# The International Physical Activity Questionnaire (IPAQ): a study of concurrent and construct validity

Maria Hagströmer[1,2,*], Pekka Oja[1] and Michael Sjöström[1]

[1]Unit for Preventive Nutrition – Novum, Department of Biosciences, Karolinska Institutet, S-141 57 Huddinge, Sweden: [2]Division of Physiotherapy, Department of Neurotec, Karolinska Institutet, Stockholm, Sweden

## Abstract

*Introduction:* The International Physical Activity Questionnaire (IPAQ) was developed to measure health-related physical activity (PA) in populations. The short version of the IPAQ has been tested extensively and is now used in many international studies. The present study aimed to explore the validity characteristics of the long-version IPAQ.

*Subjects and methods:* Forty-six voluntary healthy male and female subjects (age, mean ± standard deviation: $40.7 \pm 10.3$ years) participated in the study. PA indicators derived from the long, self-administered IPAQ were compared with data from an activity monitor and a PA log book for concurrent validity, and with aerobic fitness, body mass index (BMI) and percentage body fat for construct validity.

*Results:* Strong positive relationships were observed between the activity monitor data and the IPAQ data for total PA ($\rho = 0.55$, $P < 0.001$) and vigorous PA ($\rho = 0.71$, $P < 0.001$), but a weaker relationship for moderate PA ($\rho = 0.21$, $P = 0.051$). Calculated MET-h day$^{-1}$ from the PA log book was significantly correlated with MET-h day$^{-1}$ from the IPAQ ($\rho = 0.67$, $P < 0.001$). A weak correlation was observed between IPAQ data for total PA and both aerobic fitness ($\rho = 0.21$, $P = 0.051$) and BMI ($\rho = 0.25$, $P = 0.009$). No significant correlation was observed between percentage body fat and IPAQ variables. Bland–Altman analysis suggested that the inability of activity monitors to detect certain types of activities might introduce a source of error in criterion validation studies.

*Conclusions:* The long, self-administered IPAQ questionnaire has acceptable validity when assessing levels and patterns of PA in healthy adults.

Despite general agreement regarding the significance of physical activity (PA) for health at both individual and population level[1], there is a need for more precise understanding about the amount and the pattern of PA required for different health-enhancing effects to occur[2,3]. One of the reasons for this is the difficulty in accurately assessing total PA, and PA of different intensity levels, in day-to-day life and over extended periods of time[4,5].

PA can be assessed using subjective (questionnaires, diaries, etc.) or objective (motion sensors, heart-rate monitors, etc.) methods[4]. At population level, questionnaires are the most commonly used. Most existing questionnaires focus on PA during leisure time or at the workplace, which limits the use of these instruments. Only a few of the existing questionnaires capture PA in a variety of daily situations, such as transportation, occupation, household and family care, and leisure time[6]. No standardised instrument capturing all facets of PA has been available[2,4,7].

To overcome this problem, a consensus group of PA assessment experts met in 1998 to develop a valid and reliable questionnaire measuring health-enhancing PA covering most daily situations.

A questionnaire was designed specifically for adults (18–65 years old) and consisted of four domains: (1) during transportation, (2) at work, (3) during household and gardening tasks and (4) during leisure time, including exercise and sport participation. Two questionnaires were proposed: a short form (nine items) and a long form (31 items). Altogether eight versions differing by length, reference period and mode of administration were designed. Twelve countries participated in an evaluation of both the reliability and validity of the instrument[7]. This questionnaire is now known as the International Physical Activity Questionnaire (IPAQ).

The results from the 12-country reliability test produced repeatable data for all questionnaire versions tested (Spearman's $\rho$ of 0.81 for the long form and 0.76 for the short form), with correlated but not directly comparable data from the short and the long version[7]. Criterion validity was assessed against an accelerometer, the Manufacturing Technologies Inc. (MTI) Actigraph, model WAM 7164

*Corresponding author:* Email maria.hagstromer@prevnut.ki.se

(MTI, Fort Walton Beach, FL, USA; formerly known as the Computer Science and Application (CSA) activity monitor)[8]. The criterion validity for the total amount of weekly PA had a median correlation of 0.33 and 0.30 (Spearman's $\rho$) for the long and short form, respectively[7], which is comparable to results from previous self-reported PA validation studies using objective criterion instruments such as accelerometers[9]. The correlation between different intensity levels was not analysed in the 12-country validity study.

Although the results were promising, the IPAQ instrument needs further evaluation especially with regard to its validity. The short versions have been found suitable for use in large-scale surveys such as the Eurobarometer[10] and International Prevalence Study[11]. Interest has now been raised in the long form of the IPAQ and its suitability for use in research.

The purpose of the present study was to further evaluate the concurrent and construct validity of the self-administered, last 7 days, long form version of the IPAQ in a Swedish sample of adult men and women.

## Subjects and method

### Study design
A cross-sectional study was performed, comparing measures of PA by the long, self-administered, last 7 day version of the IPAQ with those obtained by a log book and an activity monitor for concurrent validity, and with aerobic fitness and body composition for construct validity.

Validity has been defined as the extent to which an instrument assesses the true exposure of interest[4]. This definition is frequently referred to as internal validity and implies an absolute measure of the variable of interest. However, as PA is a multidimensional exposure, it is difficult to find an exact absolute measure for it. Instead, indirect criteria of the exposure, such as data derived from activity monitors and diaries, can be used to assess concurrent validity. Another aspect of validity is construct validity, which indicates the consistency between the activity instrument and a physiological variable related to PA such as maximal aerobic power or body mass index (BMI). For example, episodes of vigorous PA should correlate with aerobic fitness[4].

### Subjects
Fifty healthy volunteers (24 men) living in the Stockholm metropolitan area of Sweden participated in the study. Four subjects (two men) were excluded from the analyses due to incomplete log book activity recordings or failure to test aerobic fitness. The remaining 46 subjects had a higher level of education than the average Swedish population, as indicated by the high prevalence of university education. The physical characteristics of the subjects are shown in Table 1.

**Table 1** Characteristics of subjects, $n = 46$ (22 males)

| | |
|---|---|
| Age (years) | 40.7 ± 10.3 |
| Height (cm) | 171.0 ± 8.9 |
| Weight (kg) | 70.4 ± 10.3 |
| Body mass index (kg m$^{-2}$) | 24.0 ± 2.4 |
| Body fat (%) | 26.6 ± 2.1 |
| Aerobic fitness (ml O$_2$ kg$^{-1}$ min$^{-1}$) | 37.7 ± 9.8 |

Values are expressed as mean ± standard deviation.

The ethics committee of Huddinge University Hospital approved the study protocol and the subjects provided written informed consent.

### Procedure
On day 1 the participants were invited to the clinic and were provided with detailed instruction on how to use the activity monitor and how to fill in the activity log book. Data on anthropometric and demographic characteristics were collected. Starting on day 2 the participants wore the activity monitor for seven consecutive days and filled in the log book at the end of each day. On day 8 the subjects returned to the clinic, filled in the IPAQ and performed an aerobic fitness test.

### The IPAQ instrument
The IPAQ was translated into Swedish and translated back into English using the instructions given in the IPAQ manual for reliability and validity[11]. The long, self-administered IPAQ covers four domains of physical activity: work-related, transportation, housework/gardening and leisure-time activity. The questionnaire also includes questions about time spent sitting as an indicator of sedentary behaviour. In each of the four domains the number of days per week and time per day spent in both moderate and vigorous activity are recorded. At work, during transportation and in leisure time, walking time is also included. Practical examples of culturally relevant activities of moderate and vigorous intensity are given. In this study, moderate intensity was defined as 3–6 MET (Metabolic Equivalent Task) and vigorous intensity was defined as >6 MET[12]. One MET is equal to energy expenditure during rest and is approximately equal to 3.5 ml O$_2$ kg$^{-1}$ min$^{-1}$ in adults.

Outcome measures used were: (1) MET hours per week and (2) hours reported in moderate- and vigorous-intensity activity per week. PA data from the questionnaire were transformed into energy expenditure estimates as MET using published values[13,14]. To calculate the weekly physical activity (MET-h week$^{-1}$), the number of hours dedicated to each activity class was multiplied by the specific MET score for that activity[7].

### Concurrent validity

#### Activity monitor
The MTI activity monitor measures accelerations ($g$) from 0.05 to 2.1 $g$ in the vertical axis. It is equipped with a

frequency bandwidth filter, which discriminates human movements from vibrations. The output from the monitor is sampled 10 times per second and summed over a selected time interval or epoch[8]. The monitor was secured directly to the skin just above the suprailiac crest using an elastic belt. It was initialised as described by the manufacturer and the 60-s epoch was used. The subjects were carefully instructed verbally and in writing on how to handle the monitor. The subjects were asked to wear the monitor during waking hours for seven consecutive days but not during water activities.

The activity monitor provides measures of PA intensity and duration and of sedentary periods. The sum of accelerations is transformed into counts. Cut-off values for moderate-intensity (3–6 MET, 1952–5723 counts) and vigorous-intensity ($>6$ MET, 5724 counts) PA were used[15] to calculate time per week (h week$^{-1}$) spent at these intensity levels and thereafter compared with the self-reported time (h week$^{-1}$) spent at these intensity levels from the IPAQ. Total activity was expressed as counts per recorded time (counts min$^{-1}$) and thereafter compared with the self-reported total activity (MET-h week$^{-1}$). In order to get a distinct measure of inactivity, sedentary time was expressed as an arbitrary cut-off of $<101$ counts[16] and compared with self-reported time spent sitting from the IPAQ.

*PA log book*
The log book consisted of one page for each day of the 7-day period in which the subject filled out the type of and time spent in PA. Each page had separate entries for the four different PA domains: physical activity at work, during transportation, household activity and leisure-time activity. The subjects were also asked to report the amount of time spent sitting. Data from the log book were given a MET value derived from the compendium of physical activities by Ainsworth *et al.*[13,14] and transformed into MET-hours per week (MET-h week$^{-1}$), as described earlier, and then compared with the questionnaire. For example, if a person reported a 20-min bike ride to work twice a week, 30 min of carrying heavy loads at work once a week and a 30-min walk in leisure time four times a week, the MET-hours per week value was calculated as $(2 \times 20 + 1 \times 30 + 4 \times 30) = 190$ MET-min week$^{-1}$ or 3.16 MET-h week$^{-1}$. Reported time spent in moderate and vigorous intensity of PA for each domain (work, transport, home and leisure-time) was also calculated.

**Construct validity**

*Aerobic fitness*
Aerobic fitness ($\dot{V}O_{2max}$, ml $O_2$ kg$^{-1}$ min$^{-1}$) was estimated using a 15-min sub-maximal treadmill walking test. The test began at a speed of 4.8 km h$^{-1}$ at zero gradient. The gradient was raised by 3% every third minute up to a 15% gradient. Heart rate (HR) was measured throughout the

test, using a Polar Vantage NV monitor (Polar Electro OY, Kempele, Finland). During the last minute of each stage perceived exhaustion was assessed using the Borg RPE 6–20 scale[17]. The maximal aerobic power was calculated from the last heart rate at the final stage using a modified Balke formula[18].

*Anthropometry*
Body weight and height were measured in light clothing using standard equipment and BMI (kg m$^{-2}$) was calculated. Four skinfold measurements (biceps, triceps, subscapular and suprailiac) were taken by the same person using a Harpenden calliper and percentage body fat (BF) was calculated according to Durnin and Womersley[19]. The intra-tester reliability was assessed by calculating the technical error of measurement (TEM). Ten subjects were measured twice with at least 3 days between each measurement. The following TEM was achieved: biceps = 6.7%, triceps = 2.7%, subscapular = 2.5% and suprailiac = 3.8%. A TEM $<10\%$ was considered as acceptable inter-tester reliability[20].

Maximal aerobic power (ml $O_2$ kg$^{-1}$ min$^{-1}$), BMI and BF (%) were considered measures of health-related fitness and thereby considered as construct validity criteria for activity. They were compared with total physical activity (MET-h day$^{-1}$) from IPAQ. Maximal aerobic power (ml $O_2$ kg$^{-1}$ min$^{-1}$) was also compared with time (min day$^{-1}$) spent in moderate- and vigorous-intensity PA from IPAQ.

**Statistical analysis**
All statistical analyses were performed using the Statistical Package for the Social Sciences for Windows, version 10.0 (SPSS Inc., Chicago, IL, USA).

The characteristics of subjects and outcomes from the IPAQ, log book and activity monitor are described as mean ± standard deviation (SD).

The non-parametric Spearman correlation coefficient ($\rho$) was calculated to assess the relationship between MET-h week$^{-1}$ from IPAQ data and MTI counts, log book outcomes, aerobic capacity and BF (%). Judgements of the significance of the correlation coefficients were made according to Altman[21]. The Bland–Altman method[21,22] was used to provide an indication of the systematic and random error and heteroscedasticity of the data, and 95% limits of agreement were used for describing the total error between the two methods. Variables used for the Bland–Altman analysis were weekly time spent in moderate and vigorous activity according to IPAQ versus MTI and the weekly total amount of PA (MET) according to IPAQ versus the log book. The level of significance was set at $P < 0.05$.

**Results**

There were no significant gender differences regarding age, BMI or time spent in different intensities of PA

measured with the accelerometer. Consequently, the two genders were combined in all further analyses.

Tables 1–3 present descriptive data of most variables used in the analysis.

### Construct validity

#### IPAQ versus the activity monitor
Table 4 shows that the reported time (h week$^{-1}$) spent in vigorous-intensity PA and the total amount of PA (MET-h week$^{-1}$) from the IPAQ were significantly correlated with the time spent in vigorous-intensity PA ($\rho = 0.71$, $P < 0.001$) and the total amount of PA (counts min$^{-1}$) from the MTI activity monitor ($\rho = 0.55$, $P < 0.001$), respectively. The correlation between reported time in moderate PA by IPAQ and time spent in moderate level from the MTI activity monitor was not significant ($\rho = 0.12$) (Table 4). Self-reported time ($52.0 \pm 16.1$ h week$^{-1}$) spent sitting according to IPAQ and the amount of time at $<101$ counts ($50.4 \pm 9.1$ h week$^{-1}$) recorded by the MTI activity monitor were not significantly correlated ($\rho = 0.17$).

The absolute difference between reported time for total PA (moderate- and vigorous-intensity) from IPAQ and measured time in moderate and vigorous intensity from the activity monitor was $1.0 \pm 8.3$ h week$^{-1}$ (not significant, NS).

A Bland−Altman plot for the activity monitor (MTI) and the IPAQ showed that the mean differences were small ($1.0$ h week$^{-1}$ or $15$ min day$^{-1}$). However, the 95% limits of agreement were wide, ranging from $-15$ to $+17$ h week$^{-1}$, mainly due to nine outliers that reported more than $12$ h week$^{-1}$ of total activity (Fig. 1).

#### IPAQ versus the log book
Calculated MET-h day$^{-1}$ from the PA log book was significantly correlated with MET-h day$^{-1}$ from IPAQ ($\rho = 0.67$, $P < 0.001$). Excluding five outliers the correlation was slightly higher ($\rho = 0.77$, $P < 0.001$). The absolute difference between reported energy expenditure (MET-h week$^{-1}$) from IPAQ and reported total physical activity (MET-h week$^{-1}$) from the log book was $-2.9 \pm 44$ MET-h week$^{-1}$ (NS).

**Table 2** Descriptive data from the International Physical Activity Questionnaire (IPAQ) and the activity monitor output ($n = 46$)

|  | IPAQ | Activity monitor |
|---|---|---|
| HEPA (h week$^{-1}$) | $7.4 \pm 9.5$ | $10.8 \pm 3.4$ |
| Moderate PA (h week$^{-1}$) | $5.1 \pm 6.9$ | $9.1 \pm 2.7$ |
| Vigorous PA (h week$^{-1}$) | $2.3 \pm 4.4$ | $1.7 \pm 1.1$ |
| Total activity† | $51 \pm 51$ | $591 \pm 219$ |
| Sedentary (h week$^{-1}$) | $52.0 \pm 16.1$ | $50.4 \pm 9.1$ |

HEPA – health-enhancing physical activity ( = moderate plus vigorous activity); PA – physical activity.
Values are expressed as mean ± standard deviation.
† Total activity in MET-h week$^{-1}$ (IPAQ) or counts min$^{-1}$ (activity monitor).

**Table 3** Descriptive data from the different domains of the International Physical Activity Questionnaire (IPAQ) and the log book ($n = 46$)

|  | IPAQ | Log book |
|---|---|---|
| PA at work (MET-h week$^{-1}$) | $15.3 \pm 30.1$ | $16.5 \pm 4.2$ |
| PA during transport (MET-h week$^{-1}$) | $5.1 \pm 4.2$ | $15.8 \pm 14.3$ |
| PA at home or in garden (MET-h week$^{-1}$) | $15.2 \pm 12.6$ | $7.4 \pm 11.6$ |
| Leisure-time PA (MET-h week$^{-1}$) | $14.5 \pm 12.8$ | $15.4 \pm 19.8$ |
| Time spent sitting (h week$^{-1}$) | $52.0 \pm 16.0$ | $44.9 \pm 15.5$ |

PA – physical activity.
Values are expressed as mean ± standard deviation.

In addition, comparisons between self-reported energy expenditure (MET-h day$^{-1}$) for each of the four different domains from the IPAQ and the log book were performed (Table 5). A significant relationship ($P < 0.001$) between the methods was observed for three (work PA, $\rho = 0.64$; home/garden PA, $\rho = 0.47$; leisure-time PA, $\rho = 0.58$) of the four domains as well as for time spent sitting ($\rho = 0.75$, $P < 0.001$).

A Bland−Altman plot for total MET-h week$^{-1}$ for the PA log book and the IPAQ showed that the mean differences were small ($2.9$ MET-h week$^{-1}$), but a few outliers affected the 95% limits of agreement (Fig. 2). Otherwise the methods were similar and there were no statistical differences between the methods.

To better understand the results, the total PA variables for the IPAQ, activity monitor and log book were split into tertiles and compared using percentage of agreement and kappa coefficient (data not shown). The results were similar to the Spearman analysis: low percentage of agreement and low kappa coefficient values for both IPAQ total activity versus MTI total activity (% agreement) and IPAQ total activity versus log book total activity (% agreement).

### Construct validity
Aerobic fitness showed a weak positive correlation with the total amount of PA (MET-h week$^{-1}$) and with the time spent in moderate-intensity PA (h week$^{-1}$) from IPAQ ($\rho = 0.21$, $P < 0.05$). There was no significant relationship between aerobic fitness and time spent in vigorous activity (h week$^{-1}$) from IPAQ ($\rho = 0.14$). There was a weak positive correlation between BMI and total amount (MET-h week$^{-1}$) of PA ($\rho = 0.25$, $P < 0.01$) and time (h week$^{-1}$) spent in moderate ($\rho = 0.27$, $P < 0.01$) and vigorous activity ($\rho = 0.17$, $P < 0.05$) from IPAQ. There was no significant relationship between BF (%) and any of the PA variables from IPAQ (Table 6).

### Discussion

The present study examined the concurrent and construct validity of the Swedish long, self-administered, last 7 days version of the IPAQ. The results indicated that IPAQ had acceptable properties for assessing PA in healthy adults.

**Table 4** Spearman correlation coefficients ($\rho$) for time spent in physical activity (PA) from the International Physical Activity Questionnaire (IPAQ) and the activity monitor ($n = 46$)

| IPAQ measure | Activity monitor measure | $\rho$ |
|---|---|---|
| Vigorous PA (h week$^{-1}$) | Vigorous PA (h week$^{-1}$) | 0.63*** |
| Moderate PA (h week$^{-1}$) | Moderate PA (h week$^{-1}$) | 0.12 |
| Vigorous and moderate PA (h week$^{-1}$) | Vigorous and moderate PA (h week$^{-1}$) | 0.36*** |
| Vigorous PA (h week$^{-1}$) | Total PA (counts min$^{-1}$) | 0.71*** |
| Moderate PA (h week$^{-1}$) | Total PA (counts min$^{-1}$) | 0.20 |
| Total PA (MET-h week$^{-1}$) | Total PA (counts min$^{-1}$) | 0.55*** |
| Time spent sitting (h week$^{-1}$) | Inactivity (h week$^{-1}$) | 0.17 |

***, $P < 0.001$.

The subjects were volunteers whose age range, 19–62 years, was similar to the population for whom the IPAQ instrument was designed. Most of the subjects lived in or just outside the Stockholm City area and had a higher level of education than the average population.

One particular issue for the validation of PA questionnaires is the choice of an appropriate criterion measure[5,23]. The gold standard for estimating energy expenditure in free-living individuals is the doubly labelled water method[24]. However, this method is expensive and cannot provide information about the intensity, frequency and pattern of PA. On the other hand, the output from the MTI activity monitor is a valid indicator of the total amount of PA[25]. Activity monitors have recently been suggested as one of the best criterion instruments when validating self-report instruments of PA[26]. In the present study the activity monitor was considered a criterion measure for concurrent validity. Furthermore, the MTI measurements were performed for the same time period as the questionnaire. We have therefore no reason to believe that our respondents did not refer to the same day when answering the questionnaire as was measured by the activity monitor.

One limitation with the use of self-report instruments is that a single estimate of the energy cost of a specific activity, taken from a published compendium, is applied to all individuals[13,14,26]. This does not allow for inter-individual variation in energy expenditure for a given intensity or through variations in mechanical and metabolic efficiency[4,27,28]. However, the published estimates of energy costs of different PA types are the only such data available.

Most correlations between the total amount of PA from IPAQ and the MTI activity monitor were moderate. Several factors might explain these correlations. First, it has been argued that it is difficult to obtain a good measure of low and moderate PA using self-administered questionnaires[4,27,28], because these activities are being accumulated throughout the day and the number and diversity of these activities is enormous, resulting in a poor recall. In contrast, high-intensity PA such as different types of exercise are much more structured and stable over time and are much easier to recall. The stronger correlations found in the present study for vigorous-intensity PA compared with moderate-intensity PA illustrate this point and agree with previous findings[4]. Second, it has been reported that people tend to overestimate time spent in high-intensity activities and underreport time spent in light and moderate-intensity activities[9,29]. The MTI activity monitor is known to underestimate PA level at specific activities and this may lead to underestimation of total energy expenditure[30,31]. Ekelund et al.[32] showed that the MTI activity monitor underestimated energy expenditure during activities like bicycling, skating and strength
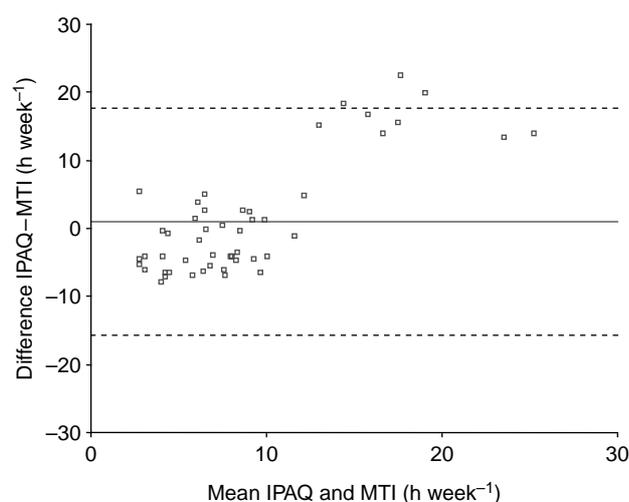


**Fig. 1** Bland–Altman plot for time spent in at least moderate physical activity (h week$^{-1}$) as assessed by the International Physical Activity Questionnaire (IPAQ) and measured using an activity monitor (Manufacturing Technologies Inc. Actigraph; MTI). Mean difference: 1.0 h week$^{-1}$ ± 2SD (standard deviations), −15.7 to 17.7 h week$^{-1}$ (not significant)

**Table 5** Spearman correlation coefficients ($\rho$) for physical activity (PA) from the International Physical Activity Questionnaire (IPAQ) and the log book ($n = 46$)

| IPAQ measure versus log book | $\rho$ |
|---|---|
| PA at work (MET-h week$^{-1}$) | 0.64*** |
| PA during transport (MET-h week$^{-1}$) | 0.18 |
| PA at home or in garden (MET-h week$^{-1}$) | 0.47** |
| Leisure-time PA (MET-h week$^{-1}$) | 0.58*** |
| Time spent sitting (h week$^{-1}$) | 0.75*** |

**, $P < 0.01$; ***, $P < 0.001$.
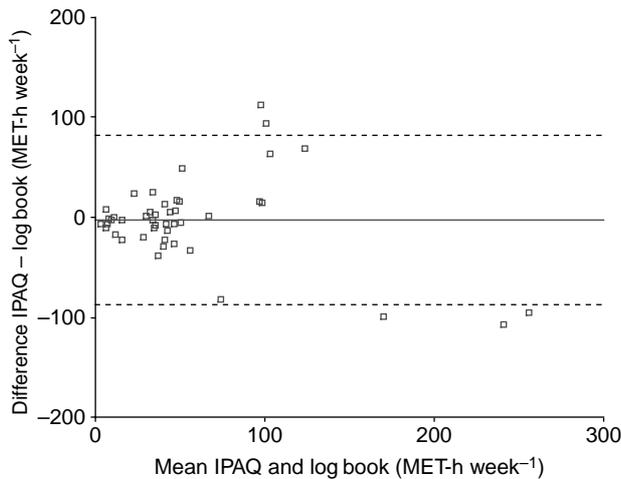
**Fig. 2** Bland−Altman plot for total physical activity (MET-h week$^{-1}$) as assessed by the International Physical Activity Questionnaire (IPAQ) and a 7-day activity log book. Mean difference: −2.9 MET-h week$^{-1}$ ± 2SD (standard deviations), −88 to 82 MET-h week$^{-1}$ (not significant)

training. Because uniaxial activity monitors, such as MTI, can only provide information about acceleration in the vertical plane, they do not yield correct information about the intensity of activities like carrying heavy loads, walking on stairs or riding a bike.

The Bland−Altman plot showed a small difference between the means of total weekly activity hours by IPAQ and by the activity monitor but the 95% limits of agreement were affected considerably by nine outliers. On further examination of the outliers it was found that three of these reported cycling to work or as mail delivery for up to 5 h per day. One subject did strength training at vigorous level for 60−90 min every day. Three subjects reported at least 60−120 min of daily vigorous activity at work, carrying heavy loads, etc. Since it is likely that these activities were not recorded by activity monitor, the IPAQ showed much higher values for these individuals. Two of the outliers reported extreme activity values at work according to IPAQ and seemed to have misunderstood the questions.

In the present study we used cut-off values for vigorous

and moderate PA obtained from an experimental study[33]. These cut-off points have been criticised for use in field studies and different cut-off points have been suggested[15,31,34,35]. It seems that when establishing cut-off points, the results are affected by the types of activities and the settings where the calibration between activity counts and energy expenditure estimates were performed[16]. It has been suggested that until cut-off points that are known to reflect normal daily life have been established, the main application of the activity monitor may be as an indicator of the total amount of PA[25].

When the IPAQ was compared with the 7-day log book, moderate to high correlations were found for the work and leisure-time domains. The correlations for total PA also supported the concurrent validity of the IPAQ. Again, low- and moderate-intensity PA had low correlations. However, they correlated better with the log book than with the MTI activity counts. It would have been interesting to compare the specific domains from IPAQ with the activity monitor. However, owing to the design of the study this could not be performed, because we had no data on actual working hours or on time in transportation and domestic work.

The Bland−Altman plot for total weekly activity reported by IPAQ and by the log book showed a small difference between the means by the two methods, but the 95% limits of agreement were affected substantially by a few outliers reporting >100 MET-h week$^{-1}$. On closer inspection of the outliers it was found that the two subjects who reported extreme values at work did not report this in the diary, and two other subjects reported extreme values of walking (4−5 h per day) in the diary but not in the IPAQ, hence accounting for the four extreme values outside the 95% confidence intervals.

A relatively weak, although significant, correlation between both total PA (MET-h week$^{-1}$) and moderate PA (h day$^{-1}$) from the IPAQ and aerobic fitness was found. The relationship between vigorous-intensity PA from IPAQ and aerobic fitness was not significant ($\rho = 0.14$). A similar correlation ($r = 0.16$) was recently reported when comparing time spent in vigorous-intensity PA with

**Table 6** Spearman rank correlation coefficients ($\rho$) for reported time of physical activity (PA) in different domains from the International Physical Activity Questionnaire (IPAQ) and measures of construct validity ($n = 46$)

| IPAQ measure | Construct measure | $\rho$ |
|---|---|---|
| Total PA (MET-h week$^{-1}$) | Aerobic fitness (ml O$_2$ kg$^{-1}$ min$^{-1}$) | 0.21* |
| Moderate PA (h week$^{-1}$) | Aerobic fitness (ml O$_2$ kg$^{-1}$ min$^{-1}$) | 0.21* |
| Vigorous PA (h week$^{-1}$) | Aerobic fitness (ml O$_2$ kg$^{-1}$ min$^{-1}$) | 0.14 |
| Total PA (MET-h week$^{-1}$) | BMI (kg m$^{-2}$) | 0.25** |
| Moderate PA (h week$^{-1}$) | BMI (kg m$^{-2}$) | 0.27** |
| Vigorous PA (h week$^{-1}$) | BMI (kg m$^{-2}$) | 0.17* |
| Total PA (MET-h week$^{-1}$) | BF (%) | 0.11 |
| Moderate PA (h week$^{-1}$) | BF (%) | 0.12 |
| Vigorous PA (h week$^{-1}$) | BF (%) | 0.09 |

BMI − body mass index; BF, body fat.
*, $P < 0.05$; **, $P < 0.01$.

aerobic fitness using another self-report instrument[36]. Pereira et al.[37] evaluated a collection of PA questionnaires and found that most of them showed a weak correlation between reported time in moderate PA and aerobic fitness. On the other hand, they reported a moderate to strong relationship between time spent in vigorous PA and aerobic fitness ($r = 0.45–0.61$). These differences between studies may be explained by the differences in the questionnaires used. The IPAQ questionnaires as well as the EPIC (European Prospective Investigation into Cancer and Nutrition) questionnaire[36] are both designed to collect data on PA in a variety of settings, whereas previous questionnaires have mainly focused on leisure-time PA and structured exercise.

Thus, it seems that instruments designed for collecting data on leisure-time PA and structured exercise demonstrate a higher correlation with aerobic fitness, whereas 'all-activity' questionnaires demonstrate a higher relationship to objectively measured total PA. This highlights the need to choose an appropriate criterion instrument when validating 'all-activity' questionnaires.

No significant correlation was found between any of the PA variables from IPAQ and BF (%). These results agree with previous findings[38,39].

In conclusion, the results from the present study show that the long-version, self-administered, last 7 days IPAQ instrument has acceptable validity properties for assessing different domains of PA, PA intensities and total PA in healthy adults. The concurrent validity provided stronger agreement with IPAQ than the construct validity, which is likely to reflect the 'all-activity' nature of the IPAQ. The findings from the Bland–Altman analysis suggest that, when using activity monitors in validation studies, the inability of the monitors to detect certain types of activities may introduce a source of measurement error.

## Acknowledgements

## References

1 Centers for Disease Control and Prevention (CDC). *Physical Activity and Health. A Report from the Surgeon General.* Pittsburg, PA: CDC, 1996.

2 Booth M. Assessment of physical activity: an international perspective. *Research Quarterly for Exercise and Sport* 2000; **71**: S114–20.

3 Oja P. Descriptive epidemiology of health-related physical activity and fitness. *Research Quarterly for Exercise and Sport* 1995; **66**: 303–12.

4 Welk G. *Physical Assessment in Health-related Research.* Leeds, UK: Human Kinetics, 2002.

5 Wareham NJ, Rennie KL. The assessment of physical activity in individuals and populations: why try to be more precise about how physical activity is assessed? *International Journal of Obesity and Related Metabolic Disorders* 1998; **22**(Suppl. 2): S30–8.

6 Kriska A, Caspersen C. Introduction. In: Pereira MA, FitzerGerald SJ, Gregg EW, Joswiak ML, Ryan WJ, Suminski RR, et al. eds. A Collection of Physical Activity Questionnaires for Health-related Research [special issue]. *Medicine and Science in Sports and Exercise* 1997; **29**: S3–9.

7 Craig CL, Marshall AL, Sjostrom M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. *Medicine and Science in Sports and Exercise* 2003; **35**: 1381–95.

8 Computer Science and Applications, Inc. (CSA). *Activity Monitor Operator's Manual, Model 7164, Version AM7164-2.2.* Shalimar, FL: CSA, 1995.

9 Sallis JF, Saelens BE. Assessment of physical activity by self-report: status, limitations, and future directions. *Research Quarterly for Exercise and Sport* 2000; **71**: S1–S14.

10 Rutten A, Abu-Omar K. Prevalence of physical activity in the European Union. *Sozial- und Präventivmedizin* 2004; **49**: 281–9.

11 The International Physical Activity Questionnaire, 2005. Available at http://www.ipaq.ki.se/

12 American College of Sports Medicine Position Stand. The recommended quantity and quality of exercise for developing and maintaining cardiorespiratory and muscular fitness, and flexibility in healthy adults. *Medicine and Science in Sports and Exercise* 1998; **30**: 975–91.

13 Ainsworth BE, Haskell WL, Leon AS, Jacobs DR Jr, Montoye HJ, Sallis JF, et al. Compendium of physical activities: classification of energy costs of human physical activities. *Medicine and Science in Sports and Exercise* 1993; **25**: 71–80.

14 Ainsworth BE, Haskell WL, Whitt MC, Irwin ML, Swartz AM, Strath SJ, et al. Compendium of physical activities: an update of activity codes and MET intensities. *Medicine and Science in Sports and Exercise* 2000; **32**: S498–S504.

15 Freedson PS, Melanson E, Sirard J. Calibration of the Computer Science and Applications, Inc. accelerometer. *Medicine and Science in Sports and Exercise* 1998; **30**: 777–81.

16 Yngve A, Nilsson A, Sjostrom M, Ekelund U. Effect of monitor placement and of activity setting on the MTI accelerometer output. *Medicine and Science in Sports and Exercise* 2003; **35**: 320–6.

17 Borg G. *Borg's Perceived Exertion and Pain Scales.* Stockholm: Human Kinetics, 1998.

18 American College of Sports Medicine. Physical fitness testing. In: *ACSM's Guidelines for Exercise Testing and Prescription.* London: Williams & Wilkins, 1995; 49–78.

19 Durnin JV, Womersley J. Body fat assessed from total body density and its estimation from skinfold thickness: measurements on 481 men and women aged from 16 to 72 years. *British Journal of Nutrition* 1974; **32**: 77–97.

20 Norton K, Olds T. *Anthropometrica: A Textbook of Body Measurement for Sports and Health Courses.* Sydney: University of New South Wales Press, 1996; 80–90.

21 Altman D. *Practical Statistics for Medical Research.* London: Chapman & Hall, 1991.

22 Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Medicine* 1998; **26**: 217–38.

23 Rennie KL, Wareham NJ. The validation of physical activity instruments for measuring energy expenditure: problems and pitfalls. *Public Health Nutrition* 1998; **1**: 265–71.

24 Schoeller DA. Measurement of energy expenditure in free-living humans by using doubly labeled water. *Journal of Nutrition* 1988; **118**: 1278–89.

25 Ekelund U, Sjostrom M, Yngve A, Poortvliet E, Nilsson A, Froberg K, et al. Physical activity assessed by activity monitor and doubly labeled water in children. *Medicine and Science in Sports and Exercise* 2001; **33**: 275–81.

26  Sirard JR, Pate RR. Physical activity assessment in children and adolescents. *Sports Medicine* 2001; **31**: 439–54.

27  Washburn RA, Heath GW, Jackson AW. Reliability and validity issues concerning large-scale surveillance of physical activity. *Research Quarterly for Exercise and Sport* 2000; **71**: S104–13.

28  Bassett DR Jr. Validity and reliability issues in objective monitoring of physical activity. *Research Quarterly for Exercise and Sport* 2000; **71**: S30–6.

29  Klesges RC, Eck LH, Mellon MW, Fulliton W, Somes GW, Hanson CL. The accuracy of self-reports of physical activity. *Medicine and Science in Sports and Exercise* 1990; **22**: 690–7.

30  Sirard JR, Melanson EL, Li L, Freedson PS. Field evaluation of the Computer Science and Applications, Inc. physical activity monitor. *Medicine and Science in Sports and Exercise* 2000; **32**: 695–700.

31  Hendelman D, Miller K, Baggett C, Debold E, Freedson P. Validity of accelerometry for the assessment of moderate intensity physical activity in the field. *Medicine and Science in Sports and Exercise* 2000; **32**: S442–9.

32  Ekelund U, Yngve A, Sjostrom M, Westerterp K. Field evaluation of the Computer Science and Application's Inc. activity monitor during running and skating training in adolescent athletes. *International Journal of Sports Medicine* 2000; **21**: 586–92.

33  Freedson P, Melanson E, Sirard J. Calibration of the Computer Science and Applications, Inc. accelerometer. *Medicine and Science in Sports and Exercise* 1998; **30**: 777–81.

34  Ainsworth BE, Bassett DR Jr, Strath SJ, Swartz AM, O'Brien WL, Thompson RW, *et al.* Comparison of three methods for measuring the time spent in physical activity. *Medicine and Science in Sports and Exercise* 2000; **32**: S457–64.

35  Nichols JF, Morgan CG, Chabot LE, Sallis JF, Calfas KJ. Assessment of physical activity with the Computer Science and Applications, Inc., accelerometer: laboratory versus field validation. *Research Quarterly for Exercise and Sport* 2000; **71**: 36–43.

36  Wareham NJ, Jakes RW, Rennie KL, Mitchell J, Hennings S, Day NE. Validity and repeatability of the EPIC–Norfolk Physical Activity Questionnaire. *International Journal of Epidemiology* 2002; **31**: 168–74.

37  Pereira MA, FitzerGerald SJ, Gregg EW, Joswiak ML, Ryan WJ, Suminski RR, *et al.*, eds. A collection of Physical Activity Questionnaires for health-related research [special issue]. *Medicine and Science in Sports and Exercise* 1997; **29**: S1–205.

38  Jacobs DR Jr, Haskell HW, Haskell L, Pirie P, Sidney S. Reliability and validity of a short physical activity history: CARDIA and Minnesota Heart Health Program. *J Cardio-pulmonary Rehabilitation* 1989; **9**: 448–59.

39  Bouchard C, Tremblay A, Leblanc C, Lortie G, Savard R, Theriault G. A method to assess energy expenditure in children and adults. *American Journal of Clinical Nutrition* 1983; **37**: 461–7.