
Tom Loney, Martyn Standage, Dylan Thompson, Simon J. Sebire, and Sean Cumming

Background: To examine the agreement between self-reported and objectively assessed physical activity (PA) according to current public health recommendations. Methods: One-hundred and fourteen British University students wore a combined accelerometer and heart rate monitor (Actiheart; AHR) to estimate 24-hour energy expenditure over 7 consecutive days. Data were extracted based on population-based MET-levels recommended to improve and maintain health. On day 8, participants were randomly assigned to complete either the short-form International Physical Activity Questionnaire (IPAQ) or the Leisure-Time Exercise Questionnaire (LTEQ). Estimates of duration (IPAQ; N = 46) and frequency (LTEQ; N = 41) of PA were compared with those recorded by the AHR. Results: Bland-Altman analysis showed the mean bias between the IPAQ and AHR to be small for moderate-intensity and total PA, however the 95% limits of agreement (LOA) were wide. The mean number of moderate bouts of PA estimated by the LTEQ was similar to those derived by the AHR but the 95% LOA between the 2 measures were large. Conclusions: Although self-report questionnaires may provide an approximation of PA at a population level, they may not determine whether an individual is participating in the type, intensity, and amount of PA advocated in current public health recommendations. Keywords: accelerometry, heart rate monitoring, limits of agreement, method-comparison

Regular physical activity (PA) confers numerous psychological and physiological health benefits.1–3 Accordingly, the accurate and valid quantification of PA is critical for researchers concerned with epidemiology, behavioral medicine, intervention evaluation, and understanding the dose-response relationship between energy expenditure and health.

A variety of methods exist for assessing PA in free-living and field settings, including self-report questionnaires, motion sensors (pedometers and accelerometers), and heart rate telemetry. Two self-report instruments that have been widely used in the extant literature are the short-form International Physical Activity Questionnaire (IPAQ)4 and the Leisure-Time Exercise Questionnaire (LTEQ).5,1 The IPAQ and LTEQ have been previously validated against objective measures of PA, including pedometers, accelerometers, and heart rate telemetry (eg,4,6–8), past work has shown that these instruments have a number of limitations (see9–11 for reviews). With this in mind, the 1-piece Actiheart unit (Actiheart [AHR], Cambridge Neurotechnology, UK) has been developed to synchronize accelerometer and heart rate measurements using a novel branched-equation model to estimate minute-by-minute energy expenditure above rest (see12 for a technical overview).

Previous research validating the AHR has demonstrated excellent estimates for total energy expenditure against indirect calorimetry during treadmill walking and running in healthy young individuals,12–14 and for a range of low-to-moderate intensity lifestyle activities, such as sweeping with a broom, digging and transferring sand into boxes, simulated watering of house plants, and folding and stacking laundry.14 Further, the branched equation modeling of simultaneous accelerometry and heart rate monitoring should improve the estimation of energy expenditure for activities involving gross arm and leg movements (eg, weight lifting, cycling, and rowing), as the sole use of either accelerometry or heart rate telemetry has been shown to underestimate or overestimate energy expenditure during nonlocomotor or upper-body activities.15,16

To better understand the validity of commonly used self-report assessments of PA behavior such as the IPAQ and LTEQ, it is vital that researchers quantify whether these subjective inventories provide the same information as objective assessments of PA. With the latter in mind, the AHR represents one of the most valid and reliable methods currently available to researchers to assess PA in free-living conditions.17 Therefore, in the current study we sought to assess the concurrent validity and agreement between 2 commonly used self-report measures of physical activity and the AHR according to current public health recommendations.
Methods

Participants
Following institutional ethical approval, a convenient sample of 114 individuals (51 male; mean age 21.8 years; SD = 2.2) was recruited from a British University population via word of mouth, e-mail, and poster announcements. All participants were fully informed of the study protocol and provided their written informed consent. Figure 1 shows the flow of participants through the study.

Protocol
Participants visited the laboratory on 2 occasions, 8 days apart. During the first visit, participants were briefed on the study procedure and then informed consent was obtained before height, body mass, and waist circumference were measured. Participants were then fitted with an AHR unit to be worn for 8 consecutive days. The 8-day PA assessment period started and finished on the same day of the week for all participants. Following the 8-day period, participants returned to the laboratory and were randomly assigned to complete either the IPAQ or the LTEQ and a measure of social desirability, before the AHR unit was removed and data downloaded. In line with previous work utilizing the AHR monitor, the first 24-hr recording was deleted to reduce the effects of any potential reactivity. As such, the analyses and results represent 7 full days of observation. In attempt to reduce any potential reactivity, participants were blinded to the functionality of the AHR and the true objective of the research. Throughout the recruitment process and study duration, all participants were told that the research was a double-blind design and that the researchers themselves were blinded to the study objectives and functionality of the new equipment.

Anthropometric Measurements
All measurements were made using standard procedures with participants wearing light clothes and without shoes. Height was measured to the nearest 0.1 cm using a portable stadiometer (SECA Leicester height measure), body...
mass was measured to the nearest 0.05 kg using electronic scales (Omega 873, Seca Ltd—Medical Class), and waist circumference (WC) was measured to the nearest 0.5 cm using a 200cm flexible nonstretch nylon tape measure.

**International Physical Activity Questionnaire**

Participants completed the short-form IPAQ, which measures the frequency (days per week), and duration (minutes per day) of moderate- and vigorous-intensity PA, in bouts of at least 10 minutes during the past 7-day period, globally in all contexts of daily life. The IPAQ also assesses the time spent walking (≥10 minutes duration) during the past 7 days and this data were merged with the total moderate-intensity PA data. In addition, the IPAQ includes an item that gauges sedentary behavior by asking respondents to report the time spent sitting (sedentary activity) on a week day during the previous week.

**Leisure-Time Exercise Questionnaire**

Godin and Shephard developed the LTEQ as a simple self-report measure of habitual PA behavior in adults. Respondents were asked to indicate the frequency (number) of discrete 15-minute bouts of moderate (not exhausting; eg, fast walking, badminton), and vigorous intensity PA (heart beats rapidly; eg, jogging, soccer, hockey) during the last 7 days.

**Social Desirability**

To control for response bias and social desirability, all participants completed the 10-item shortened version of the Marlowe-Crowne Social Desirability scale (MCSD). In the current study, the alpha coefficient was $\alpha = .50$ which is similar to that reported by Strahan and Gerbasi ($\alpha = .57$).

**Physical Activity Monitor**

Participants wore a synchronized accelerometer and heart rate monitor (Actiheart, Cambridge Neurotechnology, UK) for a period of 8 days (24-hours a day). Due to the removal of the first 24-hr recording to account for any short-term reactivity, 7 days of minute-by-minute energy expenditure were used for analysis. AHR units were programmed to use the default group calibration equations as these have been shown to provide robust estimates of energy expenditure in comparison with indirect calorimetry in a sample similar to that used in the current study. In accordance with the manufacturer’s instructions, the AHR unit was attached to 2 ECG electrodes (Red Dot 2560, 3M) which were applied to the participant’s chest (cf.) and set to record in 60 s epochs. This lower positioning of the AHR monitor (below apex of sternum) has been shown to enhance the quality of the heart rate trace and provide excellent estimates of PA intensity and energy expenditure during treadmill walking, running, and free-living activity for both males and females. Although waterproof, the AHR was not primarily designed to measure PA during aquatic-based activities. Therefore, swimmers and water polo players were excluded during recruitment.

**Data Reduction and Extraction**

AHR data for each participant was downloaded to a computer using the AHR reader interface unit and software (version 2.226.001) and then exported into Microsoft Excel for further processing. Utilizing a branched-equation model, the AHR software estimates energy expenditure above rest. As such, sleeping heart rate was used to calibrate the AHR to the participants’ resting value (cf.). The quality of heart rate data were assessed using the AHR software which records the number (%) of ECG samples determined as ‘LOST’ (where a maximum of 128 samples/sec are possible) during calculation of the average heart rate per epoch. The mean (± SD) percentage of time that heart rate data were recorded for these participants was 94 ± 6% sampled at 60-second epochs over 24-hours for 7 days. Seven days of energy expenditure were assessed and using in-house software, data were extracted based on general population MET values (eg, moderate = 3.00 to 6.00 METS; vigorous ≥ 6.00 METS). In addition, the AHR estimates an individual’s physical activity level (PAL) which is the sum of total energy expenditure divided by resting energy expenditure over a 24-hr period. The World Health Organization advocates a PAL of 1.75 to indicate a person who is physically active and the PAL values are presented in the results to provide an indication of the participants’ general activity levels.

The IPAQ requires respondents to recall the time (minutes) spent in different intensities of PA (eg, moderate-, vigorous-intensity) for at least 10 minutes during the past 7 days. Therefore, to assess the level of agreement between the IPAQ and the AHR, data were extracted based on 3 thresholds: (i) time (minutes) spent in bouts of moderate-intensity PA (3.00-6.00 METs) for ≥ 10 minutes in duration; (ii) time (minutes) spent in bouts of vigorous-intensity PA (>6.00 METs) for ≥ 10 minutes in duration; and (iii) total PA assessed by time (minutes) spent in bouts of PA (≥3.00 METs) for ≥ 10 minutes in duration.

Sedentary behavior involves activities with very low energy expenditure (eg, 1.00 to 1.80 METs) which are usually performed in a sitting position. The IPAQ contains an item which assesses the time (hours/minutes) spent sitting on a week day in the last 7 days. Weekday traces for each participant in the IPAQ group (N = 46) were visually inspected to determine wake and sleep times. Using only data during waking hours over the working week (Monday to Friday), sedentary activity was determined as any activity during waking hours between 1.00 to 1.80 METs which according to the compendium of PA this encompasses activities such as sitting quietly watching television or listening to music (1.00 MET) to sitting at a desk writing or typing (1.80 METs).
were extracted for time spent in sedentary activity for
≥ 1 minute epochs between 1.00 to 1.80 METs during waking hours on weekdays.

The LTEQ requires respondents to recall the frequency (number) of discrete 15-minute bouts of PA of varying intensity (eg, moderate, vigorous) during the past week. Therefore, to assess the level of agreement between the LTEQ and the AHR, data were extracted based on 3 thresholds (i) frequency (number) of bouts of moderate-intensity PA (3.00–6.00 METs) for ≥ 15 minutes in duration per week; (ii) frequency (number) of bouts of vigorous-intensity PA (≥6.00 METS) for ≥ 15 minutes in length per week; and (iii) frequency (number) of bouts of PA (≥3.00 METs) for ≥ 15 minutes in duration per week.

**Data Analysis**

All analyses were conducted using SPSS 16.0. Inspection of the PA data revealed that it was not normally distributed and was therefore logarithmically transformed before analyses (cf.23). To ensure common metric between AHR and questionnaire output, data were analyzed using generic population MET values (eg, moderate = 3.00–6.00 METs; vigorous ≥26.00 METs).1 As such, association and agreement between the AHR and the 2 questionnaires (IPAQ and LTEQ) for estimates of time spent in, and frequency of different PA intensities (moderate, vigorous), were explored using a variety of approaches, including t tests, correlations, and the Bland-Altman Limits of Agreement (LOA) technique.24

Paired-samples t tests were conducted to determine if any mean differences existed between self-reported and objectively assessed PA data. To interpret Hedgesg values, effect sizes were interpreted as small if g = 0.2, moderate if g ≥ 0.5, or large if g ≥ 0.8. Pearson correlations were used to examine the associations between data from pairs of measures for the IPAQ [and LTEQ] questionnaire versus the AHR monitor for absolute time of PA bouts ≥ 10 minutes [or frequency of PA bouts ≥ 15 minutes] spent in moderate-, and vigorous-intensity activities, in addition to time spent (≥1 minute) in sedentary activity over 7 days. Spearman rank order correlations were used to explore the relationship between social desirability scores and self-reported PA. The alpha level was set at .05.

Agreement between the measures of PA were assessed using LOA which compares the differences between the mean and standard deviation obtained from the IPAQ/LTEQ with the AHR unit and 95% LOA were calculated as Mdiff ± (1.96 × SDdiff).24 Values closer to 0 indicate a greater level of agreement. However, scatter diagrams of the difference against the average of the 2 measurements (Bland-Altman plots) revealed there to be a positive relationship between the differences in measurement method (IPAQ/LTEQ vs. AHR) and the size of the measurement indicating heteroscedasticity (r’s = .50–.91, P < .01). In this situation, numerous researchers (eg,24,26) advocate logarithmic transformation of the raw data as the LOA would be wider apart than necessary for small values and narrower than they should be for large values. As such and before calculation of the LOA, a natural-log transformation was applied to both the mean of the and the difference between the 2 measures (IPAQ/LTEQ and AHR).10 The log-transformed LOA were then back-transformed (antilogged), providing values that can be interpreted in relation to the original scale, and allowing the outcomes to be expressed as the mean bias ± 2SD by the 95% agreement component (random error) on the ratio scale (see26 for a discussion on this analysis).

**Results**

**Sample Exclusion and Participant Descriptives**

To be included in the final analysis participants were required to have 7 full (24-hour) days of PA assessment with a minimally interrupted heart rate trace. Six participants (1 male) dropped-out of the study and 21 participants (14 female) were removed from the final analysis due to insufficient ECG trace quality (Figure 1). Complete data were analyzed for 87 British University students (43 male) and all subject characteristics can be found in Table 1. There were no significant differences between included and excluded participants (all P’s > .05) on any variable [eg, age, body mass index (BMI), WC, self-reported PA].

The majority of the sample (N = 67; 77%) were within the normal/healthy range for BMI and WC (cf.27). With regard to the physical activity level (PAL) of the total sample, participants were, on average, classified as being physically active (cf.27). Specifically, 56 participants (64%) were deemed as being physically active (ie, PAL ≥ 1.75), 18 participants’ activity level was between the criteria for limited activity to physically active (ie, PAL = 1.61–1.74), 3 participants were classified as achieving limited activity (ie, PAL = 1.55–1.60), while 10 participants were between the sedentary to limited activity range (ie, PAL = 1.40–1.54).

**Concurrent Estimates of PA: IPAQ vs. the AHR**

As shown in Table 2, IPAQ self-report estimates of time spent in vigorous-intensity PA were significantly higher than objectively-assessed estimates (137 ± 194 min-wk−1; t(45) = 4.78). In addition, IPAQ estimates of time spent in sedentary behavior were significantly lower than those captured by the AHR (–193 ± 251 min-wk−1; t(45) = –5.21). However, estimates of time spent in moderate-intensity PA (–5 ± 301 min-wk−1; t(45) = –0.12) and total PA (–52 ± 457 min-wk−1; t(45) = 0.78) were similar between the IPAQ and AHR. There was a moderate strength association between the IPAQ and the AHR estimates of time (min-wk−1) spent in vigorous-intensity PA (r = .63, N = 46, P < .001) but not for time spent in moderate-intensity PA (r = –.063, N = 46, P > .05), total
PA (r = .251, N = 46, P > .05), or sedentary activity (r = .182, N = 46, P > .05). As shown in Table 3, the mean bias between the IPAQ and AHR for moderate-intensity and total PA were relatively small. However, the B-A analysis showed that there was poor agreement between the AHR and IPAQ estimates of time spent in both moderate- and vigorous-intensity PA, in addition to total PA and time spent in sedentary behavior by revealing very large LOA between the AHR and IPAQ.

**Concurrent Estimates of PA: LTEQ vs. the AHR**

As shown in Table 2, self-reported (LTEQ) number of vigorous-intensity PA bouts (≥15 minutes) per week were significantly higher than objectively-derived frequency of vigorous-intensity PA ≥ 15 minutes per week (1.3 ± 2.1 bouts-wk⁻¹; r(40) = 3.86). However, self-reported estimates for the total number of PA bouts (>3.00 METs for ≥15 minutes) were significantly lower when compared with the AHR over the 7-day period (–2.5 ± 5.0 bouts-wk⁻¹; r(40) = 3.22). Although not statistically significant, estimates of the frequency of moderate-intensity PA bouts ≥ 15 minutes were similar between the LTEQ and AHR (–0.5 ± 4.2 bouts-wk⁻¹; r(40) = –0.70) (Table 2). There was a moderate strength association between the LTEQ and AHR estimates of the frequency (bouts-wk⁻¹) of vigorous-intensity PA (r = .441, N = 41, P < .01) and total PA bouts (≥3.00 METs) per week (r = .443, N = 41, P < .01) but not for the frequency of moderate-intensity PA bouts (r = .087, N = 41, P > .05). Similar to the IPAQ findings, although the mean bias between the LTEQ and AHR for frequency of moderate-intensity PA bouts was small, the B-A analysis showed that there was poor agreement (large LOA) between the LTEQ and AHR estimates of the frequency of both moderate- and vigorous-intensity PA bouts, in addition to the total number of PA bouts (Table 3).

---

**Table 1** Descriptive Characteristics of Participants

<table>
<thead>
<tr>
<th></th>
<th>Males (N = 43)</th>
<th>Females (N = 44)</th>
<th>All (N = 87)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>SD</td>
<td>Min-Max</td>
<td>Mean</td>
</tr>
<tr>
<td>Age (yr)</td>
<td>22.1</td>
<td>1.8</td>
<td>18.5–28.1</td>
</tr>
<tr>
<td>Body mass (kg)</td>
<td>76.4</td>
<td>11.3</td>
<td>61.1–117.2</td>
</tr>
<tr>
<td>Height (m)</td>
<td>1.79</td>
<td>0.07</td>
<td>1.63–1.96</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>23.8</td>
<td>3.4</td>
<td>19.0–39.6</td>
</tr>
<tr>
<td>WC (cm)</td>
<td>82.4</td>
<td>8.4</td>
<td>71.3–119.0</td>
</tr>
<tr>
<td>PAL</td>
<td>1.97</td>
<td>0.24</td>
<td>1.43–2.50</td>
</tr>
</tbody>
</table>

Abbreviations: BMI, body mass index; WC, waist circumference; PAL, physical activity level.

**Table 2** Comparison of Estimates for Weekly Duration (IPAQ) and Frequency (LTEQ) of Physical Activity in Different Intensities According to Questionnaire and AHR Data for the Whole Sample

<table>
<thead>
<tr>
<th></th>
<th>IPAQ (min-wk⁻¹)</th>
<th>AHR (min-wk⁻¹)</th>
<th>LTEQ (frequency-wk⁻¹)</th>
<th>AHR (frequency-wk⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
<td>P</td>
</tr>
<tr>
<td>Total sample (N = 46)</td>
<td>(N = 46)</td>
<td>(N = 41)</td>
<td>(N = 41)</td>
<td></td>
</tr>
<tr>
<td>Moderate (3.00-6.00 METs)</td>
<td>180</td>
<td>227</td>
<td>185</td>
<td>168</td>
</tr>
<tr>
<td>Vigorous (&gt;6.00 METs)</td>
<td>237</td>
<td>199</td>
<td>100</td>
<td>97</td>
</tr>
<tr>
<td>Total activitya (&gt;3.00 METs)</td>
<td>417</td>
<td>379</td>
<td>469</td>
<td>289</td>
</tr>
<tr>
<td>Sedentary activityb (1.00-1.80 METs)</td>
<td>426</td>
<td>171</td>
<td>619</td>
<td>221</td>
</tr>
</tbody>
</table>

Abbreviations: AHR, Actiheart; IPAQ, International Physical Activity Questionnaire; LTEQ, Leisure-Time Exercise Questionnaire.

a AHR data compared with IPAQ and LTEQ (moderate + vigorous values). To assess the agreement between the IPAQ (LTEQ) and AHR, data are presented as absolute duration ≥ 10 minutes [frequency of PA bouts ≥ 15 minutes] spent in the different PA intensities.

b The IPAQ assesses time spent sitting during a typical week day (Monday to Friday), therefore, AHR data were extracted for time spent in sedentary activity for ≥ 1 minute epochs between 1.00 to 1.80 METs during waking hours on weekdays and the units are minutes per day.

It is worthwhile noting that analyses of the raw untransformed data exhibited the same pattern of findings as the log-transformed data. As such, we can have confidence in the raw metric of our data (cf.37).
Table 3  Bland-Altman Analyses of the Duration and Frequency of Physical Activity in Different Intensities as Measured by the IPAQ and LTEQ Compared With Actiheart

<table>
<thead>
<tr>
<th>Method compared with AHR</th>
<th>Raw data</th>
<th>Logarithmic transformed data</th>
<th>Anti-logged data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean bias (SD)</td>
<td>95% LOA</td>
<td>Mean bias (SD)</td>
</tr>
<tr>
<td></td>
<td>(min·wk⁻¹) Upper</td>
<td>Lower</td>
<td>(min·wk⁻¹) Upper</td>
</tr>
<tr>
<td>IPAQ (N = 46)</td>
<td>Moderate (3.00–6.00 METs)</td>
<td>–5 (301)</td>
<td>585</td>
</tr>
<tr>
<td></td>
<td>Vigorous (&gt;6.00 METs)</td>
<td>137 (193)</td>
<td>515</td>
</tr>
<tr>
<td></td>
<td>Totala (&gt;3.00 METs)</td>
<td>–52 (457)</td>
<td>844</td>
</tr>
<tr>
<td></td>
<td>Sedentary activityb (1.00–1.80 METs)</td>
<td>–193 (251)</td>
<td>299</td>
</tr>
<tr>
<td></td>
<td>LTEQ (N = 41)</td>
<td>Moderate (3.00–6.00 METs)</td>
<td>–0.5 (4.2)</td>
</tr>
<tr>
<td></td>
<td>Vigorous (&gt;6.00 METs)</td>
<td>1.3 (2.1)</td>
<td>5.4</td>
</tr>
<tr>
<td></td>
<td>Totalb (&gt;3.00 METs)</td>
<td>–2.5 (5.0)</td>
<td>7.3</td>
</tr>
</tbody>
</table>

Abbreviations: AHR, Actiheart; IPAQ, International Physical Activity Questionnaire; LOA, limits of agreement; LTEQ, Leisure-Time Exercise Questionnaire.

a AHR data compared with IPAQ and LTEQ (moderate + vigorous values). Negative values indicate under-estimation of PA by IPAQ/LTEQ.
b The IPAQ assesses the time spent sitting during a typical week day (Monday to Friday), therefore, AHR data are presented as time (minutes per day) spent in sedentary activity during waking hours on weekdays.

Social Desirability. Bivariate correlations revealed that there was no significant relationship between total scores on the MCSD scale and self-reported time spent in moderate-intensity PA \( r = .20, N = 46, P > .05 \), vigorous-intensity PA \( r = .29, N = 46, P > .05 \), or time spent in sedentary activity \( r = .16, N = 46, P > .05 \) as measured by the IPAQ. There was also no significant relationship between total scores on the MCSD scale and self-reported frequency of moderate-intensity \( r = .13, N = 41, P > .05 \) or vigorous bouts of PA \( r = .28, N = 41, P > .05 \) as assessed by the LTEQ.

Discussion

We examined the agreement between 2 widely used self-report PA questionnaires (viz. IPAQ and LTEQ) and objectively-determined PA estimates captured using a combined accelerometer and heart rate monitor (AHR). According to current PA recommendations for improving and maintaining public health, objectively-observed PA was higher both for moderate-intensity \( r = .20, N = 46, P > .05 \) and vigorous-intensity \( r = .29, N = 46, P > .05 \) as measured by the IPAQ. The difference between the 2 questionnaires was small for the moderate-intensity and total PA raw data, the agreement between the 2 questionnaires compared with the AHR was still good. Using the IPAQ-AHR vigorous-intensity PA data as an example, the mean bias ratio was 0.39 and the agreement ratio was ×/± 5.18, which is 95% of the ratios for the sample should be contained between the values of 2.0 (0.39 × 5.18) and 0.07 (0.39 ÷ 5.18). To place the ratio LOA in a practical context, if an individual from the study sample wore an AHR monitor for 7 days which recorded that they participated in 120 min·wk⁻¹ of vigorous-intensity (>6.00 METs) PA of at least 10 minutes in duration; it is possible that the same subject completing the IPAQ questionnaire could obtain a self-reported estimate between 240 (120 × 2.0) min·wk⁻¹ and 9 (120 × 0.07) min·wk⁻¹ of vigorous-intensity (>6.00 METs) PA lasting 10 minutes or more. Similarly, using the time spent in sedentary behavior from the IPAQ-AHR data, the mean bias ratio was 1.52 and the agreement ratio was ×/± 1.68.

To our knowledge there are no published position statements for acceptable bias limits in PA measurement. However, it would seem that the LOA for the IPAQ and LTEQ compared with the AHR are too wide to be accepted for the accurate measurement of total PA and its subcomponents at the individual level. However, the mean scores between the self-report measures and the AHR were often comparable and hence these self-report measures may be more useful when used for group
comparisons. In agreement with previous research, this study found minimal evidence to suggest that social desirability influences self-reported PA in young adults.

An important consideration for interpreting the IPAQ/LTEQ-AHR agreement data is the potential for analytical error. The AHR is a precise and accurate instrument that estimates energy expenditure, but data then have to be extracted to categorize activities as ‘moderate’ or ‘vigorous’ intensity PA using general population MET values for public health. Therefore, some of the lack of agreement between the self-report PA questionnaires and the AHR may be derived from the way the AHR output is dissected postmeasurement. For example, we used the absolute MET values advocated in current public health recommendations. In a recent study that examined the commonality in the classification of an individual’s PA ‘status’ between 12 variations of PA recommendations (using different PA guidelines but the same raw AHR data), between 11 and 98% of the sample could be defined as meeting ‘recommended’ levels of activity. Therefore, even small variations in the expression and interpretation of PA recommendations will have an enormous impact on the interpretation of objectively-derived estimates of physical activity. In the context of the current study, it is possible that some of the lack of agreement between self-report and objectively measured physical activity is because current physical activity recommendations are anchored to rigid thresholds (eg, absolute intensity expressed in METs).

As a corollary, the PA measures used in the current study are all anchored by subjective perceptions of fitness as both the IPAQ and the LTEQ use physiological descriptors (eg, “makes you breathe much harder than normal” and “heart beats rapidly”) to help the respondent recall and code the correct amount of PA completed during the past 7 days. Due to individual differences in cardiovascular fitness and technique (or mechanical efficiency), some of the lack of agreement might be due to the fact that for some people a self-reported level of PA perceived and classified as ‘vigorous’ may be captured by the AHR as some other MET level (eg, ‘moderate-intensity’ PA). As such, self-reports of an individual’s perceived intensity of PA may not reflect the absolute intensity across all age and sex groups required for converting the perceived levels of PA intensity into METs. For example, a recent study validated the IPAQ short version against maximal oxygen uptake and found that 10% of young men had poor fitness yet reported very high PA on the IPAQ.

In addition, a marked proportion of individuals who reported high PA on the IPAQ had poor muscular fitness as indexed by scores on the 60-second push-up, sit-up and repeated squats tests. With this in mind, self-report PA instruments might be more useful when employed in conjunction with physiologic measures of cardiovascular fitness. Moreover, since it is the interplay between PA and health status that concerns most scientists and practitioners, future research should explore the predictive utility of these PA measures with health-related outcomes such as blood pressure, serum lipids and inflammatory markers (eg, C-reactive protein).

The B-A method does not indicate which of the 2 techniques is superior (IPAQ/LTEQ vs. AHR) but merely assesses agreement between methods. Therefore, it is important to consider the purpose of the PA measure in question. The primary purpose of the IPAQ short-form is population surveillance of PA at the individual level and it was not designed to be used as a precise evaluation tool in intervention studies. Similarly, the LTEQ was designed as a simple and easy-to-administer self-report PA measure. Eisenmann and colleagues highlighted that there is generally an inverse relationship between the validity and feasibility of PA assessment tools. For example, while self-report questionnaires are the most practical method in large epidemiological studies they will not be as valid as a criterion measure of energy expenditure, such as doubly labeled water or possibly AHR, which would be impractical and not currently cost-effective in large population-based studies. Nevertheless, our results suggest that these self-report PA questionnaires have limited utility when researchers and scientists are interested in assessing individual PA levels according to the current public health recommendations. However, based on the current analyses, it is not possible to elucidate whether this is because these self-report measures fail to capture the critical components of PA behavior or because of difficulties associated with the interpretation of objectively measured PA in the context of current physical activity recommendations.

Although PA is linked with a multitude of positive health benefits, antithetically, sedentary behavior is associated with the development of numerous chronic degenerative diseases. Previous work has documented that individuals tend to over-report PA and underestimate sedentary behavior such as television viewing. In the current study we also found that individuals underestimated the time spent in sedentary activity and although the cut-offs used in the current study (1.00–1.80 METs) might have captured standing activity this is unlikely to account for the large difference in objectively assessed and self-reported sedentary behavior. The assessment of sedentary behavior is an emerging field of research and future work should attempt to disentangle the relationship between PA, sedentary behavior and health outcomes, to affirm whether it is a lack of moderate-to-vigorous physical activity that is associated with poor health or whether it is prolonged sedentary behavior, or both.

Currently available methods of PA assessment have both strengths and limitations with respect to accuracy, cost-effectiveness, practicality and degree of invasive-ness. Self-report questionnaires are generally inexpensive, relatively easy to administer, have low participant burden and are useful tools for monitoring PA in large cohorts of free-living humans. However, self-report questionnaires provide fairly poor estimates of PA anchored to public health recommendations at the individual level, most likely due to recall bias and coding errors involving
individuals misclassifying the intensity, duration and frequency of PA bouts. Indeed, subjective measures of PA appear to quantify an individual’s perception of PA, rather than their actual PA, hence, current self-report methods for surveillance of the amount and intensity of PA may provide different estimates compared with objectively-assessed PA.35 While the AHR is not the panacea for PA assessment in free-living situations it may prove more useful when the precise quantification of PA is required, such as intervention evaluation. Although the AHR monitor is lightweight (<10g) and fairly unobtrusive, its utility is constrained by factors including cost, technical expertise required for fitting the monitor (e.g., skin preparation, positioning) and the time needed for more labor-intensive data analysis. Hence, choice of appropriate instrument will depend on the research objectives, study design, available resources and physiological characteristics of the target study population.

Numerous factors may contribute to the lack of agreement between self-reported and objectively assessed PA, including several limitations in the current study. Similar to past work utilizing convenient samples (e.g.,28,32), a potential limitation of the current work is that we recruited an available cohort of young, educated and reasonably active British university students. At a population level therefore, our sample of self-selected young adults introduces bias if we wish to extrapolate these results to the British population. A lack of sample heterogeneity limits the applications of the findings to wider populations and we do not know whether the agreement between self-reported and objective estimates of PA would be better, worse or the same in different or more heterogeneous populations. Although the cohort included a variety of young individuals, both males and females, with a wide range of body composition and PAL, future research using more diverse samples is needed to determine the tenability of the present findings to the general population.

Conclusion

The findings from this study suggest that the IPAQ and LTEQ questionnaires do not provide the same information as objectively-assessed PA according to current public health recommendations. Despite acceptable mean bias between the self-report measures and the AHR, the LOA between both the subcomponents (moderate- and vigorous-intensity activity) and total PA were large. With this in mind, future work should seek to clarify the source of bias between measures as it is currently unclear if the error is associated with the subjective perceptions of intensity from the PA questionnaires or the classification of METs using the AHR. Although the IPAQ and LTEQ questionnaires may provide an approximation of PA at a population level, these measures will not inform you whether a given individual is participating in the type, intensity, and amount of physical activity advocated in current public health recommendations.

References


