Author's response to reviews

Title: Validity of the International Physical Activity Questionnaire and the Singapore Prospective Study Program physical activity questionnaire in a multiethnic urban Asian population

Authors:

Ei Ei Khaing Nang (ephneek@nus.edu.sg)
Susan Ayuko GITAU NGUNJIRI (ayukogita@yahoo.com)
Yi Wu (yi_wu@nuhs.edu.sg)
Agus SALIM (agus_salim@nuhs.edu.sg)
E Shyong TAI (eshyong@pacific.net.sg)
Jeannette LEE (jeannette_lee@nuhs.edu.sg)
Rob M. VAN DAM (ephrmvd@nus.edu.sg)

Version: 2 Date: 24 August 2011

Author's response to reviews: see over
Dear Editor,

Thank you for providing us with the opportunity to revise manuscript number 1845403602539899 – “Validity of the International Physical Activity Questionnaire and the Singapore Prospective Study Program physical activity questionnaire in a multiethnic urban Asian population”. We thank the reviewers for their constructive and useful comments and are encouraged that they feel that this manuscript is important. We have addressed these comments below on a point-by-point basis.

We look forward to your response.

Yours sincerely,

Rob M. van Dam, PhD
Response to Reviewer 1

This manuscript is methodological study on validity of IPAQ and another physical activity questionnaire (SP2PAQ) in a multiethnic urban Asian population. Since IPAQ was designed to be suitable in various cultures and countries, repeated validations are kind of warranted. The present paper is easy to read, but I have several concerns and remarks related to the design and analyses.

1). A major limitation in the present study is that only an accelerometer was used to validate the questionnaires.

We acknowledged that using only accelerometer to validate the questionnaires was a limitation of our study. However, the accelerometer can provide the frequency, duration and intensity of free living physical activity to obtain a good estimate of energy expenditure and has been recommended as an objective method of choice to use in validating questionnaires or studying patterns of physical activity [6, 51, 52]. (Discussion section, page 18, line 415-418).

2). A general question is that SP2PAQ is quite likely unfamiliar to most outside Singapore. Therefore, it would be important to have to questionnaire presented as an appendix and/or Internet material.

As recommended, we have now attached SP2PAQ as an internet material (please see additional file 2).

3). Another general finding from the results (Figures 1 and 2) is that there is a very strong linear relation between the size of measurement and difference between two methods. It is unfortunately impossible to see from the figure if this relation is only observed beyond a certain point, but it is nevertheless an intriguing finding. The strong linear relation suggest that at least after this magical point one of the methods method gives constant results while the other starts to over- or underestimate EE. The authors have already commented this, but I would like to have even more insight into this strange result.
We have now discussed this result in more detail: “When we compared questionnaire and accelerometer estimates of physical activity using Bland-Altman plots, greater differences between the two methods were observed with increasing means of measurements for both moderate and vigorous activity. This may be due to either the questionnaire over-estimating activity increasingly with increasing activity or the accelerometer under-estimating activity increasingly with increasing activity. Klesges et al reported that people tend to overestimate the duration of their physical activity, especially for aerobic activities [45]. In addition, the Actical accelerometer may have underestimated energy expenditure, especially at high levels of energy expenditure [46]. The accelerometer is known to substantially underestimate energy expenditure for specific activities [47]. For example, accelerometers have limitations in detecting activities where the body is mostly stationary such as when cycling or weight lifting [47]. Moreover, in our study, the accelerometer was taken off during water-based activities. This may have reduced the amount of activity detected by the accelerometer as compared with the questionnaire although only five participants reported swimming during the period in which they wore the accelerometer. The combination of over-estimation by the questionnaires and under-estimation by the accelerometer may have given rise to the observation that the difference between these methods was greater at higher levels of activity. (Discussion section, page16, line 353-368).

4).p. 6, para 2: Interval between the test-retest measurement was variable and for me (too?) long. If the reliability of the assessment/method per se is tested, all variations in activity should be negligible. Otherwise it is not clear, if the correlation is measuring variation in the measurement or in activity.

We realize that real within-person variation in activity that is not captured by the method of physical activity assessment may have affected the test-retest reliability. We have now included this as a potential limitation of our study as follows: “Finally, the interval between the test and retest measurements was rather long. We realize that as a result the reliability estimates are affected by both measurement error related to the assessment of short-term activity and real changes in
activity habits of participants over time. However, in epidemiological studies we are generally interested in habitual activity over years as this is most relevant for the development of chronic diseases. For this application, an inability of assessment methods to capture long-term physical activity is therefore a limitation and long-term reproducibility, part of which may be due to real changes in physical activity, is most relevant.” (Discussion section, page 18-19, line 419-426).

5). p. 8, last line: Are the CV’s (10.2 to 16.6%) good, bad or acceptable?
CV’s (10.2 to 16.6%) are acceptable as all were below 20%. We now mention this in the manuscript (Page 9, line 187).

6). p. 9, para 3: Now I am a little confused. The results are shown as kcal/week or did I not understand this correctly? What is shown as EE/5 days? And eventually, why not showing everything as daily EE?
We now show the results in Kcal/day and consistently used this unit throughout the manuscript.

7). p. 9, last para: Is anything known about the validity of accelerometer’s cut-off points for activity of different intensities? In general, one of the main problems with accelerometer data has been that the counts/min cut-off points are not consistently established and variation in these naturally affects the results.
The physical activity intensity prediction of the Actical accelerometers was validated with a room calorimeter. This showed that differences between the measurements of the Actical accelerometer and the calorimeter for the time spent in each moderate and vigorous intensity activity was < 2% (Rothney MP, Schaefer EV, Neumann MM, Choi L, Chen KY. Validity of physical activity intensity predictions by ActiGraph, Actical, and RT3 accelerometers. Obesity. [Comparative Study Research Support, N.I.H., Extramural Validation Studies]. 2008 Aug;16(8):1946-52.). We have now included this information in the manuscript (page 8, line 175-177).
8). p. 10, line 1: Is the minimum collection time (8 h) long enough? It is only 50% of hours awake. In many cases 10 or 12 h is used, I would personally find 12 h to be a compromise between loosing data during hours without data collection vs. having enough days with acceptable nr of hours.

We revised the analysis and have now excluded a total of 8 participants with <10 hrs wearing time/day for any day of the study (Statistical analysis section, page 10, line 224-225). Although some of the remaining participants have <12 h/day of wearing time, this was mostly only on Sunday. Hence, the average wearing time for 5 days was >12h/day for all participants currently included in the analysis.

9). p. 11, para 2: How representative are the subjects in relations to the entire population? Is the gender and age distribution similar through all three ethnic groups? Likewise, is the age and ethnic distribution similar in both genders? The two last questions are important in order to understand if sub-group analyses can be done without confounders biasing the results.

Our participants represent a wide range of ages and education levels and include large proportions of men and women and the three main ethnic groups in Singapore. However, as our participants were mainly from a university and hospital setting, the study population may not have been completely representative of the entire population. Although the distributions of age, gender and ethnicity were not exactly the same across the sub-groups, these differences in distribution were not statistically significant. We have now mentioned this more explicitly as limitation of our study in the Discussion section (page 18, line 406-411).

Below, we present the distribution of gender and age across the ethnic groups and also the distribution of age and ethnicity in men and women. Malays were older and higher percentage of female and fewer percentage of male than the other 2 ethnic groups. This might result in higher correlation for moderate activity and lower correlation for vigorous activity for Malay group for SP2PAQ. However, this finding was not consistent in IPAQ where the correlation in Malay was lower than that of Chinese. Similarly, although female has higher percentage of Malay and older
than male, the correlations for female were not consistently higher or lower than male in both questionnaires.

Distribution of age and gender by ethnicity.

<table>
<thead>
<tr>
<th></th>
<th>Chinese</th>
<th>Malay</th>
<th>Indian</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>30(45.45%)</td>
<td>11(32.35%)</td>
<td>23(44.23%)</td>
<td>0.4</td>
</tr>
<tr>
<td>Female</td>
<td>36(54.55%)</td>
<td>23(67.65)</td>
<td>29(55.77%)</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>37.62±13.94</td>
<td>41.01±11.90</td>
<td>37.39±11.99</td>
<td>0.4</td>
</tr>
</tbody>
</table>

Distribution of age and ethnicity by gender

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th>Female</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chinese</td>
<td>30(46.88%)</td>
<td>36(40.91%)</td>
<td>0.4</td>
</tr>
<tr>
<td>Malay</td>
<td>11(17.19%)</td>
<td>23(26.14%)</td>
<td></td>
</tr>
<tr>
<td>Indian</td>
<td>23(35.94%)</td>
<td>29(32.95%)</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>36.31±12.63</td>
<td>39.75±12.9</td>
<td>0.1</td>
</tr>
</tbody>
</table>

10). p. 12, para 3: Why not presenting results also for total physical activity EE? Figure 1 and 2: The scales are too large and it is not possible to really see the distribution of results close to the zero line. I would remove the outlier from both IPAQ analyses. Since a similar outlier was not seen with SP2PAQ, it is clear that the gross overestimation was only seen in IPAQ. After removing the outlier, the scales can be made more appropriate for the data.
Although the SP2PAQ recorded all light, moderate and vigorous intensity activities, the IPAQ did not recorded light activity. Hence, we could not present results for total physical activity EE. For Figure 1 and 2, we removed the outliers from the data, “In addition, 3 participants who reported the sum total of all walking, moderate and vigorous time more than 16 hours per day in IPAQ were treated as outliers and excluded from analysis according to IPAQ data processing rule [34]” (page 10, line 225-228) and revised the scales of the figures.
Response to Reviewer 2

I have with interest read your manuscript entitled "Validity of the International Physical Activity Questionnaire and the Singapore Prospective Study Program physical activity questionnaire in a multiethnic urban Asian population". Overall it is an interesting and well conducted study but I have a few concerns that I feel needs to be addressed.

Major compulsory revisions

1) It is unclear how the IPAQ data was treated, was the IPAQ scoring protocol followed during data cleaning (http://www.ipaq.ki.se/scoring.pdf)? When looking at the figures it looks as if there may be some outliers that potentially is a result of not following them.

We have revised the analysis, treating data according to IPAQ data processing rules as suggested. We removed the outliers as described in the revised manuscript: “In addition, 3 participants who reported the sum total of all walking, moderate and vigorous time more than 16 hours per day in IPAQ were treated as outliers and excluded from analysis according to IPAQ data processing rule [34]” (page 10, line 225-228) and also truncated the data: “The data was truncated at 21 hours/week for each walking activity, moderate activity and vigorous activity according to IPAQ data processing rules” (page 9, line 190-192).

2) The figures (1 & 2) are not very good, for example units on the axises are lacking. Please tidy them up.

We have added the units on both axis and revised the figures.

Minor essential revisions

1) Throughout the paper the authors need to make a better job in defining physical activity, physical inactivity and energy expenditure and to be consistent in how they use the words.
We revised the wording and consistently used “the energy expenditure from physical activity” wherever possible throughout the manuscript.

2) The accelerometer needs to be better described. It is very confusing to read row 5-8 on page 8. There are for example no anatomical plane called single plane and the word sensitive occurs five times in four lines.

We revised the paragraph and removed the term “single plane” (page 8, line 165-169).

3) On the bottom of the same page the abbreviation CV is introduced without having been explained.

We added the explanation “coefficient of variation” before the abbreviation “CV” (page 9, line 184).

Discretionary Revisions

1. The authors may be interested in the following article:


We read the article with great interest. The findings in that study was similar as in our study and we included the study as a reference in our article as follows: “Similar findings were reported for a nationally representative sample of the Swedish population, where the difference between the IPAQ and accelerometer measurements of time spent on physical activity was larger at higher activity levels reported by the IPAQ [48].” (Discussion section, page 16, line 368-370).