

# International Physical Activity Questionnaire: Validity against Fitness

MIKAEL FOGELHOLM<sup>1,2</sup>, JARMO MALMBERG<sup>1</sup>, JAANA SUNI<sup>1</sup>, MATTI SANTTILA<sup>3</sup>, HEIKKI KYRÖLÄINEN<sup>4</sup>, MATTI MÄNTYSAARI<sup>5</sup>, and PEKKA OJA<sup>1</sup>

<sup>1</sup>The UKK Institute for Health Promotion Research, Tampere, FINLAND; <sup>2</sup>Research Unit, Pirkanmaa Hospital District, Tampere, FINLAND; <sup>3</sup>The Training Division of the Defence Staff, Finnish Defence Forces, Helsinki, FINLAND; <sup>4</sup>Department of Biology of Physical Activity, University of Jyväskylä, FINLAND; and <sup>5</sup>Research Institute of Military Medicine, Central Military Hospital, Helsinki, FINLAND

## ABSTRACT

FOGELHOLM M., J. MALMBERG, J. SUNI, M. SANTTILA, H. KYRÖLÄINEN, M. MÄNTYSAARI, and P. OJA. International Physical Activity Questionnaire: Validity against Fitness. *Med. Sci. Sports Exerc.*, Vol. 38, No. 4, pp. 753–760, 2006. **Purpose:** This study was conducted to criterion-validate the short format of International Physical Activity Questionnaire (IPAQ) against health-related fitness. **Methods:** Participants included 951 men, aged 21–43 yr.  $\dot{V}O_{2\max}$  by ergometer was used to measure cardiorespiratory fitness. Muscular fitness tests included the number of sit-ups, push-ups, and squats performed during 60 s. Physical activity was assessed from IPAQ by categories for health (health-enhancing physical activity (HEPA)) and by calculating total and vigorous MET-hours per week. The participants were classified in five groups of almost equal size by the IPAQ outcomes. Moreover, the frequency of vigorous activity was assessed by one precoded question (single-item question on leisure-time vigorous physical activity (SIVAQ)). **Results:** The mean fitness improved from the first through the fourth IPAQ group.  $\dot{V}O_{2\max}$  in the most active IPAQ group, however, was lower than in the fourth group ( $P < 0.05$ ), both with HEPA categories and vigorous METs as the outcome. In contrast, the weekly frequency of vigorous physical activity showed a positive association with fitness through all six categories. The 65 sedentary (according to SIVAQ) and unfit individuals who belonged to the highest 20% (by IPAQ total METs) were older ( $30.6 \pm 5.1$  vs  $28.9 \pm 3.8$ ,  $P = 0.02$ ), more obese (waist  $93 \pm 12$  vs  $89 \pm 10$  cm), more often current smokers ( $59.4$  vs  $42.0\%$ ;  $P = 0.03$ ), and less educated ( $<12$  yr of education  $70.8$  vs  $52.0\%$ ;  $P = 0.02$ ), compared with the others in the highest IPAQ group. **Conclusions:** Almost 10% of young men had poor fitness and apparently low physical activity, but they reported very high physical activity by IPAQ. An evident need therefore exists to develop IPAQ further to solve the apparent overreporting by a considerable proportion of sedentary individuals. **Key Words:** PHYSICAL ACTIVITY ASSESSMENT, SURVEILLANCE, CARDIORESPIRATORY FITNESS, MUSCULAR FITNESS

Because of the growing concern of decreasing physical activities all around the world, an evident need is seen for comparable data on the amount of physical activity in different populations. The International Physical Activity Questionnaire (IPAQ) was designed by a multinational working group to provide common instruments that can be used internationally to obtain comparable population estimates of health-enhancing physical activity from surveillance system data (7). IPAQ has two versions, namely a short (9 items) and a long (31 items) format. Both formats are designed to assess physical activity during the last 7 d or during a “typical week,” and they can be administered by telephone interview or self-administered.

The initial validation of the two IPAQ formats showed good test–retest repeatability (repeatability coefficient for the pooled data from all 14 study centers was 0.81 for the long and 0.76 for the short format) and reasonably good between-format comparability (concurrent validity, the pooled coefficient for comparison between the long and short formats was 0.67) when assessed by Spearman correlation coefficients (7). The criterion validity against accelerometer was modest, with pooled coefficients of 0.33 and 0.30 for the long and short formats, respectively. Reliability and criterion validity against accelerometer for a computerized version of the IPAQ long format were comparable to those reported earlier for the “paper and pencil” format (21).

Some recent reports have been concerned about the validity of IPAQ. The European Physical Activity Surveillance System found poor comparability between traditional physical activity indicators (various single-item survey questions on the frequency of physical activity) used in Europe and the short format of IPAQ (17). In the same paper, the national indicators seemed to predict perceived health better than IPAQ. In a Brazilian study, the limits of agreement between the long and short formats (concurrent validity) of IPAQ were not good: standard deviation (SD)

Address for correspondence: Mikael Fogelholm, The UKK Institute for Health Promotion Research, POB 30, 33501 Tampere, Finland; E-mail: mikael.fogelholm@uta.fi.

Submitted for publication June 2005.

Accepted for publication October 2005.

0195-9131/06/3804-0753/0

MEDICINE & SCIENCE IN SPORTS & EXERCISE,

Copyright © 2006 by the American College of Sports Medicine

DOI: 10.1249/01.mss.0000194075.16960.20

for the difference for time being at least moderately physical active was  $629 \text{ min}\cdot\text{wk}^{-1}$ , with extreme individual differences amounting to more than  $3000 \text{ min}\cdot\text{wk}^{-1}$  (9). Moreover, the long format seemed to produce much higher estimations for physical activity. Two studies suggested that even the IPAQ short format may overestimate total physical activity, at least when compared against a more detailed (probe protocol) report (18) or against two other survey tools (6). Evidently, more research on the validity of IPAQ is warranted.

To our best knowledge, IPAQ has not been criterion-validated against physical fitness or health-related fitness. Health-related fitness consists of cardiorespiratory, metabolic, muscular, motor, and morphological components (5). Habitual physical activity can improve health-related fitness, which in turn is associated with lower mortality and morbidity. IPAQ has been designed to measure particularly activities related to cardiorespiratory fitness (CRF) (7). Therefore, if IPAQ is validated against fitness, maximal oxygen uptake ( $\dot{V}O_{2\text{max}}$ ) may be regarded as the primary criterion variable.

The main objective of the present study was to validate the self-administered IPAQ short format against CRF in young, adult men. Although IPAQ has been mainly designed to assess physical activities related to CRF, several of the activity examples given in the questionnaire are also related to muscular fitness (e.g., heavy lifting, digging, carrying light loads). A secondary objective, therefore, was to criterion-validate the results of IPAQ against tests of muscular fitness.

## METHODS

**Participants.** The participants were healthy Finnish men, who were invited to refresher training organized by the Finnish Defence Forces (Table 1). By law, the period of liability for compulsory military service starts at the beginning of the year in which a young man reaches his 18th birthday and continues until the end of the year in which he turns 60. After the military service of 6, 9, or 12 months, the men are invited about one to three times to a 4- to 7-d refresher training session. Some of the trainings are just for specialized tasks (e.g., signalists, engineers, leaders), whereas most of them are arranged for less specialized troops. Physical fitness is routinely tested during the military service and occasionally during refresher training. The study was accepted by the review board of the Finnish Defence Forces.

All men invited to six nonspecialized troop refresher training sessions ( $N = 1472$ ) during the year 2003 were also invited to participate in the assessments of health-related fitness. The trainings were geographically distributed over southern, eastern, and northern regions of Finland (two training sessions in each region). The number of non-respondents was 498 (33.8%). Almost all of the non-respondents declined to participate in the refresher training because of personal reasons, and only five directly refused

TABLE 1. Descriptive characteristics of the participants.

	<i>N</i>	Mean	SD	Min	Max
Age (yr)	967	29.0	3.9	21	43
Height (cm)	951	179	6	158	198
Weight (kg)	951	81.1	13.5	48	137
Body mass index (BMI)	951	25.3	3.8	16.3	42.4
Waist (cm)	949	91	11	55	166

to participate in the fitness tests. The percentage of refusals was close to the long-term average.

**Data collection.** The data collection consisted of three phases: filling in a questionnaire, pretest health screening (screening for health limitations affecting the eligibility for participation in fitness tests), and fitness tests. All participants answered questions on, for example, place of residence (urban, semiurban, rural), years of education ( $\leq 12$  yr, 13–15 yr,  $\geq 16$  yr), and smoking habits (never, quit more than 6 months ago, quit less than 6 months ago, current). The participants answered the questionnaire in a classroom setting, and they were allowed to ask for assistance if needed. The study protocol was explained verbally to the participants before they gave their written consent.

Physical activity was assessed by the self-administered short version of IPAQ, covering the previous 7 d (7). We report the IPAQ outcome in three different ways:

1. IPAQ categories for health, from the IPAQ scoring protocol ([www.ipaq.ki.se](http://www.ipaq.ki.se)): HEPA (reaching recommendations for health-enhancing physical activity) active: vigorous activity  $\geq 3 \text{ d}\cdot\text{wk}^{-1}$ , totaling  $\geq 1500 \text{ MET}\cdot\text{min}\cdot\text{wk}^{-1}$ , or  $\geq 7 \text{ d}\cdot\text{wk}^{-1}$  of any combination of walking, moderate-intensity, or vigorous activities, totaling  $> 3000 \text{ MET}\cdot\text{min}\cdot\text{wk}^{-1}$ . Minimally active: not HEPA active, but  $\geq 3 \text{ d}\cdot\text{wk}^{-1}$  of vigorous activity of  $\geq 20 \text{ min}\cdot\text{d}^{-1}$ , or  $\geq 5 \text{ d}\cdot\text{wk}^{-1}$  of moderate-intensity activity or walking  $\geq 30 \text{ min}\cdot\text{d}^{-1}$ , or  $\geq 5 \text{ d}\cdot\text{wk}^{-1}$  of any combination of walking, moderate-intensity, or vigorous activities, totaling  $\geq 600 \text{ MET}\cdot\text{min}\cdot\text{wk}^{-1}$ . Insufficiently active: not belonging to either of the above categories.
2. Total METs (continuous score from the IPAQ scoring protocol) were calculated as follows: (daily minutes of walking  $\times$  days per week with walking  $\times 3.3$ ) + (daily minutes of moderate-intensity activity  $\times$  days per week with moderate-intensity activity  $\times 4.0$ ) + (daily minutes of vigorous activity  $\times$  days per week with vigorous activity  $\times 8.0$ ). The MET values were derived from the IPAQ validity and reliability study (7).
3. Vigorous activity METs were calculated from the equation daily minutes of vigorous activity  $\times$  days per week with vigorous activity  $\times 8.0$ .

In addition, we calculated truncated MET-minutes per week, in which all daily minutes exceeding 120 min were truncated to 120 min. This rule has been proposed in the "Guidelines of Data Processing and Analysis of IPAQ Short Version" with the attempt to normalize skewed population data. Because the validation results were almost

identical to those with total MET-minutes, the truncated results are not shown.

Pretesting health assessment included measures of systolic and diastolic blood pressure, a modified physical activity readiness questionnaire, and a question on perceived health status (19). One of the questions (single-item question on leisure-time vigorous physical activity (SIVAQ)) used on this questionnaire was, "Think of the previous 3 months and consider all leisure-time physical activity with duration of at least 20 min. How frequently were you physically active?" The choices were 1) less than once a week; 2) no vigorous activities, but light or moderate physical activity at least once a week; 3) vigorous activity once a week; 4) vigorous activity twice a week; 5) vigorous activity three times a week; and 6) vigorous activity at least four times a week. Body weight and height, in very light clothing and without shoes, were measured during the health assessment. The results from SIVAQ were used in this study for comparative purposes. Waist circumference was measured three times at the midway between the lowest rib and iliac crest after a normal outbreathing, and the mean value was used in further analyses. The health assessment was carried out by a physician and military nurses.

The participants performed 10 fitness tests. To limit the published data, six tests were excluded from this paper. In preliminary analyses, these tests were either not related to physical activity (grip strength, trunk side-bending) or they did not add any new insight to the remaining four tests (backward walking, vertical jump, static back extension, pull-ups). The fitness tests used in the present paper were done in the following order:

1. The maximal oxygen uptake (aerobic capacity ( $\text{mL}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ )) test was performed on an ergometer cycle (Ergoline, Ergoline GmbH, Bitz, Germany). The handlebars and seat were individually adjusted. A pedaling rate of 60 rpm was maintained during the test. The test started with a power output of 75 W, which was increased by 25 W every 2 min. Heart rate (HR) was recorded continuously (Polar S 610, Polar Vantage NV). Verbal encouragement was given during the final minutes of the test. The test was terminated at volitional exhaustion, including a drop of the pedaling frequency below 50 rpm. Maximal oxygen uptake was estimated from HR and power output data (1) by using Fitware software (Fitware Oy, Mikkeli, Finland) for maximal test. The reliability and accuracy of this protocol has been studied in women (11) and in men (Keskinen et al., personal communication, 2005). Compared with indirect calorimetry, the Fitware protocol overestimated  $\dot{V}\text{O}_{2\text{max}}$  by 7.3% (SD 8.2) in women and underestimated  $\dot{V}\text{O}_{2\text{max}}$  by 7.0% (SD 5.7) in men. Correlations between these two methods were  $r = 0.87$  (women) and  $r = 0.94$  (men). The test-retest repeatability was  $r = 0.89$  and  $0.96$  for women and men, respectively.
2. Sit-ups measure performance of the abdominal and hip flexor muscles. In the starting position, the participant was lying on the floor, keeping hands behind the neck

and directing elbows forward. The knees were flexed at an angle of  $90^\circ$ , legs were slightly abducted, and the assistant supported the ankles. During the movement, subjects lifted their upper body and touched their knee by the elbows. Result of this test was expressed as the number of sit-ups during 60 s (22).

3. Push-ups measure performance of arm and shoulder extensor muscles. The test was started from the lowest face-down position. The hands were kept shoulder-wide and level. Fingers were forward directed and legs were kept parallel near by each other. During the movement, arms were fully extended and torso was straightly tensioned. Then the torso was lowered down to the elbow angle of  $90^\circ$ . Result of this test was expressed as the number of push-ups during 60 s (12,16).
4. Repeated squats measure muscular endurance of extensor muscles of lower limbs, mainly quadriceps femoris. The test was started from a natural standing position, hands hanging freely. During the test, the participants were asked to squat as many times as possible during 60 s so that their fingertips touched the floor (15).

Each participant performed a warm-up protocol, including 10 min of cycling and stretching before the test session. They were given a 30-min break after the maximal oxygen uptake test. The recovery time between each muscular fitness test was at least 3 min. All tests were taught, supervised, and controlled by experienced, professional testers of the Finnish Defence Forces.

**Statistical analyses.** The physical activity level of the participants was classified using the IPAQ HEPA categorical score (insufficiently, minimally, and HEPA high active), by quintile classification of the two continuous IPAQ outcomes (total METs, vigorous METs) and the outcome of the SIVAQ, also referred to as "vigorous physical activity frequency"). Because the proportion (55.9%) of participants in the HEPA active IPAQ category was very large, this category was further divided into three equal groups by total METs tertile classification (cut points 4170 and 8094  $\text{MET}\cdot\text{min}\cdot\text{wk}^{-1}$ ), yielding five groups of almost equal size, as when quintile classification was used.

The distribution of vigorous physical activity frequency (SIVAQ) was cross-tabulated against the five IPAQ groups. The association between physical activity and fitness was tested by analysis of covariance, with *post hoc* Tukey's tests. HEPA category, total MET category, and vigorous MET category from IPAQ, and vigorous physical activity frequency from SIVAQ, were all in turn used as grouping (independent) variables with fitness test result as the dependent variable. Waist circumference, age, smoking (current or quit smoking less than 6 months ago: no smoking as reference) and education ( $\leq 12$  yr, 13–15 yr,  $\geq 16$  yr as reference) were used as covariates. The motivation for education as a covariate is that education might be associated with an individual's ability to understand the questions and to answer IPAQ, and that education might also be a proxy for a lifestyle that could affect health-related fitness. Body mass index (BMI) was not included in the equation, because waist

TABLE 2. Descriptive results of single-item question on leisure-time vigorous physical activity questionnaire (SIVAQ), International Physical Activity Questionnaire (IPAQ), and the tests used for validation.

	N	%			
<b>SIVAQ categories</b>					
No physical activity	141	14.6			
No vigorous, some moderate	250	25.9			
Vigorous 1 × wk <sup>-1</sup>	118	12.2			
Vigorous 2 × wk <sup>-1</sup>	186	19.2			
Vigorous 3 × wk <sup>-1</sup>	165	17.1			
Vigorous ≥ 4 × wk <sup>-1</sup>	80	8.2			
<b>IPAQ categories</b>					
Insufficiently active	207	21.4			
Minimally active	216	22.3			
Health-enhancing physical activity (HEPA) active	541	65.9			
	<b>N</b>	<b>Mean</b>	<b>SD</b>	<b>Minimum</b>	<b>Maximum</b>
IPAQ (total MET·min·min <sup>-1</sup> )	847	5035	4949	0	35,805
IPAQ (vigorous MET·min·min <sup>-1</sup> )	909	2511	2939	0	19,600
$\dot{V}O_{2max}$ (mL·kg <sup>-1</sup> ·min <sup>-1</sup> )	891	42.7	7.2	23.3	64.4
Sit-ups (reps during 1 min)	895	34	10	3	61
Push-ups (reps during 1 min)	893	24	11	0	69
Repeated squat (reps during 1 min)	928	32	7	12	68

circumference was found to be a stronger predictor of physical fitness in a preliminary analysis (results not shown). SPSS version 12.1. was used for all statistical analyses.

## RESULTS

The descriptive results of SIVAQ, IPAQ, and the physical fitness test results (mean, SD, range) are shown in Table 2. The MET-minute-per-week distributions were extremely skewed to the right, and also the proportion of individuals scored as HEPA active was considerable. The associations between

IPAQ groupings and fitness were consistent, regardless of the way IPAQ was used (HEPA categories, total METs, vigorous): almost all outcome variables showed increasing fitness from the first (least active) through the fourth group, but at least numerically lower (compared with the fourth group) mean fitness level for those who reported the highest level of activity (Figs. 1–4). The results for estimated  $\dot{V}O_{2max}$  were significantly different ( $P < 0.05$ ) between the fourth and fifth group, when the classification was done by HEPA categories or by IPAQ vigorous activity METs. Moreover, the fourth vigorous MET quintile had better ( $P < 0.05$ ) results in repeated squats, compared with the fifth quintile by the same grouping. The weekly frequency of vigorous physical activity (SIVAQ) was consistently and positively associated with  $\dot{V}O_{2max}$ , sit-ups, push-ups, and repeated squatting through all six precoded classes. The highest class (frequency four to seven times a week) had significantly ( $P < 0.05$ ) higher  $\dot{V}O_{2max}$  and push-ups result, compared with the next class (three times a week).

The distribution of reported weekly vigorous activity frequency by IPAQ groupings are shown in Figure 5. Especially the proportion of those with vigorous activity less than once a week showed a consistent pattern: the proportion decreased gradually when moving from first to the fourth group. However, this association was lost in the fifth (highest) group when the proportion of sedentary individuals again rose. An almost opposite pattern was seen for the proportion of physically active individuals.

## DISCUSSION

In the present study, the fitness level in IPAQ categories 1 and 2 (insufficiently and minimally active) were not different, whereas those classified as HEPA active had

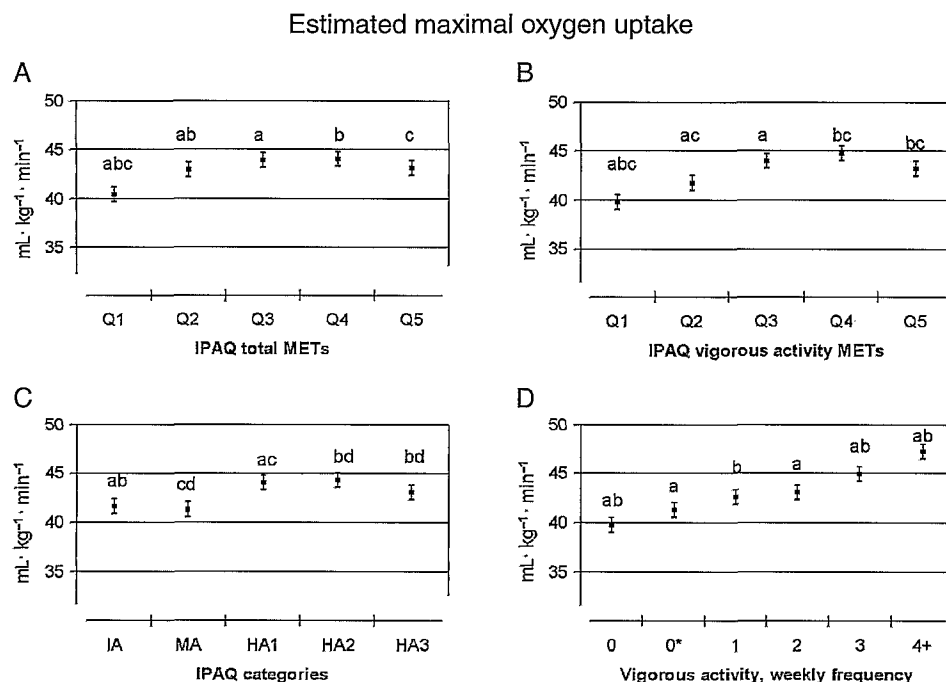
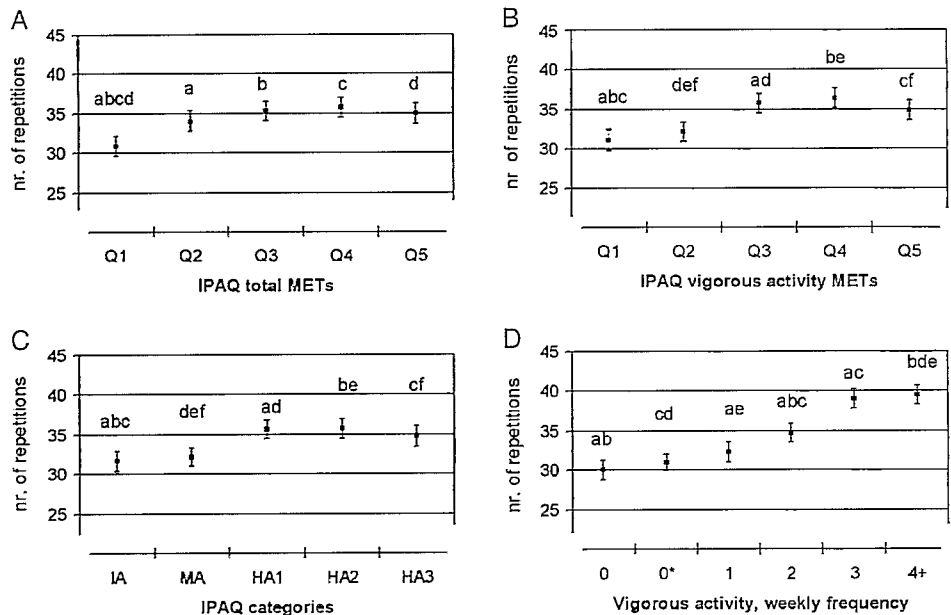


FIGURE 1—Estimated maximal oxygen uptake (mean, 95% confidence interval for the mean) in young adult men, grouped according to their reported physical activity by the International Physical Activity Questionnaire (IPAQ) or a single-item question on weekly vigorous activity frequency (SIVAQ). Results are adjusted to age, waist circumference, education, and smoking. Q1–Q5, groups obtained by quintile classification of IPAQ continuous score; IA, inactive; MA, minimally active; HA1–3, HEPA (sufficiently) active, three equal groups obtained by tertile classification. 0\*, no vigorous activity, but at least some regular moderate physical activity. Significant between-group differences are shown by identical letters (a–f).

## Sit-ups

**FIGURE 2**—Number of sit-ups during 60 s (mean, 95% confidence interval for the mean) in young adult men, grouped according to their reported physical activity by the International Physical Activity Questionnaire (IPAQ) or a single-item question on weekly vigorous activity frequency (SIVAQ). Results are adjusted to age, waist circumference, education, and smoking. Q1–Q5, groups obtained by quintile classification of IPAQ continuous score; IA, inactive; MA, minimally active; HA1–3, HEPA (sufficiently) active, three equal groups obtained by tertile classification. 0\*, no vigorous activity, but at least some regular moderate physical activity. Significant between-group differences are shown by identical letters (a–f).

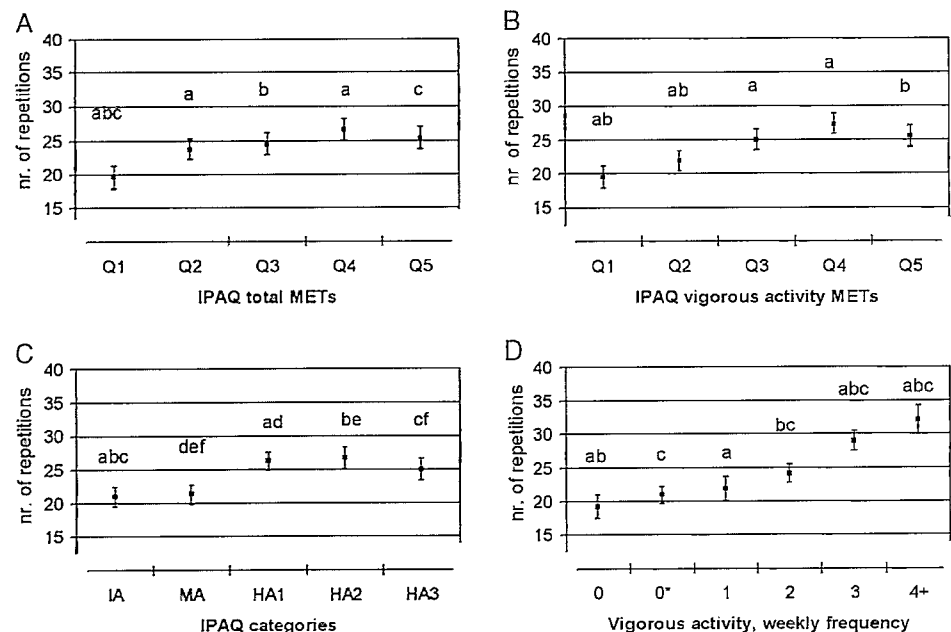


superior fitness. When the HEPA active category was further divided into three equal groups, the fitness results of the three different IPAQ outcomes (HEPA categories, total METs, and vigorous activity METs) were more similar: the dose–response relation between physical activity and CRF was as expected (higher activity is associated with better CRF) from lowest up to the fourth group, but surprisingly the CRF in the most active 20% (as calculated from IPAQ) was at least numerically lower than in the previous group, with lower total and vigorous MET-minutes per week. Although most of the comparisons between the fourth and fifth group were not statistically significant, the lack of consistently improving fitness was

unexpected. The present participants were ordinary young men and the highest estimated  $\dot{V}O_{2max}$  ( $64.4 \text{ mL}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ ) was clearly lower compared with elite endurance athletes. It was therefore anticipated that higher physical activity levels would have been associated with better CRF over the entire range of activity in the present population.

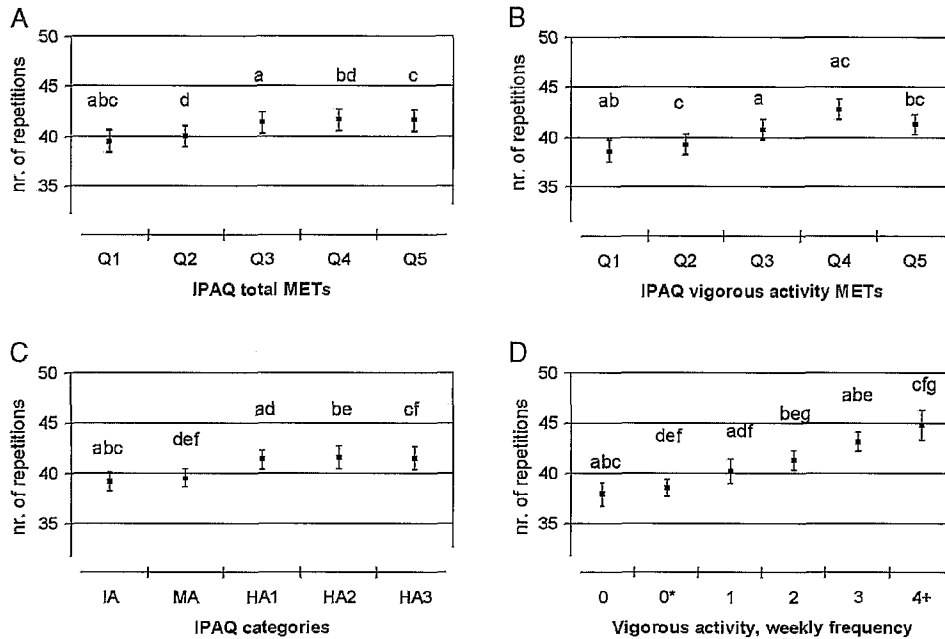
An explanation for the unexpected finding could be that several individuals in the most active 20% perform only low-intensity physical activities (e.g., walking). This was apparently not the case, however, because the vigorous activity METs showed a similar association with CRF than total METs. In fact, the only significant difference between the fourth and fifth quintile (the fifth had significantly

**FIGURE 3**—Number of push-ups during 60 s (mean, 95% confidence interval for the mean) in young adult men, grouped according to their reported physical activity by the International Physical Activity Questionnaire (IPAQ) or a single-item question on weekly vigorous activity frequency (SIVAQ). Results are adjusted to age, waist circumference, education and smoking. Q1–Q5, groups obtained by quintile classification of IPAQ continuous score. IA, inactive; MA, minimally active; HA1–3, HEPA (sufficiently) active, three equal groups obtained by tertile classification; 0\*, no vigorous activity, but at least some regular moderate physical activity. Significant between-group differences are shown by identical letters (a–f).



## Push-ups

## Repeated squats



**FIGURE 4**—Number of repeated squats during 60 s (mean, 95% confidence interval for the mean) in young adult men, grouped according to their reported physical activity by the International Physical Activity Questionnaire (IPAQ) or a single-item question on weekly vigorous activity frequency (SIVAQ). Results are adjusted to age, waist circumference, education, and smoking. Q1–Q5, groups obtained by quintile classification of IPAQ continuous score. IA, inactive; MA, minimally active; HA1–3, HEPA (sufficiently) active, three equal groups obtained by tertile classification; 0\*, no vigorous activity, but at least some regular moderate physical activity. Significant between-group differences are shown by identical letters (a–g).

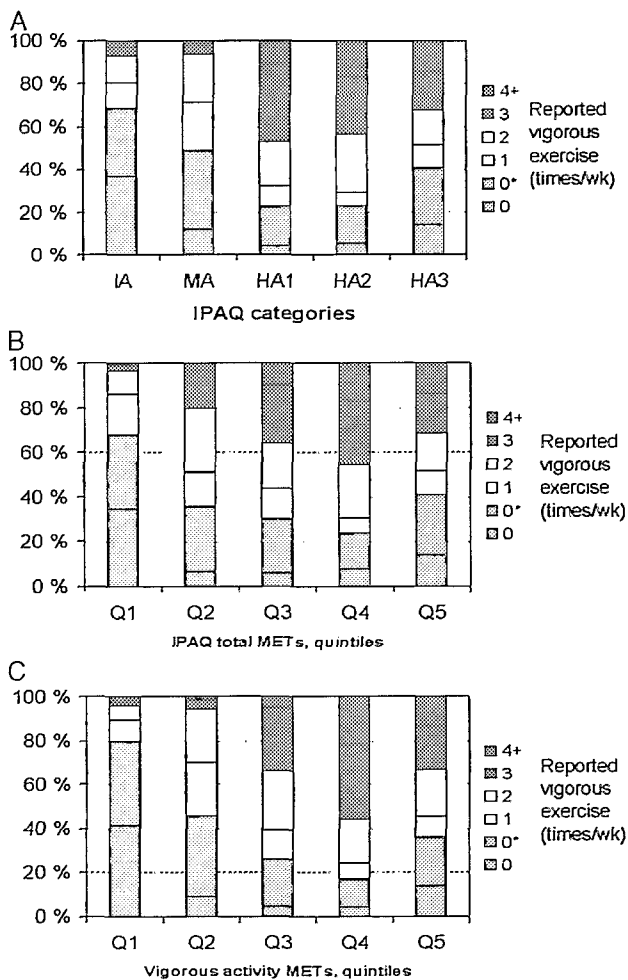
lower CRF) was observed when the classification was based on vigorous activity METs. Consequently, it is very likely that a significant proportion of those who were classified to the highest 20% by IPAQ were in fact much less active. This was also indicated by the result that a remarkable proportion of those in the highest group of IPAQ reported no vigorous activity by using the SIVAQ, regardless of the way IPAQ results were assessed. In contrast, the criterion validity of the single-item question on weekly frequency of vigorous activity against CRF was consistent from the lowest to the highest frequency.

Maximal oxygen uptake was chosen as the primary outcome test, because IPAQ has mainly been designed to assess activities that are related to cardiovascular fitness. The three other tests assessed muscular fitness. Sit-ups and push-ups are good measures of upper- and central-body muscular endurance. Repeated squats measure endurance and anaerobic capacity of the lower extremities. The tests were chosen to indicate muscular fitness especially related to occupational ability and functional capacity in daily life (15,20). The conclusions from the muscular fitness tests were similar to those from CRF: a marked proportion of individuals in the highest activity group of IPAQ had poor muscular fitness, suggesting again overreporting of physical activity. In contrast, the association between the weekly frequency of vigorous physical activity and muscular fitness seemed more consistent. The results show that the concerns related to IPAQ were not only restricted to CRF as the outcome variable.

We made a *post hoc* analysis by comparing apparent extreme overreporters (65 individuals with no reported vigorous or moderate activity by the single-item question, but who belonged to the highest 20% by IPAQ total METs) against the remaining participants ( $N = 102$ ) of the high-

activity reporters (highest 20% by IPAQ total METs). Selected results are presented in Table 3. The overreporters were older, more abdominally obese, had clearly poorer physical fitness, smoked more, and were less educated. When the fitness results are compared against those presented in Figures 1–4, it can be seen that the apparent overreporters' fitness was numerically very close to those of the lowest 20% by IPAQ. The remaining high-activity reporters' fitness was numerically (although not significantly) higher compared with those belonging to the fourth quintile of IPAQ total METs. Hence, if the apparent overreporters are removed from the highest 20% of IPAQ, the association between IPAQ quintiles and fitness becomes much more logical; that is, at least numerically improving fitness for each 20%. A hypothesis is that less educated individuals with very low and infrequent physical activity might think of weekly physical activity, when in fact daily duration is queried. Because these individuals reported low physical activity in the frequency question, a deliberate overreporting in IPAQ is not a likely hypothesis.

In an earlier study, physical activity in middle-aged men, assessed by SIVAQ, was repeated after a 5-yr follow-up (8). About 80% of those who answered "no vigorous activity" (categories 1 or 2) in the first examination gave the same answer after 5 yr. Moreover, about 60% of those reporting vigorous activity at least once a week (categories 3–6) in the beginning of the follow-up gave the same answer in the end. Because of the long time interval between the measurements, the above study was not a real validation study, so we do not know, for instance, the test-retest repeatability of SIVAQ. Nevertheless, these kinds of simple questions are often used in national surveys and other large epidemiological studies. Rütten et al. (17) compared IPAQ with single questions on physical activity



**FIGURE 5**—Distribution of reported frequency of a single item on weekly vigorous physical activity (SIVAQ) in each of the five groups obtained from International Physical Activity Questionnaire. Q1–Q5, groups obtained by quintile classification of IPAQ continuous score. IA, inactive; MA, minimally active; HA1–3, HEPA (sufficiently) active, three equal groups obtained by tertile classification; 0\*, no vigorous activity, but at least some regular moderate physical activity.

frequency used in national surveys. Interestingly, they reported that the single questions were more closely related to perceived health.

Rzewnicki et al. (18) reported that about 5% of respondents in a national survey in Belgium gave extremely high (not credible or even impossible) levels of activity by IPAQ, and that overreporting to a lesser degree was

indicated in a much larger proportion. Their conclusion was based on a probe control (respondents were asked to explain their responses during the interview and to give more detailed reports for the last 7 d) to systematically examine the accuracy of IPAQ responses. The authors concluded that “many people from the general population do not understand the IPAQ consistently.” In another study, Brown et al. (6) used four different instruments to measure time spent in physical activity. IPAQ gave clearly the highest estimations of physical activity, which obviously is only a suggestion of overreporting.

Obvious concerns are seen in choosing fitness as a criterion variable for physical activity. The health-related fitness concept (5) provides a rational theory for validation of IPAQ against fitness: physical activity may improve different components of health-related fitness, and fitness, in turn, may be positively associated with different aspects of health.

Physical activity, in particular aerobic activity, is related to CRF ( $\dot{V}O_{2max}$ ) in a positive dose–response manner (14). Although genes have an influence on the individual variation of CRF response to a given dose (duration, intensity, and type of physical activity), the dose–response relationship still remains over a broad range (4). This was one of the reasons for choosing quintile classification (analysis of covariance) rather than individual values (correlation) as the base for dose–response analysis.

In addition to genetics, some other factors may also modify the relation between physical activity and fitness. Several studies show that CRF, when measured as maximal oxygen uptake ( $mL \cdot kg^{-1} \cdot min^{-1}$ ), is poor in individuals with obesity (3,10,13). Smoking and older age are two other factors with a negative influence on CRF (2,10). Education, on the other hand, might be related to the way IPAQ is understood and answered. In our models, all of the above variables (obesity, smoking, age, and education) were significantly related to CRF (results not shown). The choice of covariates in the model, therefore, was justified based both on the literature and on the actual outcome itself.

The present study is the first to validate IPAQ by using components of health-related fitness as criterion variables. We found that almost 10% of young men participating in the present study had poor physical fitness and reported no vigorous activity in a single-item question on physical activity, but they nevertheless reported very high physical activity

**TABLE 3** Body mass index (BMI), waist circumference, and physical fitness (mean, 95% confidence interval for the mean) for individuals reporting high physical activity by International Physical Activity Questionnaire (IPAQ) (highest 20% for total METs): comparison of apparent overreporters (no vigorous activity according to a single-item question on the frequency of vigorous physical activity) and the remaining group.

	Apparent Overreporters (N = 65)	Remaining Group (N = 102)	P
Age (yr)	30.6 (29.3–31.8)	28.9 (28.2–29.6)	0.02
BMI	25.7 (25.4–26.8)	24.7 (24.0–25.4)	0.13
Waist circumference (cm)	92.7 (89.8–95.5)	88.5 (86.6–90.5)	0.02
$\dot{V}O_{2max}$ ( $mL \cdot kg^{-1} \cdot min^{-1}$ )	40.5 (38.6–42.3)	44.5 (43.3–45.7)	<0.001
Sit-ups (reps during 1 min)	30.1 (27.8–32.3)	36.2 (34.4–38.1)	<0.001
Push-ups (reps during 1 min)	22.0 (19.5–24.5)	26.8 (24.6–29.1)	0.005
Repeated squat (reps during 1 min)	39.3 (37.7–40.9)	42.2 (40.6–43.7)	0.01
Current smokers (%)	59.5	42.0	0.03
Low education (<12 yr)	70.8	52.0	0.02

by using IPAQ. As a consequence, the most active 20% (fifth group) by IPAQ had numerically (in three comparisons also significantly) lower  $\dot{V}O_{2\max}$  and number of sit-ups, push-ups, and squats during 60 s, compared with those belonging to the lower 20% (fourth group). In contrast, a single question on the weekly frequency of vigorous physical activity showed a more consistent dose-response relationship with cardiorespiratory and muscular fitness.

## REFERENCES

1. ÅSTRAND, P. O., and I. RHYMING. A nomogram for calculation of aerobic capacity (physical fitness) from pulse rate during submaximal work. *J. Appl. Physiol.* 7:218-221, 1954.
2. BERNAARDS, C. M., J. W. TWISK, W. VAN MECHELEN, J. SNEL, and H. C. KEMPER. A longitudinal study on smoking in relationship to fitness and heart rate response. *Med. Sci. Sports Exerc.* 35: 793-800, 2003.
3. BERTOLI, A., N. DI DANIELE, M. CECCOBELLI, A. FICARA, C. GIRASOLI, and A. DE LORENZO. Lipid profile, BMI, body fat distribution, and aerobic fitness in men with metabolic syndrome. *Acta Diabetol.* 40(1 Suppl):S130-S133, 2003.
4. BOUCHARD, C., and T. RANKINEN. Individual differences in response to regular physical activity. *Med. Sci. Sports Exerc.* 33(6 Suppl):S446-S451, 2001.
5. BOUCHARD, C., and R. J. SHEPHARD. Physical activity, fitness, and health: the model and key concepts.. In: *Physical Activity, Fitness and Health*, C. Bouchard, R. J. Shephard, and T. Stephens (Eds.). Champaign, IL: Human Kinetics. 1994, pp. 77-88.
6. BROWN, W., A. BAYMAN, T. CHEST, S. TROST, and K. MUMMERY. Comparison of surveys used to measure physical activity. *Aust. N. Z. J. Public Health* 28:128-134, 2004.
7. CRAIG, C. L., A. L. MARSHALL, M. SJÖSTRÖM, et al. International Physical Activity Questionnaire: 12-country reliability and validity. *Med. Sci. Sports Exerc.* 35:1381-1395, 2003.
8. HAAPANEN, N., S. MIILUNPALO, I. VUORI, P. OJA, and M. PASANEN. Characteristics of leisure-time physical activity associated with decreased risk of premature all-cause and cardiovascular disease mortality in middle-aged men. *Am. J. Epidemiol.* 143:870-880, 1996.
9. HALLAL, P. C., C. G. VICTORA, J. C. K. WELLS, R. C. LIMA, and N. J. VALLE. Comparison of short and full-length International Physical Activity Questionnaires. *J. Phys. Activity Health* 1(3): 227-234, 2004.
10. HULENS, M., G. VANSANT, A. L. CLAESSENS, R. LYSSENS, and E. MULS. Predictors of 6-min walk test in lean, obese and morbidly obese women. *Scand. J. Med. Sci. Sports* 13:98-105, 2003.
11. KESKINEN, K., O. KESKINEN, T. TAKALO, and K. HÄKKINEN. Comparison between straight measurement and two concurrent protocols to predict maximal oxygen uptake. *Med. Sci. Sports Exerc.* 34:S271, 2002.
12. MEIR, R., R. NEWTON, E. CURTIS, M. FARDELL, and B. BUTLER. Physical fitness qualities of professional rugby league football players: determination of positional differences. *J. Strength Cond. Res.* 15:450-458, 2001.
13. MIYATAKE, N., S. TAKANAMI, Y. KAWASAKI, and M. FUJII. Relationship between visceral fat accumulation and physical fitness in Japanese women. *Diabetes Res. Clin. Pract.* 64:173-179, 2004.
14. OJA, P. Dose response between total volume of physical activity and health and fitness. *Med. Sci. Sports Exerc.* 33 (6 Suppl):S428-S437, 2001.
15. POHJONEN, T. Age-related physical fitness and the predictive values of fitness tests for work ability in home care work. *J. Occup. Environ. Med.* 43(8):723-730, 2001.
16. ROBERTS, M. A., J. O'DEA, A. BOYCE, and E. T. MANNIX. Fitness levels of firefighter recruits before and after a supervised exercise training program. *J. Strength Cond. Res.* 16(2):271-277, 2002.
17. RÜTTEN, A., A. VUILLEMIN, W. T. M. OUIJENDIJK, et al. Physical activity monitoring in Europe. The European Physical Activity Surveillance System (EUPASS) approach and indicator testing. *Public Health Nutr.* 6:377-384, 2003.
18. RZEWNICKI, R., Y. VANDEN AUWEELE, and I. DE BOURDEAUDHUI. Addressing overreporting on the International Physical Activity Questionnaire (IPAQ) telephone survey with a population sample. *Public Health Nutr.* 6:299-305, 2003.
19. SUNI, J. H., S. I. MIILUNPALO, T.-M. ASIKAINEN, et al. Safety and feasibility of a health-related fitness test battery for adults. *Phys. Ther.* 78:134-148, 1998.
20. SUNI, J. H., P. OJA, S. I. MIILUNPALO, M. E. PASANEN, I. M. VUORI, and K. BÖS. Health-related fitness test battery for adults: associations with perceived health, mobility, and back function and symptoms. *Arch. Phys. Med. Rehabil.* 79:559-569, 1998.
21. VANDELANOTTE, C., I. DE BOURDEAUDHUI, R. PHILIPPAERTS, M. SJÖSTRÖM, and J. SALLIS. Reliability and validity of a computerized and Dutch version of the International Physical Activity Questionnaire (IPAQ). *J. Phys. Activity Health* 2:63-75, 2005.
22. VIILANEN, T., J. T. VITASALO, and U. M. KUJALA. Strength characteristics of a healthy urban adult population. *Eur. J. Appl. Physiol.* 63:43-47, 1991.



A vertical bar on the left side of the page, consisting of a series of horizontal segments in shades of yellow and orange, with a small red diamond at the top.

COPYRIGHT INFORMATION

TITLE: Interational Physical Activity Questionnaire: Validity  
against Fitness

SOURCE: Med Sci Sports Exercise 38 no4 Ap 2006

WN: 0609101727021

The magazine publisher is the copyright holder of this article and it is reproduced with permission. Further reproduction of this article in violation of the copyright is prohibited.

Copyright 1982-2006 The H.W. Wilson Company. All rights reserved.