Review

Scand J Work Environ Health 2011;37(1):6-29

doi:10.5271/sjweh.3085

The repeatability and validity of questionnaires assessing occupational physical activity – a systematic review

by Kwak L, Proper KI, Hagströmer M, Sjöström M

Affiliation: Unit for Preventive Nutrition, Department of Biosciences and Nutrition, Karolinska Institute, Hälsovägen 7, 141 83, Huddinge, Sweden. lydia.kwak@ki.se

Refers to the following texts of the Journal: 2010;36(5):366-372 2009;35(6):466-474 2007;33(6):405-424

The following articles refer to this text: 2012;38(6):582-589; 2016;42(2):125-134; 2017;43(2):127-135; 2017;43(3):269-278; 2018;44(1):96-105; 2018;44(5):530-538; [online first; 23 April 2019]

Key terms: assessment; construct; criterion; evidence; methodological quality; occupational physical activity; physical activity; questionnaire; repeatability; review; survey; systematic review; validity; work-related

This article in PubMed: www.ncbi.nlm.nih.gov/pubmed/20802979

Review

Scand J Work Environ Health 2011;37(1):6–29

The repeatability and validity of questionnaires assessing occupational physical activity – a systematic review

by Lydia Kwak, PhD,[1] Karin I Proper, PhD,[2] Maria Hagströmer, PhD,[1, 3] Michael Sjöström, PhD [1]

Kwak L, Proper KI, Hagströmer M, Sjöström M. The repeatability and validity of questionnaires assessing occupational physical activity – a systematic review. *Scand J Work Environ Health.* 2011;37(1):6–29.

Objectives This study aims to review systematically the repeatability and validity of questionnaires used to assess occupational physical activity among healthy adults.

Methods We searched PubMed and Embase using occupational, work-related, job-related, physical activity, motor activity, and questionnaires as keywords. Two reviewers independently performed article selection, data extraction, and quality assessment. The methodological quality and results of the studies were evaluated based on an existing checklist. The level of evidence and repeatability, criterion, and construct validity were rated.

Results We included 31 papers describing 30 questionnaires in the review. Repeatability was assessed in 22 studies, 11 used appropriate measures to assess 12 questionnaires. Intra-class correlation coefficients and weighted Cohen's kappa ranged between 0.43–0.95. Six studies used appropriate measures to assess criterion validity of 13 questionnaires. One questionnaire, the Tecumseh Self Administered Occupational Physical Activity Questionnaire (TOQ), showed good criterion validity against a physical activity (PA) record. Eighteen studies used appropriate measures to assess the construct validity of 23 questionnaires. Comparison included those against accelerometers, maximal oxygen uptake, questionnaires, and body composition measures. None showed good construct validity.

Conclusions There is strong evidence for good reliability of four questionnaires. None of the reviewed questionnaires showed good criterion validity compared to objective measures. Compared to PA records, moderate-to-good validity was observed for two questionnaires. Objective measures of occupational PA are needed.

Key terms assessment; construct; criterion; evidence; methodological quality; survey; work-related.

Regular physical activity (PA) has shown to provide a variety of health benefits including a reduction in the risk of morbidity, such as cardiovascular disease, diabetes, high blood pressure, and obesity as well as a reduction in the risk of premature mortality (1). The results of recent studies have led to the suggestion that the health benefits of PA might differ for different domains of PA (2–4). The adult population spends most time in the work domain (5). The PA performed in this domain is referred to as occupational or work-related PA and includes all PA done as part of a job (6). Little is, however, known regarding the health effects of occupational PA, as few studies to date have adequately examined the contribution of occupational PA when studying the health benefits of PA (7). Available data provide conflicting information; while some studies observe the protective effects of occupational PA against, for example, cardiovascular disease (3, 8, 9), others show no or negative associations (4, 10, 11). Moreover, recent studies have shown contrasting cardiovascular effects of PA performed in different domains, such as during leisure time and work (3, 9, 11, 12).

In order to draw any conclusions regarding the amount of occupational PA and the influence of occupational PA on health, it is essential to have a reliable and valid measurement instrument. Several methods are available for assessing PA, for example accelerometers, pedometers, observations, and questionnaires (13), with the latter frequently used in surveys or studies. There are numerous different questionnaires that assess occupational PA, some of which have been tested on repeatability and/or validity. To date, an overview of the measurement properties of the questionnaires assessing occupational PA is, however, lacking. As the contradictory findings with respect to the relation between occupational PA and health could be

1 Unit for Preventive Nutrition, Department of Biosciences and Nutrition, Karolinska Institutet, Huddinge, Sweden.
2 Department of Public and Occupational Health, EMGO Institute, Vrije Universiteit Medical Centre, Amsterdam, the Netherlands.
3 Division of Physiotherapy, Department of Neurobiology, Care Sciences and Society (NVS), Karolinska Institutet, Huddinge, Sweden.

Corresponding author: Dr L Kwak, Unit for Preventive Nutrition, Department of Biosciences and Nutrition, Karolinska Institutet, Hälsovägen 7, 141 83, Huddinge, Sweden. [E-mail: lydia.kwak@ki.se]

due to unreliable and invalid questionnaires, a review of the repeatability and validity of questionnaires measuring occupational PA is needed. The purpose of our study was to conduct a systematic review of published evidence on the repeatability and validity of questionnaires used to assess occupational PA among healthy adults.

Methods

Literature search

In March 2009, we searched relevant peer-reviewed English-language papers in the PubMed electronic bibliographic database (complete database until 28 February 2009). Subsequently, additional unique papers were searched in Embase following a screening of the reference lists of retrieved articles. The following full search strategy in PubMed was used: [(occupational OR "work related" OR "job related") AND ("physical activity"[tiab] OR "motor activity"[mesh]) AND (questionnaire[mesh] OR questionnaire*[tiab]) AND (English[lang])]. In Embase, the Emtree terms of "physical activity" AND occupation AND questionnaire AND [English]/lim were used.

Eligibility criteria

We screened all hits for possible inclusion based on the title and abstract. The following inclusion criteria were used: (i) the study was a validation and/or repeatability study of one or more questionnaires measuring occupational PA, which included the validation and/or repeatability of occupational PA questions; (ii) the questionnaire could be used to measure occupational PA in the general population; (iii) information on (at least one of) the measurement properties of the questionnaire should be provided; (iv) the article should be published in the English language; and (v) the study was published before March 2009. Occupational PA was defined as a type of physical activity performed that is related to energy expended during work. Studies that were performed among a specific population, such as patients or pregnant women were excluded, as were studies that measured occupational PA in relation to specific disorders and/or symptoms (ie, back pain). Finally, we excluded studies that lacked sufficient information on the protocol used to examine the validity and/or repeatability of the questionnaire.

Data extraction

Two reviewers independently performed data extraction and quality assessment. A description of the questionnaires and the protocols used in the studies was extracted from the included papers using a standardized data extraction form. Data extracted included: (i) sample characteristics (ie, sample size, age, gender, employment status); (ii) information on the protocol used [ie, methods, time interval between test and re-test, reference method, type of administration (self, interview)]; (iii) description of the questionnaire studied [ie, unit of measurement (energy expenditure, work index), number of occupational PA questions, scoring protocol]; (iv) statistical information (tests performed, package used); and (v) results of repeatability and validity.

Quality assessment of the studies and results

A slightly modified version of the checklist developed by de Vries and colleagues (14, 15) was used to assess the methodological quality and results of the studies included [see appendix (table A)]. For the assessment of the methodological quality of the study, information was extracted and evaluated regarding the study design (sample characteristics, protocol, measurements, and statistical analyses). All items were scored 0, 0.5 or 1 point and summed per study. Accordingly, repeatability, criterion, and construct validity were rated depending on the results of the study (see below).

Repeatability

Repeatability concerns the degree to which repeated measurements among stable persons (test–retest) provide similar answers (16). The use of intra-class correlation coefficients (ICC) and weighted Cohen's kappa's (K_w) were considered appropriate methods to quantify repeatability with regard to continuous (17) and ordinal measures (18), respectively. An ICC or K_w ≥0.70 was rated positively (+), an ICC or K_w of 0.40–0.70 was rated as moderate (±), and an ICC or K_w <0.40 was scored negatively (–).

One point was given when a study assessed repeatability and an additional point assigned if the ICC or K_w was used. These points were added to the study design score. On the basis of the total score three levels of evidence were formulated: strong evidence (≥4.0 points), moderate evidence (2.0–4.0 points) and poor evidence (<2.0 points).

Criterion validity

Criterion validity refers to the extent to which scores on a particular instrument relate to a gold standard (ie, an instrument that measures the same construct) (18). Comparisons with accelerometers, when limited to occupational time, were considered as appropriate methods for objective criterion validity. Comparisons with PA records, diaries, and logbooks measuring occupational

PA, were considered as appropriate methods of subjective criterion validity. Correlation coefficients >0.75 were scored positively (+), correlations of 0.50–0.75 moderately (±), and a correlation <0.50 was scored negatively (–) (14, 16). Correlation coefficients >0.75 were scored positively as this indicates that the occupational measure and the criterion measure share >50% (~56%) of the variance in common (16).

One point was given when a study assessed criterion validity and an additional point if the sensitivity, specificity, Pearson's product moment, or Spearman's rank correlation coefficients was used; 0.5 points were given if a Bland Altman plot was used. These points were added to the study design score. On the basis of the total score, three levels of evidence were formulated: strong evidence (≥4.0 points), moderate evidence (2.0–4.0 points) and poor evidence (<2.0 points).

Construct validity

Construct validity refers to the extent to which scores on a particular instrument relate to other measures in a manner that is consistent with theoretically derived hypotheses concerning the concepts that are being measured (17, 19). Methods that can be used to measure the same or similar aspects of occupational PA, or aspects that are related to occupational PA, were considered appropriate for the assessment of construct validity (ie, accelerometers, doubly labeled water, fitness-test, body composition measurements etc). A positive score (+) was given if the correlation coefficient was >0.60, a moderate score (±) if the correlation was 0.30–0.60, and a negative score (–) for correlations <0.30 (14, 16). Correlation coefficients >0.60 indicate that the occupational measure and the comparison measure share 36% of the variance in common.

One point was given when a study assessed criterion validity and an additional point allocated if the Pearson's product moment, Spearman's rank correlation coefficients, t-test, Mann-Whitney U-test or chi-square test was used; 0.5 points were given if a Bland Altman plot was used. These points were added to the study design score. On the basis of the total score three levels of evidence were formulated: strong evidence (≥4.0 points), moderate evidence (2.0–4.0 points) and poor evidence (<2.0 points).

Results

The literature search resulted in 962 hits, of which 55 were selected on the basis of relevant titles and/or abstracts (figure 1). Of the fulltext articles, 24 were excluded after reading the article; the main reasons for exclusion were:

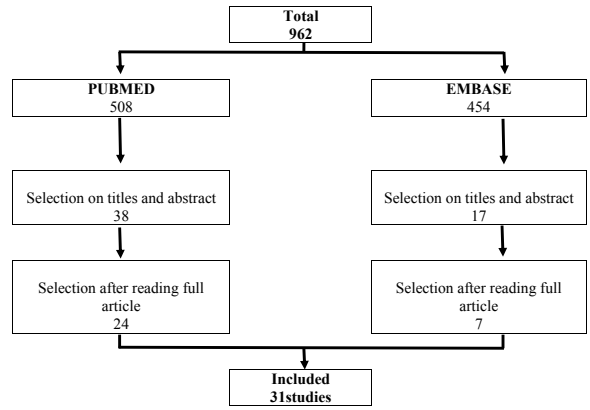

Figure 1: Flowchart of literature search and paper selection

(i) lack of information on occupational PA measures and/or results, (ii) focusing on physical inactivity, or (iii) executed in a specific sample (ie, pregnant women). Finally, 31 papers (20–50) describing 30 questionnaires were included in the review (see table B in the Appendix). Two studies assessed the accuracy of several questionnaires simultaneously (20, 33). Of the included questionnaires, Kuopio Ischemic Heart Disease Occupational Physical Activity Interview (KIDH-O), Occupational Physical Activity Questionnaire (OPAQ), Saltin & Grimby Lifetime Occupational Activity (SGLOA), Saltin & Grimby Present Occupational Activity (SGPOA), and Tecumseh Self-Administered Occupational Physical Activity Questionnaire (TOQ) assessed "only" occupational PA. The remaining questionnaires covered the entire range of PA, including occupational PA/measure diverse domains of PA, including occupational PA. A full description of the included questionnaires is provided in table 1.

Repeatability

In total, 22 studies assessed the repeatability of 26 questionnaires, of which 11 studies used appropriate measures to quantify the repeatability of 12 questionnaires (table 2). The repeatability of each questionnaire was only assessed in one study each. The level of evidence was strong for all studies (mean 5.2±0.5, range 4.5–6.0). Six of the studies were conducted in mixed-gender samples (27, 29, 43, 44, 47, 48), three among males (34, 35, 39) and two among females (22, 38). The average age of the study populations ranged from ~31–65 years, the sample sizes varied from 39–399 individuals (average N=132). The time intervals between the test and re-test varied from one week to one year. ICC and K_w were observed between 0.43–0.95. The Baecke, Behavioural Risk Factor Surveillance System (BRFSS), Nord-Trondelag Health Study (HUNT-2), Kaiser Physical Activity Survey (KPAS), and Tecumseh Community Health

Table 1. Description of physical activity questionnaires. For full names of questionnaires, see appendix. [OPA=occupational physical activity; M=male; F=female; MET=metabolic equivalent task; EE=energy expenditure; PA=physical activity; S=self-report; I=interview administered]

Questionnaire	Study	Adapted from	Population	Number of OPA questions	Type of work	Unit of measurement	Recall period	Type of administration	Time (minutes)
A-IPAQ-L	Graff-Iversen et al, 2007 (30)	IPAQ-L	M and F (adults)	7	All work types	MET minutes per week^{-1}	Usual week	S	..
ARIC/ Baecke	Ainsworth et al, 1993 (20)	Baecke	M and F (adults)	8	All work types	Work index (1–5)	Usual activity	S	..
Baecke	Baecke et al, 1982 (23); Jacobs Jr et al, 1993 (33); Philippaerts & Lefevre, 1998 (39); Philippaerts et al, 1999 (40); Philippaerts et al, 2001 (41)	·	M and F (adults)	8	All work types	Work index (1–5)	Usual activity	S	..
BRFSS	Evenson & McGinn, 2005 (27)	·	M and F (adults)	1	All work types	Activity score (1–3)	Usual activity	I	1
Cadbury	Batty, 2000 (24)	·	M (adults)	2	All work types	Work description (1–4); overall work activity (1–4)	Usual week past year	I	..
CARDIA-PAQ	Ainsworth et al, 1993 (20)	·	M and F (young adults)	1	All work types	Activity unit score	Past year	S	8 [a]
CARDIA-SDR	Ainsworth et al, 1993 (20); Ainsworth et al, 1999 (21); Jacobs Jr et al, 1993 (33)	SDR (Stanford Five-City Project)	M and F (adults)	5	All work types	MET minutes per day^{-1}	Past week	S	12 [a]
EPAQ2	Wareham et al, 2002 (48)	Modified Tecumseh Occupational Activity	M and F (adults)	13	All work types	MET hours per week^{-1}	Past year	S	..
EPIC	Cust et al, 2008 (26); Wareham et al, 2003 (49)	·	M and F (adults)	1	All work types	Activity index (1–4)	Past year	S	..
HIP	Ainsworth et al, 1993 (20); Jacobs Jr et al, 1993 (33); Shapiro et al, 1965 (45)	·	M and F (adults)	6	All work types	Work index	Current	S	..
HUNT 2	Kurtze et al, 2007 (34)	·	M and F (adults)	1	All work types	Activity score (1–4)	..	S	..
IPAQ-L	Hagströmer et al, 2006 (31); Vandelanotte et al, 2005 (47)	·	M and F (adults)	7	All work types	MET minutes per week^{-1}	Past 7 days	S	..
KIHD-O	Lakka & Salonen, 1992 (35); Lakka & Salonen, 1992 (36)	·	M (middle-aged)	11	All work types	MET hours per day^{-1}	Typical work day	I	..
KPAS	Ainsworth et al, 2000 (22)	Baecke	F (adults)	8	All work types	Work index (1–5)	Current	S	20 [a]
LRC	Ainsworth et al, 1993 (20)	·	M and F (adults)	1	All work types	Activity score (1–5)	Usual	I	..
LTPAQ	Friedenreich et al, 1998 (28)	·	M and F (adults)	Open	All work types	MET hours per week^{-1}	Lifetime	I	20–30 [a]
MAQ	Jacobi et al, 2009 (32)	Prima Indian Physical Activity Questionnaire	M and F (adults)	1	All work types	MET hours per week^{-1}	Average day past year	S	..
MHHP	Ainsworth et al, 1993 (20); Jacobs Jr et al, 1993 (33)	·	M and F (adults)	2	All work types	Work score	Usual	S	..
MOSPA-Q	Roeykens et al, 1998 (44)	·	M and F (adults)	6	All work types	~ weekly time, EE	Average week	S/I	..
OPAQ	Reis et al, 2005 (43)	·	M and F (adults)	7	All work types	Work score hours per week^{-1}; MET minutes per week^{-1}	Usual week	I	<5
PAQ	Norman et al, 2001 (37)	·	M (middle-aged, older)	1	All work types	MET hours per day^{-1}	Past year	S	..
pre-EPIC	Pols et al, 1997 (42)	·	M and F (adults)	1	All work types	24 hours-EE	Past year	S	15
PYTPAQ	Friedenreich et al, 2006 (29)	LTPAQ	M and F (adults)	open	All work types	MET minutes per week^{-1}	Past year	S	..
QAPSE	Bonnefoy et al, 1996 (25)	·	M and F (adults)	7	All work types	MET minutes per week^{-1}	1 week	S/I	20–30 [a]

(continued)

Table 1. Continued

Questionnaire	Study	Adapted from	Population	Number of OPA questions	Type of work	Unit of measurement	Recall period	Type of administration	Time (minutes)
SGLOA	Wilbur et al, 1989 (50)	·	M (middle-aged)	1	All work types	Work score	Lifetime	S	··
SGPOA	Wilbur et al, 1989 (50)	TOQ	M (middle-aged)	17	All work types	MET hours per day^{-1}	Past year	S	··
SMC-PAQ	Orsini et al, 2007 (38)	PAQ	M and F (adults)	1	All work types	MET hours per day^{-1}	Current, historical PA	S	··
SSAAQ	Sobngwi et al 2001 (46)	Prima Indian Physical Activity Questionnaire	M and F (Sub-Saharan adults)	2	All work types	MET hours per day^{-1}	Past year	I	··
TCQ	Philippaerts & Lefevre, 1998 (39); Philippaerts et al, 1999 (40); Philippaerts et al, 2001 (41)	Tecumseh Community Health Study	M and F (adults)	··		EE, work index	Average week last year	I	··
TOQ	Ainsworth et al, 1993 (20); Ainsworth et al, 1993 (21); Wilbur et al, 1989 (50)	Tecumseh Structured Physical Activity Interview	M and F (adults)	29	All work types	MET minutes per week^{-1}	Past year	S	15

[a] Completion time for total questionnaire

Study Questionnaire (TCQ) were all assessed with regard to the repeatability of their *work index/activity score*. All showed good repeatability: the BRFSS among a mixed-gender sample (ICC 0.82) and among males (ICC 0.86) and Caucasians separately (ICC 0.93) (27); the Baecke (ICC 0.95) (39), the HUNT-2 (K_w 0.80) (34), the TCQ among males (ICC 0.88) (39); and the KPAS among females (ICC 0.85) (22). Moderate repeatability was observed for the BRFSS among females (ICC 0.67) and African Americans (ICC 0.67) (27).

The European Prospective Investigation into Cancer Physical Activity Questionnaire (EPAQ2), International Physical Activity Questionnaire-long version (IPAQ-L), KIHD-O, MONICA Optimal Study of Physical Activity (MOSPA-Q), Past Year Total Physical Activity Questionnaire (PYTPAQ), Swedish Mammography Cohort-Physical Activity Questionnaire (SMC-PAQ), and TCQ were all assessed with regard to the repeatability of *energy expenditure*. Five showed good repeatability: the IPAQ-L (ICC 0.80) (47) and MOSPA-Q (ICC 0.78, ~146 days; ICC 0.85, >3 months) (44) in mixed-gender samples; the EPAQ-2 among males and females (K_w 0.79 and 0.82, respectively) (48); the TCQ among males (ICC 0.90) (39) and the SMC-PAQ among females but only for historical PA at ages 15, 30 and 50 years (ICC 0.73–0.75) and in some of the body mass index (BMI) and age subgroups (38). Moderate repeatability was observed for two questionnaires: the PYTPAQ [except among those with high levels of PA (ICC=0.78), for which it showed good repeatability (29)] and the KIHD-O among males (ICC=0.69) (35).

The repeatability of *duration of activity* was assessed for the IPAQ-L, MOSPA-Q, OPAQ, and TCQ. All showed good repeatability: the IPAQ-L (ICC 0.83) (47), the MOSPA-Q (ICC 0.77, ~146 days; ICC 0.82, >3 months) (44) and the OPAQ in mixed gender samples (ICC 0.76) (43), and the TCQ (ICC 0.93) among males (39). In addition, the OPAQ showed good repeatability for all its individual items (ICC 0.78–0.91), except walking, which showed moderate repeatability (43).

Validity

Criterion validity. Six studies used appropriate measures to assess criterion validity of, in total, 13 questionnaires. In two studies, criterion validity was assessed by validation against accelerometer data limited to occupational time only (41, 43) (table 3), with both studies providing strong levels of evidence (mean 5.3±0.4, range 5.0–5.5). One of these studies was conducted in a mixed-gender sample (N=166) (43), the other in an all male sample (N=41) (41). The average age of the study populations was around 39 years for both studies. Observed correlations (r) varied between -0.20–0.50.

The objective criterion validity of the *work index* of the Baecke and the TCQ were assessed among males and both showed poor validity against an accelerometer (41). The TCQ did, however, show moderate objective criterion validity with regard to *energy expenditure* in the same sample of males, but only when validated against the sum of work counts of the accelerometer (r=0.50) (41). Poor criterion validity was observed for OPAQ's

Table 2. Repeatability of questionnaires assessing occupational physical activity. For full names of questionnaires, see appendix. Evidence was rated as follows: strong evidence (3), moderate evidence (2) and poor evidence (1). Acceptable levels of repeatability rated as: intra-class correlation coefficients (ICC)≥0.70 (+), 0.40≤ICC<0.70 (±), ICC<0.40 (–). [M=male; F=female; C=Caucasian; AA=African American; K_w=weighted kappa; MET=metabolic equivalent task; EE=energy expenditure; PA=physical activity; S=self-report; I=interview administered; HPA=high physical activity, LMPA=low -to-moderate physical activity; BMI=body mass index; kcal = kilocalories; kg=kilograms.]

Questionnaire	Study	Study population				Interval	Type of administration	Unit of measurement	Repeatability	95% confidence interval	Rating
		Sample size	Age (years)	Educational level	Country						
Work index/activity score											
Baecke	Phippaerts & Lefevre, 1998 (39)	M: 90	35	..	Belgium	1 month	I	Work index (1–5)	ICC=0.95		3+
BRFSS	Evenson & McGinn, 2005 (27)	106	47	39.6% college	US	Median 16 days (range 9–30 days)	I	Activity score (1–3)	ICC=0.82	0.72–0.89	3+
									M: ICC=0.86; F: ICC=0.67	0.73–0.93 0.42–0.82	M: 3+; F: 3±
									C: ICC=0.93; AA: ICC=0.67	0.86–0.97 0.42–0.83	C: 3+; AA: 3±
HUNT-2	Kurtze et al, 2007 (34)	M: 108	32.4	..	Norway	1 week	S	Activity score (1–4)	K_w=0.80	0.71–0.89	3+
KPAS	Ainsworth et al, 2000 (22)	F: 45	39.1	78% college	US	1 month	S	Work index (1–5)	ICC=0.85 [a]		3+
TCQ	Phippaerts & Lefevre, 1998 (39)	M: 90	35	..	Belgium	1 month	I	Work index (MET)	ICC=0.88		3+
Energy expenditure											
EPAQ2	Wareham et al, 2002 (48)	399	M: 65; F: 63.8	..	UK	3 months	S	Work activity (MET hours per week^{-1})	M: K_w=0.79; F: K_w=0.82		M: 3+; F: 3+
IPAQ-L	Vandelanotte et al, 2005 (47)	53	30.9	88.5% high level	Belgium	1 week+ 11–14 days	S	Work activity (kcal)	ICC=0.80		3+
KIHD-O	Lakka & Salonen, 1992 (35)	M: 39	54	..	Finland	12 month	S/I	Work activity (MET hours per day^{-1})	ICC=0.69		3±
MOSPA-Q	Roeykens et al, 1998 (44)	156	M: 35.9; F: 35.7	..	Belgium	Average 146 days	S	Work EE (kcal per week^{-1})	ICC=0.78; <3 months: ICC=0.67; >3 months: ICC=0.85		3+; <3 months: 3±; >3 months: 3+
PYTPAQ	Friedenreich et al, 2006 (29)	154	48	56.5% university	Canada	Average 9 weeks	S	Work activity (MET hours per week^{-1})	ICC=0.58	0.47–0.68	3±
									M: ICC=0.43; F: ICC=0.69	0.22–0.60 0.55–0.79	M: 3±; F: 3±
									<50 years of age: ICC=0.54; >50 years of age: ICC=0.63	0.36–0.69 0.47–0.74	<50 years: 3±; >50 years: 3±
									HPA: ICC=0.78; LMPA: ICC=0.48	0.64–0.87 0.32–0.62	HPA: 3+; LMPA: 3±
									BMI<25: ICC=0.61; BMI>25: ICC=0.56	0.41–0.75 0.40–0.68	BMI<25: 3±; BMI>25: 3±
SMC-PAQ	Orsini et al, 2007 (38)	F: 303	65.2	46% ≥9 years	Sweden	1 year	S	Work (MET hours per day^{-1})			
									Current age (past year)		
									ICC=0.59	0.51–0.66	3±
									BMI<25: ICC=0.66; BMI>25: ICC=0.51		BMI<25: 3±; BMI>25: 3±
									<65 years of age: ICC=0.72; >65 years of age: ICC=0.43		<65 years: 3+; >65 years: 3±

(continued)

Table 2. Continued

Questionnaire	Study	Study population				Interval	Type of administration	Unit of measurement	Repeatability	95% confidence interval	Rating
		Sample size	Age (years)	Educational level	Country						
									At 50 years of age		
									ICC=0.73	0.67–0.79	3+
									BMI<25: ICC=0.77; BMI>25: ICC=0.69		BMI<25: 3+; BMI>25: 3±
									<65 years of age: ICC=0.82; >65 years of age: ICC=0.62		<65 years: 3+; >65 years: 3±
									At 30 years of age		
									ICC=0.75	0.69–0.80	3+
									BMI<25: ICC=0.79; BMI>25: ICC=0.70		BMI<25:3+; BMI>25:3±
									<65 years of age: ICC=0.86; >65 years of age: ICC=0.59		<65 years: 3+; >65 years: 3±
									At 15 years of age		
									ICC=0.73	0.67–0.79	3+
									BMI<25: ICC=0.70; BMI>25: ICC=0.77		BMI<25: 3±; BMI>25: 3+
									<65 years of age: ICC=0.68; >65 years of age: ICC=0.74		<65 years: 3±; >65 years: 3+
TCQ	Phippaerts & Lefevre, 1998 (39)	M: 90	35	..	Belgium	1 month	I	EE work (kcal per kg^{-1} per $week^{-1}$)	ICC=0.90		3+
Duration of activity											
IPAQ-L	Vandelanotte et al, 2005 (47)	53;	30.9	88.5% high level	Belgium	1 week	S	Work activity (minutes)	ICC=0.83		3+
MOSPA-Q	Roeykens et al, 1998 (44)	156;	M: 35.9; F: 35.7	..	Belgium	Average 146 days	S	Work score (minutes per $week^{-1}$)	ICC=0.77; <3 months: ICC=0.68; >3 months: ICC=0.82		3+; <3 months: 3±; >3 months: 3+
OPAQ	Reis et al, 2005 (43)	41	38.8	65.9% college	US	2 weeks	I	Work activity (hours per $week^{-1}$)	ICC=0.76	0.59–0.86	3+
								Sitting/standing (hours per $week^{-1}$)	ICC=0.78	0.62–0.87	3+
								Walking (hours per $week^{-1}$)	ICC=0.55	0.30–0.73	3±
								Heavy labor (hours per $week^{-1}$)	ICC=0.89	0.81–0.94	3+
								Employed for wages (hours per $week^{-1}$)	ICC=0.91	0.84–0.95	3+
TCQ	Phippaerts & Lefevre, 1998 (39)	M: 90	35	..	Belgium	1 month	I	Time work (hours per $week^{-1}$)	ICC=0.93		3+

[a] P<0.001.

Table 3. Criterion validity of questionnaires assessing occupational physical activity (OPA). For full names of questionnaires, see appendix. Evidence was rated as follows: strong evidence (3), moderate evidence (2) and poor evidence (1). Acceptable level of criterion validity rated as: correlation coefficient (r)≥0.75 (+), 0.50≤r<0.75 (±) and r<0.50 (–). [M=male; F=female; PA=physical activity; MET=metabolic equivalent task; EE=energy expenditure; kcal=kilocalories; kg=kilograms.]

Questionnaire	Study	Study population				Comparison measure	Unit of measurement	Criterion validity	Rating
		Sample size	Age (years)	Educational level	Country				
Work index/activity score									
ARIC/ Beacke	Ainsworth et al, 1993 (20)	75	37.3	71% college	US	OPA record book (activity units per week^{-1})	Work index	r=0.04	3-
Baecke	Philippaerts et al, 2001 (41)	M: 166	40	..	Belgium	Tracmor (work sum; work mean)	Work index	Sum: r=0.42 [a]; mean: r=0.33 [a]	Sum: 2-; mean: 2-
CARDIA-PAQ	Ainsworth et al, 1993 (20)	75	37.3	71% college	US	OPA record book (activity units per week^{-1})	Activity unit score	r=-0.05	3-
HIP	Ainsworth et al, 1993 (20)	75	37.3	71% college	US	OPA record book (activity units per week)	Work index	r=0.10	3-
KPAS	Ainsworth et al, 2000 (22)	F: 50	39.1	78% college	US	PA record: work score (MET minutes per day^{-1})	Work index	r=0.35 [b]	3-
						PA record sitting (MET minutes per day^{-1})	Sitting	r= 0.58 [a]	3±
						PA record standing (MET minutes per day^{-1})	Standing	r= 0.46 [a]	3-
						PA record walking (MET minutes per day^{-1})	Walking	r= 0.50 [a]	3±
						PA record heavy occupation (>6.5 MET)	Lift heavy loads	r= 0.32 [b]	3-
							Sweating/exertion	r= 0.33 [b]	3-
							Tired at the end of the day	r= 0.14	3-
LRC	Ainsworth et al, 1993 (20)	75	37.3	71% college	US	OPA record book (activity units per week^{-1})	Activity score	r=0.09	3-
MHHP	Ainsworth et al, 1993 (20)	75	37.3	71% college	US	OPA record book (activity units per week^{-1})	Work score	r=0.33 [b]	3-
TCQ	Philippaerts et al, 2001 (41)	M: 166	40	..	Belgium	Tracmor (work sum; work mean)	Work index (MET)	Sum: r=0.39 [a], mean: r=0.34 [a]	Sum: 2-, mean: 2-
Energy expenditure									
CARDIA-SDR	Ainsworth et al, 1993 (20)	75	37.3	71% college	US	PA record (MET)	Work (MET)	r=0.30 [b]	3-
						OPA record book (activity units per week^{-1})	Work (MET)	r=0.44 [a]	3-
IPAQ-L	Hagströmer et al, 2006 (31)	46	40.7	Mostly university	Sweden	Logbook work activity (MET hours per week^{-1})	Work activity (MET hours per week^{-1})	r=0.64 [c]	3±
PAQ	Norman et al, 2001 (37)	M: 111	63.2	..	Sweden	PA record work	Work (MET hours per day^{-1})	r=0.38	2-
							Work (MET hours per day^{-1})	<65 years of age: r=0.49, >65 years of age: r=0.43	<65: 2-; >65: 2-
TOQ	Ainsworth et al, 1993 (20)	75	37.3	71% college	US	PA record (MET)	Work (MET)	r=0.52 [a]	3±
						OPA record book (activity units per week^{-1})	Work score (activity units per week^{-1})	r=0.92 [c]	3+
						PA record total (activity units per week^{-1})	Work score (activity units per week^{-1})	r=0.92 [c]	3+
						PA record sitting, light (activity units per week^{-1})	Sitting, light (activity units per week^{-1})	r=0.77 [c]	3+
						Standing, light (activity units per week^{-1})	Standing, light (activity units per week^{-1})	r=0.57 [a]	3±
						Walking (activity units per week^{-1})	Walking (activity units per week^{-1})	r=0.41 [a]	3-
						Other (activity units per week^{-1})	Other (activity units per week^{-1})	r=-0.05	3-

(continued)

Table 3. Continued

Questionnaire	Study	Study population				Comparison measure	Unit of measurement	Criterion validity	Rating
		Sample size	Age (years)	Educational level	Country				
TCQ	Philippaerts et al, 2001 (41)	M: 166	40	..	Belgium	Tracmor (work sum; work mean)	EE work (kcal per kg^{-1} per week^{-1})	Sum: r=0.50 [a]; mean: r=0.26 [b]	Sum: 2±; mean: 2-
Duration of activity									
CARDIA-SDR	Ainsworth et al, 1993 (20)	75	37.3	71% college	US	PA record (hours per day^{-1})	Work (hours per week^{-1})	r=0.16	3-
OPAQ	Reis et al, 2005 (43)	41	38.8	65.9% college	US	OPA record total (hours per week^{-1})	Work activity (hours per week^{-1})	r=0.19	3-
						OPA record sitting/standing (hours per week^{-1})	Sitting/standing (hours per week^{-1})	r=0.37	3-
						OPA record walking (hours per week^{-1})	Walking (hours per week^{-1})	r=0.74	3±
						OPA record heavy labor (hours per week^{-1})	Heavy labor (hours per week^{-1})	r=0.31	3-
						OPA record light (hours per week^{-1})	Sitting/standing (hours per week^{-1})	r=0.12	3-
						OPA record moderate (hours per week^{-1})	Walking (hours per week^{-1})	r=0.41	3-
						OPA record vigorous (hours per week^{-1})	Heavy labor (hours per week^{-1})	r=0.24	3-
						Accelerometer light (hours per week^{-1})	Sitting/standing (hours per week^{-1})	r=-0.20	3-
						Accelerometer moderate (hours per week^{-1})	Walking (hours per week^{-1})	r=0.21	3-
						Accelerometer vigorous (hours per week^{-1})	Heavy labor (hours per week^{-1})	r=-0.17	3-
TCQ	Philippaerts et al, 2001 (41)	M: 166	40	..	Belgium	Activity log (hours per week^{-1})	Time work (hours per week^{-1})	r=0.58 [a]	2±
TOQ	Ainsworth et al, 1993 (20)	75	37.3	71% college	US	PA record (hours per day^{-1})	Work (hours per week^{-1})	r=0.11	3-
						Sitting, light (hours per week^{-1})	Sitting, light (hours per week^{-1})	r=0.82 [c]	3+
						Standing, light (hours per week^{-1})	Standing, light (hours per week^{-1})	r=0.57 [a]	3±
						Walking (hours per week^{-1})	Walking (hours per week^{-1})	r=0.38 [a]	3-
						Other (hours per week^{-1})	Other (hours per week^{-1})	r=-0.04	3-

[a] P<0.01.
[b] P<0.05.
[c] P<0.001.

items: "sitting" (hours per week^{-1}), "walking" (hours per week^{-1}), "heavy labor" (hours per week^{-1}), when compared "against light" (hours per week^{-1}), "moderate occupational PA" (hours per week^{-1}) and "heavy occupational PA" (hours per week^{-1}), respectively, assessed with an accelerometer in a mixed-gender sample (43).

In six studies (20, 22, 31, 37, 41, 43), subjective criterion validity was assessed by validation against a PA record, diary, or logbook. Three of the studies were conducted among a mixed-gender sample (20, 31, 43), two in all-male samples (37, 41), and one among a female sample (22). Except for the studies conducted among males only, all provided strong levels of evidence for their findings (mean 5.5±0.6, range 4–6). The average age of the study populations varied between ~37–63 years, the study samples ranged from 41–166 individuals. Correlations were observed between -0.05–0.92.

The subjective criterion validity of the *work index/activity score* was assessed for the following questionnaires: Atherosclerosis Risk In Community Study/Baecke (ARIC/Baecke), Coronary Artery Risk Development in Young Adults-Physical Activity Questionnaire (CARDIA-PAQ), Health Insurance Plan (HIP), KPAS, Lipid Research Clinics Physical Activity Questionnaire

(LRC), and Minnesota Heart Health Program (MHHP). None showed good criterion validity. The KPAS showed moderate criterion validity among females but only for two items [sitting (r=0.58) and walking (r=0.50)] when validated against similar categories of the PA record (22).

With regard to *energy expenditure*, the subjective criterion validity was assessed for the Coronary Artery Risk Development in Young Adults-Seven Day Recall (CARDIA-SDR), IPAQ-L, PAQ, and TOQ. The work score of the TOQ (activity units per week^{-1}) showed good criterion validity against a PA record (r=0.92) in a mixed-gender sample, in addition to its item "sitting at work" (activity units per week^{-1}; r=0.77) (20). Moderate criterion validity was observed in mixed-gender samples for: the IPAQ-L "work activity" (MET hours per day^{-1}) against a PA logbook (r=0.64) (31) and the TOQ "work" (MET; r=0.52) and "standing at work" (activity units per week^{-1}; r=0.57) against similar activities of the PA record (20).

The subjective criterion validity of *duration of activity* was assessed for the CARDIA-SDR, OPAQ, TOQ, and TCQ. None showed good criterion validity. Individual items of the OPAQ and TOQ showed good and moderate criterion validity against similar activities of a PA record in mixed-gender samples: TOQ "sitting" (hours per week^{-1}) showed good criterion validity (r=0.82) while TOQ "standing" (hours per week^{-1}; r=0.57) (20) and OPAQ "walking" (hours per week^{-1}; r=0.74) (43) showed moderate criterion validity. The TCQ "time at work" (hours per week^{-1}) showed moderate criterion validity against an activity log (r=0.58) among males (41).

Construct validity. Eighteen studies used appropriate measures to assess the construct validity of 23 questionnaires (table 4); the validity of most questionnaires was only assessed in one study each. Eleven studies were conducted in a mixed-gender sample (20, 25, 26, 30, 32, 33, 43, 44, 46–48), four among a male sample (34, 36, 40, 41) and three studied females only (21, 22, 50). The average age of the study populations ranged from ~16–71 years, samples sizes varied from 19–2240. Comparison included those against accelerometers (N=11), VO_2Max (N=10), other questionnaires (N=8), and body composition measures (N=7). Many studies compared the questionnaire against more than one construct validity measure.

The construct validity of the *work index/activity score* was assessed for the ARIC/Baecke, Baecke, CARDIA-PAQ, European Prospective Investigation into Cancer (EPIC), HIP, HUNT-2, KPAS, LRC, MHHP, SGLOA, and SGPOA. Overall, none showed good construct validity. Four questionnaires showed moderate construct validity for their work index/activity score: (i) the EPIC against an accelerometer in a mixed-gender sample (MET hours per week^{-1}; r=0.37) (26); (ii) the Baecke against doubly labeled water (DLW) among males [average daily metabolic rate (ADMR) and physical activity level (PAL): r=0.37 and r=0.52, respectively)] (40); (iii) the HUNT-2 against several accelerometer measures among males (r=-0.45–0.48) (34); and (iv) the SGLOA against several comparison measures among females (50). The KPAS item "compared to others" showed moderate construct validity among females against a PA record (r=0.41) (22). Moderate-to-poor validity was found among females for the SGPOA (50). The *work index/activity score* of the ARIC/Baecke, HIP, LRC, and MHHP showed moderate construct validity against some of their comparison methods, but overall showed poor construct validity (20).

The construct validity of *energy expenditure* was assessed for the adapted IPAQ long version (A-IPAQ-L), CARDIA-SDR, EPAQ2, KIDH-O, IPAQ-L, Modifiable Activity Questionnaire (MAQ), MOSPAQ, Questionnaire d'Activité Physique Saint-Etienne (QAPSE), Sub-Saharan Africa Activity Questionnaire (SSAAQ), and TOQ. None of the questionnaires showed good construct validity. The SSAAQ work (MET per day^{-1}) showed good construct validity but only among a subsample of urban females when validated against a heart rate monitor (r=0.72), in other subsamples it showed moderate construct validity (46). Among females, moderate construct validity was observed for the MAQ "work" (MET hours per week^{-1}) when validated against accelerometer total activity (counts per day^{-1}; r=0.43), but poor against accelerometer sedentary time (hours per week^{-1}; r=-0.19) (32).

The construct validity of *duration of activity* was assessed for the CARDIA-SDR, IPAQ-L, MAQ, MOSPAQ, OPAQ, TOQ, and TCQ. None showed good construct validity. Moderate construct validity was observed in a mixed-gender sample for both the CARDIA-SDR and TOQ when validated against each other (hours per week^{-1}; r=0.47) (20) and for OPAQ's item "heavy labor" (hours per week^{-1}), "against light" and "moderate activities" assessed with a PA record (r=-0.35 and r=0.32, respectively) and OPAQ's item "sitting/standing" against submax HR (r=0.32) (43). Among males, the TCQ's "time work" (hours per week^{-1}) showed moderate construct validity against DLW (ADMR and PAL; r=0.33 and r=0.55 respectively) (40).

Discussion

This is the first systematic review of studies assessing the measurement properties of occupational PA questionnaires, in which both the results and methodological

quality of the included studies have been taken into account. In general, the quality of the studies was not very high, mostly as a result of the use of inadequate measures, such as the use of Pearson and/or Spearman correlation coefficients instead of ICC or K_w as repeatability measures and the use of inadequate reference measures for validation. Moreover, our results show that few questionnaires were tested in more than one study, or tested for both repeatability and validity.

The quality of the studies assessing repeatability was in general poor (50% calculated Pearson or Spearman correlation coefficients instead of ICC or K_w); moreover, the repeatability of only six questionnaires was assessed in a mixed-gender sample. Of these, the *work index* of the BRFSS showed good repeatability, as did *energy expenditure* and *duration of activity* of the IPAQ-L and MOSPA-Q, and *duration of activity* of several of OPAQ's items (ICC 0.76–0.83), all based on a strong level of evidence. The six questionnaires are, however, all very different, varying in length, completion time, form of administration and occupational PA outcome. The choice of any of them depends largely on the purpose for which it will be used. The BRFSS is a single-item surveillance measure, designed to categorize occupational PA into three components and can provide a rapid assessment of occupational PA levels (27, 51). When a more comprehensive analysis of time spent in various occupational categories is desired, the multiple-item OPAQ or MOSPA-Q are more suitable (43, 44). Both are relatively short and take less than five minutes to complete. An advantage of the MOSPA-Q over the OPAQ is that it contains questions on several PA domains and proved to be reliable in assessing both *duration of activity* and *energy expenditure* (44). The OPAQ on the other hand was additionally tested for *duration of activity* of each separate item: "sitting/standing", "heavy labor", and "employed for wages" all showed good repeatability. Like the MOSPA-Q, the IPAQ-L occupational PA was also reliable when expressed in both *duration of activity* and *energy expenditure* and contains questions on several PA domains. However its aim is to assess health-enhancing PA, therefore focusing on moderate and vigorous intensity PA, walking at work is included, but not sitting (31, 47).

Three studies objectively assessed criterion validity in a mixed-gender sample (by comparing occupational PA with similar activities from either an accelerometer or PA record) and provided a high level of evidence for their findings. This small number of studies results from the fact that many of the studies were not primarily designed to assess the validity of the occupational PA measures, but of the PA questionnaires overall. The TOQ was the only questionnaire that showed good criterion validity, namely for *energy expenditure*, for "sitting", both when expressed in *energy expenditure* and *duration of activity* and "standing" when expressed in *duration of activity*, all validated against a PA record. Moreover, moderate construct validity was observed for the TOQ "work" when expressed in *duration of activity* (20). The TOQ, however, proved less valid for assessing "time spent at work" and "walking at work". The TOQ, which is a 29-item self-administered questionnaire and assesses time spent in various types of occupational PA for 3 jobs in the previous year, can be seen as a useful questionnaire for identifying specific occupational PA habits that differ by type and intensity over time (20). However, when time and space allow only a brief assessment, it might be a less suitable questionnaire (the TOQ takes about 20 minutes to complete). Moreover, its questions are limited to occupational PA questions only. An additional questionnaire which showed moderate criterion validity with regard to energy expenditure is the IPAQ-L, when validated against a PA logbook, it also showed good repeatability (31, 47). The IPAQ-L might be a more suitable questionnaire when one is also interested in assessing different domains of PA; however its main focus is on health-enhancing PA.

An important finding of our review was that the criterion- and construct-related validity correlation coefficients were generally low for most comparisons. This could be due to shortcomings of the questionnaires. However, it is not necessarily an indication of poor validity; it might also be a reflection of the complexity of assessing the validity of (occupational) PA questionnaires. One particular problem is the choice of an appropriate comparison instrument. Even though accelerometry was considered an objective criterion standard, accelerometers may not be sensitive enough for evaluating the criterion validity of occupational questionnaires (ie, as a result of their inability to detect upper-body movement while a person is sitting or standing) (43). Using PA records, diaries, or other forms of self-report methods as a validation method might produce higher correlations; however, the risk of correlated error is also higher as both the method under scrutiny and the validation method are subject to the same forms of bias (48). An additional objective criterion measure for occupational PA is observation, which is likely to cover those PA that cannot all be captured with an accelerometer. However, in this review, no observation studies were found in our search. Validating occupational PA measures against physiological constructs (ie, cardiovascular fitness, overall energy expenditure, VO_2Max and body composition), which are related to PA behavior in general, might also not be optimal. As the short time spent in heavy occupational PA and the longer hours spent in sitting and standing activities at work might not be sufficient to result in any changes in the above-mentioned

Table 4. Construct validity of questionnaires assessing occupational physical activity. For full names of questionnaires, see appendix. Evidence was rated as follows: strong evidence (3), moderate evidence (2) and poor evidence (1). Acceptable level of criterion validity rated as: correlation coefficient (r)≥0.60 (+), 0.30≤r<0.60 (±) and r<0.30 (–). Inadequate measure used to assess criterion validity (0). [VO_2Max=maximal oxygen consumption; NS=non-significant; MET=metabolic equivalent task; ADMR=average daily metabolic rate; PAL=physical activity level; AUC=area under the curve; PAR=physical activity ratio; EE=energy expenditure; kcal=kilocalories; kg=kilograms; KJ=kilojoules; km=kilometer; BMI=body mass index; M=males; F=female, 95% CI=95% confidence interval.]

Questionnaire	Study	Study population				Comparison measure	Unit of measurement	Construct validity	Rating
		Sample size	Age (years)	Educational level	Country				
Work index/activity score									
ARIC/ Beacke	Ainsworth et al, 1993 (20)	75	37.3	71% college	US	SDR (activity units per week^{-1})	Work index	r=0.15	3-
						TOQ work score (activity units per week^{-1})	Work index	r=0.09	3-
						LRC (work question points)	Work index	r=0.45 [a]	3±
						MHHP (work question points)	Work index	r=0.16	3-
						HIP (work index activity units)	Work index	r=0.38 [a]	3±
						CARDIA work (activity unit score)	Work index	r=0.02	3-
						VO_2Max	Work index	NS	3-
						% body fat	Work index	NS	3-
						Pulmonary function	Work index	NS	3-
Baecke	Jacobs Jr et al, 1993 (33)	78	37.3	71% college	US	Accelerometer (kcal per day^{-1}; MET minutes per day^{-1})	Work index	kcal: r=-0.03; MET: r=0.11	kcal: 3-; MET: 3-
						4-week histories MET minutes per day^{-1} (total)	Work index	r=0.05	3-
						4-week histories MET minutes per day^{-1} (light; moderate)	Work index	Light: r=-0.07; moderate: r=-0.07	Light: 3-; moderate: 3-
						4-week histories MET minutes per day^{-1} (heavy; household)	Work index	Heavy: r=0.07; household: r=0.01	Heavy: 3-; household: 3-
						Workload 160	Work index	r=0.06	3-
						% body fat	Work index	r=-0.17	3-
						Forced expiratory volume	Work index	r=-0.16	3-
						VO_2 Max	Work index	r=0.23 [b]	3-
	Philippaerts et al, 1999 (40)	M: 19	40	·	Belgium	Doubly labeled water (ADMR; PAL)	Work index	ADMR: r=0.37; PAL r=0.52 [b]	ADMR: 2±; PAL: 2±
	Philippaerts et al, 2001 (41)	M: 166	40	·	Belgium	VO_2-peak (ml per minute^{-1} per kg^{-1})	Work index	r=0.11	2-
						Tracmor (total mean)	Work index	r=0.26 [b]	2-
						Tracmor active leisure (sum; mean)	Work index	Sum: r=-0.06; mean r=0.04	Sum: 2-; mean 2-
						Tracmor quiet leisure (sum; mean)	Work index	Sum: r=0.02; mean r=0.03	Sum: 2-; mean 2-
CARDIA-PAQ	Ainsworth et al, 1993 (20)	75	37.3	71% college	US	SDR (activity units per week^{-1})	Activity unit score	r=-0.06	3-
						TOQ work score (activity units per week^{-1})	Activity unit score	r=0.02	3-
						LRC (work question points)	Activity unit score	r=-0.02	3-
						MHHP (work question points)	Activity unit score	r=0.10	3-
						HIP (work index activity units)	Activity unit score	r=0.12	3-

(continued)

Table 4. Continued

Questionnaire	Study	Study population				Comparison measure	Unit of measurement	Construct validity	Rating
		Sample size	Age (years)	Educational level	Country				
						ARIC/Baecke work index	Activity unit score	r=0.02	3-
						VO_2Max	Activity unit score	NS	3-
						% body fat	Activity unit score	NS	3-
						Pulmonary function	Activity unit score	NS	3-
EPIC	Cust et al, 2008 (26)	182	57.2	·	Australia	Accelerometer MET hours per week^{-1}	Work index	r=0.37 (95% CI 0.22–0.51)	3±
HIP	Ainsworth et al, 1993 (20)	75	37.3	71% college	US	SDR (activity units per week^{-1})	Work index	r=-0.20	3-
						TOQ work score (activity units per week^{-1})	Work index	r=0.15	3-
						LRC (work question points)	Work index	r=0.31 [b]	3±
						MHHP (work question points)	Work index	r=0.26 [b]	3-
						CARDIA work (activity units)	Work index	r=0.12	3-
						ARIC/Baecke work index	Work index	r=0.38 [a]	3±
						VO_2 Max	Work index	NS	3-
						% body fat	Work index	NS	3-
						Pulmonary function	Work index	NS	3-
						Caltrac (kcal per day; MET)	Work index	NS	3-
	Jacobs Jr et al, 1993 (33)	78	37.3	71% college	US	Accelerometer (kcal per day^{-1}; MET minutes per day^{-1})	Work index	kcal: r=0.07; MET: r= 0.14	kcal: 3-; MET: 3-
						4-week histories MET minutes per day^{-1} (total)	Work index	r=0.00	3-
						4-week histories MET minutes per day^{-1} (light; moderate)	Work index	Light: r=0.19; moderate r=-0.17	Light: 3-; moderate 3-
						4-week histories MET minutes per day^{-1} (heavy; household)	Work index	Heavy r=0.04; household r=-0.06	Heavy 3-; household 3-
						Workload 160	Work index	r=0.23 [b]	3-
						% body fat	Work index	r=-0.03	3-
						Forced expiratory volume	Work index	r=0.07	3-
HUNT-2	Kurtze et al, 2007 (34)	M: 108	32.4	·	Norway	VO_2 Max	Work index	r=-0.06	3-
						ActiReg (EE; PAL)	Work index	EE: r=0.39 [a]; PAL: r=0.38 [a]	EE: 3±; PAL: 3±
						ActiReg (MET1–3; MET 3–6; MET 6+)	Work index	MET 1–3: r=-0.45 [a]; MET 3–6: r=0.48 [a]; MET 6+: r=-0.14	MET 1–3: 3±; MET 3–6: 3±; MET 6+: 3-
						IPAQ-s (total vigorous; total moderate)	Work index	Vigorous: r=0.17; moderate: r=0.27 [a]	Vigorous; 3-; moderate: 3-
						IPAQ-s (walking; sitting)	Work index	Walking: r=0.22 [b]; sitting: r=-0.57 [a]	Walking: 3-; sitting: 3±
						IPAQ-s (MET)	Work index	r=0.34 [a]	3±
KPAS	Ainsworth et al, 2000 (22)	F: 50	39.1	78% college	US	Accelerometer (kcal per day^{-1}; MET minutes per day^{-1})	Work index	kcal: r=0.30 [b]; MET: r= 0.16	kcal: 3± MET: 3-
						VO_{2}-peak (ml per minute^{-1} per kg^{-1})	Work index	r= 0.04	3-
						% body fat	Work index	r= -0.06	3-

(continued)

Table 4. Continued

Question-naire	Study	Study population				Comparison measure	Unit of measurement	Construct validity	Rating
		Sample size	Age (years)	Educa-tional level	Country				
						PA record moderate and heavy (>3 MET)	Compared with others	r= 0.41 [a]	3±
LRC	Ainsworth et al, 1993 (20)	75	37.3	71% college	US	SDR (activity units per week^{-1})	Activity score	r=0.20	3-
						TOQ work score (activity units per week^{-1})	Activity score	r=0.13	3-
						HIP (work index activity units)	Activity score	r=0.31 [b]	3±
						MHHP (work question points)	Activity score	r=0.01	3-
						CARDIA work (activity units)	Activity score	r=-0.02	3-
						ARIC/Baecke work index	Activity score	r=0.45 [a]	3±
						VO_2 Max	Activity score	NS	3-
						% body fat	Activity score	NS	3-
						Pulmonary function	Activity score	NS	3-
						Caltrac (kcal per day; MET)	Activity score	NS	3-
MHHP	Ainsworth et al, 1993 (20)	75	37.3	71% college	US	SDR (activity units per week^{-1})	Work score	r=0.11	3-
						TOQ work score (activity units per week^{-1})	Work score	r=0.34 [b]	3±
						HIP (work index activity units)	Work score	r=0.26 [b]	3-
						LRC (work question points)	Work score	r=0.01	3-
						CARDIA work (activity units)	Work score	r=0.10	3-
						ARIC/Baecke work index	Work score	r=0.16	3-
						VO_2Max	Work score	NS	3-
						% body fat	Work score	NS	3-
						Pulmonary function	Work score	NS	3-
						Caltrac (kcal per day; MET)	Work score	NS	3-
	Jacobs Jr et al, 1993 (33)	78	37.3	71% college	US	Accelerometer (kcal per day^{-1}; MET minutes per day^{-1})	Work index	kcal: r=-0.01, MET: r=0.04	kcal: 3-; MET: 3-
						4-week histories MET minutes per day^{-1} (total)	Work index	Total: r=-0.09	Total: 3-
						4-week histories MET minutes per day^{-1} (light; moderate)	Work index	Light: r=0.06; moderate: r=0.07	Light: 3-; moderate: 3-
						4-week histories MET minutes per day^{-1} (heavy; household)	Work index	Heavy: r=-0.07; household: r=-0.10	Heavy:3-; household: 3-
						Workload 160	Work index	r=0.02	3-
						% body fat	Work index	r=0.07	3-
						Forced expiratory volume	Work index	r=-0.04	3-
						VO_2Max	Work index	r=0.00	3-
SGLOA	Wilbur et al, 1989 (50)	F: 43	47 years	·	US	SGPOA	Work score	r=0.73 [c]	1+
						TOQ (MET per hour^{-1})	Work score	r=0.45 [a]	1±
						Beeper work (MET per hour^{-1})	Work score	r=0.449 [a]	1±

(continued)

Table 4. Continued

Questionnaire	Study	Study population				Comparison measure	Unit of measurement	Construct validity	Rating
		Sample size	Age (years)	Educational level	Country				
SGPOA	Wilbur et al, 1989 (50)	F: 43	47	·	US	SGLOA (4-point)	Work score	r=0.73 [c]	1+
						TOQ (MET per hour^{-1})	Work score	r=0.29 [b]	1-
						Beeper work (MET per hour^{-1})	Work score	r=0.29 [b]	1-
Energy expenditure									
A-IPAQ-L	Graff-Iversen et al, 2007 (30)	2240	47.1	12 years	Norway	BMI	Work (MET minutes)	M: r=-0.03 F: r=0.02	M: 2-; F: 2 -
						Waist/hip	Work (MET minutes)	M: r=0.09 [a], F: r=-0.02	M: 2-; F: 2 -
CARDIA-SDR	Ainsworth et al, 1993 (20)	75	37.3	71% college	US	TOQ (MET)	Work (MET)	r=0.24 [b]	3-
						TOQ work score (activity units per week^{-1})	Work (MET)	r=0.38 [a]	3-
						HIP work index (activity units)	Work (MET)	r=-0.20	3-
						LRC (work question points)	Work (MET)	r=0.20	3-
						MHHP (work question points)	Work (MET)	r=0.11	3-
						CARDIA (work question activity units)	Work (MET)	r=-0.06	3-
						ARIC/Baecke work index	Work (MET)	r=0.15	3-
						VO_2Max	Work (MET)	NS	3-
						% body fat	Work (MET)	NS	3-
						Pulmonary function	Work (MET)	NS	3-
						Caltrac (kcal per day, MET)	Work (MET)	NS	3-
	Jacobs Jr et al, 1993 (33)	78	37.3	71% college	US	Accelerometer (kcal per day^{-1}; MET minutes per day^{-1})	Work (MET minutes per day)	kcal: r=-0.01; MET: r=0.10	kcal: 3-; MET: 3-
						4-week histories MET minutes per day^{-1} (total)	Work (MET minutes per day^{-1})	Total: r=-0.16	Total: 3-
						4-week histories MET minutes per day^{-1} (light; moderate)	Work (MET minutes per day^{-1})	Light: r=-0.13; moderate r=-0.16	Light: 3-; moderate 3-
						4-week histories MET minutes per day^{-1} (heavy; household)	Work (MET minutes per day^{-1})	Heavy: r=-0.13; household r=-0.06	Heavy: 3-; household 3-
						Workload 160	Work (MET minutes per day^{-1})	r=0.17	3-
						% body fat	Work (MET minutes per day^{-1})	r=0.05	3-
						Forced expiratory volume	Work (MET minutes per day^{-1})	r=-0.00	3-
						VO_2Max	Work (MET minutes per day^{-1})	r=0.09	3-
	Ainsworth et al, 1999 (21)	F: 46	39.4	78% college	US	Caltrac (kcal, MET minutes)	Work score (MET minutes per week^{-1})	NS	3-
						VO_2Max	Work score (MET minutes per week^{-1})	NS	3-
						% body fat	Work score (MET minutes per week^{-1})	NS	3-
						Caltrac (kcal, MET minutes)	Work <3 MET (MET minutes per week^{-1})	NS	3-

(continued)

Table 4. Continued

Question-naire	Study	Study population				Comparison measure	Unit of measurement	Construct validity	Rating
		Sample size	Age (years)	Educa-tional level	Country				
						VO_2Max	Work <3 MET (MET minutes per week^{-1})	NS	3-
						% body fat	Work <3 MET (MET minutes per week^{-1})	NS	3-
						Caltrac (kcal, MET minutes)	Work 3–6 MET (MET minutes per week^{-1})	NS	3-
						VO_2Max	Work 3–6 MET (MET minutes per week^{-1})	NS	3-
						% body fat	Work 3–6 MET (MET minutes per week^{-1})	NS	3-
EPAQ2	Wareham et al, 2002 (48)	173	M: 58.8; F: 55.5	·	UK	Mean daytime EE (KJ per hour^{-1})	Work activity (MET hours per week^{-1})	r=0.17 [b]	2-
						Mean VO_2Max (ml per minutes^{-1} per kg^{-1})	Work activity (MET hours per week^{-1})	r=0.01	2-
						PAR ≥5(%)	Work activity (MET hours per week^{-1})	r=-0.07	2-
KIHD-O	Lakka & Salonen, 1992 (36)	M: 1834	53.1		Finland	Total activity (MET hours per day^{-1})	Work activity (MET hours per day^{-1})	r=0.516 [c]	2±
						7-day leisure time activity (MET hours per day^{-1})	Work activity (MET hours per day^{-1})	r=-0.127 [c]	2-
						12-month leisure time activity (MET hours per day^{-1})	Work activity (MET hours per day^{-1})	r=-0.139 [c]	2-
						Conditioning activity (km per week^{-1})	Work activity (MET hours per day^{-1})	r=-0.144 [c]	2-
						Conditioning activity (mean MET)	Work activity (MET hours per day^{-1})	r=-0.084 [a]	2-
						VO_2Max (ml per kg^{-1} per minutes^{-1})	Work activity (MET hours per day^{-1})	r=-0.003	2-
IPAQ-L	Vandelanotte et al, 2005 (47)	53	30.9	88.5% high level	Belgium	PA diary (kcal)	Work activity (kcal)	r=0.23	3-
MAQ	Jacobi et al, 2009 (32)	160	41	51% university	France	Accelerometer total activity (counts per day^{-1})	Work (MET hours per week^{-1})	r=0.22 [a]	3-
						Accelerometer total activity (counts per day^{-1})	Work (MET hours per week^{-1})	M: r=0.43 [a] F: r=0.13	M: 3±; F: 3-
						Accelerometer sedentary time (hours per week^{-1})	Work (MET hours per week^{-1})	r=-0.08	3-
						Accelerometer sedentary time (hours per week^{-1})	VO_2Max	M: r=-0.19: F: r=-0.10	M: 3-; F: 3-
MOSPAQ	Roeykens et al, 1998 (44)	167	35	·	Belgium	BMI (kg/m^2)	Work (kcal per week^{-1})	r=0.081	2-
						VO_2-peak (ml per kg^{-1} per minutes^{-1})	Work (kcal per week^{-1})	r=0.216 [a]	2-
						Lean body mass (kg)	Work (kcal per week^{-1})	r=0.199 [b]	2-
						% body fat	Work (kcal per week^{-1})	r=-0.167 [b]	2-
QAPSE	Bonnefoy et al, 1996 (25)	65	M: 71.1; F: 70.5	Middle/ high level	France	Body mass	Work activity (KJ per day^{-1})	r=0.249 [b]	2-
						Skinfold	Work activity (KJ per day^{-1})	r=0.181	2-

(continued)

Table 4. Continued

Questionnaire	Study	Study population				Comparison measure	Unit of measurement	Construct validity	Rating
		Sample size	Age (years)	Educational level	Country				
						Fat free mass	Work activity (KJ per day^{-1})	r=0.264 [b]	2-
						Body fat %	Work activity (KJ per day^{-1})	r=-0.066	2-
						VO_2Max	Work activity (KJ per day^{-1})	r=-0.018	2-
SSAAQ	Sobngwi et al, 2001 (46)	89	19–68	·	Cameroon	Accelerometer (MET per day^{-1})	Work (MET per day^{-1})	F: r=0.42 [b] (rural); M: r=0.40 [b] (rural)	F: 2± (rural) M: 2± (rural)
						Heart rate monitor (AUC beats per day^{-1})	Work (MET per day^{-1})	M: r=0.72 [a] (urban); M: r=0.44 [b] (rural)	M: 2+(urban) M: 2± (rural)
						Heart rate monitor (AUC beats per day^{-1})	Work (MET per day^{-1})	M: r=0.47 [b] (urban); M: r=0.49 [b] (rural)	M: 2±(urban) M: 2± (rural)
TOQ	Ainsworth et al, 1999 (21)	F: 46	39.4	78% college	US	Caltrac (kcal; MET minutes)	Total work (MET minutes per week^{-1})	kcal: NS; MET: NS	kcal: 3-; MET: 3-
						VO_2Max	Total work (MET minutes per week^{-1})	NS	3-
						% body fat	Total work (MET minutes per week^{-1})	NS	3-
						Caltrac (kcal; MET minutes)	Sitting 1.5 MET (MET minutes per week^{-1})	kcal: NS; MET: NS	kcal: 3-; MET: 3-
						VO_2Max	Sitting 1.5 MET (MET minutes per week^{-1})	NS	3-
						% body fat	Sitting 1.5 MET (MET minutes per week^{-1})	NS	3-
						Caltrac (kcal; MET minutes)	Standing 3 MET (MET minutes per week^{-1})	kcal r=0.34 [b]; MET: NS	kcal 3±; MET: 3-
						VO_2Max	Standing 3 MET (MET minutes per week^{-1})	r=0.29 [b]	3-
						% body fat	Standing 3 MET (MET minutes per week^{-1})	NS	3-
						Caltrac (kcal; MET minutes)	Walking 3.5 MET (MET minutes per week^{-1})	kcal: NS; MET: NS	kcal: 3-; MET: 3-
						VO_2Max	Walking 3.5 MET (MET minutes per week^{-1})	NS	3-
						% body fat	Walking 3.5 MET (MET minutes per week^{-1})	NS	3-
						Caltrac (kcal; MET minutes)	<3 MET (MET minutes per week^{-1})	kcal: NS; MET: NS	kcal: 3-; MET: 3-
						VO_2Max	<3 MET (MET minutes per week^{-1})	NS	3-
						% body fat	<3 MET (MET minutes per week^{-1})	NS	3-
						Caltrac (kcal; MET minutes)	3–6 MET (MET minutes per week^{-1})	kcal: NS; MET: NS	kcal: 3-; MET: 3-
						VO_2Max	3–6 MET (MET minutes per week^{-1})	NS	3-
						% body fat	3–6 MET (MET minutes per week^{-1})	NS	3-

(continued)

Table 4. Continued

Question-naire	Study	Study population				Comparison measure	Unit of measurement	Construct validity	Rating
		Sample size	Age (years)	Educa-tional level	Country				
TOQ	Ainsworth et al, 1993 (20)	75	37.3	71% college	US	CARDIA SDR (MET)	Work (MET)	r=0.24 [b]	3-
						SDR (activity units per week^{-1})	Work score (activity units per week^{-1})	r=0.38 [a]	3±
						HIP (work index activity units)	Work score (activity units per week^{-1})	r=0.15	3-
						LRC (work question points)	Work score (activity units per week^{-1})	r=0.13	3-
						MHHP (work question points)	Work score activity units per week-1)	r=0.34 [a]	3±
						CARDIA (work question activity units)	Work score (activity units per week^{-1})	r=0.02	3-
						ARIC/Baecke work index	Work score (activity units per week^{-1})	r=0.09	3-
						VO_2Max	Work score (activity units per week^{-1})	NS	3-
						% body fat	Work score (activity units per week^{-1})	NS	3-
						Pulmonary function	Work score (activity units per week^{-1})	NS	3-
						Caltrac (kcal per day; MET)	Work score	NS	3-
	Wilbur et al, 1989 (50)	F: 43	47	·	US	SGLOA (4-point)	TOQ (MET per hour^{-1})	r=0.45 [a]	1±
						SGPOA	TOQ (MET per hour^{-1})	r=0.29 [b]	1-
						Beeper work (MET per hour^{-1})	TOQ (MET per hour^{-1})	r=0.29 [b]	1-
Duration of activity									
CARDIA-SDR	Ainsworth et al, 1993 (20)	75	37.3	71% college	US	TOQ (hours per week^{-1})	Work (hour per week^{-1})	r=0.47 [a]	3±
IPAQ-L	Vandelanotte et al, 2005 (47)	53	30.9	88.5% high level	Belgium	PA diary (minutes)	Work activity (minute)	r=0.22	3-
MAQ	Jacobi et al, 2009 (32)	160	41	51% university	France	Accelerometer total activity (counts per day^{-1})	Work activity (hour per week^{-1})	r=0.06	3-
							Work activity (hour per week^{-1})	M: r=0.13; F: r=-0.03	M: 3-; F: 3-
						Accelerometer sedentary time (hours per week^{-1})	Work activity (hour per week^{-1})	r=0.20 [b]	3-
							Work activity (hour per week^{-1})	M: r=0.04; F: r=0.14	M: 3-; F: 3-
MOSPAQ	Roeykens et al, 1998 (44)	167	35	·	Belgium	BMI (kg/m^2)	Work (minute per week^{-1})	r=0.081	2-
						VO_2-peak (ml per kg per minutes)	Work (minute per week^{-1})	r=0.210 [a]	2-
						Lean body mass (kg)	Work (minute per week^{-1})	r=0.201 [a]	2-
						% body fat	Work (minute per week^{-1})	r=-0.166 [b]	2-
OPAQ	Reis et al, 2005 (43)	41	38.8	65.9% college	US	OPA record hours per week^{-1} (sitting/standing; walking; heavy)	Work activity (hour per week^{-1})	Sitting/standing: r=0.09; walking: r=0.18; heavy: r=0.19	Sitting/standing: 3-; walking: 3-; heavy: 3-

(continued)

Table 4. Continued

Questionnaire	Study	Study population				Comparison measure	Unit of measurement	Construct validity	Rating
		Sample size	Age (years)	Educational level	Country				
						OPA record hours per week^{-1} (light; moderate; vigorous)	Work activity (hours per week^{-1})	Light: r=0.09; moderate: r=0.09; vigorous: r=0.12	Light: 3-; moderate: 3-; vigorous: 3-
						Accelerometer hours per week^{-1} (light; moderate; vigorous)	Work activity (hours per week^{-1})	Light: r=0.22; moderate: r=0.12; vigorous: r=-0.02	Light: 3-; moderate: 3-; vigorous: 3-
						Submax heart rate	Work activity (hours per week^{-1})	r=0.23	3-
						% body fat	Work activity (hours per week^{-1})	r=-0.05	3-
						OPA record hours per week^{-1} (walking; heavy labor; total)	Sitting/standing (hours per week^{-1})	Walking: r=-0.49; heavy labor r=-0.14; total r=-0.20	Walking: 3±; heavy labor 3-; total 3-
						OPA record hours per week^{-1} (moderate; vigorous)	Sitting/standing (hours per week^{-1})	Moderate: r=-0.33; vigorous: r=-0.09	Moderate: 3± ; vigorous: 3-
						Accelerometer hours per week^{-1} (moderate; vigorous)	Sitting/standing (hours per week^{-1})	Moderate: r=-0.01; vigorous r=0.25	Moderate: 3-; vigorous 3-
						Submax heart rate	Sitting/standing (hours per week^{-1})	r=0.32	3±
						% body fat	Sitting/standing (hours per week^{-1})	r=0.04	3-
						OPA record hours per week^{-1} (sitting/standing; heavy labor, total)	Walking (hours per week^{-1})	Sitting/standing: r=-0.27; heavy labor: r=0.18; total: r=0.43	Sitting/standing: 3-; heavy labor: 3-; total: 3±
						OPA record hours per week^{-1} (light; vigorous)	Walking (hours per week^{-1})	Light: r=0.12; vigorous: r=0.15	Light: 3-; vigorous: 3-
						Accelerometer hours per week^{-1} (light; vigorous)	Walking (hours per week^{-1})	Light: r=0.41; vigorous: r=-0.18	Light: 3±; vigorous: 3-
						Submax heart rate	Walking (hours per week^{-1})	r=-0.17	3-
						% body fat	Walking (hours per week^{-1})	r=-0.11	3-
						OPA record hours per week^{-1} (sitting/standing; walking; total)	Heavy labor (hours per week^{-1})	Sitting/standing: r=-0.13; walking: r=0.20; total: r=0.13	Sitting/standing: 3-; walking: 3-; total: 3-
						OPA record hours per week^{-1} (light; moderate)	Heavy labor (hours per week^{-1})	Light: r=-0.35; moderate: r=0.32	Light: 3±; moderate: 3±
						Accelerometer hours per week^{-1} (light; moderate)	Heavy labor (hours per week^{-1})	Light: r=0.46; moderate: r=0.14	Light: 3±; moderate: 3-
						Submax heart rate	Heavy labor (hours per week^{-1})	r=0.24	3-
						% body fat	Heavy labor (hours per week^{-1})	r=0.10	3-
TOQ	Ainsworth et al, 1999 (21)	F: 46	39.4	78% college	US	Caltrac (kcal; MET minutes)	Sitting 1.5 MET (hours per week^{-1})	kcal: NS; MET: NS	kcal: 3-; MET: 3-
						VO_2Max	Sitting 1.5 MET (hours per week^{-1})	NS	3-
						% body fat	Sitting 1.5 MET (hours per week^{-1})	NS	3-
						Caltrac (kcal; MET minutes)	Standing 3 MET (hours per week^{-1})	kcal: NS; MET: NS	kcal: 3-; MET: 3-

(continued)

Table 4. Continued

Question-naire	Study	Study population				Comparison measure	Unit of measurement	Construct validity	Rating
		Sample size	Age (years)	Educa-tion	Country				
						VO_2Max	Standing 3 MET (hours per week^{-1})	NS	3-
						% body fat	Standing 3 MET (hours per week^{-1})	NS	3-
						Caltrac (kcal; MET minutes)	Walking 3.5 MET (hours per week^{-1})	kcal: NS; MET: NS	kcal: 3-; MET: 3-
						VO_2Max	Walking 3.5 MET (hours per week^{-1})	NS	3-
						% body fat	Walking 3.5 MET (hours per week^{-1})	NS	3-
						Caltrac (kcal; MET minutes)	<3 MET (hours per week^{-1})	kcal: NS; MET: NS	kcal: 3-; MET: 3-
						VO_2Max	<3 MET (hours per week^{-1})	NS	3-
						% body fat	<3 MET (hours per week^{-1})	NS	3-
						Caltrac (kcal; MET minutes)	3–6 MET (hours per week^{-1})	kcal: NS; MET: NS	kcal: 3-; MET: 3-
						VO_2Max	3–6 MET (hours per week^{-1})	NS	3-
						% body fat	3–6 MET (hours per week^{-1})	NS	3-
						Caltrac (kcal; MET minutes)	Total work (hours per week^{-1})	kcal: NS; MET: NS	kcal: 3-; MET: 3-
						VO_2Max	Total work (hours per week^{-1})	NS	3-
						% body fat	Total work (hours per week^{-1})	NS	3-
	Ainsworth et al, 1993 (20)	75	37.3	71% college	US	CARDIA SDR (hours per week)	Work (hours per week^{-1})	r=0.47 [a]	3±
TCQ	Philippaerts et al, 1999 (40)	M: 19	40	·	Belgium	Doubly labeled water (ADMR; PAL)	Time work (hours per week^{-1})	ADMR: r=0.33; PAL: r=0.55 [b]	ADMR: 2± PAL: 2±
						Doubly labeled water (ADMR; PAL)	EE work (kcal per kg per week^{-1})	ADMR: r=0.34; PAL: r=0.55 [b]	ADMR: 2±; PAL:2±
						Doubly labeled water (ADMR; PAL)	Work index (MET)	ADMR: r=0.29; PAL: r=0.44	ADMR; 2-; PAL: 2±
	Philippaerts et al, 2001 (41)	M: 166	40	·	Belgium	VO_2-peak (ml per kg^{-1} per minutes^{-1})	EE work (kcal per kg^{-1} per week^{-1})	r=0.00	2-
						Tracmor (total mean)	EE work (kcal per kg^{-1} per week^{-1})	r=0.28 [b]	2-
						Active leisure	EE work (kcal per kg^{-1} per week^{-1})	Sum: r=0.02; mean: r=0.15	Sum: 2-; mean: 2-
						Quiet leisure	EE work (kcal per kg^{-1} per week^{-1})	Sum: r=-0.23 [b]; mean: r=-0.04	Sum: 2-; mean: 2-
						VO_2-peak (ml per kg^{-1} per minutes^{-1})	Work index (MET)	r=0.14	2-
						Tracmor (total mean)	Work index (MET)	r=0.27 [b]	2-
						Active leisure	Work index (MET)	Sum: r=0.02; mean: r=0.00	Sum: 2-; mean: 2-
						Quiet leisure	Work index (MET)	Sum: r=-0.03; mean: r=-0.04	Sum: 2-; mean: 2-

[a] P<0.01.
[b] P<0.05.
[c] P<0.001.

physiological constructs (21, 52). Moreover a perfect correlation cannot be expected, as many other factors besides PA (including genetic predisposition and environmental conditions) influence changes in constructs, such as body composition and cardiovascular fitness (22). Thus, the observed criterion- and construct-related validity correlations coefficients should be interpreted in light of these considerations. Finally, as no standardized method for assessing PA exists, it is necessary to evaluate several measures of occupational PA.

Limitations of the review

One limitation of this review is a potential publication bias, as our search strategy only located articles that were published in peer-reviewed journals and referenced in electronic databases. Moreover, the inclusion of only English-language articles, may have discarded some studies that could have added relevant information regarding the repeatability and/or validity of (occupational) PA questionnaires, originating from a greater diversity of countries. In addition, despite the abundance of studies that have evaluated the repeatability and/or validity of measures for physical working tasks and/or positions, these are not included in the review. A second limitation concerns the checklist that was used to score the results and methodological quality of the included studies. Others might have chosen different cut-off points for scoring negative or positive on repeatability or validity, for scoring the study design, or for scoring the use of Spearman correlation coefficients in repeatability studies. However, we feel confident with the choices made in our study as they are based on an existing checklist (14, 15). We believe that using the ICC as sole adequate method to quantify repeatability with regard to continuous measures as proposed in the checklist is appropriate, as ICC reduces the risk of overestimation of repeatability, which might occur when using the Pearson correlation coefficient, as the latter does not take systematic differences between the two measurements into account (17). The information provided in the tables makes it possible for the reader to interpret the findings using their own insights.

Recommendations for future research

Our results indicated that few questionnaires were examined for repeatability and/or validity in more than one study and often inadequate measures were used to determine the repeatability or validity. In our opinion, future studies are needed to test the reliability and/or validity of existing questionnaires. Moreover, in order to enhance the quality of occupational research with respect to physical work exposures and health, more insight is needed into optimal comparison methods for validation and optimal ways of assessing occupational PA, by using both self-reported and objective measures.

Concluding remarks

In conclusion, based on our review of the literature on measurement properties of questionnaires measuring occupational PA, there is strong evidence for: (i) good repeatability of the *work index* of the BRFSS, *energy expenditure* and *duration of activity* of the IPAQ-L and MOSPA-Q, and *duration of activity* of the OPAQ and (ii) moderate-to-good validity of *energy expenditure* of the TOQ and IPAQ-L. However, because of the great diversity of the questionnaires and the purpose for which questionnaires will be used, we feel that that no further conclusion can be drawn regarding the best questionnaire. Finally, as a result of the poor criterion validity of the questionnaires, objective measures of occupational PA are needed.

Acknowledgements

This manuscript arises from the project Instruments for Assessing Levels of Physical Activity and Related Health Determinants (ALPHA), which has received funding from the European Union in the framework of the Public Health Programme. Additional personal funding was received from the Swedish Council for Working Life and Social Research (FAS). None of the authors have any professional relationships with companies or manufacturers who will benefit from the results of the present study. The study sponsors had no involvement in the study design or the collection, analysis, and interpretation of data. The authors declare no potential conflicts of interest, real or perceived.

References

1. Warburton DE, Nicol CW, Bredin SS. Health benefits of physical activity: the evidence. CMAJ. 2006;174:801–9.
2. Andersen LB, Schnohr P, Schroll M, Hein HO. All-cause mortality associated with physical activity during leisure time, work, sports, and cycling to work. Arch Intern Med. 2000;160:1621–8.
3. Hu G, Jousilahti P, Borodulin K, Barengo NC, Lakka TA, Nissinen A, et al. Occupational, commuting and leisure-time physical activity in relation to coronary heart disease among middle-aged Finnish men and women. Atherosclerosis. 2007;194:490–7.
4. Kristal-Boneh E, Harari G, Melamed S, Froom P. Association of physical activity at work with mortality in Israeli industrial

employees: the CORDIS study. J Occup Environ Med. 2000;42:127–35.

5. Kruger J, Yore MM, Ainsworth BE, Macera CA. Is participation in occupational physical activity associated with lifestyle physical activity levels? J Occup Environ Med. 2006;48:1143–8.

6. Howley ET. Type of activity: resistance, aerobic and leisure versus occupational physical activity. Med Sci Sports Exerc. 2001;33:S364–9; discussion S419–20.

7. Probert AW, Tremblay MS, Gorber SC. Desk potatoes: the importance of occupational physical activity on health. Can J Public Health. 2008;99:311–8.

8. Greendale GA, Bodin-Dunn L, Ingles S, Haile R, Barrett-Connor E. Leisure, home, and occupational physical activity and cardiovascular risk factors in postmenopausal women: The Postmenopausal Estrogens/Progestins Intervention (PEPI) Study. Arch Intern Med. 1996;156:418–24.

9. Barengo NC, Hu G, Lakka TA, Pekkarinen H, Nissinen A, Tuomilehto J. Low physical activity as a predictor for total and cardiovascular disease mortality in middle-aged men and women in Finland. Eur Heart J. 2004;25:2204–11.

10. Menotti A, Seccareccia F. Physical activity at work and job responsibility as risk factors for fatal coronary heart disease and other causes of death. J Epidemiol Community Health. 1985;39:325–9.

11. Krause N, Brand RJ, Kaplan GA, Kauhanen J, Malla S, Tuomainen T-P, et al. Occupational physical activity, energy expenditure and 11-year progression of carotid atherosclerosis. Scand J Work Environ Health. 2007;33:405–24.

12. Holtermann A, Mortensen OS, Burr H, Søgaard K, Gyntelberg F, Suadicani P. The interplay between physical activity at work and during leisure time – risk of ischemic heart disease and all-cause mortality in middle-aged Caucasian men. Scand J Work Environ Health. 2009;35:466–74.

13. Lamonte MJ, Ainsworth BE. Quantifying energy expenditure and physical activity in the context of dose response. Med Sci Sports Exerc. 2001;33:S370–8; discussion S419–20.

14. De Vries SI, Bakker I, Hopman-Rock M, Hirasing RA, van Mechelen F. Clinimetric review of motion sensors in children and adolescents. J Clin Epidemiol. 2006;59:670–80.

15. De Vries SI, Van Hirtum HW, Bakker I, Hopman-Rock M, Hirasing RA, Van Mechelen F. Validity and reproducibility of motion sensors in youth: a systematic update. Med Sci Sports Exerc. 2009;41:818–27.

16. Welk GJ. Physical activity assessments for health-related research. Champaign (IL): Human Kinetics Publishers, Inc; 2002.

17. Streiner D, Norman G. Health measurement scales: a practical guide to their development and use. New York (NY): Oxford University Press; 2003.

18. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007;60:34–42.

19. Kirshner B, Guyatt G. A methodological framework for assessing health indices. J Chronic Dis. 1985;38:27–36.

20. Ainsworth BE, Jacobs DR Jr, Leon AS, Richardson MT, Montoye HJ. Assessment of the accuracy of physical activity questionnaire occupational data. J Occup Med. 1993;35:1017–27.

21. Ainsworth BE, Richardson MT, Jacobs DR Jr, Leon AS, Sternfeld B. Accuracy of recall of occupational physical activity by questionnaire. J Clin Epidemiol. 1999;52:219–27.

22. Ainsworth BE, Sternfeld B, Richardson MT, Jackson K. Evaluation of the kaiser physical activity survey in women. Med Sci Sports Exerc. 2000;32:1327–38.

23. Baecke JA, Burema J, Frijters JE. A short questionnaire for the measurement of habitual physical activity in epidemiological studies. Am J Clin Nutr. 1982;36:936–42.

24. Batty D. Reliability of a physical activity questionnaire in middle-aged men. Public Health. 2000;114:474–6.

25. Bonnefoy M, Kostka T, Berthouze SE, Lacour JR. Validation of a physical activity questionnaire in the elderly. Eur J Appl Physiol Occup Physiol. 1996;74:528–33.

26. Cust AE, Smith BJ, Chau J, van der Ploeg HP, Friedenreich CM, Amstrong BK, et al. Validity and repeatability of the EPIC physical activity questionnaire: a validation study using accelerometers as an objective measure. Int J Behav Nutr Phys Act. 2008;5:33.

27. Evenson KR, McGinn AP. Test-retest reliability of adult surveillance measures for physical activity and inactivity. Am J Prev Med. 2005;28:470–8.

28. Friedenreich CM, Courneya KS, Bryant HE. The lifetime total physical activity questionnaire: development and reliability. Med Sci Sports Exerc. 1998;30:266–74.

29. Friedenreich CM, Courneya KS, Neilson HK, Matthews CE, Willis G, Irwin M, et al. Reliability and validity of the Past Year Total Physical Activity Questionnaire. Am J Epidemiol. 2006;163:959–70.

30. Graff-Iversen S, Anderssen SA, Holme IM, Jenum AK, Raastad T. An adapted version of the long International Physical Activity Questionnaire (IPAQ-L): construct validity in a low-income, multiethnic population study from Oslo, Norway. Int J Behav Nutr Phys Act. 2007;4:13.

31. Hagströmer M, Oja P, Sjöström M. The International Physical Activity Questionnaire (IPAQ): a study of concurrent and construct validity. Public Health Nutr. 2006;9:755–62.

32. Jacobi D, Charles MA, Tafflet M, Lommez A, Borys JM, Oppert JM. Relationships of self-reported physical activity domains with accelerometry recordings in French adults. Eur J Epidemiol. 2009;24:171–9.

33. Jacobs DR Jr, Ainsworth BE, Hartman TJ, Leon AS. A simultaneous evaluation of 10 commonly used physical activity questionnaires. Med Sci Sports Exerc. 1993;25:81–91.

34. Kurtze N, Rangul V, Hustvedt BE, Flanders WD. Reliability and validity of self-reported physical activity in the Nord-Trondelag Health Study (HUNT 2). Eur J Epidemiol. 2007;22:379–87.

35. Lakka TA, Salonen JT. Intra-person variability of various physical activity assessments in the Kuopio Ischaemic Heart Disease Risk Factor Study. Int J Epidemiol. 1992;21:467–72.

36. Lakka TA, Salonen JT. Physical activity and serum lipids: a cross-sectional population study in eastern Finnish men. Am J Epidemiol. 1992;136:806–18.

37. Norman A, Bellocco R, Bergström A, Wolk A. Validity and reproducibility of self-reported total physical activity – differences by relative weight. Int J Obes Relat Metab Disord. 2001;25:682–8.

38. Orsini N, Bellocco R, Bottai M, Pagano M, Wolk A. Reproducibility of the past year and historical self-administered total physical activity questionnaire among older women. Eur J Epidemiol. 2007;22:363–8.

39. Philippaerts RM, Lefevre J. Reliability and validity of three physical activity questionnaires in Flemish males. Am J Epidemiol. 1998;147:982–90.

40. Philippaerts RM, Westerterp KR, Lefevre J. Doubly labelled water validation of three physical activity questionnaires. Int J Sports Med. 1999;20:284–9.

41. Philippaerts RM, Westerterp KR, Lefevre J. Comparison of two questionnaires with a tri-axial accelerometer to assess physical activity patterns. Int J Sports Med. 2001;22:34–9.

42. Pols MA, Peeters PH, Ocké MC, Bueno-de-Mesquita HB, Slimani N, Kemper HC, et al. Relative validity and repeatability of a new questionnaire on physical activity. Prev Med. 1997;26:37–43.

43. Reis JP, Dubose KD, Ainsworth BE, Macera CA, Yore MM. Reliability and validity of the occupational physical activity questionnaire. Med Sci Sports Exerc. 2005;37:2075–83.

44. Roeykens J, Rogers R, Meeusen R, Magnus L, Borms J, de Meirleir K. Validity and reliability in a Flemish population of the WHO-MONICA Optional Study of Physical Activity Questionnaire. Med Sci Sports Exerc. 1998;30:1071–5.

45. Shapiro S, Weinblatt E, Frank CW, Sager RV. The H.I.P. Study of Incidence and Prognosis of Coronary Heart Disease; preliminary findings on incidence of myocardial infarction and angina. J Chronic Dis. 1965;18:527–58.

46. Sobngwi E, Mbanya JC, Unwin NC, Aspray TJ, Alberti KG. Development and validation of a questionnaire for the assessment of physical activity in epidemiological studies in Sub-Saharan Africa. Int J Epidemiol. 2001;30:1361–8.

47. Vandelanotte C, De Bourdeaudhuij I, Philippaerts RM, Sjöström M, Sallis JF. Reliability and validity of a computerized and Dutch version of the International Physical Activity Questionnaire (IPAQ). J Phys Act Health. 2005:63–75.

48. Wareham NJ, Jakes RW, Rennie KL, Mitchell J, Hennings S, Day NE. Validity and repeatability of the EPIC-Norfolk Physical Activity Questionnaire. Int J Epidemiol. 2002;31:168–74.

49. Wareham NJ, Jakes RW, Rennie KL, Schuit J, Mitchell J, Hennings S, et al. Validity and repeatability of a simple index derived from the short physical activity questionnaire used in the European Prospective Investigation into Cancer and Nutrition (EPIC) study. Public Health Nutr. 2003;6:407–13.

50. Wilbur J, Miller A, Dan AJ, Holm K. Measuring physical activity in midlife women. Public Health Nurs. 1989;6:120–8.

51. Yore MM, Bowles HR, Ainsworth BE, Macera CA, Kohl HWI. Single versus multiple item questions on occupational physical activity. J Phys Act Health. 2006;3:102–11.

52. Holtermann A, Mortensen OS, Burr H, Søgaard K, Gyntelberg F, Suadicani P. Fitness, work and leisure-time physical activity, and ischaemic heart disease and all-cause mortality among men with pre-existing cardiovascular disease. Scand J Work Environ Health. 2010;36(5):366–372.

Received for publication: 31 March 2010

Appendix

Table A. Evaluation checklist for (occupational) physical activity questionnaires. The table was adapted from De Vries et al (14, 15). [r=correlation coefficient]

S1. Sample characteristics (number, gender, age, body mass index, fat%, VO_2max, education, employment)	
≥6 described (at least: number, gender, age)	1
4–5 described	0.5
≤3 described	0
S2. Protocol	
Information on interval (reliability) or comparison measure (validity), setting, and detailed procedure	1
Some information on protocol is missing	0.5
Not clear at all	0
S3. Measurements	
Complete information on questionnaires (occupational outcome), scoring, questions, domains, form of administration, and reference to detailed description if missing	1
Some information on questionnaires is missing	0.5
Very limited information on questionnaires	0
S4. Statistical analyses	
Complete information on statistics (tests), software package, and P-value	1
Some information on statistics (tests), software package, and P-value	0.5
Very limited information on statistics (tests), software package, and P-value	0
R1. Is intra-instrument reliability reported?	
Yes	1
No	0
R2. Is there an adequate measure of intra-instrument reliability?	
Intra-class correlation	1
95% limits of agreement (Bland Altman)	1
Weighted kappa	1
Pearson's product-moment correlation coefficient	0
Spearman's rank correlation coefficient	0
Kendall's tau	0
Other measures	0

(continued)

Table A. Continued

R3. Is there an acceptable level of intra-instrument reliability?	
ICC ≥0.70 (or weighted kappa)	+
0.40≤ICC<0.70	±
ICC<0.40	-
V1. Is 'criterion' validity reported?	
Yes	1
No	0
V2. Is there an adequate measure of criterion validity?	
Sensitivity	1
Specificity	1
Pearson's product-moment correlation coefficient	1
Spearman's rank correlation coefficient	1
95% limits of agreement (Bland Altman)	0.5
Other measures	0
V3. Is there an acceptable level of criterion validity?	
r≥0.75	+
0.50≤ r<0.75	±
r<0.50	-
V4. Is construct validity reported?	
Yes	1
No	0
V5. Is there an adequate measure of construct validity?	
Pearson's product-moment correlation coefficient	1
Spearman's rank correlation coefficient	1
t-test	1
Mann-Whitney U-test	1
Chi-square test	1
95% limits of agreement (Bland Altman)	0.5
Other measures	0
V5. Is there an acceptable level of construct validity?	
r ≥0.60	+
0.30≤r<0.60	±
r<0.30	-

Table B. List of evaluated questionnaires

Abbreviation	Full name
A-IPAQ-L	Adapted International Physical Activity Questionnaire-long version
ARIC/Baecke	Atherosclerosis Risk In Community Study/Baecke
Baecke	Beacke
BRFSS	Behavioural Risk Factor Surveillance System
Cadbury	Cadbury
CARDIA-PAQ	Coronary Artery Risk Development In Young Adults-Physical Activity Questionnaire
CARDIA-SDR	Coronary Artery Risk Development In Young Adults-Seven Day Recall
EPAQ2	EPIC Physical Activity Questionnaire
EPIC	European Prospective Investigation into Cancer
HIP	Health Insurance Plan
HUNT 2	Nord-Trondelag Health Study
IPAQ-L	International Physical Activity Questionnaire -long version
KIHD-O	Kuopio Ischemic Heart Disease Occupational Physical Activity Interview
KPAS	Kaiser Physical Activity Survey
LRC	Lipid Research Clinics Physical Activity Questionnaire
LTPAQ	Lifetime Total Physical Activity Questionnaire
MAQ	Modifiable Activity Questionnaire
MHHP	Minnesota Heart Health Program
MOSPA-Q	MONICA Optimal Study of Physical Activity
OPAQ	Occupational Physical Activity Questionnaire
PAQ	Physical Activity Questionnaire
pre-EPIC	Pre-European Prospective Investigation into Cancer
PYTPAQ	Past Year Total Physical Activity Questionnaire
QAPSE	Questionnaire d'Activité Physique Saint-Etienne
SGLOA	Saltin & Grimby Lifetime Occupational Activity
SGPOA	Saltin & Grimby Present Occupational Activity
SMC-PAQ	Swedish Mammography Cohort-Physical Activity Questionnaire
SSAAQ	Sub-Saharan Africa Activity Questionnaire
TOQ	Tecumseh Self-Administered Occupational Physical Activity Questionnaire
TCQ	Tecumseh Community Health Study Questionnaire