Estimating physical activity energy expenditure, sedentary time, and physical activity intensity by self-report in adults

Hervé Besson, Søren Brage, Rupert W Jakes, Ulf Ekelund, and Nicholas J Wareham

ABSTRACT

Background: Few questionnaires that assess usual physical activity have been reported to be valid for all different subdimensions of physical activity.

Objective: The objective was to assess the validity and reliability of the Recent Physical Activity Questionnaire (RPAQ), which assesses usual physical activity (PA) in 4 domains (work, travel, recreation, and domestic life).

Design: Total energy expenditure (TEE) was measured for 14 d by using the doubly labeled water technique combined with a measure of resting metabolic rate to yield PA energy expenditure (PAEE) in 25 men and 25 women. Simultaneously, intensity of activity was measured by using combined heart rate and movement sensing for 11 d. Repeatability of the RPAQ was assessed in an independent sample of 71 women and 60 men aged 31–57 y.

Results: Estimated TEE and PAEE were significantly associated with criterion measures (TEE: \(r = 0.67\); PAEE: \(r = 0.39\)) with mean (±SD) biases of \(-3452 ± 2025\) kJ/d and \(-13 ± 24\) kJ \(\cdot\) d\(^{-1}\) \(\cdot\) kg\(^{-1}\). The correlation between self-reported and measured time spent was significant for vigorous PA \((r = 0.70)\) and marginally insignificant for sedentary time \((r = 0.27, P = 0.06)\). The mean biases were relatively small for sedentary time and vigorous PA: \(0.7 ± 2.8\) h/d and \(-12 ± 24\) min/d, respectively. The intraclass correlation coefficient for repeatability of total PAEE (kJ/d) was 0.76 \((P < 0.0001)\).

Conclusion: The RPAQ is the first questionnaire with demonstrated validity for ranking individuals according to their time spent at vigorous-intensity activity and overall energy expenditure. Am J Clin Nutr 2010;91:106–14.

INTRODUCTION

Physical activity (PA) is a complex behavior that is difficult to assess accurately, partly because of its several dimensions. For practicality reasons, most large-scale epidemiologic studies rely on questionnaires (1), which are routinely used as surveillance tools (2, 3) or in etiological studies either focusing directly on activity or using activity as an important confounding factor. In this context, PA indexes may also be used to normalize self-reported dietary energy intake (4). A critical dimension of PA is overall energy expenditure (EE), but this has been proven difficult to quantify by simple questionnaire assessment (5). It is therefore important to validate and improve the accuracy of questionnaires aimed at assessing PA in a multidimensional manner. We previously developed the EPIC–Norfolk Physical Activity Questionnaire (EPAQ2) using the past year as the frame of reference. Like many other questionnaires of this type, its test-retest reliability is high, but its validity compared with an objective assessment of usual EE is somewhat limited (6), which may be explained by the relatively long reference period. Therefore, we redesigned the EPAQ2 questionnaire, and, as a consequence, we elected to shorten the frame of reference and developed the Recent Physical Activity Questionnaire (RPAQ), which is aimed at assessing usual PA in the past month. In the present study, we sought to examine the validity of RPAQ for estimating PAEE and its combination with a predicted value of resting EE (REE) for estimating total EE (TEE) against objective measures of EE (TEE) from the doubly labeled water (DLW) method (7, 8).

Although DLW is considered the gold standard for measuring total free-living EE (9), few questionnaires have been validated with this method as the criterion method (5). Thus, to date, there is limited knowledge on whether questionnaires can accurately estimate PAEE and TEE. In addition to estimating TEE, there is interest in assessing the absolute validity of questionnaires for estimating time spent at different intensity levels. For example, there is increasing interest in the separate and combined effects of vigorous- and moderately vigorous-intensity activity and sedentary time on health outcomes (10). We therefore investigated the validity of the present questionnaire for estimating the intensity of PA compared with an objective measurement of intensity by the validated combined accelerometry and heart rate monitoring (Acc+HR) approach (11). Because vigorous PA and sedentary time are both associated with health (10), it is crucial to validate a questionnaire that collects valid data for these 2 subdimensions. Consequently, the current study examined concomitantly the validity of the newly developed questionnaire for estimating PAEE, TEE, and time spent at different intensity levels. Finally, the RPAQ was completed twice, approximately 2 wk apart, to examine its repeatability.

1 From the Medical Research Council, Epidemiology Unit, Cambridge, United Kingdom.
2 Supported by the Medical Research Council and the Wellcome Trust.
3 Address correspondence to H Besson, MRC Epidemiology Unit, Institute of Metabolic Science, Box 285, Addenbrooke’s Hospital, Hills Road, Cambridge, United Kingdom. E-mail: herve.besson@mrc-epid.cam.ac.uk.
Received July 23, 2009. Accepted for publication October 1, 2009.

SUBJECTS AND METHODS

Validity study

Validity study population

Between February and June 2004, a total of 26 female and 25 male healthy volunteers aged 21–55 y living in the Cambridge (United Kingdom) area were recruited by advertisement. One female participant was excluded from the analysis because we were not able to calculate TEE from the urine samples provided. One female and one male participant did not provide Acc+HR data and were excluded from the intensity analyses. All participants completed a consent form after being informed of the procedures of the present study, which was approved by the Cambridge Local Research Ethics Committee.

Anthropometric characterization

The volunteers were asked not to eat or exercise 2 h before arrival at the test center. Height and weight were measured by standard anthropometric techniques. Body mass index (BMI; in kg/m²) was calculated as body weight divided by the square of height. Percentage body fat was assessed with the 4-compartment model (12) by using dual-energy X-ray absorptiometry (Lunar Prodigy, GE Healthcare, Piscataway, NJ) to measure bone mineral density, air-displacement plethysmography (BodPod; Life Measurement Inc, Concord, CA) to measure body density, and stable-isotope dilution (174 mg H 2 18O and 70 mg 2H2O/kg body weight) to measure total body water.

Resting energy expenditure

REE (kJ/min) was measured after the subjects fasted for ≥3 h for 45 min by indirect calorimetry with a ventilated hood (Oxycon Pro; Jaeger GmbH, Wurzburg, Germany). REE was averaged, disregarding the first 7 min, and then multiplied by [8 × 0.095 +16] × 60] to provide a 24-h average of REE (kJ/d) under the assumption that REE during 8 h/d is 5% lower than awake REE (13). REE was also predicted by using the Oxford equations (14).

Intensity of physical activity by combined acceleration and heart rate monitoring

Metabolic rate (J · kg⁻¹ · min⁻¹) and heart rate (HR) were measured during individual calibration tests. Briefly, after measurement of the resting metabolic rate (RMR), participants underwent a ramped treadmill test lasting ≤20.5 min if 90% of age-predicted HR was reached. PA intensity (PAI) was defined as total metabolic rate minus RMR and was expressed relative to body weight. HR and acceleration were measured in all tests with a combined sensor (Actiheart; Cambridge Neurotechnology Ltd, Papworth, United Kingdom) attached to standard electrocardiogram electrodes (Red Dot 2570; 3M, Loughborough, United Kingdom) at the level just below the apex of the sternum (11). Participants were asked to wear the combined movement and HR sensor for 24 h/d for the following 11 d (maximal memory capacity of the monitor). Monitors recorded minute-by-minute acceleration, trimmed average heart rate, 2 fastest and 2 slowest heart beats, average electrocardiogram voltage level, and the fraction of time during which the monitor firmware could not detect an HR between 30 and 250 bpm as described elsewhere (11).

HR data were processed by using a Bayesian approach as described in detail elsewhere (15). Accelerometry data were analyzed in their raw form, although segments of data with continuous zero acceleration lasting >60 min were treated as “monitor not worn” if also accompanied by an unphysiologic HR trace. PAI was predicted by using a branched equation model and individually calibrated HR (16) and expressed as multiples of predicted RMR (14). From this, the criterion intensity distribution was derived as the fraction of time spent at 22 levels of intensity (see details for intensity categories below).

Free-living PAEE by the DLW method

TEE was measured for 14 d by using the DLW method. The volunteers provided 2 baseline urine samples and thereafter drank a body weight–dependent dose (174 mg H 2 18O and 70 mg 2H2O/kg body weight) of DLW. Additional urine samples were collected every day (any void but the first in the morning) for the next 14 d. Participants were instructed to record the time of each sample on both the sampling bottle and a separate recording sheet and then to keep all samples refrigerated until collection by the research team at the end of the 14-d study period. At this time, a validated food-frequency questionnaire (FFQ) was administered (17). The disappearance of the 1H and 18O isotopes was analyzed in duplicate in all urine samples with an isotope ratio mass spectrometer as described in detail elsewhere (18). Carbon dioxide production was calculated according to the equation of Schoeller (19). TEE was calculated according to the Weir equation (20) by using the observed food quotient from the FFQ as a proxy for the respiratory quotient and was expressed in kJ/d. Diet-induced thermogenesis was calculated from the macronutrient composition of the diet (21) as assessed by the FFQ but normalized by DLW-based TEE. PAEE was calculated as TEE minus REE minus DIT and expressed in kJ · kg⁻¹ · d⁻¹. PA level (PAL) was calculated as the ratio of TEE to REE.

Reliability study

Reliability study population

A total of 60 male and 71 female healthy individuals aged 31–57 y completed the RPAQ twice, approximately 2 wk apart during the months of October and November 2006. These volunteers are participants of the ongoing Medical Research Council (MRC) Fenland Study, which is a population-based cohort study enrolling individuals born between 1950 and 1975 registered at surgeries in East Cambridgeshire.

The RPAQ

At the end of the Acc+HR and DLW measurements, participants reported their last month PA by using the RPAQ (see Appendix A under “Supplemental data” in the online issue). The RPAQ is a slightly shorter and further developed version of the Second EPAQ (EPAQ2) (6). Apart from a different time frame, the RPAQ is similar to the EPAQ2 in that data on self-reported PA behavior is collected in a disaggregated way such that the information may be summarized according to the dimension of PA of interest. RPAQ contains questions about PA in 4 main domains: activity at home, during work, during transport, and during leisure time. A specificity of this questionnaire is that information related to housework activity was not collected.
This was guided by validation of the EPAQ2, in which house-
work was inversely correlated with objective measurements of
PA (6). In each domain, questions are closed rather than open-
ended to make them easy to complete and to facilitate large-
scale data entry. Although this questionnaire has closed questions
with ordered categories of continuous variables that the volun-
tee has to select from, the selection of these categories was
determined by the range of responses in early versions in which
questions were open-framed. The questions in the leisure-time
domain, adapted from the Minnesota Leisure Time Activity
questionnaire (22, 23), ask about activities frequently perfo-
rmed in the UK population (24). Activities are further described by the
intensity. In the case of running, the more vigorous option (eg,
competitive running) is asked before the less vigorous (jogging).
The categorization of occupational activity was derived from the
Modified Tecumseh Occupational Activity questionnaire that
has been validated elsewhere (25).

**Questionnaire calculations**

Before calculating self-reported PAEE, the following assump-
tions were made:

1) All individuals reporting they had a job were assigned
37.5 h work/wk because no information on the number of
working hours was obtained.

2) Distance from home to work was calculated by entry of
work and home addresses in an Internet router (26). For
individuals reporting walking or cycling between home and
work but not reporting the address of their work, dis-
tances of 2 and 3 miles (1 mile = 1609 m) were assumed
(corresponding to the mean distance for people reporting
walking or cycling between home and work in the current
sample).

Estimates of PAEE for the 4 different domains (home, work,
travel, and leisure time) were calculated by multiplying partici-
pation (h/d) by the metabolic cost of each activity, expressed in
metabolic equivalents (METs) obtained from the PA compen-
dium (27, 28). Energy expenditure for transportation was cal-
culated by multiplying the energy cost of the mode by the
distance from home to work and by the number of trips per week
divided by the speed of the transport. This product was weighted
by a frequency coefficient of use of the given transport:

\[
\text{Energy expenditure for transport (MET-h/d)} = \sum k_i n_i d_i s_i^{\frac{1}{2}} c_i + \sum k_j n_j d_j s_j^{\frac{1}{3}} c_j + \sum k_w n_w d_w s_w^{\frac{1}{2}} c_w
\]  

(1)

where \( k \) is the frequency coefficient of use \((k < 1)\), \( n \) is the
number of trips per week, \( d \) is distance (miles), \( s_i \) is 45
miles/h, \( s_j \) is 10 miles/h, \( s_w \) is 3 miles/h, \( c_i \) is 2 METs
(code 16010), \( c_j \) is 8 METs (code 01015), \( c_w \) is 3.3 METs
(code 17190), \( i \) is motorized vehicle, \( j \) is cycle, and \( w \) is walk.
The frequency coefficient was normalized by dividing the reported
frequency coefficient by the summation of the frequency coef-
ficients reported by the participant.

It was assumed that the energetic equivalent of 1 L oxygen is
equal to 20.3 kJ (29) and that 3.5 mL O\textsubscript{2} \cdot kg\textsuperscript{-1} \cdot min\textsuperscript{-1} cor-
responds to 1 MET in the PA compendium (27). For each ac-
tivity, PAEE was derived by subtracting the estimated RMR (14)
weighted by the daily duration of the activity:

PAEE for a given activity (kJ/d) = score of the activity \([\text{MET-h/d}} \times \text{weight (kg)} \times (3.5 \times 20.3 \times 60/10000) - (\text{RMR predicted} \\
\times \text{duration of the activity/24}) \]  

(2)

Total PAEE was calculated by summing up PAEE assessed for
each domain. A second score for total PAEE was also calculated
by adding time not accounted for by the EPAQ, assuming 8 h
sleep/d.

Time not accounted for by the EPAQ (h) = 16 – \( \text{duration}_w \\
+ \text{duration}_t + \text{duration}_r + \text{duration}_r \) 

(3)

where \( w \) is work, \( t \) is transport, \( h \) is home, and \( r \) is recreation. For
individuals reporting their main mode of transportation when
“getting about” as walking or bicycling \((n = 24)\), a 1.2 MET
score was assigned to the time not accounted for by the self-
report. This MET score \((1.2)\) is the threshold between resting
and sedentary activities (27).

Estimates of time (h/d) spent for sedentary \(<2\) MET), light
\((2–3.5\) METs), moderate \((3.6–6\) METs), and vigorous \((>6
METs)\) PA were calculated by using 1 MET = RMR\textsubscript{predicted}. To
evaluate the assessment of activity intensity distribution by
EPAQ, time spent at 22 different intensities was assessed. We
assumed 8 h sleep/d and for the time not accounted for by the
EPAQ we assigned a 1.2 MET score for individuals reporting
their main mode of transportation when “getting about” as
walking or bicycling. Time spent at different intensities was
categorized as follows: 1 category <1.25 MET, 7 categories
in 0.25-MET wide bins in the 1.25–3 MET interval, 6 categories
in 0.5-MET wide bins in the 3–6 MET interval, 7 categories
in 1-MET wide bins in the 6–13 MET interval, and >13 METs.
(See Appendix B under “Supplemental data” in the online issue
for further details.)

**Statistical analysis**

**Validity study**

Descriptive results are reported as mean ± SD. Differences
between sexes in anthropometric measures and main indexes of
PA from the 3 methods were compared by using a Kruskal-
Wallis test.

Associations between objectively measured TEE and PAEE
per kilogram body weight with self-reported EE estimates by
domain of activity and all domains combined were assessed by
using Spearman correlation coefficients. Multiple rank regression
analyses were also performed by domain to assess Spearman
partial correlation, adjusting for PAEE from the other domains.

Absolute validity of the EPAQ for estimating PAEE and the
combined ability of EPAQ and REE prediction to estimate TEE
were assessed by the degree of agreement with the DLW esti-
mates according to the Bland-Altman technique. A mean bias
was defined as a significant mean difference obtained by sub-
tracting the EE measured by DLW from the EE estimated by
The error was defined as 2 SD of the mean bias (30, 31). In the Bland-Altman plots, the x axis represents measures for the DLW, and the y axis represents the difference between measures from RPAQ and DLW. The 95% limits of agreement as well as the trend line were plotted by using the slope and the intercept obtained by regressing on TEE or PAEE the differences between the self-reported estimates and the DLW-measured values.

Associations between self-reported and objectively measured subdomains of PA (ie, sedentary time and times spent at light, moderate, and vigorous PA) were assessed by using Spearman correlation coefficients. For these 4 intensity levels, a mean bias was calculated by averaging the difference obtained by subtracting the time measured by combined Acc+HR from the time estimated by RPAQ. Average time estimates spent at 22 different intensity levels according to the RPAQ and combined Acc+HR were plotted. Average differences between estimated and measured times spent for these 22 different intensity levels were also plotted. Spearman partial correlation between measured and self-reported time spent was assessed across the 22 different intensity levels by running a rank regression analysis adjusted for the intensity.

Reliability study

The repeatability analysis was based on data from the 131 individuals completing the RPAQ twice, approximately 2 wk apart. To test the reliability of the questionnaire, we compared scores for total PA, scores by each domain of activity (home, work, transport, and leisure time), and scores by time spent at different intensity levels (sedentary, light, moderate, and vigorous activities). The repeatability of the RPAQ was described by using intraclass correlation coefficients (ICCs). ICCs were calculated by using log-transformed variables, and zero values were converted to the value 10⁻⁹. Multiple linear regression analyses with their corresponding interactions terms were used to test for heterogeneity by sex, age, and elapsed time between the 2 completions of the questionnaire. The Bland-Altman method was used to illustrate the degree of agreement between time points for self-reported overall PA. Analyses were conducted by using Statistical Analysis System software version 9.1 (32).

RESULTS

Validity study

Age, main anthropometric characteristics, and objective assessment of PAL are shown by sex in Table 1. BMI was not significantly different between men and women, but percentage body fat was significantly lower (P < 0.0001) in men than in women. PAL ranged from 1.34 to 2.62 and did not differ between men and women.

Distributions of objective measures of TEE, PAEE, and time spent at 4 different intensity levels (ie, sedentary, light, moderate, and vigorous) as well as the mean biases of self-reported measures are displayed in Table 2. Self-reported TEE was significantly correlated with objectively measured TEE (r = 0.67, P < 0.0001), but the RPAQ significantly underestimated TEE (mean bias = −3451.9 ± 2025.1 kJ/d). After assumptions of 8 h of sleep and assigning a 1 or 1.2 MET value to the time not accounted for by the questionnaire, the correlation between self-reported and measured TEE improved (r = 0.78, P < 0.0001), and the underestimation of TEE from the RPAQ was approximately one-third of that calculated by using the original method.

The RPAQ significantly underestimated PAEE (mean bias = −12.9 ± 23.9 kJ · d⁻¹ · kg⁻¹), but self-reported PAEE was significantly correlated with objectively measured PAEE (r = 0.39, P = 0.0004). After assumptions of 8 h of sleep and assigning a 1 or 1.2 MET value to the time not accounted for by the questionnaire, the correlation between self-reported and measured PAEE improved (r = 0.46, P = 0.0007), and the underestimation of PAEE from the RPAQ was reduced to −5.2 ± 23.3 kJ · d⁻¹ · kg⁻¹ (P = 0.12). Self-reported transport-related PAEE and recreational PA were significantly correlated with PAEE measured by DLW [r = 0.29 (P = 0.04) and r = 0.44 (P = 0.001)], respectively (data not shown). In contrast, self-reported PAEE for the home and work domains were not significantly correlated with measured PAEE (data not shown). Interaction terms with sex, age, and BMI were not significant in any of the regression models (data not shown).

The questionnaire overestimated the time spent sedentary by only 0.7 ± 2.8 h/d, and the correlation between self-reported and objectively measured sedentary time was marginally insignificant (r = 0.27, P = 0.06). Self-reported duration of vigorous PA strongly correlated with objectively measured time spent at vigorous intensity (r = 0.70, P < 0.0001), and the mean bias was only 0.2 ± 0.4 h/d. No significant correlations were observed between self-reported and objectively measured time spent at light- and moderate-intensity PA. After correcting the intensity of each activity relative to the participant’s estimated RMR, the correlations between self-reported and objectively measured times spent at different intensity levels did not improve.

Bland-Altman plots illustrating the difference between self-reported TEE estimates and that measured by DLW plotted against TEE measured by DLW are shown in Figure 1. Trends (proportional error) in the scatter plots indicated that the underestimation from the RPAQ increased with increasing amounts of TEE. The correlation between the difference in the methods and the measured values was significant for the initial scoring of TEE (r = −0.38, P = 0.007) but was marginally significant for the scoring of TEE including time not accounted by the RPAQ (r = −0.29, P = 0.04).

Bland-Altman plots illustrating the difference between self-reported PAEE estimates and that measured by DLW plotted against PAEE measured by DLW are shown in Figure 2. Trends (proportional error) in the scatter plots indicated that the underestimation from the RPAQ increased with increasing levels of PA. The correlations between the difference in the methods and

<table>
<thead>
<tr>
<th>TABLE 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristics of the validity study population (n = 50)¹</td>
</tr>
<tr>
<td>Men</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>Age (y)</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
</tr>
<tr>
<td>Percentage body fat</td>
</tr>
<tr>
<td>Physical activity level³</td>
</tr>
</tbody>
</table>

¹ All values are means ± SDs.
² Obtained by using the Kruskal-Wallis test.
³ Measurement by the doubly labeled water method.

Downloaded from https://academic.oup.com/ajcn/article-abstract/91/1/106/4597031 on 17 April 2018
the measured values were significant for both the initial scoring of PAEE ($r = -0.57$, $P < 0.0001$) and the scoring of PAEE including time not accounted by the RPAQ ($r = -0.54$, $P < 0.0001$).

We thereafter performed analyses across 22 different intensity levels. The relation between measured and self-reported duration across the 22 different intensity levels assessed by running a rank regression model adjusted for intensity was not significant ($P = 0.17$). The duration of time spent at these 22 intensity levels by self-report and combined sensing, respectively, is shown in Figure 3. The differences (mean ± SD) between self-reported and measured time spent at these 22 different intensity levels are shown in Figure 4. Mean (SD) Kullback Leibler distance between intensity distributions assessed by RPAQ and combined sensing was 6.5 (±1.6) MET-h.

### Reliability study

The age of the participants ranged from 31 to 57 y and was not significantly different between men and women. The interval of time between the 2 completions of the questionnaire was longer for men than for women (16.4 ± 5.9 compared with 14.3 ± 3.7 d; $P = 0.02$).

The Bland-Altman plot illustrating the difference between the 2 administrations of RPAQ is shown in Figure 5. A description
of PA self-reported from the 2 administrations of the RPAQ and their ICCs are shown in Table 3. The ICC for total PAEE (kJ/d) was 0.76 ($P < 0.0001$), and the highest levels of repeatability in any subdomain of activity were observed for work-related (ICC = 0.85, $P < 0.0001$) and recreational (ICC = 0.66, $P < 0.0001$) domains. PAEE calculated based on activities performed at home was not reproducible (ICC = 0.11, $P = 0.10$). Transport-related PAEE had a low level of repeatability (ICC = 0.32, $P < 0.0001$), and for this domain reliability interacted negatively ($P = 0.03$) with time elapsed between the 2 completions of the questionnaire. For the other domains and all domains combined, no heterogeneity in reliability was observed with increasing duration of time between completions. Repeatability results based on PA expressed either in MET-h or in hours were highly significant across and within domains (data not shown). The repeatability for time spent at different intensity levels was very similar across intensity levels, with ICCs ranging from 0.74 ($P < 0.0001$) for moderate to 0.86 ($P < 0.0001$) for light PA. When age and sex were included in the multilinear models, the effects of these variables were not significant, and no significant interaction between activity and sex were observed (data not shown).

DISCUSSION

The results from this study, based on a sample of active and inactive men and women, suggests that the RPAQ is a valid instrument for ranking individuals according to TEE, PAEE, sedentary time, and time spent at vigorous-intensity PA. The results suggest that the RPAQ may be used to estimate absolute TEE, PAEE, and time spent sedentary and in vigorous PA in groups of healthy adults. The estimation of EE, however, required some assumptions. This study was unique because it used 2 different gold standard methods (DLW and combined heart rate and movement sensing) to report on the validity of a PA questionnaire for the estimation of PA by domain, TEE, PAEE, and time spent sedentary and at other intensity levels.

Several questionnaires have been examined for their validity to accurately assess TEE, but according to a recent review (5) only 4 reported correlations $>0.60$ against DLW measurements. Previous questionnaires validated against measurements of PAEE by DLW (5, 33, 34) have failed to report evidence of both absolute validity and high correlation coefficients. One of these studies reported high absolute validity, but was limited in sample size ($n = 9$) and did not report whether the instrument was valid for ranking individuals (34). In contrast, others have reported correlations coefficients between self-reported and objectively measured PA $>0.50$ (33, 35, 36); however, none of these studies measured PAEE and were therefore not able to examine the absolute validity of their instruments.

The literature review performed by Neilson et al (5) suggests that most questionnaires underestimate PAEE with an increased and systematic underestimation for questionnaires with a time frame of $>7$ d (5). The RPAQ also underestimates the PAEE, but to a lesser extent than other questionnaires of similar time frame. Furthermore, the RPAQ underestimated both TEE and PAEE, suggesting that EE estimated from the questionnaire is not mainly explained by the body weight of the participant, which could have been the case if the questionnaire had simultaneously...
underestimated PAEE and overestimated TEE. In parallel, whereas all of the sections of the RPAQ essentially report on nonsedentary activities, the home section is mainly based on reporting sedentary activities.

The RPAQ did not provide valid assessments of the absolute times spent at light- or moderate-intensity activities. This is consistent with previous results derived from questionnaires with similar time frames (6, 37) and by questionnaires with shorter time frames (38–40). Corroborating the present results, most previous validation studies have reported higher validity for time spent at vigorous intensity than at lower intensity levels (6, 37–40). The high correlation coefficient observed between self-reported and objectively measured vigorous PA in the present study is in agreement or higher than those previously reported.

Consistent with previous studies investigating reproducibility of similar questionnaires (6, 37), the present study reported high repeatability results across domains of activity and within home-related, work-related, and recreational activities. The transport domain was less reliable than other sections of the questionnaire, but the reliability increased with decreasing time between the 2 administrations of the questionnaire, suggesting that the PA undertaken for transportation may truly have changed. In addition, PA was expressed in kJ. This may have decreased the reliability level compared with results from studies expressing PA in hours or MET-h when assessing test-retest reliability.

The repeatability was consistently high across intensity levels of PA. This is encouraging because most previous studies have reported very low repeatability for light and moderate PA and acceptable repeatability for vigorous PA (38, 41, 42). The concepts of light, moderate, and vigorous PA are unfamiliar to many participants, and the high reproducibility of each intensity level for the RPAQ might have been due to the structure of the questionnaire by domains rather than by intensity levels. This is supported by other questionnaires with acceptable repeatability for these intensity levels also being structured by domains of activity (37, 42).

The strengths of our study include the use of 2 different criterion methods for measuring EE and time spent in different intensity levels. These methods, the DLW method for measuring EE and combined heart rate and movement sensing for measuring time spent at different intensity levels, are probably the most accurate criterion methods available. By using these methods in a combined manner, all the light activities entailing low EE and low level intensity body movement are captured.

Moreover, most previous studies have assessed the validity of shorter (usually 7 d) PA questionnaires. The results of the present study are therefore encouraging given the longer time frame of the RPAQ aimed at capturing usual PA. The associations reported here are stronger than those previously reported for similar instruments using a 1-y time frame (6, 37). The RPAQ captures both intensity and EE and also information about different domains of PA.

The following limitations must, however, be considered when interpreting the results from our study. First, we do not know the extent to which the educational level of participants may influence the validity of the instrument because data on education or socioeconomic level were not available. In addition, the absence of interaction with BMI, age, and sex may be the results of inadequate statistical power. Moreover, the age range of the participants was narrow, which limits the generalizability of the results to older populations.

Second, because the home domain asks questions mainly on sedentary behaviors, this section of the RPAQ may not be appropriately constructed to calculate PAEE following the principle of summing the EE of each activity. To compensate for the lack of information on the other activities usually performed at home, we developed and imputed a MET score for the remaining time of the day not covered by the questionnaire. This imputation was a priori based on the self-reported mode of transportation and justified because individuals who reported walking and bicycling as their main mode of transportation had a significantly higher PAEE (measured by DLW) than did those who reported motorized transportation as their main mode of transportation. Third, compared with other individuals, those with a self-reported estimated sedentary time >18 h/d were characterized by a lower total daily time covered by the questionnaire. This suggests that the questionnaire does not cover all PAs performed, and the lack of information related to these activities may have contributed to the overestimation of sedentary time for some of our participants.

There was a tendency for underestimation of self-reported EE with increasing levels of EE. This phenomenon was probably due to the inability of the questionnaire to assess all types of activities, which is likely to have a pronounced effect in active individuals.

TABLE 3
Characteristics of the repeatability study population and intraclass correlation coefficients between the 2 administrations of the Recent Physical Activity Questionnaire (RPAQ) (n = 131)

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>Follow-up</th>
<th>Intraclass correlation coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAEE (kJ/d)</td>
<td>9823 ± 4703</td>
<td>9128 ± 4453</td>
<td>0.76²</td>
</tr>
<tr>
<td>PAEE at home (kJ/d)</td>
<td>1884 ± 786</td>
<td>1896 ± 868</td>
<td>0.62²</td>
</tr>
<tr>
<td>PAEE at work (kJ/d)</td>
<td>3671 ± 3064</td>
<td>3497 ± 2882</td>
<td>0.85²</td>
</tr>
<tr>
<td>PAEE for transport (kJ/d)</td>
<td>504 ± 666</td>
<td>370 ± 813</td>
<td>0.32²</td>
</tr>
<tr>
<td>PAEE for recreations (kJ/d)</td>
<td>3764 ± 3018</td>
<td>3365 ± 2848</td>
<td>0.66²</td>
</tr>
<tr>
<td>Sedentary time, &lt;2 METs (h/d)</td>
<td>16.3 ± 3.8</td>
<td>16.8 ± 3.7</td>
<td>0.76²</td>
</tr>
<tr>
<td>Light PA, 2–3.5 METs (h/d)</td>
<td>5.2 ± 3.3</td>
<td>5.1 ± 3.2</td>
<td>0.86²</td>
</tr>
<tr>
<td>Moderate PA, 3.6–6 METs (h/d)</td>
<td>2.0 ± 1.3</td>
<td>1.7 ± 1.4</td>
<td>0.74²</td>
</tr>
<tr>
<td>Vigorous PA, &gt;6 METs (h/d)</td>
<td>0.5 ± 1.4</td>
<td>0.4 ± 1.1</td>
<td>0.81²</td>
</tr>
</tbody>
</table>

1 METs, metabolic equivalent tasks; PA, physical activity; PAEE, physical activity energy expenditure.
2 P < 0.001.
The observed tendency for heteroscedasticity may also have been due to minimization of the social desirability response bias by the home section of the questionnaire, because sedentary individuals are more prone to this type of reporting bias (43). Finally, it has been suggested that self-reports of recreational activities are significantly lower compared with interviewer assessments of the same activities (44). The most active individuals being more frequently engaged in leisure activities, this could have led to underestimation of their EE.

In conclusion, the RPAQ appears to be valid for estimating TEE, PAEE, and time spent in vigorous PA in healthy adults and for ranking individuals according to these 3 indexes. RPAQ could therefore be a very convenient tool for describing PALs in Western populations. The RPAQ also provides a description of activity performed at home, at work, and in leisure time. This easy and quickly self-administered questionnaire seems to be feasible for use in large-scale epidemiologic studies. The use of the RPAQ in future research will test its descriptive ability in surveillance studies and its predictive ability in intervention and etiologic studies.

We acknowledge Susie Hennings, Rosalind Stevenson, Emanuella de Lucia-Rolfe (MRC Epidemiology Unit, Cambridge), Samantha Northcott, and Peter R Murgatroyd (Wellcome Trust Clinical Research Facility, Addenbrooke’s Hospital, Cambridge, United Kingdom) for assistance with the data collection and equipment checks; Andy Coward, Antony Wright, and Cheryl Kidney (MRC Collaborative Centre for Human Nutrition Research, Cambridge, United Kingdom) for stable-isotope preparation and analysis; and Jian’an Luan (MRC Epidemiology Unit, Cambridge, United Kingdom) for statistical assistance.

The authors’ responsibilities were as follows—HB: performed the data analyses and drafted the manuscript; SB, UE, and NJW: provided critical input for the conception of this manuscript and on the data analyses and assisted with editing the manuscript; SB (Principal Investigator): was responsible for the overall design of this validation study; and RWJ: designed the questionnaire. All authors took part in the discussion of the results and approved the final version of the manuscript. None of the authors had any conflicts of interest.

REFERENCES


