Reviewer’s report

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

Version: 1 Date: 3 May 2011

Reviewer: Mikael Fogelholm

Reviewer’s report:

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.

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

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.

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.

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.

p. 8, last line: Are the CV’s (10.2 to 16.6%) good, bad or acceptable?

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?

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.
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.

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.

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.

Level of interest: An article whose findings are important to those with closely
related research interests

Quality of written English: Acceptable

Statistical review: No, the manuscript does not need to be seen by a
statistician.

Declaration of competing interests:

No COI