Comparison of a Subjective and an Objective Measure of Physical Activity in a Population Sample

Maria Hagstromer, Barbara E. Ainsworth, Pekka Oja, and Michael Sjostrom

Background: The aim of this study was to compare physical activity components in the long, self-administrated version of IPAQ with an accelerometer in a population sample. Methods: In total 980 subjects (18-65 years) wore an accelerometer (Actigraph) for 7 consecutive days and thereafter filled in the IPAQ. Measures of total physical activity, time spent in moderate and in vigorous activity as well as time spent sitting as assessed by the IPAQ and the Actigraph were compared. Results: The results showed significant low to moderate correlations (Rs = 0.07–0.36) between the 2 instruments and significantly (P < .001) higher values for sitting and vigorous intensity physical activity from the IPAQ compared with the Actigraph. The higher the values reported by the IPAQ the bigger differences were seen between the instruments. Comparison between the tertiles of total physical activity by the 2 instruments showed significant overall association with consistent agreement in the low and the high tertiles. Conclusion: The long form of IPAQ is a valid measure of physical activity in population research. However, the IPAQ likely overestimates actual physical activity as shown by its limited ability to classify adults into low and high categories of physical activity based on accelerometer data.

Keywords: accelerometry, MET-minute, monitoring, physical activity, self-report

Researchers interested in describing and evaluating the relationship between physical activity and health require accurate measures of physical activity with a high degree of standardization in physical activity definition and assessment methods. To collect internationally comparable data,1,2 a standardized questionnaire, the International Physical Activity Questionnaire (IPAQ), was developed for use across different cultural milieus and geographic locations. Two versions of the IPAQ exist: a long quantitative history for use in research studies, and a short recall form for use in epidemiological surveillance settings. For ease of administration 2 formats (self-administered and interviewer administered) of the long and the short IPAQ are available with both using a 7-day recall period. During the development process the questionnaires was evaluated for their validity and reliability in 14 locations in 10 countries.3 The results show the IPAQ has similar measurement properties for validity and reliability as other questionnaires used in epidemiological research, and that the questionnaires are acceptable for use in diverse cultural settings and in different languages.4 The 7-day recall versions of the short and long IPAQ in various languages have been published on a webpage for convenient access (www.ipaq.ki.se).

The short version has been further tested for validity in several countries and settings.5-10 To date there has been limited evaluation of the long version of the IPAQ11-13 and none using nationally representative population samples. For example, the subjects used in the international validation study were mainly convenience samples of men and women, aged 18 to 65 with most samples being predominantly middle aged. As IPAQ is developed for use on population level it is of high importance that its validity is tested on that population. Therefore, the aim of this study was to compare physical activity scores arising from the long self-administrated version of IPAQ with an objective measure, an accelerometer, of physical activity in a nationally representative sample.

Material and Methods

Subjects

Subjects were recruited from 3300 men and women, age 18 to 70 years, identified at simple random sampling from the Swedish population register and contacted by a telemarketing company. The sampling scheme required contacting approximately 40 subjects by telephone each week from throughout a calendar year. Once contacted, subjects were informed of the study and asked for their willingness to participate in the study. Of the 3300 men and women, 2262 (69%) had a telephone number registered and could be contacted. Among those contacted, 1556 (69%) agreed to participate in the study. Subjects with incomplete physical activity data or technical errors in the instrument (n = 442), or subjects older than 65 years or retired (n = 132) were excluded. Complete data
were obtained from 980 subjects, age 18 to 65 years, reflecting a completion rate of 62% of subjects who agreed to participate in the study. The representativeness of the sample is described elsewhere. The study was approved by the Huddinge University Hospital ethical board (# 378/02) and all subjects gave verbal consent to participate in the study.

Procedure
Following agreement to participate in the study, the self-administered long-IPAQ, a demographic questionnaire, and an Actigraph accelerometer were mailed to each subject through the Swedish postal system. Before mailing the materials, the Actigraph was initialized as described by the manufacturer using a 60-second epoch for recording movement. Subjects were instructed to wear the accelerometer from when they woke up until going to bed during 7 consecutive days, except when in water, starting on the Monday morning following the receipt of the study materials. Subjects were given instructions to wear the accelerometer attached to a belt and secured directly next to their skin at their lower back, close to the center of gravity. On the 8th day, subjects were directly next to their skin at their lower back, close to wear the accelerometer attached to a belt and secured directly next to their skin at their lower back, close to the center of gravity. On the 8th day, subjects were instructed to complete the demographic questionnaire and the long-IPAQ self-administered questionnaire and return the questionnaires and the accelerometer to the study center in a prepaid mailing envelope. Subjects failing to return the materials within 1 week of completing the data collection were called to return the materials. Upon receipt of the materials at the study center, the accelerometer and the questionnaire data were downloaded into a Microsoft Access database for later scoring and data analysis.

Questionnaires
Demographic data included questions about age in years, sex, height in cm, weight in kg, and educational attainment. Body Mass Index (BMI) was computed as weight in kilograms divided by height in meters squared. Age and BMI were categorized into 3 levels (age: 18–34, 35–50, 51–65 years; BMI: <25, 25–30, >30). Educational attainments were categorized into 3 levels (compulsory school, at least 2 years of high school, and college/university).

The self-administered format for the long-IPAQ was translated into Swedish in accordance with the IPAQ translation guidelines published on the IPAQ website (www.ipaq.ki.se). The long-IPAQ consists of 27 items that identify the frequency (times per week) and duration (minutes or hours per day) of physical activity performed in the activity domains of occupation (7 items), transportation (7 items), housework, house maintenance, and family care (6 items), recreation, sport and leisure (6 items), and time spent sitting (minutes or hours per day) in a weekday and in a weekend day (2 items). For all physical activity domains, except sitting, participation in vigorous and moderate intensity physical activity was obtained. Examples of culturally relevant moderate and vigorous intensity activities for Swedish adults were listed with instructions for each IPAQ physical activity domain. According to the IPAQ guidelines moderate intensity was assigned 4 METs (Metabolic Equivalent Task) vigorous intensity as 8 METs, and walking 3.3 METs. One MET is equal to the energy expenditure during rest and is approximately equal to 3.5 ml O₂·kg⁻¹·hr⁻¹ as the oxygen cost of activity and 1 kcal·kg⁻¹·hr⁻¹ as the caloric equivalent for adults.

Outcome measures from the IPAQ were (1) total physical activity expressed as MET-minutes per day and minutes reported in (2) vigorous, (3) moderate intensity activity, and (4) in sitting per day. MET-min was computed by multiplying METs by minutes of participation in vigorous and moderate intensity physical activities and walking. Sitting was expressed as min/day. The methods used to score the long-IPAQ are presented in detail on the IPAQ website as well (www.ipaq.ki.se). The scoring protocol was followed for cleaning and truncation.

Actigraph Accelerometer
The Actigraph MTI model 7164 (Manufacturing Technology Inc, Fort Walton Beach, FL, USA) was used to provide objective measures of physical activity. The technical specifications and performance characteristics of the Actigraph have been described elsewhere. Seven days of consecutive data were collected from study subjects. Data were deemed complete if subjects had accelerometer counts for at least 10 hours per day, for at least 4 days, including at least 1 weekend day. A minimum of 4 recording days is recommended by Trost et al to reflect 1-week of physical activity in children and adults, respectively. The threshold criteria used in this study for the time required for wearing the Actigraph is also consistent with the procedures used in the international IPAQ validity study and in other studies using. The Actigraph data were uploaded onto a computer and analyzed by software based on Microsoft Access. Data with periods of 0 values for more than 20 minutes were excluded from the analysis. Accelerometer malfunction was identified as having counts per minute greater than 20,000. All spurious data were set to missing values and excluded from the analyses. The data were then analyzed for minutes per day spent sedentary as well as in physical activity of low, lifestyle and moderate or higher intensity is based on application of count thresholds corresponding to these activities. Time spent on vigorous intensity ambulatory activities was estimated as the amount of time accumulated above 5724 counts/min. Minutes per day spent on lifestyle moderate intensity activities was estimated as the amount of time accumulated between 760 and 5724 counts/minute. Minutes spent sitting were estimated as the amount of time accumulated below 100 counts/minute during periods when the accelerometer have been worn. Time spent on low intensity physical activity was consequently defined as minutes accumulated between 100 and 760 counts/minute. The time spent on each activity level is presented...
as every minute that meets the specific criteria. Average intensity was calculated taking the total counts divided by the recorded time (counts·min⁻¹) and this was considered as a weighted measure of total physical activity.

**Data Analyses**

All statistical analyses were performed using SPSS (Statistical Package for the Social Science for Windows, 14.0, 2004, SPSS Inc., Chicago, IL). The characteristics of subjects were described with frequency distribution and outcomes from IPAQ and activity monitor were described by mean, standard deviation (SD) and interquartile ranges in total and by sex, age, BMI and education.

Comparison between the long-IPAQ and the Actigraph scores was done using Spearman’s rank correlation coefficients (Rs). The Bland and Altman method²⁴,²⁵ was used to provide an indication of the systematic and random error and the heteroscedasticity of the IPAQ measures of physical activity and 95% limits of agreement were used for describing the total error between the 2 methods. Variables used for the Bland & Altman analysis were daily time spent in moderate and in vigorous intensity activity and in sitting according to IPAQ and the Actigraph. The differences between the instruments were tested with a paired nonparametric Wilcoxon test.

To show concordance between classification of the IPAQ and Actigraph data, tertiles based on the distribution of the data for total physical activity from the IPAQ (Q1 = 0–269, Q2 = 270–713, Q3 = 714–5086 MET-min·d⁻¹) and average intensity as measured by the Actigraph (Q1 = 105–308, Q2 = 309–416, Q3 = 417–1010 counts·min⁻¹) were compared using Kendall’s tau-b.

For all analyses, the level of significance was set to $P < .05$.

**Results**

Table 1 shows the subject characteristics for the 537 women and 443 men included in this study. There was a normal distribution of subjects by age, and most subjects were in the BMI categories of < 30. About one third of the subjects had a University education.

Table 2 shows descriptive data for the IPAQ and the Actigraph instruments stratified by sex, age, BMI and education. On average the accelerometers were worn for 6.6 (1) days, and on those days they wore it for 830 (84) minutes per day, equaling approximately 14 hours. Across all stratified variables, the IPAQ showed significantly higher minutes per day in vigorous intensity activities and in sitting compared with the Actigraph ($P < .001$). Men, adults age 18 to 34 and those with the highest education spent most time in vigorous physical activity. This was consistent for the IPAQ and the Actigraph. Moderate intensity activity from the Actigraph showed significantly higher minutes per day compared with the sum of walking and moderate intensity activity from IPAQ ($P < .001$).

Table 3 shows the Spearman rank order correlations for the time spent in physical activity from the long-IPAQ with time spent in similar activities from the Actigraph by sex. Modest, but significant correlations were observed for vigorous, moderate and walking minutes per day between the IPAQ and the Actigraph and correlations were also statistically significant between the IPAQ MET-minutes per day and the Actigraph total counts per minute. Similar, low to modest correlations were found when using the raw not truncated IPAQ data, results not shown.

Table 4 shows the median (interquartile range) of IPAQ per domain was 0 (0; 1590), 396 (0; 1125) 375 (0; 1320) 480 (0; 1386) for work, transportation, household chores and leisure time, respectively. Table 4 shows the
### Table 2  Descriptive PA Data From IPAQ and Actigraph, by Sex, Age, BMI and Education; Mean (SD) and Interquartile Ranges

<table>
<thead>
<tr>
<th></th>
<th>Total PA</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IPAQ&lt;sup&gt;a&lt;/sup&gt; MET min·d&lt;sup&gt;−1&lt;/sup&gt;</td>
<td>Actigraph&lt;sup&gt;b&lt;/sup&gt; counts.min&lt;sup&gt;−1&lt;/sup&gt;</td>
<td>IPAQ&lt;sup&gt;a&lt;/sup&gt; MET min·d&lt;sup&gt;−1&lt;/sup&gt;</td>
<td>Actigraph&lt;sup&gt;b&lt;/sup&gt; min·d&lt;sup&gt;−1&lt;/sup&gt;</td>
<td>IPAQ&lt;sup&gt;a&lt;/sup&gt; MET min·d&lt;sup&gt;−1&lt;/sup&gt;</td>
<td>IPAQ&lt;sup&gt;a&lt;/sup&gt; MET min·d&lt;sup&gt;−1&lt;/sup&gt;</td>
<td>IPAQ&lt;sup&gt;a&lt;/sup&gt; MET min·d&lt;sup&gt;−1&lt;/sup&gt;</td>
<td>Actigraph&lt;sup&gt;b&lt;/sup&gt; min·d&lt;sup&gt;−1&lt;/sup&gt;</td>
<td>IPAQ&lt;sup&gt;a&lt;/sup&gt; MET min·d&lt;sup&gt;−1&lt;/sup&gt;</td>
<td>Actigraph&lt;sup&gt;b&lt;/sup&gt; min·d&lt;sup&gt;−1&lt;/sup&gt;</td>
<td>IPAQ&lt;sup&gt;a&lt;/sup&gt; MET min·d&lt;sup&gt;−1&lt;/sup&gt;</td>
<td>Actigraph&lt;sup&gt;b&lt;/sup&gt; min·d&lt;sup&gt;−1&lt;/sup&gt;</td>
<td>Sitting</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All</td>
<td>648 (629)</td>
<td>382 (141)</td>
<td>19 (33)</td>
<td>5 (15)</td>
<td>56 (45)</td>
<td>50 (44)</td>
<td>106 (71)</td>
<td>133 (58)</td>
<td>416;188</td>
<td>133;162</td>
<td>432;626</td>
<td>403;527</td>
<td>590 (208)</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>614 (613)</td>
<td>375 (129)</td>
<td>14 (29)</td>
<td>4 (18)</td>
<td>56 (44)</td>
<td>52 (44)</td>
<td>108 (72)</td>
<td>132 (52)</td>
<td>413;545</td>
<td>413;545</td>
<td>457 (80)</td>
<td>398;508</td>
<td>557 (211)</td>
</tr>
<tr>
<td>Men</td>
<td>701 (675)</td>
<td>390 (154)</td>
<td>23 (36)</td>
<td>5 (9)</td>
<td>56 (46)</td>
<td>47 (44)</td>
<td>103 (71)</td>
<td>135 (64)</td>
<td>413;545</td>
<td>413;545</td>
<td>632 (195)</td>
<td>477 (94)</td>
<td>632 (195)</td>
</tr>
<tr>
<td>Age (years)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18–34</td>
<td>705 (698)</td>
<td>411 (147)</td>
<td>27 (38)</td>
<td>7 (12)</td>
<td>51 (45)</td>
<td>48 (44)</td>
<td>99 (70)</td>
<td>133 (58)</td>
<td>141;74</td>
<td>141;74</td>
<td>605 (207)</td>
<td>472 (92)</td>
<td>605 (207)</td>
</tr>
<tr>
<td>35–50</td>
<td>682 (654)</td>
<td>391 (143)</td>
<td>18 (33)</td>
<td>4 (6)</td>
<td>62 (44)</td>
<td>52 (45)</td>
<td>111 (73)</td>
<td>142 (61)</td>
<td>485 (91)</td>
<td>485 (91)</td>
<td>606 (202)</td>
<td>458 (91)</td>
<td>606 (202)</td>
</tr>
<tr>
<td>51–65</td>
<td>581 (573)</td>
<td>347 (125)</td>
<td>10 (24)</td>
<td>4 (23)</td>
<td>54 (45)</td>
<td>47 (42)</td>
<td>106 (70)</td>
<td>124 (51)</td>
<td>393;527</td>
<td>393;527</td>
<td>563 (211)</td>
<td>470 (77)</td>
<td>563 (211)</td>
</tr>
<tr>
<td>BMI (m·kg&lt;sup&gt;−1&lt;/sup&gt;)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;25</td>
<td>622 (619)</td>
<td>406 (146)</td>
<td>17 (30)</td>
<td>5 (8)</td>
<td>54 (44)</td>
<td>48 (43)</td>
<td>102 (70)</td>
<td>138 (59)</td>
<td>405;134</td>
<td>405;134</td>
<td>587 (211)</td>
<td>465 (92)</td>
<td>587 (211)</td>
</tr>
<tr>
<td>25–29.9</td>
<td>683 (658)</td>
<td>360 (134)</td>
<td>18 (34)</td>
<td>4 (22)</td>
<td>57 (45)</td>
<td>53 (45)</td>
<td>110 (72)</td>
<td>129 (58)</td>
<td>439;528</td>
<td>439;528</td>
<td>591 (204)</td>
<td>467 (79)</td>
<td>591 (204)</td>
</tr>
<tr>
<td>&gt;30</td>
<td>653 (592)</td>
<td>320 (106)</td>
<td>18 (35)</td>
<td>3 (15)</td>
<td>60 (47)</td>
<td>45 (41)</td>
<td>105 (69)</td>
<td>116 (50)</td>
<td>414;527</td>
<td>414;527</td>
<td>610 (195)</td>
<td>472 (77)</td>
<td>610 (195)</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Compulsory (9 yrs)</td>
<td>671 (627)</td>
<td>349 (134)</td>
<td>16 (35)</td>
<td>2 (5)</td>
<td>62 (46)</td>
<td>51 (44)</td>
<td>112 (73)</td>
<td>124 (61)</td>
<td>427;540</td>
<td>427;540</td>
<td>564 (218)</td>
<td>446 (77)</td>
<td>564 (218)</td>
</tr>
<tr>
<td>High school</td>
<td>740 (741)</td>
<td>389 (142)</td>
<td>19 (36)</td>
<td>3 (6)</td>
<td>63 (46)</td>
<td>54 (47)</td>
<td>109 (72)</td>
<td>144 (67)</td>
<td>395;492</td>
<td>395;492</td>
<td>541 (222)</td>
<td>446 (87)</td>
<td>541 (222)</td>
</tr>
<tr>
<td>University</td>
<td>541 (529)</td>
<td>387 (131)</td>
<td>13 (24)</td>
<td>7 (24)</td>
<td>48 (43)</td>
<td>47 (41)</td>
<td>94 (66)</td>
<td>130 (50)</td>
<td>427;540</td>
<td>427;540</td>
<td>620 (195)</td>
<td>482 (82)</td>
<td>620 (195)</td>
</tr>
</tbody>
</table>

<sup>a</sup> Difference IPAQ - Actigraph tested for total PA, vigorous, moderate, and sitting, respectively, using paired nonparametric Wilcoxon test, \( P < .001 \).

<sup>b</sup> Cut-off values for sitting, moderate and vigorous were <100, 761–5724, and >5724 counts, respectively.
Concurrent Validity of Long IPAQ

Spearman rank order correlations for each domain from the long-IPAQ and time spent in at least moderate intensity activity and Average intensity from the Actigraph, by sex. The highest correlations were found for leisure time physical activity.

Figures 1 to 3 shows the IPAQ and the Actigraph minutes per day Bland & Altman analyses for sitting, moderate and vigorous intensity physical activity, respectively. For sitting the mean (SD) difference was 130 (207) min·d⁻¹, \( P < .001 \). Differences (ie, error) between the IPAQ and Actigraph scores increased as the minutes per day in sitting reported in the IPAQ increased (\( R^2 = .50 \)).

The Bland-Altman plot for the moderate intensity physical activity minutes per day from the IPAQ and the Actigraph showed that a mean difference of 2 (32) min·d⁻¹, \( P < .001 \). Differences (ie, error) between the IPAQ and Actigraph scores increased as the minutes per day in moderate intensity physical activity reported in the IPAQ increased (\( R^2 = .42 \)).

The Bland-Altman plot for vigorous intensity activities from the IPAQ and the Actigraph showed that a mean difference of 15 (32) min·d⁻¹, \( P < .001 \). Differences (ie, error) between the IPAQ and Actigraph scores increased as the minutes per day in vigorous intensity physical activity reported in the IPAQ increased (\( R^2 = .90 \)).

Figure 4 shows the concordance in tertiles of total physical activity scores between the IPAQ MET min/day and the Actigraph counts per minute. A significant, but weak, overall association (Kendall’s tau-b, 0.210, \( P < .001 \)) was found between the instruments. The figure shows consistent associations between the 2 methods at low and high tertiles of physical activity. However, only half of the Actigraph counts were misclassified within each IPAQ tertile.

<table>
<thead>
<tr>
<th>IPAQ (min·d⁻¹)</th>
<th>Actigraph* (min·d⁻¹)</th>
<th>All Rs</th>
<th>Men Rs</th>
<th>Women Rs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vigorous</td>
<td>Vigorous</td>
<td>0.31**</td>
<td>0.31**</td>
<td>0.29**</td>
</tr>
<tr>
<td>Moderate</td>
<td>Moderate</td>
<td>0.27**</td>
<td>0.32*</td>
<td>0.22**</td>
</tr>
<tr>
<td>Walking</td>
<td>Moderate</td>
<td>0.25**</td>
<td>0.17**</td>
<td>0.15**</td>
</tr>
<tr>
<td>Moderate and Walking</td>
<td>Moderate</td>
<td>0.29**</td>
<td>0.35**</td>
<td>0.23**</td>
</tr>
<tr>
<td>Moderate</td>
<td>Low</td>
<td>0.27**</td>
<td>0.29**</td>
<td>0.25**</td>
</tr>
<tr>
<td>Walking</td>
<td>Low</td>
<td>0.08*</td>
<td>0.07*</td>
<td>0.07*</td>
</tr>
<tr>
<td>Total time IPAQ</td>
<td>Total time Actigraph</td>
<td>0.28**</td>
<td>0.30**</td>
<td>0.27**</td>
</tr>
<tr>
<td>Vigorous</td>
<td>Average intensity</td>
<td>0.28**</td>
<td>0.31**</td>
<td>0.26**</td>
</tr>
<tr>
<td>Moderate</td>
<td>Average intensity</td>
<td>0.16**</td>
<td>0.21**</td>
<td>0.11*</td>
</tr>
<tr>
<td>Walking</td>
<td>Average intensity</td>
<td>0.22**</td>
<td>0.22**</td>
<td>0.23**</td>
</tr>
<tr>
<td>Total PAb</td>
<td>Average intensityc</td>
<td>0.30**</td>
<td>0.36**</td>
<td>0.30**</td>
</tr>
<tr>
<td>Time spent sitting</td>
<td>Sitting</td>
<td>0.23**</td>
<td>0.13*</td>
<td>0.26**</td>
</tr>
</tbody>
</table>

* \( P < .05 \); ** \( P < .01 \).

Table 3 Spearman’s Rank Correlations Coefficient (Rs) for Total Physical Activity and Time Spent in Physical Activity From IPAQ and Actigraph (n = 980)

<table>
<thead>
<tr>
<th>IPAQ</th>
<th>Actigraph</th>
<th>All Rs</th>
<th>Men Rs</th>
<th>Women Rs</th>
</tr>
</thead>
<tbody>
<tr>
<td>PA at work</td>
<td>MVPAa</td>
<td>0.15**</td>
<td>0.17**</td>
<td>0.11*</td>
</tr>
<tr>
<td></td>
<td>Average intensity</td>
<td>0.21**</td>
<td>0.25**</td>
<td>0.18**</td>
</tr>
<tr>
<td>PA during transportation</td>
<td>MVPA</td>
<td>0.08*</td>
<td>0.04</td>
<td>0.13**</td>
</tr>
<tr>
<td></td>
<td>Average intensity</td>
<td>0.08*</td>
<td>0.06</td>
<td>0.11*</td>
</tr>
<tr>
<td>PA at home or in garden</td>
<td>MVPA</td>
<td>–0.02</td>
<td>0.01</td>
<td>–0.06</td>
</tr>
<tr>
<td></td>
<td>Average intensity</td>
<td>0.05</td>
<td>0.09</td>
<td>0.02</td>
</tr>
<tr>
<td>Leisure time PA</td>
<td>MVPA</td>
<td>0.26**</td>
<td>0.29**</td>
<td>0.23**</td>
</tr>
<tr>
<td></td>
<td>Average intensity</td>
<td>0.24**</td>
<td>0.25**</td>
<td>0.22**</td>
</tr>
</tbody>
</table>

* \( P < .05 \); ** \( P < .01 \).

Table 4 The Correlation Between the IPAQ Domains (MET min·d⁻¹) and the Actigraph

* Spearman rank order correlations for each domain from the long-IPAQ and time spent in at least moderate intensity activity and Average intensity from the Actigraph, by sex. The highest correlations were found for leisure time physical activity.

Figures 1 to 3 shows the IPAQ and the Actigraph minutes per day Bland & Altman analyses for sitting, moderate and vigorous intensity physical activity, respectively. For sitting the mean (SD) difference was 130 (207) min·d⁻¹, \( P < .001 \). Differences (ie, error) between the IPAQ and Actigraph scores increased as the minutes per day in sitting reported in the IPAQ increased (\( R^2 = .50 \)).

The Bland-Altman plot for the moderate intensity physical activity minutes per day from the IPAQ and the Actigraph showed that a mean difference of 2 (32) min·d⁻¹, \( P < .001 \). Differences (ie, error) between the IPAQ and Actigraph scores increased as the minutes per day in moderate intensity physical activity reported in the IPAQ increased (\( R^2 = .42 \)).

The Bland-Altman plot for vigorous intensity activities from the IPAQ and the Actigraph showed that a mean difference of 15 (32) min·d⁻¹, \( P < .001 \). Differences (ie, error) between the IPAQ and Actigraph scores increased as the minutes per day in vigorous intensity physical activity reported in the IPAQ increased (\( R^2 = .90 \)).

Figure 4 shows the concordance in tertiles of total physical activity scores between the IPAQ MET min/day and the Actigraph counts per minute. A significant, but weak, overall association (Kendall’s tau-b, 0.210, \( P < .001 \)) was found between the instruments. The figure shows consistent associations between the 2 methods at low and high tertiles of physical activity. However, only half of the Actigraph counts were misclassified within each IPAQ tertile.

* Spearman rank order correlations for each domain from the long-IPAQ and time spent in at least moderate intensity activity and Average intensity from the Actigraph, by sex. The highest correlations were found for leisure time physical activity.

Figures 1 to 3 shows the IPAQ and the Actigraph minutes per day Bland & Altman analyses for sitting, moderate and vigorous intensity physical activity, respectively. For sitting the mean (SD) difference was 130 (207) min·d⁻¹, \( P < .001 \). Differences (ie, error) between the IPAQ and Actigraph scores increased as the minutes per day in sitting reported in the IPAQ increased (\( R^2 = .50 \)).

The Bland-Altman plot for the moderate intensity physical activity minutes per day from the IPAQ and the Actigraph showed that a mean difference of 2 (32) min·d⁻¹, \( P < .001 \). Differences (ie, error) between the IPAQ and Actigraph scores increased as the minutes per day in moderate intensity physical activity reported in the IPAQ increased (\( R^2 = .42 \)).

The Bland-Altman plot for vigorous intensity activities from the IPAQ and the Actigraph showed that a mean difference of 15 (32) min·d⁻¹, \( P < .001 \). Differences (ie, error) between the IPAQ and Actigraph scores increased as the minutes per day in vigorous intensity physical activity reported in the IPAQ increased (\( R^2 = .90 \)).

Figure 4 shows the concordance in tertiles of total physical activity scores between the IPAQ MET min/day and the Actigraph counts per minute. A significant, but weak, overall association (Kendall’s tau-b, 0.210, \( P < .001 \)) was found between the instruments. The figure shows consistent associations between the 2 methods at low and high tertiles of physical activity. However, only half of the Actigraph counts were misclassified within each IPAQ tertile.
Figure 1 — Bland-Altman plot for min·d⁻¹ sitting from the IPAQ and Actigraph, defined as <100 counts·min⁻¹. Mean difference and (SD) for the total sample, 130 (207) min·d⁻¹.

Figure 2 — Bland-Altman plot for min·d⁻¹ reported in walking + moderate intensity from the IPAQ and moderate intensity minutes per day Actigraph, defined as 761 to 5724 counts·min⁻¹. Mean difference and (SD) for the total sample 2 (121) min·d⁻¹.
Figure 3 — Bland-Altman plot for min·d⁻¹ reported in vigorous intensity from the IPAQ and vigorous intensity from the Actigraph, defined as > 5724 counts·min⁻¹. Mean difference and (SD) for the total sample 15 (32) min·d⁻¹.

Figure 4 — Tertiles of total PA from IPAQ (MET-min/day) and average intensity from Actigraph, defined as total counts divided by recorded time (counts/min), by percent of Actigraph counts/min classified as low, moderate, and high intensity. Kendal Tau-b = 0.21 (P < .01).
Discussion

In this cross-sectional study of a simple random nationally sample of Swedish adults, the long form of IPAQ was compared with an objective measure of physical activity to determine the validity of the IPAQ using accelerometry. The results showed significant but low to moderate correlations between the IPAQ and the Actigraph accelerometer. Significantly more time ($P < .001$) was reported in vigorous intensity activity and sitting reported on the IPAQ as compared with the Actigraph.

A strength of the study is the large population sample size and the response rate. The sample is considered to represent the Swedish population with regard to sex, age and where in Sweden they live according to a previously published study. The age range was 18 to 65, which is the same as for whom the IPAQ instrument is designed. Information was sent by mail, which is commonly used in population based studies. Therefore we believe that our results are valid. Another strength of this study is the use of an objective measure to evaluate the long-IPAQ. The subjects were asked to wear the accelerometer for 1 consecutive week and thereafter answer the questionnaire, including IPAQ.

The Spearman rank correlation coefficients ($Rs$) between the IPAQ and Actigraph scores ranged from 0.07 to 0.36, indicating a low to moderate correlation between the 2 instruments. These correlations for total activity are similar to those obtained in the 12-country reliability and validity study reported by Craig et al. Women had lower correlations for all comparisons than men except for sitting and walking, which were higher. In a validation study of the short form of IPAQ in Sweden, no differences in correlations between men and women were found. No other IPAQ validation studies are reported that stratify the results by sex. However, validation studies on other physical activity questionnaires have found higher correlations for men than for women.

Accelerometers have been suggested as one of the best criterion measures for validation of self-report instruments of physical activity. However, this practice has been criticized by others that accelerometers and self-reports measure different things. It has been recommended that a range of physical activity measures, such as both objective and subjective measures, are needed to most accurately quantify the relationship between physical activity and health. The Actigraph measures body movement while questionnaires often ask for respondents to rate activities related to effort. Further, activities reported in IPAQ such as heavy manual work, household scores, bicycling and weight lifting are not possible to capture with the accelerometer as the Actigraph only captures locomotor activities. To enhance the assessment of various physical activities with the Actigraph, we chose cut-points recommended by Matthews et al. that are derived from both ambulatory and nonambulatory activities. Comparisons of accelerometer based physical activity estimates using different cut-off levels have shown that they produce varying results with none of the cut-off points being ideal. Thus, the observed agreement between IPAQ (self-reported) and the accelerometer (objectively measured) needs to be interpreted, bearing this in mind.

The IPAQ and the Actigraph yielded similar patterns of activity by sex, age and BMI. For example, on the IPAQ men reported more physical activity than women, which was confirmed by the Actigraph. Further, with increasing age less vigorous physical activity was reported on the IPAQ, which was also confirmed by the Actigraph. Thus, the use of IPAQ to study patterns of activity on group level is possible. However, for education level, the pattern of activity was different with more activity reported in the IPAQ for those with only compulsory school compared with those with a university degree. The opposite was found for the Actigraph. It is possible that low educated people engaged in manual work activities and the types of activities performed during work might not be recorded well using the Actigraph. There might also be a differential bias between educational groups in how they answered the IPAQ.

The Bland-Altman plots showed that IPAQ gave significantly higher estimates of sitting and vigorous physical activity than the Actigraph. For moderate intensity physical activity a slightly but significant higher minutes per day was found on the Actigraph than on the IPAQ. The difference was larger with higher activity levels reported on the IPAQ. In another validation study of the long IPAQ in a Swedish sample, no significant differences were observed between the IPAQ and the accelerometer scores. That study was done on a smaller group of volunteers who answered the questionnaire at the clinic as opposed to the current population based sample who filled in the IPAQ at home without assistance. Given the large differences between the absolute minutes per day derived from the accelerometer and from the IPAQ in the current study, it is likely that by including more items in the long form of the IPAQ, subjects’ over-report of the time and frequency of activities is not reflected in actual movement data as compared with the short IPAQ. This is coupled with the observation that the Actigraph does not measure all moderate and vigorous activities as reported in other studies.

The differences in daily physical activity minutes reported on the IPAQ and presented on the Actigraph increased as with higher IPAQ scores. This may reflect over-reporting, failure to recall time well, or rounding up of time for the IPAQ data. In questionnaires, when a person reports different tasks such as cleaning, she/he might report 1 hour, but probably not the whole hour was moderate intensity cleaning. The same is true for structured training such as aerobic, gym classes in which warm up, stretching is included. Therefore, in this case the IPAQ would yield higher values compared with the accelerometer. Further, IPAQ asks for time and frequency spent in moderate and vigorous intensity physical activities, which is a subjective rating of intensity. This, perception of intensity may be higher for a person with a higher body weight and/or low aerobic capacity.
Alternatively, the Actigraph may misclassify physical activity performed by not measuring nonambulatory activity. However, it is unlikely this would account for the large differences noted.

When creating tertiles of the total reported physical activity from IPAQ and comparing those with tertiles from the Actigraph data, a significant, but weak, association between the instruments was found. However, only half of subjects Actigraph scores were correctly classified within each tertile of the IPAQ total minutes. The agreement is better in the low and high end of activity, which shows that the IPAQ instrument can to a good extent detect the low active people (ie, those needed to be reached in interventions).

The long version of IPAQ seems to yield considerably higher estimates of physical activity compared with the short IPAQ. This is likely to be due to the fact that the more domains or possibilities to report the more activity are reported. There are studies that have found that IPAQ over-reports physical activity, but they have compared the results with another self-report measure, which is also connected to recall bias. IPAQ has also been compared with health related fitness tests and the results showed that the unfit subjects were more likely to over-report in the IPAQ. However, this does not necessarily reflect the current physical activity level.

The Actigraph measure number of physical activity minutes per day was significantly lower with increasing age and BMI, but this was not seen in the IPAQ self-reported data. This can be a differential bias, but it is also likely that as the IPAQ classifies activities by effort it takes more effort to do a given ambulatory activity by an obese compared with a normal weight person. In a validation study comparing aerobic capacity with the IPAQ it was found that subjects with low aerobic capacity were more likely to over-report their physical activities.

In conclusion, our results show low to moderate correlation between the long, last 7-days form of the IPAQ and the Actigraph among adults. The results indicate that the long form of IPAQ is a valid measure of physical activity in population research. However, IPAQ likely overestimates actual physical activity as shown by its limited ability to classify adults into low and high categories of physical activity based on accelerometer data.

Acknowledgments

This study has been carried out by a grant from the Stockholm County Council and arises also from the project ALPHA, which has received funding from the European Union, in the framework of the Public Health Programme.

References

2. WHA. World Health Assembly 57.17. Global Strategy on diet, physical activity and health; 2004 May.