the authors themselves remark that also the accelerometer measure is not without errors. But fine enough, we will investigate how accurate AAS measures the values obtained by the accelerometer; that has possibly some merit in itself. However, this I will call accuracy, not validity.

We agree that we should have been more specific with our use of the term ‘validity’ and have corrected this to ‘criterion validity’ where needed throughout the manuscript. We also understand that accelerometry is not without measurement error. Nonetheless, it is a useful measure of activity in a free-living environment, which the Active Australia Survey seeks to capture. Additionally, accelerometers are commonly used to assess the criterion validity of physical activity questionnaires – a method which is well accepted within the field. We have edited the background section of the abstract so that it is clear to the reader in the first sentence of the paper that we are using accelerometry as the criterion:

This study examined the criterion validity of the online Active Australia Survey, using accelerometry as the criterion, and whether self-report bias was related to level of activity, age, sex, education, body mass index and health-related quality of life. (page 2, lines 2-4)

2. In the introduction, and, later on, commenting on the results of the paper, a large number of Spearman's rho measures of association (mostly the authors are correct in writing "association", but at points "correlation" is used, Spearman's rho is strictly not a correlation coefficient). These are freely compared between populations, e.g. in the first line of the discussion: "lower than reported in most previous studies". But Spearman's rho (and quite some other measures of association) is dependent on the marginal distributions of the sample! They are not to be compared between populations! It is even hard to gauge what a rho that is not (close to) 0 or 1 really means: rho=0.5, is this a strong association?

We have replaced ‘correlation’ with ‘association’ throughout the manuscript. We have removed the problematic sentence in the first line of the discussion on page 11, and edited a key sentence in the introduction to remove direct comparison between coefficients:

Preliminary research, outlined below, suggests the online AAS may be unreliable. (page 3, lines 17-18).
We would argue, however, that Spearman’s rho does provide a useful indication of the strength of the monotonic rank association between 2 variables. In the introduction, we therefore provide our interpretation of the strength of the coefficients but also provide the exact values so that the reader can make their own interpretations. This approach is consistent with previous studies examining the criterion validity of the AAS. We understand that Spearman’s rho is not a measure of the strength of agreement between variables; therefore, we have also examined agreement using bias scores and a Bland-Altman plot.

3. The AAS has eight items, but in this paper only four of them are used (I read on page 6, line 12). Hence, only the "minutes" subscale of AAS is "validated" in this paper; the reason for this (page 6, line 13) is "recommendation for calculating sufficient MVPA" which phrase I do not understand. The "number of times" subscale/items is ignored.

The Reviewer is correct; the paper sought to examine minutes of MVPA on the AAS and did not examine the number of MVPA sessions. We have made this clearer in the aims and have also clarified the method of calculation of MVPA:

The objectives of this study were therefore to (1) examine the criterion validity of minutes of MVPA on the online AAS relative to accelerometry… (page 5, line 10-11)

Weekly MVPA was calculated as the sum of minutes spent walking, in moderate activity, and in vigorous activity. This follows the standard method for calculating total activity [5], except that vigorous activity was not weighted to enable comparison with accelerometer-recorded MVPA minutes. Daily MVPA was calculated as weekly MVPA divided by 7. (page 7, lines 11-13)

4. Accelerometer measures are only included in the study when the device is used more than a given amount of time in the week; often not the full 7 days, 24 hours. Hence, the accelerometer measure is always biased downward! Maybe an idea to add an indicator variable to the bias predictor list whether the accelerometer was used (almost) the full week, or less. The accelerometer is not used in water-based activities, which is slightly funny given the popular image of Australians as surfers, etc.

Daytime wear compliance was excellent, therefore we determined that we did not require inclusion of an indicator of wear time in the model. We now include information on compliance in the methods:

Participants had valid accelerometry data for a median of 7 days (IQR 6–7) and wore the accelerometer for a median of 16.9 hours per day (IQR 16.0–17.7). Once sleep logs were included for participants who did not wear the accelerometer overnight, data summed to 23.7 hours (IQR 23.5–24.0), indicating there was little missing activity data. (page 6, line 25 – page 7, line 2)
We now include as a limitation that the accelerometer was not worn during water-based activities:

Additionally, the accelerometer could not be worn during water-based activities. (page 14, line 5)

5. What is a relative association (page 7 line 12)? The description of the Bland-Altman plot is a little confusing (page 7 line 14). The absolute difference is the same as the bias scores I figure.

We have re-written this section to improve clarity:

To examine objective 1 (criterion validity of the online AAS relative to accelerometry), a Spearman’s rank coefficient described the association between AAS-derived and accelerometer-derived MVPA minutes. Bias scores were calculated for each participant as AAS-reported - accelerometer-recorded MVPA. A Bland-Altman plot presented all bias scores and indicated the mean bias and the limits of agreement (±1.96SD; an interval within which 95% of the bias scores lie) [20]. Bias scores, which were approximately normally distributed, were plotted against the average of the two measures as a proxy for the ‘true’ level of MVPA (page 7, line 21 – page 8, line 7)

6. There is higher variance for higher values of activity: nobody thought about using a logarithm transform, i.e. using a measure of bias that reports on the percentage of the bias relative to the mean of the AAS and accelerometer measure? And instead of just the accelerometer measure, the mean of AAS and the accelerometer measure may be used as bias predictor.

We did not consider log transformation to be required as bias scores were normally distributed and back transformation would be required for interpretation. However, we appreciate the suggestion to additionally examine bias scores using the % of bias relative to the mean of the measures. Therefore, to further examine the nature of bias, an additional plot examined the difference between AAS-reported and accelerometer-recorded daily minutes of MVPA, expressed as a % of the average of the two measures. On average, participants reported 1% less MVPA per day on the AAS (M = -1.3, SD = 98.5) than was recorded by accelerometry, with the limits of agreement ranging from -194.4% to 191.8%. The mean bias percentage was not significantly different from 0 (t(343) = -0.24, p = .81). As the results of this analysis were similar to the previous analysis, we have elected not to include this in the revised manuscript.

Although we used the average of the two measures for the x-axis of the Bland-Altman plot, we chose to use accelerometry (rather than the average of the two measures) as a predictor of bias to enable comparison to only previous study examining predictors of the association between the online AAS and accelerometry [11].

7. Page 8 line 2: how adjusts this estimator for missing data?
We now include this information in the manuscript:

The full information maximum likelihood estimator was used to enable inclusion of incomplete cases by estimating parameters using all available data points (note that missing data was minimal: 3 participants had missing data for BMI, with 2 of these participants also missing data for age). (page 8, lines 16-18)

8. How are all these information criteria used?

Although we were not aiming to specifically identify a model of best fit, we now explain:

Fit indices were inspected to confirm the inclusion of additional parameters (i.e. in models 2 and 3) did not considerably reduce model fit. (page 8, lines 20-21)

9. Only in the end of the discussion it is clearly written that the AAS assessment and the accelerometer measure are not over the same period, but different periods rather some time apart! This could have been written more clearly in the methods section.

We agree this would be helpful to have in the methods section. On reflection, we also realise that we could provide more detailed information about this. We have therefore included the following in the ‘Participants and design’ section of the methods:

Assessment times varied; 31% completed the AAS (which assesses the previous 7 days) after day 7 of the accelerometry assessment (median 11 days, IQR 5–20) and 69% completed the AAS before day 7 of the accelerometry assessment (median 13 days, IQR 8–17). (page 6, lines 8-11)

Reviewer 2 Sakine Gocer Sahin

1. Please make clear why the authors used Spearman rho instead of Pearson correlation?

We now explain in the analysis section that physical activity data (MVPA from accelerometry and the AAS) was not normally distributed:

Spearman’s coefficient was used because physical activity data were not normally distributed, and a bivariate scatterplot indicated a monotonic but potentially non-linear association. Spearman’s rho has been used widely in validation studies, including previous studies of the AAS [6–11]. (page 7, line 25 – page 8, line 2)

2. What kind of correlation did you use to examine the relationship between sex and other variables? It should be point biserial correlation since sex is naturally dichotomous. However, I couldn’t find any information about this.
Thank you for highlighting this. We now examine the relationship between sex and other variables using the point biserial correlation. Please see Table 1 (page 10). This correction means that gender is now significantly associated with AAS-reported MVPA (see also page 9, line 1).

3. I have never heard/used ‘differential bias’ term. In psychometrics, there are two terms related to this point: Differential item function and bias. It sounds like the authors combined these both terms. Please, use bias instead of differential bias.

As suggested, we now use ‘bias’ instead of ‘differential bias’ throughout the manuscript.

4. Correlation and agreement are different terms. If you measure two different constructs and examine relationship between them, you should use correlation, if the constructs are the same then you should use agreement. If you are examining the validity of AAS, which is one of objectives of your paper, that means you are still not sure if it’s valid or not (measuring the same thing or not). So, if you aren’t sure they both measure same thing, how can you use agreement?

Thank you for this suggestion. Reviewer 1 prefers the term ‘association’ rather than ‘correlation’. In correcting our terminology, we have therefore changed ‘agreement’ to ‘association’ where relevant.

5. In page 4, line 5, you indicated that “a recent study found little evidence to support the validity of the online AAS 4 compared to accelerometry (MVPA rs = .23) among adults enrolled in a physical activity randomised controlled trial (RCT)”. This looks like the first of objective of your study. If the validity of AAS is already known, what makes your research novel? Please provide more information, what makes this research different than others? Why is this study novel?

We have now included:

This study therefore adds to the literature by examining the association between MVPA reported on the online AAS and recorded via accelerometer in a new sample. Additionally, it extends previous research by examining self-report bias (the difference between daily minutes of MVPA reported on the online AAS and daily minutes of MVPA derived from accelerometer). While previous research has examined the association between measurement methods in different subgroups [11], this is the first study to examine predictors of bias on the online AAS. The objectives of this study were to (1) evaluate the criterion validity of minutes of MVPA on the online AAS relative to accelerometry by examining both the association between methods and the difference between daily minutes of MVPA reported on the online AAS and daily minutes of MVPA derived from accelerometer, and (2) examine whether participant factors (activity level, age, sex, education, body mass index (BMI) and general health perception) were associated with bias. (page 5, lines 5-15)