REVIEW ARTICLE

Sports Med 2010; 40 (7): 539-563
0112-1642/10/0007-0539/$49.95/0

Physical Activity Questionnaires for Youth

A Systematic Review of Measurement Properties

Mai J.M. Chinapaw,[1] *Lidwine B. Mokkink,*[2] *Mireille N.M. van Poppel,*[1] *Willem van Mechelen*[1] and *Caroline B. Terwee*[2]

1 Department of Public and Occupational Health, the EMGO Institute for Health and Care Research, VU University Medical Center, Amsterdam, the Netherlands

2 Department of Epidemiology and Biostatistics, the EMGO Institute for Health and Care Research, VU University Medical Center, Amsterdam, the Netherlands

Contents

Abstract

Because of the diversity in available questionnaires, it is not easy for researchers to decide which instrument is most suitable for his or her specific demands. Therefore, we systematically summarized and appraised studies examining measurement properties of self-administered and proxy-reported physical activity (PA) questionnaires in youth.

Literature was identified through searching electronic databases (PubMed, EMBASE using 'EMBASE only' and SportDiscus®) until May 2009. Studies were included if they reported on the measurement properties of self-administered and proxy-reported PA questionnaires in youth (mean age <18 years) and were published in the English language. Methodological quality and results of

included studies was appraised using a standardized checklist (qualitative attributes and measurement properties of PA questionnaires [QAPAQ]).

We included 54 manuscripts examining 61 versions of questionnaires. None of the included questionnaires showed both acceptable reliability and validity. Only seven questionnaires received a positive rating for reliability. Reported validity varied, with correlations between PA questionnaires and accelerometers ranging from very low to high (previous day PA recall: correlation coefficient [r] = 0.77). In general, PA questionnaires for adolescents correlated better with accelerometer scores than did those for children.

From this systematic review, we conclude that no questionnaires were available with both acceptable reliability and validity. Considerably more high-quality research is required to examine the validity and reliability of promising PA questionnaires for youth.

Physical activity (PA) is an important behaviour related to a number of health outcomes in children and adolescents.[1,2] Accurate assessment of PA levels is important, not only to understand the association between PA and health, but also to monitor secular trends in behaviour and to evaluate the effectiveness of interventions.[3] Therefore, valid, reliable and responsive instruments that measure PA are needed. Questionnaires are a commonly used method to estimate (change in) total amount of daily or weekly PA.[4,5] Other popular PA measures include movement counters and heart-rate monitoring.

PA questionnaires are easy to administer, relatively inexpensive and acceptable to study participants.[6] Furthermore, in some situations, self-reports may be the only feasible method to be used in large-scale population surveys due to available resources. While objective methods such as heart-rate monitoring and accelerometry may better capture the duration and intensity of PA, they provide no information about the type of PA behaviour or in what context and where the activity was performed (e.g. active transport, sports, school).

In past decades, numerous questionnaires have been developed for different populations, including children and adolescents, with major differences in length, type of activities and recall period used. Recalling PA is a highly complex cognitive task requesting information about PA performed at some point in the past, with recall periods varying from 1 day to 1 week or 'a usual week'. Youth are less likely to make accurate self-report assessments than adults because of developmental differences, especially in the ability to think abstractly and perform detailed recall.[7,8] In addition, youth have an activity pattern that is much more variable and intermittent than that of adults.[9] Therefore, PA questionnaires may suffer from recall bias, especially in youth.

Selection of an appropriate PA questionnaire depends not only on the specific purpose of the study (e.g. discrimination, evaluation, prediction), but also the characteristics of the population and the outcome of interest. Other critical considerations in the choice of a questionnaire are the relative importance of practical issues such as study size and budget, as well as reliability, validity and responsiveness. Because of the diversity in available questionnaires, it is not easy for researchers to decide which instrument is most suitable for his or her specific demands.

Therefore, the aim of the present review is to summarize primary studies on measurement properties (reliability, construct validity and responsiveness) of self-report questionnaires that have been developed or modified for assessing PA in children and adolescents. This is one of a series of articles on measurement properties of PA questionnaires published in *Sports Medicine*.

1. Methods

1.1 Literature Search

Literature searches were performed in PubMed, EMBASE (using 'EMBASE only') and

in SportDiscus® (complete databases up until May 2009) on the topic of self-report questionnaires of PA.

The full search strategy in PubMed is presented as follows: 'exercise'[mesh] OR 'physical activity'[tiab] OR 'motor activity'[mesh] AND 'questionnaire'[mesh] OR 'questionnaire*'[tiab]. Limits: 'humans'.

In EMBASE and SportDiscus®, 'physical activity' and 'questionnaire' were used as free-text words, and in EMBASE this was complemented with the EMTREE term 'exercise'.

1.2 Eligibility Criteria

We used the following inclusion criteria: (i) the aim of the study should be to evaluate the measurement properties of a self-report questionnaire; (ii) the aim of the questionnaire should be to measure PA in youth (average age of the study population <18 years) [PA was defined as any bodily movement produced by skeletal muscles that results in energy expenditure above resting level]; (iii) the questionnaire could be used to measure PA in youth in the general population; (iv) the article was published in the English language; and (v) information should be provided on at least one of the measurement properties of the self-report questionnaire in a sample of youth.

We included information on measurement properties only if it was intentionally collected or calculated to assess the measurement properties of the particular self-report questionnaire. We included proxy-report questionnaires but excluded PA interviews or diaries. We also excluded studies that evaluated the measurement properties of a self-report questionnaire in a specific population, such as patients or obese youth.

1.3 Selection of Papers

Two independent reviewers (MC and LM) performed abstract selection, selection of full-text articles, data extraction and quality assessment. Disagreements were discussed and resolved. We retrieved the full-text paper of all abstracts that fulfilled the inclusion criteria and of abstracts that did not contain measurement properties but indicated that they were presented in the full-text paper.

1.4 Data Extraction

1.4.1 Description of Questionnaires

We extracted data from the included papers using a standardized data-extraction form based on a standard checklist for appraising the qualitative attributes and measurement properties of PA questionnaires (QAPAQ).[10] The following data were extracted: (i) the target population for which the questionnaire was developed; (ii) the construct(s) that the questionnaire intends to measure (e.g. habitual PA); (iii) the dimensions of PA that the questionnaire is measuring (e.g. frequency, duration and intensity); (iv) the type of activities that the questionnaire is measuring (e.g. sport, recreational, transport, school, household activities and other); (v) the number of questions; (vi) the recall period that the questions refer to; and (vii) the scoring algorithm (which includes the type and number of scores that were calculated, e.g. total energy expenditure or minutes of activity per day).

In addition, we extracted and rated the methods and results based on the QAPAQ. Reliability, validity and responsiveness depend on the setting and the population in which they are assessed. Therefore, a clear description of the design of each individual primary study, including characteristics of the study population, design issues such as time interval, sample size and data analysis, was required in order to receive a positive rating. Furthermore, if any methodological weakness in the design or execution of the primary study was found (e.g. small sample size, inadequate time interval between test and retest), the evaluated measurement property was rated as 'indeterminate'.

1.4.2 Reliability

Reliability was rated as positive (+), negative (–) or indeterminate (?), depending on the methods and results of the primary studies. According to the QAPAQ checklist, the preferred method is the intraclass correlation coefficient (ICC), or Kappa for dichotomous data or weighted Kappa for ordinal data. An ICC of >0.70 is considered

acceptable.[11] The use of Pearson correlation coefficients is considered inadequate because systematic errors are neglected;[12,13] however, most studies included in this review did calculate Pearson correlation coefficients. We considered it too conservative to rate all these studies as 'indeterminate', as Pearson correlations >0.80 would likely result in ICCs >0.70. Therefore, we decided to rate studies that reported a Pearson correlation >0.80 as positive. Pearson correlations <0.70 would never result in ICCs ≥0.70, and were consequently rated as 'negative'. The time interval between the test and retest should have been described and should be short enough to ensure that subjects had not changed their PA levels, but long enough to prevent recalling the previous answers. We defined an adequate time interval as follows:

- >1 day but <3 months for questionnaires recalling a usual week;
- >1 day but <2 weeks for questionnaires recalling the previous week;
- >1 day but <1 week for questionnaires recalling the previous day, assuming that the two tests recall the exact same day.

A positive score was given if the study population consisted of at least 50 participants; the ICC or Kappa or Pearson correlation was above the specified cut-off point (ICC >0.70; Kappa >0.70; Pearson >0.80) and the time interval between test and retest was adequate. If the correlation was below the specified cut-off point, a negative score was given. If the sample size was <50 participants or the time interval inadequate, the score was rated as 'indeterminate'. We sorted the questionnaires based on (i) outcome measures (ICC, Kappa, correlation) – highest to lowest; and (ii) sample size ≥50 and <50 (table I).

1.4.3 Construct Validity

We initially intended to use the preferred method for assessing construct validity from the QAPAQ, stating that hypotheses about expected correlations between the questionnaire under study and other measures, or about expected differences in scores on the questionnaire between specific groups of subjects, should be defined in advance when testing validity. Almost none of the studies included in this review formulated hypotheses *a priori*. The 'best' method to use for comparison depends on what the questionnaire is aiming to measure. Instead of rating all questionnaires as 'indeterminate', we did not rate the questionnaires but instead sorted the studies based on (i) the comparison instrument (accelerometer, doubly labelled water, direct observation, pedometer, heart rate monitor, other); (ii) the outcome measures – highest to lowest; and (iii) the sample size (table II).

1.4.4 Responsiveness

Responsiveness refers to the ability of an instrument to detect change over time in the construct to be measured.[64] It should be considered an aspect of validity in a longitudinal setting.[64] Since we included only one study reporting on responsiveness, this study was not rated, but is briefly described in the results section.

2. Results

The literature search yielded a total of 21 891 hits: 9733 in PubMed, 7601 in EMBASE and 4284 in SportDiscus®. We included 54 manuscripts examining 61 versions of questionnaires (see figure 1).

2.1 Description of Questionnaires

Table III presents a description of the included questionnaires. We sorted the questionnaires on target population: preschool children (mean age <6 years); children (mean age >6 and <12 years); and adolescents (mean age >12 and <18 years). We found six questionnaires for assessing PA in preschool children, all of which had been completed by proxy report, 25 questionnaires for children and 31 that had been developed for adolescents. The construct of what the questionnaire intends to measure was mostly broadly described as 'physical activity', sometimes limited to certain types of activities. The dimensions that were measured were duration, frequency, intensity or a combination of these dimensions. Because of the different dimensions used, the unit of measurement of the questionnaires differed.

Table I. Reliability of physical activity (PA) questionnaires for youth sorted by level of evidence

Questionnaire[a]	Study population[b]	Time interval	Results	Rating
Preschoolers (mean age <6 y)				
CLASS (proxy)[14]	n = 58 Sex: 63% ♀ Age: 5.3 (0.5) [5–6]	At least 14 d	MVPA/VPA/total PA frequency: proxy 5–6 y: ICC = 0.74/0.87/0.83 MVPA/VPA/total PA duration: proxy 5–6 y: ICC = 0.49/0.81/0.76 % agreement total PA 89.2; total VPA 58.6; total MPA 84.2	Proxy +
NPAQ (proxy)[15]	n = 72 Sex: 55% ♀ Age: 5.7 (0.5) [NR]	2–8 wk	NPAQ total (collapsed into low, moderate, high) weighted Kappa: 0.39 (0.22–0.56); ICC: 0.70 (0.58–0.87); Spearman r = 0.61	+/–
CPAR (proxy)[16]	n = 27 Sex: 38% ♀ Age: 4.9 (0.7) [4–5]	7 d	ICC (one-way factor ANOVA): MVPA 0.39; PAEE 0.25	Ind
Children (mean age >6 and <12 y)				
PAQ-C[17]	Study 1 n = 215 Sex: 42% ♀ Age: [9–15] Study 2 n = 84 Sex: 51% ♀ Age: [9–14]	1 wk	ICC: ♂ 0.75; ♀ 0.82	+
GAQ[18]	n = 68 Sex: 100% ♀ Age 9.0 (0.6) [8–10]	4 d	ICC 28 activities: yesterday 0.78; usual 0.82 18 activities: yesterday 0.70; usual 0.79	+
CLASS (self-report and proxy)[14]	n = 111 Sex: 27% ♀ Age: 10.6 (0.8) [10–12]	Children 7 d, proxy at least 14 d	ICC MVPA/VPA/total PA frequency: self-report 0.75/0.42/0.36; proxy 10–12 y 0.67/0.75/0.69 MVPA/VPA/total PA duration: self-report 0.37/0.41/0.24; proxy 10–12 y 0.58/0.62/0.74 % agreement total PA 89.2; total VPA 58.6; total MPA 84.2	Self-report and proxy +/–
CLASS (self-report and proxy)[19]	n = 112 Sex: 63% ♀ Age: 10.6 (0.76) [9–13]	Children within 7 d, proxy ±14 d	ICC proxy report/self-report (in 10–12 y only) Frequency MPA: ICC = 0.75/0.97; frequency VPA: ICC = 0.74/0.58; duration MPA: ICC = 0.73/0.43; duration VPA: ICC = 0.31/0.29; proxy/self-report: 10/7 of 30 PA items Kappa >0.70	Proxy +/–; self-report –
GAQ[20]	n = 172 Sex: 100% ♀ Age: 8.8 (0.8) [8–10] Race: African American	12 wk	ICC: 28 activities: yesterday 0.59; usual 0.59 18 activities: yesterday 0.57; usual 0.55	Yesterday Ind; usual –
Daughter questionnaire[21]	n = 69 Sex: 100% ♀ Age: 9.9 (8.5–12.7)	Girls: 12–16 d	ICC: walking schoolday/weekend 0.48/0.32; exercise schoolday/weekend 0.36/0.32	–
Modified SAPAC[22]	n = 103 Sex: 50% ♀ Age: 11.7 (0.5)	Minimum 5 d	Total PA ICC ♂ 0.20; ♀ 0.19	–
Stairs score[23]	n = 84 Sex: 100% ♀ Age: 11.1 (1.54) [7–15]	11 mo	Spearman r: 0.59 (95% CI 0.43–0.71)	–

Continued next page

Table I. Contd

Questionnaire[a]	Study population[b]	Time interval	Results	Rating
Specific activity score[23]	n = 84 Sex: 100% ♀ Age: 11.1 (1.54) [7–15]	11 mo	Spearman r: 0.53 (0.36–0.67)	–
Godin-Shephard (proxy)[23]	n = 84 Sex: 100% ♀ Age: 11.1 (1.54) [7–15]	11 mo	Spearman r: 0.48 (0.30–0.63)	–
Activity-rating instrument[24]	n = 30 Sex: 50% ♀ Age: 11.2 (2.0) [7–15]	1 mo	ICC: PA rating 0.85	Ind
CPAR[25]	n = 22 Sex: 50% ♀ Age: 11.8 (1.0)	1–2 wk	ICC: total EE 0.95; activity EE 0.82	Ind
Mother and father Questionnaire[21]	n = 47/35 mother/father Sex: 100% ♀ Age: 9.9 (8.5–12.7)	12–28 d	ICC: mother/father schoolday light 0.11/0.32; MPA 0.28/0.24; VPA 0.72/0.75 ICC: mother/father weekend light 0.32/0.12; MPA 0.33/0.13; VPA 0.65/0.72	Ind Ind
PAQ[26]	n = 24 children Age: 8–11 y (3rd–5th grade)	Ind	Frequency section: short PAQ r = 0.82; long PAQ r = 0.49	Ind
Older children and adolescents (mean age >12 y)				
QAPACE[27]	n = 121 Sex: 54% ♀ Age: [8–16]	6 wk	Pearson ICC: 0.96 (0.95–0.97); LOA −515.5 and 532.5 kJ × 24^{-1} h, mean figures 7566 kJ/d^{-1} h	+
OPAQ[28]	n = 87 Sex: 45% ♀ Age: 13.1 (0.9)	1 wk	ICC: MPA 0.76; VPA 0.80; MVPA 0.91	+
Refined 60-min MVPA screening measure[29]	n = 73 Sex: 65% ♀ Age: 12.1 (0.9)	Same d up to 1 mo	ICC: 0.77 (0.76 with time to retest as a co-variate); same day 0.88 (n = 42), up to 1 mo 0.53 (n = 31); Kappa: 61%, same day 84%, up to 1 mo 36%	+
WHO HBSC[30]	n = 71 Sex: 56% ♀ Age: 14.9 (1.6) [13–18]	8–12 d	Frequency: ICC = 0.73; duration: ICC = 0.71	+
Epidemiological questionnaire[31]	n = 100 Sex: 53% ♀ Age: [15–18]	1 mo	Spearman rank r 1 mo: h/wk 0.79; MET h/wk 0.85; VPA h/wk 0.91 (1 y: 0.66, 0.72, 0.72, respectively)	1 mo +; 1 y –
3DPAR[32]	n = 71 Sex: 68% ♀ Age: 12.5 (1.1)	1 d	Pearson r MVPA 0.68; VPA 0.83 % agreement in activities mentioned: ♂ 51%; ♀ 47%	MVPA –; VPA +
APARQ[33]	Sample 1 n = 121 Sex: 48% ♀ Age: 13.7 (0.4) Sample 2 n = 105 Sex: 29% ♀ Age: 15.7 (0.4)	2 wk	ICC total EE ♂ grade 8/10: summer 0.30/0.79, winter 0.49/0.52 ♀ grade 8/10: summer 0.52/0.86, winter 0.36/0.91 Weighted Kappa (vigorous, adequate and inactive) ♂ grade 8/10: summer 0.33/0.62, winter 0.39/0.59 ♀ grade 8/10: summer 0.55/0.71, winter 0.71/0.58	Grade 8 –; grade 10 +/–
PA screening measure[29]	n = 250 Sex: 56% ♀ Age: 14.6 (1.4)	2 wk	Nine scores: 20-min bout typical wk/past 7 d/composite; accumulate 30-min typical wk/past 7 d/composite; accumulate 60-min typical wk/past 7 d/composite: ICC range 0.55–0.79; Kappa % 45–61	–

Continued next page

Table I. Contd

Questionnaire[a]	Study population[b]	Time interval	Results	Rating
Fels PAQ[34]	n = 229 Sex: 57% ♀ Age: [7–19]	6 d	ICC Total 0.48–0.68; sport 0.62–0.71; leisure 0.60–0.76; work/chore 0.49–0.65	–
SAPAC[32]	n = 66 Sex: 71% ♀ Age: 12.5 (1.1)	1 d	Pearson r MVPA 0.67; VPA 0.63 % agreement in activities mentioned: ♂ 34%; ♀ 42%	–
IPAQ[35]	n = 200 Age: 16 (0.4)	2 wk	Spearman r and ICC Total 0.45/0.37; LPA 0.44/0.28; MPA 0.33/0.15; VPA 0.52/0.40	–
PAQA[35]	n = 158 Age: 16 (0.4)	2 wk	Spearman r and ICC Total 0.48/0.40; LPA 0.22/0.28; MPA 0.32/0.12; VPA 0.57/0.40	–
IPAQ[30]	n = 71 Sex: 56% ♀ Age: 14.9 (1.6) [13–18]	8–12 d	VPA: d/wk: ICC = 0.54; and min/d: ICC = 0.30 MPA: d/wk: ICC = 0.55; and min/d: ICC = 0.34; walking: d/wk: ICC = 0.62; and min/d: ICC = 0.10; sitting: min/d: ICC = 0.27	–
WHO HBSC[36]	Sample 1 n = 121 Sex: 48% ♀ Age: 13.7 (0.4) Sample 2 n = 105 Sex: 29% ♀ Age: 15.7 (0.4)	2 wk	Kappa: frequency 0.36–0.60; duration 0.22–0.58; combination 0.12–0.70 (1 × Kappa = 0.70 boys, y 10, two categories)	–
YPAQ[16]	Sample 1 n = 25 Sex: 30% ♀ Age: 13.1 (0.3) [12–13] Sample 2 n = 24 Sex: 70% ♀ Age: 17.1 (0.6) [16–17]	7 d	ICC (one-way factor ANOVA) Group 12–13 y: MVPA 0.92; PAEE 0.86 Group 16–17 y: MVPA 0.73; PAEE 0.79	Ind
CHASE[16]	n = 25 Sex: 30% ♀ Age: 13.1 (0.3) [12–13]	7 d	ICC (one-way factor ANOVA): lifestyle score 0.64	Ind
SWAPAQ[16]	n = 24 Sex: 70% ♀ Age: 17.1 (0.6) [16–17]	7 d	ICC (one-way factor ANOVA) MVPA 0.05; PA EE 0.02	Ind

a See table IV for definitions of questionnaire names/acronyms.

b Age is presented as mean years (SD) [range].

EE = energy expenditure; **ICC** = intraclass correlation coefficient; **Ind** = indeterminate score; **LOA** = limits of agreement; **LPA** = light-intensity PA; **MPA** = moderate-intensity PA; **MVPA** = moderate to vigorous PA; **NR** = not reported; **r** = correlation coefficient; **VPA** = vigorous PA; + indicates positive score; – indicates negative score; +/– indicates some positive and some negative scores; ♀ indicates female; ♂ indicates male.

The recall period was variously 'the previous day', 'a usual day or week' or 'the past year'.

2.2 Reliability

Table I summarizes the reliability studies. Thirty-five questionnaires were tested for reliability (three among preschoolers, 14 among children and 17 among adolescents). The time interval between the first and second administration varied from the same day to 11 months. Only seven questionnaires[14,17,18,27-30] received a positive rating for reliability: in preschoolers, the CLASS (see table IV for a full list of definitions

Table II. Construct validity of physical activity (PA) questionnaires for children and adolescents sorted by comparison measure, outcome and sample size

Questionnaire[a]	Study population[b]	Comparison measure	Results
Preschool children (mean age <6 y)			
CPAQ (proxy)[16]	n = 27 Sex: 38% ♀ Age: 4.9 (0.7) [4–5]	Accelerometer (MVPA); DLW (PA EE)	Accelerometer Spearman r = 0.42; DLW Spearman r = 0.22, wide ratio LOA
NPAQ (proxy)[15]	n = 204 Sex: 55% ♀ Age: 5.7 (0.5)	Accelerometer (MTI)	Accelerometer total/vigorous counts: rho = 0.33/0.36
Parental report – outdoor time checklist (proxy)[37]	n = 250 Sex: 43% ♀ Age: 44 mo [29–52] Country: USA	Accelerometer (RT3 triaxial research tracker); recall questionnaire	Accelerometer: r = 0.33; recall: r = 0.57
Parental report – outdoor time recall questionnaire (proxy)[37]	n = 250 Sex: 43% ♀ Age: 44 mo [29–52] Country: USA	Accelerometer (RT3 triaxial research tracker); checklist questionnaire	Accelerometer: r = 0.20; checklist: r = 0.57
CLASS (proxy)[14]	n = 58 Sex: 63% ♀ Age: 5.3 (0.5) [5–6]	Accelerometer (MTI)	Accelerometer MPA/VPA/total PA/total counts/d: r = –0.06/–0.04/–0.04/0.05
Questionnaire to teachers (proxy)[38]	n = 49 Sex: 51% ♀ Age: [5–6]	Direct observation; pedometer	Direct observation: r = –0.19–0.27; pedometer: r = 0.25
Questionnaire to mothers (proxy)[38]	n = 49 Sex: 51% ♀ Age: [5–6]	Direct observation; pedometer	Direct observation: r = –0.14–0.12; pedometer: r = 0.14
Primary school children (mean age >6 and <12 y)			
SNAP[39]	n = 121 Sex: 60% ♀ Age: 10.7 (2.2)	Accelerometer (GT1M)	Mean difference between SNAP and accelerometer: –9 min (–23, 5) mean difference in proportions complying to 60 min/d MVPA guideline 0.02; 90% CI –0.08, 0.12)
PA Questionnaire for parents and teachers[40]	n = 62 Sex: 48% ♀ Age: 7.0 (0.7)	Accelerometer (Caltrac); other (HR monitor)	Accelerometer: r = 0.53; HR: r = 0.40
ACTIVITY[41]	n = 47 Sex: 60% ♀ Age: 7.7 (0.45)	Accelerometer (Caltrac); HR monitor	Accelerometer (CNTSMIN): r = 0.40; HR: average activity/50%; HR reserve: 0.17/0.51
CPAR[25]	n = 45 Sex: 56% ♀ Age: 11.8 (1.0)	Accelerometer (Tritrac)	TEE/AEE vs accelerometer: r = 0.51/0.20; % agreement = 78%; Kappa = 0.398 categorizing in active/inactive
MARCA[42]	n = 66 Sex: 50% ♀ Age: 11.6 (0.8)	Accelerometer (MTI)	6 of 7 hypotheses correct PAL/VPA/min locomotion: r = 0.45/0.35/0.37
PAQ-C[43]	n = 97 Sex: 58% ♀ Age: 11.3 (1.39) [9–14]	Accelerometer (Caltrac); questionnaire (7-d recall interview, activity rating; leisure-time exercise (Godin 1 and 2); fitness test (Chester step test)	Five hypotheses: moderate correlations with all measures Accelerometer: r = 0.39; 7-d recall interview: r = 0.46, 0.43; activity rating: r = 0.57; leisure-time exercise (Godin 1 and 2): r = 0.41, –0.57; fitness test: r = 0.28; sex differences: none

Continued next page

Table II. Contd

Questionnaire[a]	Study population[b]	Comparison measure	Results
PAQ-C[44]	n = 449 Sex: 52% ♀ Age: 11.2 (0.3)	Accelerometer (Actigraph, MVPA); hip BMC; spine BMC; whole body BMC	Spearman r: accelerometer: r = 0.38 (boys) and r = 0.24 (girls); Partial Spearman rank r: hip BMC: r = 0.28 (boys), r = 0.08 (girls); spine BMC r = 0.19 (boys), r < 0.01 (girls); whole body BMC: r = 0.22 (boys), r = 0.08 (girls)
SAPAC[45]	n = 125 Sex: 56% ♀ Age: 10.9 (0.53)	Accelerometer; HR monitor; PA interview	Accelerometer min MVPA/MVPA METs/weighted; MVPA METs/no. of activities: r = 0.30/0.32/0.32/0.02; HR min MVPA/MVPA METs/weighted; MVPA METs/no. of activities: r = 0.58/0.60/0.59/0.28; PACI min MVPA/MVPA METs/weighted; MVPA METs/no. activities: r = 0.64/0.65/0.65/0.47
GAQ[18]	n = 68 Sex: 100% ♀ Age: 9.0 (0.6) [8–10]	Accelerometer (MTI/Computer Science and Applications, Inc.)	Accelerometer 18 activities yesterday: r = 0.28; 18 activities usual d: r = 0.30
PAQ[7]	n = 52, grade 3 Sex: % ♀ Age: NR Race: American Indian	Accelerometer (Tritac)	Accelerometer before and after school: r = 0.15; during school r = 0 0.41
OPAQ[28]	n = 51 Sex: 47% ♀ Age: 12.6 (0.5)	Accelerometer (Caltrac)	Spearman r: MPA: 0.01; VPA: 0.33; MVPA: 0.32
PAQ[26]	n = 24 Sex: NR Age: grade 3–5 (NR) [NR]	Accelerometer (Caltrac)	Short PAQ vs accelerometer: r = 0.27; long PAQ vs accelerometer: r = 0.13
Health Survey for England PA Questionnaire (proxy)[46]	n = 130 Sex: 51% ♀ Age: 7.0 (0.3) [6–7]	Accelerometer	LOA: −131–376 min/d Spearman r: 0.16
Self-report PA Questionnaire for Schoolchildren[47]	n = 34 Sex: 100% ♂ Age: 10.8 (0.8) Country: Japan	Accelerometer (Actiwatch); other (life recorder)	Accelerometer: regression coefficients for counts/d ranging from −0.25 to 0.07
CLASS (self-report and proxy)[19]	n = 112 Sex: 63% ♀ Age: 10.6 (0.76) [9–13]	Accelerometer (MTI)	Proxy vs accelerometer: MVPA r = 0.01/0.18; self-report vs accelerometer MPA/VPA: −0.11/0.15
CLASS (self-report and proxy)[14]	n = 111 Sex: 27% ♀ Age: 10.6 (0.8) [10–12]	Accelerometer (MTI)	Proxy vs accelerometer MPA/VPA/total PA/total counts/d: r = 0.07/0.24/0.09/0.11; self-report vs accelerometer MPA/VPA/total PA/total counts/d: r = 0.02/−0.04/−0.04/0.06
Modified Godin-Shephard[48]	n = 24 Sex: 50% ♀ Age: [10–13]	Accelerometer (Caltrac)	School d: questionnaire = 17.2 + 1.16 × Caltrac (±98.5); weekend d: questionnaire = 68.5 + 1.26 × Caltrac (±164.8)
GAQ[20]	n = 172 Sex: 100% ♀ Age: 8.8 (0.8) [8–10] Race: African American	Accelerometer (Computer Science and Applications, Inc.)	18 activities yesterday vs accelerometer: 06:00–12:00/12:00–18:00: r = 0.06/0.03; 18 activities usual d vs accelerometer: 06:00–12:00/12:00–18:00; r = 0.12/0.11

Continued next page

Table II. Contd

Questionnaire[a]	Study population[b]	Comparison measure	Results
Activity rating instrument[24]	n = 30 Sex: 50% ♀ Age: 11.2 (2.0) [7–15]	Accelerometer (Computer Science and Applications, Inc.); other questionnaire	Accelerometer average movement count/frequency VPA: r = –0.03/0.04
PAQ-C[49]	Study 1 n = 991 Sex: 49% ♀ Age: 10.7 (0.5) Study 2 n = 414 Sex: 49% ♀ Age: 8.7 (0.57)	Cardiovascular fitness (modified Harvard Step test)	Study 1: fitness test: r = 0.08; Study 2: fitness test: European American/African American: 0.30/0.02
PAQ-C[43]	n = 89 Sex: 57% ♀ Age: 11.06 (1.46) [8–13]	Questionnaire (athletic competence); other (behavioural conduct, activity rating, teacher's rating, MVPA)	Athletic competence: r = 0.48; behavioural conduct: r = 0.16, activity rating: MVPA: 0.53, MVPA >10 min: 0.41, teacher's rating: r = 0.45, PAQ-C + MVPA: r = 0.53, 0.41; sex differences: p < 0.05
Modified Godin-Shephard (proxy)[23]	n = 479 Sex: 100% ♀ Age: 11.1 (1.54) [7–15]	Other questionnaires	Perspire score: r = 0.4; Stair score: r = 0.2; specific activity score: r = 0.38
Specific activity score[23]	n = 471 Sex: 100% ♀ Age: 11.1 (1.54) [7–15]	Other questionnaires	Perspire score/Stair score/Godin-Shephard: r = 0.3/0.10/0.38
Stairs score[23]	n = 479 Sex: 100% ♀ Age: 11.1 (1.54) [7–15]	Other questionnaires	Perspire score: r = 0.17; Godin-Shephard: r = 0.2; Specific activity score: r = 0.1
Daughter Questionnaire[21]	n = 69 Sex: 100% ♀ Age: 9.9 [8.5–12.7]	Questionnaire (mother and father version); activity diary	Activity diary: ICC = 0.19–0.52 (for 10 sub-scores)
Older children and adolescents (mean age >12 y)			
PDPAR[50]	Sample 1 n = 48 Age: grade 7–12 Sample 2 n = 26 Sex: 46% ♀ Age: [15–18]	Accelerometer (Caltrac); pedometer; HR monitor	Caltrac r = 0.77 (n = 48); pedometer: r = 0.88 (n = 48); HR: r = 0.63 (n = 26)
SAPAQ[51]	n = 50 Sex: 62% ♀ Age: 16.8 (0.4)	Accelerometer (MTI)	Total volume of self-reported PA (total MET-min); time spent in PA/total counts/total counts per min per d/time spent sedentary: 0.51/0.49/0.45/–0.45
SHAPES[52]	n = 53 Sex: 53% ♀ Age: [6–12] Country: Canada	Accelerometer (MTI)	MPA = 0.31; VPA = 0.25; MVPA = 0.44; EE from MVPA = 0.44
PAQ-A[53]	Study 1 n = 49 Sex: 43% ♀ Age: 13.5 (0.3) Study 2 n = 210 Sex: % ♀ Age: at test 11 y, at retest 13 y (same children different measures)	Study 1: activity monitor (Actigraph): total PA, percent d MVPA; Study 2: PAQ-C	Study 1: Spearman r: total PA/percent d MVPA, original score: 0.47/0.49, rescaled score: 0.56/0.63; Study 2: Spearman r; PAQ-C: original/rescaled score: 0.30/0.39

Continued next page

Table II. Contd

Questionnaire[a]	Study population[b]	Comparison measure	Results
Refined 60-min MVPA[29]	n = 138 Sex: 65% ♀ Age: 12.1 (0.9)	Accelerometer (Computer Science and Applications, Inc.)	Accelerometer r = 0.40
PAQ-A[54]	n = 85 Sex: 52% ♀ Age: 16.25 (1.51) [13–20]	Accelerometer (Caltrac), questionnaire (7-d recall interview, activity rating, leisure-time exercise)	Hypotheses: PAQ-A would be moderately correlated with all other PA measures Accelerometer: r = 0.33 7-d recall interview (PAR and PAR h): r = 0.59/0.51; activity rating: r = 0.73; leisure-time exercise (Godin 1/2): 0.57/–0.62
PA screening measure[29]	n = 57 Sex: 65% ♀ Age: 13.9 (1.7)	Accelerometer (Computer Science and Applications, Inc.)	VPA (typical wk/past 7-d/composite): r = 0.31/0.36/0.37; 30-min MPA: r = 0.20/0.26/0.26 (NS); 60-min MPA: r = 0.46/0.37/0.47
Fels PAQ[34]	n = 229 Sex: 57% ♀ Age: [7–19] Country: USA	Accelerometer (Actiwatch)	Elementary/middle/high school Total: r = 0.32/0.12/0.11; sport: r = 0.32/0.07/0.34; leisure: r = 0.28/0.28/0.20; work: r = 0.08/–0.13/–0.08
SAPAC[32]	n = 107 Sex: 70% ♀ Age: 12.5 (1.1)	Accelerometer (Actigraph)	MVPA/VPA: r = 0.24/0.28
3DPAR[32]	n = 130 Sex: 66% ♀ Age: 12.5 (1.1)	Accelerometer (Actigraph); questionnaire (SAPAC)	Different cut-off points: MVPA/VPA: r = 0.28 (0.31)/0.16 (0.19)
Modified Godin-Shephard[55]	n = 114 Sex: 60 ♀ Age: grades 6–8 Country: USA	Accelerometer (MTI)	Strenuous/moderate: r = 0.23/0.13
PAQA[35]	n = 188 Sex: NR Age: 16 (0.4)	Accelerometer (MTI)	Spearman r: total: 0.27; LPA: 0.20; MPA: 0.18; VPA: 0.24
IPAQ[35]	n = 188 Sex: NR Age: 16 (0.4)	Accelerometer (MTI)	Spearman r: total: 0.21; LPA: 0.14; MPA: –0.01; VPA: 0.29
FPACQ[56]	n = 33 Sex: 70% ♀ Age: 14.4 (1.4) [12–18]	Accelerometer (Computer Science and Applications, Inc.)	Ranging from r = –0.22 (sports participation at school) to r = 0.78 (frequency hard activities)
YRBS[55]	n = 114 Sex: 60% ♀ Age: grades 6–8 Country: USA	Accelerometer (MTI)	Accelerometer: r = 0.10
HAQ[57]	n = 683 Sex: 100% ♀ Age: [9–19]	Accelerometer (Caltrac)	Caltrac (past 3 d): r = 0.09
PAQ[58]	n = 260 Sex: 100% ♀ Age: 13.4 (1.1) [11–15]	Accelerometer (Caltrac); 3-d diary	Caltrac: r = 0.12/0.26, Kappa: 1/0, % agreement: 33/48%; 3 d diary: r = 0.57/0.16, Kappa: 0.15/0, % agreement: 43/26%

Continued next page

Table II. Contd

Questionnaire[a]	Study population[b]	Comparison measure	Results
SWAPAQ[16]	n = 24 Sex: 70% ♀ Age: 17.1 (0.6) [16–17]	Accelerometer (MTI) and DLW	Accelerometer: Spearman r = 0.23 DLW: Spearman r = 0.40, wide ratio LOA
YPAQ[16]	n = 25 Sex: 30% ♀ Age: 13.1 (0.3) [12–13] n = 24 Sex: 70% ♀ Age: 17.1 (0.6) [16–17]	Accelerometer (MTI) and DLW	12–13 y: accelerometer/DLW: Spearman r = 0.42/0.09; 16–17 y: accelerometer/DLW: Spearman r = 0.11/0.46, wide ratio LOA
CHASE[16]	n = 25 Sex: 30% ♀ Age: 13.1 (0.3) [12–13]	Accelerometer (MTI) and DLW	MVPA-accelerometer/DLW: Spearman r = 0.12/0.45, wide ratio LOA
IPAQ-A (long version)[59]	n = 248 Sex: 49% ♀ Age: [12–17]	Accelerometer (Actigraph, MTI)	MPA/total Actigraph Spearman rank r = 0.15/0.20; MPA: LOA: 12–14 y: −283–149 min/d; 15–17: −186–170 min/d; VPA: LOA: 12–14 y: −120–64 min/d; 15–17: −101–59 min/d
OPAQ[28]	n = 51 Sex: 47% ♀ Age: 12.6 (0.5)	Accelerometer (Caltrac)	Spearman rank-order correlation: MPA 0.01; VPA 0.33; MVPA 0.32
7D-PAR[60]	n = 27 Sex: 48% ♀ Age: 13.0 (1.2) [12–15]	Continuous monitoring of the HR (Polar Precision Performance 3.0 HR monitor) HR >140 bpm = MPA, HR >160 bpm = VPA	MPA: Kappa = 0.02; Pearson r = 0.05 VPA: Kappa = 0.20; Pearson r = 0.37 (n = 25)
7-d recall questionnaire[61]	n = 93 Sex: 51% ♀ Age: 12.2 (0.3)	HR monitor; 7-d interview; Godin-Shephard questionnaire	HR >159 bpm MPA/VPA: 0.30/0.34; interview: little concordance; modified Godin-Shephard: r = 0.38
MONICA survey[62]	n = 125–223 Sex: ? % ♀ Age: [9–19]; Pedometer sample n = 223; sport act n = 125; BMI n = 221; MONICA n = 220	Pedometer (Pedoboy); $\dot{V}O_{2max}$	Pedoboy: r = 0.22, n = 223; $\dot{V}O_{2max}$: r = 0.17, n = 220; weekly sports act in club: r = 0.55, n = 125
QAPACE[27]	n = 36 Sex: 50% ♀ Age: 12(2.6) [8–16]	Aerobic fitness: indirect $\dot{V}O_{2peak}$, by Lèger test, direct $\dot{V}O_{2peak}$, by ergo-spirometry	DEE vs indirect/direct $\dot{V}O_{2peak}$: ICC = 0.56/0.69
APARQ[33]	n = 1072 Sex: 48% ♀ Age: 13.1 n = 954 Sex: 45% ♀ Age: 15.1	20 metre shuttle run test	Grade 8: ♂/♀: r = 0.15/0.21; grade 10: ♂/♀: r = 0.14/0.39
PA and Exercise questionnaire[63]	n = 745 Sex: 54% ♀ Age: 14.3 (1.2)	2.4 km walk-run test	Walk-run test: r = 0.21
Epidemiological questionnaire[31]	n = 100 Sex: 53% ♀ Age: [15–18]	Fitness tests (BMI, 1 mile run, sit and reach, pull-ups, grip strength); 4 × past wk questionnaire; roster	Fitness tests ranging from −0.47 to 0.25; H/wk: questionnaire: r = 0.63/0.76 ('92, '93); MET-h/wk, questionnaire: r = 0.68/0.83 ('92, '93); VPA h/wk questionnaire: r = 0.76/0.84 ('92, '93)

Continued next page

Table II. Contd

Questionnaire[a]	Study population[b]	Comparison measure	Results
WHO HBSC[36]	Sample 1 n = 1072 Sex: 48% ♀ Age: 13.1 Sample 2 n = 954 Sex: 45% ♀ Age: 15.1	20 metre shuttle run test	Active group had significantly higher aerobic fitness than inactive group
Modified Godin-Shephard (leisure-time exercise questionnaire)[61]	n = 93 Sex: 51% ♀ Age: 12.2 (0.3)	HR monitor; other questionnaires	MPA modified Godin-Shephard 7-d recall: r = 0.38; other correlations: low

a See table IV for definitions of questionnaire names/acronyms.

b Age is presented as mean years (SD) [range].

AEE = activity-related energy expenditure; **b** = regression coefficient; **BMC** = bone mineral content; **BMI** = body mass index; **bpm** = beats per minute; **CNTSMIN** = counts per minute; **DEE** = daily energy expenditure; **DLW** = doubly labelled water; **EE** = energy expenditure; **HR** = heart rate; **ICC** = intraclass correlation coefficient; **LOA** = limits of agreement; **LPA** = light-intensity PA; **MET** = metabolic equivalent; **MPA** = moderate-intensity PA; **MVPA** = moderate- to vigorous-intensity PA; **NR** = not reported; **NS** = not significant; **PAL** = physical activity level; **PAR** = 7-day PA recall kilocalorie energy expenditure index; **r** = correlation coefficient; **TEE** = total energy expenditure; $\dot{V}O_{2max}$ = maximum oxygen uptake; $\dot{V}O_{2peak}$ = peak oxygen uptake; **VPA** = vigorous-intensity PA; ? indicates unknown or unclear; ♀ indicates female; ♂ indicates male.

for all questionnaire acronyms mentioned throughout this article) questionnaire (ICC = 0.49–0.87)[14] was the most reliable; in children, the most reliable questionnaires were the GAQ,[18] which recalled 28 activities in a usual week (ICC = 0.82), and the PAQ-C (ICC = 0.75 and 0.82 for boys and girls, respectively);[17] and, in adolescents, the most reliable instruments were

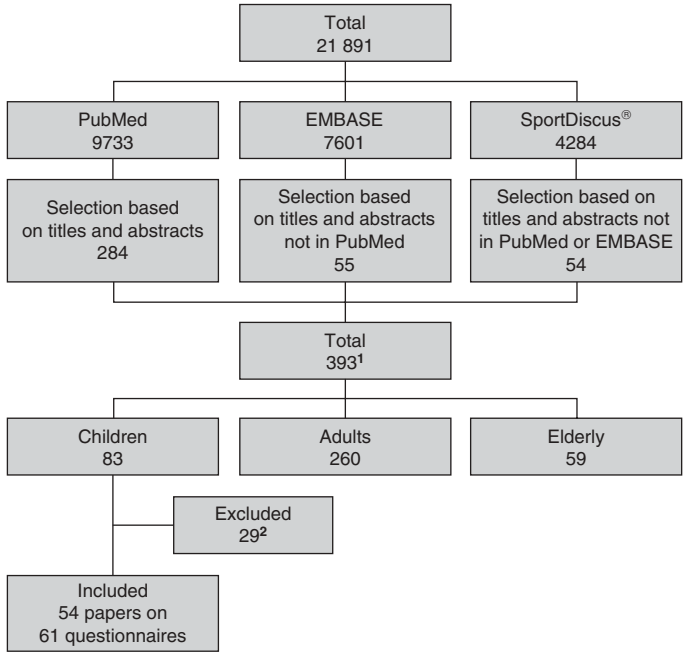

Fig. 1. Flowchart of study inclusion. **1** One paper appears in both the review for adults and for the elderly; **2** The main reason for exclusion was an interview instead of self-report.

Table III. Description of physical activity (PA) questionnaires for youth

Questionnaire[a]	Target population	Construct			Format		
		construct	setting	recall period	dimensions	no. of questions	scores
Preschoolers (mean age <6 y)							
Questionnaire to mothers (proxy),[38] V	Kindergarten children	Habitual activity level	All	In general	None	1	Activity score (1–5)
Questionnaire to teachers (proxy),[38] V	Kindergarten children	Habitual activity level	All	In general	None	1	Activity score (1–5)
Parental report – outdoor time checklist (proxy),[37] V	Preschoolers	Time playing outdoors	Recr, school	24 h (wake-up to bedtime)	D	2	Min activities Range: 0–24 min
Parental report – outdoor time recall questionnaire (proxy),[37] V	Preschoolers	Time playing outdoors	Recr, school	Typical wk/ weekend d in the last mo	D	2	Average daily time (in min) spent playing outdoors
NPAQ (proxy parents or teachers),[15] R, V	Young children (4–7 y)	Usual activity patterns	All	Previous 6 mo	None	7 (+1 on TV/ watching video	Activity score (0–5), watching TV
CPAQ (proxy),[16] R, V	Children (4–5 y)	Mode, frequency and duration of PA and sedentary activities	All	Past 7 d	F, D	?	MVPA, PA EE
Children (mean age >6 and <12 y)							
PA questionnaire for parents and teachers (proxy),[40] V	Preadolescent children (4–8 y)	PA	All	Previous d	D	12 parents, 15 and 16 for teachers	Amount of daily MVPA in min
SAPAC,[45] V	Fifth graders (10–11 y)	MVPA	All	Previous d	F, D	Checklist format (21 activities)	No. of activities, min of PA, volume of PA, volume of PA including intensity ratings, min sedentary pursuits, min MVPA, PA MET scores, weighted MET score
Modified SAPAC,[22] R	Primary school children	PA	Sport, recr, school	Previous d	D	Checklist format (24 activities)	Light, MPA, VPA, total PA + TV/video, computer use, total sedentary activity
ACTIVITY,[41] V	Young children (<10 y)	PA	All	Previous school d	I	10	Activity score: potential range 0–1396

Continued next page

Table III. Contd

Questionnaire[a]	Target population	Construct			Format		
		construct	setting	recall period	dimensions	no. of questions	scores
GAQ,[20] R, V	African American girls (8–10 y)	PA	Sport, recr, trans, school, home	Yesterday and usual d	F, D	28×4+sedentary activities	Activity score
GAQ,[18] R, V			All			4×8+7 questions about sedentary activities	Total PA score, (weighted) MET values, GAQ summary score
Daughter questionnaire (proxy),[21] R, V	Non-obese preadolescent girls	Patterns of PA	All	Typical school d and typical weekend d (24 h)	D	Three timetables (school, weekend, TV)	Hrs/d sitting, standing, walking, exercising, TV/VCR/video games
Mother and Father questionnaire (proxy),[21] R	Non-obese preadolescent girls	Daily activity level	All	Typical school d and typical weekend d (24 h)	D	?	Hrs/d sleeping, sitting, light PA/MPA/VPA, TV/VCR/video games
PAQ-C modified,[44,53] V	Children (8–14 y)	MVPA	Sports and leisure	Previous 7 d	F	9 (1–5 scale), 28 activities	Original PAQ-C summary score (averaged of the sum of the nine items); rescaled PAQ-C summary
PAQ-C,[43,49] V	Children (9–15 y)	MVPA	Sport, recr, school	Previous 7 d	F	9	Activity score
PAQ-C,[17] R	Older children (9–15 y/grades ≥4)	Habitual MVPA	All	Previous 7 d	F	10	Activity score+checklist, PE class, recess, lunch, after school, evening, weekend, described best, wk summary
MARCA,[42] V	Children and adolescents	Activity behaviour, i.e. use of time and daily EE	All	1 d recall	D, I	Segmented-d format (web-based)	PAL, time spent above a given MET level, time spent lying down, sitting, standing or in locomotion, no. of min and estimated energy cost for any activity or set of activities, time distribution of any activity or set of activities
Self-report PA Questionnaire for Schoolchildren,[47] V	Primary school children (9–11 y)	PA and outdoor playing	Sport, recr (playing outdoors)	General wk	F, I	4+section on TV watching and video games	Participation in sports club, PA intensity, frequency of PA, preferences for PA
CLASS,[14,19] R, V	Primary school children	Usual PA	All	Usual weekday and weekend d, typical wk	F, D	30 activities+6	Frequency MPA, frequency VPA, duration MPA, duration VPA intensity (min/wk)
Modified Godin-Shephard,[48] V	Schoolchildren (10–13 y)	PA	All	Previous d	D	Checklist format	TEE, kcal without the resting metabolic rate
Modified Godin-Shephard (proxy),[23] R, V	Schoolchildren	Habitual PA	Sport, recr, exercise during free time	Past y, usual wk	F, I	1	Weekly average of the no. of times they engaged in strenuous, moderate or mild exercise for >15 min during their free time over the last y

Continued next page

Table III. Contd

Questionnaire[a]	Target population	Construct			Format		
		construct	setting	recall period	dimensions	no. of questions	scores
Modified Godin-Shephard (leisure-time exercise questionnaire),[61] V	Schoolchildren	Leisure-time PA	All	Previous wk	F	?	Frequency hard, moderate, easy
Activity-rating instrument,[24] R, V	Children (7–15 y)	Usual PA in last 3 mo	?	3 mo	None	1	Activity level (1–7), activity level compared with peers
Specific activity score,[23] R, V	Girls (7–15 y)	Habitual PA	Sport (11 types)	Past y	F, D, I	?	Average weekly TEE over past y
PAQ,[26] R, V	Elementary school children and their parents	Usual activity patterns	?	?	F, D	Checklist with 22 activities	Activity score
Stairs score,[23] R, V	Girls (7–15 y)	No. of flights of stairs climbed daily	Trans.	Past y	F	1	5-point scale/no.
CHASE,[16] R, V	Primary school children living in the UK	Mode and frequency of PA and sedentary activities	All	?	F	25	Lifestyle score
Health Survey for England PA Questionnaire (parent-report),[46] V	Parents of British children	Habitual level of MVPA	All outside school	Previous 7 d	?	?	MVPA min per d
SNAP,[39] V	Children and adolescents	Physical and sedentary activities	Sedentary, structured, household chores and play, trans	Previous 24 h	D, I	Web-based, segmented-d format	MVPA
Older children and adolescents (mean age >12 y)							
Self-administered 7-day recall questionnaire,[61] V	Modified for children (12 y)	MPA/VPA	All	Normal 7-d period	D	?	Vigorous and moderate no. of h
SHAPES,[52] V	Schoolchildren	MVPA	All	Previous 7 d	F, D	10	Min/d VPA/MPA, MVPA, PAL, weekly screen time, EE on MVPA
Pathway PA recall questionnaire (PAQ),[7] R, V	Children and adolescents	PA	All (standard list of common activities)	Previous 24 h	F	Checklist format	No. of activities reported, frequencies of different types of activities, intensity

Continued next page

Table III. Contd

Questionnaire[a]	Target population	Construct			Format		
		construct	setting	recall period	dimensions	no. of questions	scores
CPAR,[25] R, V	Youth (middle school)	Sedentary and PA	All	Previous d	D	Checklist	Min activities/d, activity-related EE
PDPAR,[50] V	Youth (high school)	PA	Sports, recr, trans, home, (after school h, i.e. 1500–2330)	Previous d	I	35 activities to be filled in 30-min blocks	TEE, EE during specific periods of time, EE in specific activities, no. of 30-min block >4 MET
3DPAR,[32] R, V	Adolescents	Daily PA patterns	All	Previous 3 d	F	50 activities with main activity to be filled in 30-min blocks	No. of blocks MVPA (≥3 METs) or VPA (≥6 METs) per d
SAPAC,[32] R, V	Adolescents	Daily PA patterns	All	Previous 3 d	D	50 activities	No. of min MVPA (≥3 METs) or VPA (≥6 METs) per d
PAQ-A,[54] V	Adolescents	General levels of PA during the school y	Sports, recr, school (PE and lunch recess)	Last 7 d	F	9	Range: 1–5
SWAPAQ,[16,51] R, V	Adolescents	PA	Leisure time, trans, school	Last 7 d	F, D, I	25	Total min of self-reported PA and total MET min, MVPA
YRBS,[55] V	Youth	Participation in strenuous PA	All	Previous wk	F	1	No. of d
APARQ,[33] R, V	Adolescents	PA	Sport, recr, trans	Normal wk	F, D	4	EE and activity score
FPACQ,[56] V	Adolescents aged 12–18 y	PA	All, except PE	Usual wk	F, D	?	H/d and MET-hrs trans and sports; h/wk using TV and computer; sport-intensity index (MET); F/wk VPA; d/wk MPA
Modified Godin - Shephard Questionnaire,[55] V	Middle school-aged children	Participation in leisure-time exercise	All	Average wk	F	3	D/wk strenuous, moderate and mild PA during school y and summer
WHO HBSC,[30,36] R, V	Schoolchildren, children and adolescents	PA, time spent being vigorously active outside school h	Sports, recr (outside school h)	Usually (in a wk)	F, D	2	Frequency score, duration score, combination score
MONICA survey,[62] V	Children and adolescents (9–19 y)	Habitual PA	All	Previous wk/past 12 mo	F, D	?	No. of sport activities/sessions performed in last wk/no. of min of PA inducing sweating per d

Continued next page

Table III. Contd

Questionnaire[a]	Target population	Construct			Format		
		construct	setting	recall period	dimensions	no. of questions	scores
PA and Exercise Questionnaire,[63] V	Singapore primary and secondary schoolchildren	PA patterns	Sport, recr, all	Current PA level, previous 14 d PA level, annual sports participation/events	F, D	5 (1 + 4 multiple-choice questions)	Activity scores: d of hard exercise, d of easy exercise, TV, video computer h, no. of sports played (annual), activity grouping
Fels PAQ,[34] R, V	Children (7–19 y)	Habitual PA	Sport, recr, trans, home	Past y	F	8	Activity score, and sport, leisure and work index
HAQ,[57] V	Girls (10–18/19 y)	Habitual PA	Sports, recr, school sports	Past y	F	?	Activity score, MET times/wk
Epidemiological questionnaire,[31] R, V	Adolescents	Leisure-time PA	Sports, recr (leisure time)	Past y	F, D	Table format	MET h/wk, VPA h/wk
MVPA screening measure,[29] R, V	Adolescents in primary-care setting	Meeting guideline for PA	All	Previous 7 d and typical wk	F, D, I	6 (2 VPA, 4 MPA)	Meeting guidelines for healthy activity/fitness
Refined 60-min MVPA,[29] R, V	Adolescents	Meeting guideline for PA	All (not described)	Previous 7 d, usual/typical wk	F	2	Meeting guidelines for healthy activity (d/wk)
Weight-bearing PAQ,[58] V	Girls (11–15 y)	Level of weight-bearing activities	Sport, recr, school, home	Average weekly time in previous mo	D	58	(Corrected) energy score (min * METS and weight-bearing score (min * WEIGHT factor), and high active/medium active/low active
QAPACE,[27] R, V	Youngsters in Bogota	Daily PA	All	Past y	F, D	18	Daily energy expenditure
IPAQ (short version),[30,35] R, V	Adults	PA	VPA, MPA, walking	Habitual or past wk	F, D	?	VPA: d/wk and min/d MPA: d/wk and min/d Walking: d/wk and min/d MET min/d
PAQA,[35] R	Adolescents	PA	?	Habitual wk	F, D	?	MET min/d, LPA (sitting/sleeping), MPA/d, VPA/d
PAQ-A,[53] V	High school students (14–18 y)	MVPA	Sports and leisure	Past 7 d	F	9 (1–5 scale), 28 activities	PAQ-A summary score (original or rescaled)
YPAQ,[16] R, V	Schoolchildren (12–17 y)	Mode, frequency and duration of PA and sedentary activities	All	Past 7 d	F, D	47 activities	MVPA, PA EE

Continued next page

Table III. Contd

Questionnaire[a]	Target population	Construct			Format		
		construct	setting	recall period	dimensions	no. of questions	scores
SNAP,[39] R	Children and adolescents	Physical and sedentary activities	Sedentary, structured, home play and trans	24-h recall	D, I	Web-based, segmented-d format	MVPA
IPAQ-A (long version)[59] V	Adolescents	All dimensions of health-enhancing PA	School, trans, home, recr	Last 7 d	F, D, context	?	D/wk and F/d walking, MPA and VPA; min/d walking, VPA, MPA; MET min/d as a measure of total health-enhancing activity; daily PA (MET min/d)
7D-PAR,[60] V	Children and adolescents	PA	All	Previous 7 d	D	?	?
OPAQ,[28] R, V	Secondary school students	MVPA	All	Previous wk	D, F	Timetable format	MPA, VPA, MVPA

a See table IV for definitions of questionnaire names/acronyms.

D = duration; **EE** = energy expenditure; **F** = frequency; **home** = home-based activities (household and gardening); **I** = intensity; **kcal** = kilocalories; **LPA** = light-intensity PA; **MET** = metabolic equivalent; **MPA** = moderate-intensity PA; **MVPA** = moderate- to vigorous-intensity PA; **PAL** = physical activity level; **PE** = physical education; **R** = reliability data available; **recr** = recreational; **trans** = transport; **TEE** = total energy expenditure; **TV** = television; **V** = validity data available; **VCR** = video cassette recorder; **VPA** = vigorous-intensity PA; ? indicates not specified or unclear; * indicates multiplication.

the QAPACE (ICC = 0.96)[27] and the OPAQ (ICC = 0.76–0.91).[28]

2.3 Construct Validity

Table II summarizes the studies on construct validity. Construct validity was assessed for seven questionnaires among proxies of preschoolers, 25 questionnaires among children and 31 among adolescents. Construct validity was mostly evaluated by correlations between the questionnaire and accelerometers (n = 46). In preschool children, the highest correlation with accelerometers was found for the CPAQ (r = 0.42)[16] and the NPAQ (r = 0.33 and 0.36 for total activity and vigorous activity, respectively).[15] In primary school children, the highest correlations with an accelerometer were found for the Physical Activity Questionnaire for Parents and Teachers[40] (r = 0.53) and the ACTIVITY[41] (r = 0.40). Another questionnaire, the SNAP,[39] found a mean difference of –9 minutes between the SNAP and an accelerometer. In adolescents, the highest correlations with an accelerometer were found for the PDPAR (r = 0.77)[50] and the SAPAC (r = 0.51).[51]

2.4 Responsiveness

Responsiveness of PA questionnaires was studied for only one questionnaire: the HAQ.[57] For this questionnaire, there was a parallel trend in the pattern of the decline in activity among the HAQ, an activity diary and a Caltrac accelerometer over a period of 3 years. From years 3 to 5 (ages 11–12 to 13–14 years), the diary score decreased by 22%, whereas both the HAQ and Caltrac declined by 21%.

3. Discussion

A wide variety of PA questionnaires are available for youth of varying age recalling different dimensions of PA. Few have been examined for use in preschool children. None of the questionnaires included in our review showed acceptable reliability and acceptable validity. Reported reliability and validity varied, with test-retest correlations ranging from 0.02 to 0.96, and correlations between activity questionnaires and

Table IV. Full list of questionnaire acronyms and their corresponding definitions

Questionnaire acronym	Definition
3DPAR	3-Day Physical Activity Recall
7D-PAR	7-Day Physical Activity Recall Questionnaire
ACTIVITY	Assessment of Young Children's Activity using Video Technology
APARQ	Adolescent PA Recall Questionnaire
CHASE	Child Heart and Health Study in England Questionnaire
CLASS	Children's Leisure Activities Study Survey
CPAQ	Children's Physical Activity questionnaire
CPAR	Computerized PA Recall
Fels PAQ	Fels PA Questionnaire for Children
FPACQ	Flemish PA computer questionnaire
GAQ	Girls health Enrichment Multisite Study Activity Questionnaire
HAQ	Habitual Activity Questionnaire
IPAQ	International PA Questionnaire
IPAQ-A	International PA Questionnaire-modified for adolescents
MARCA	Multimedia activity recall for children and adolescents
MONICA	Monitoring instrument for cardiovascular disease survey
NPAQ	Netherlands Physical Activity Questionnaire for Young Children
OPAQ	Oxford Physical Activity Questionnaire
PAQ	Physical Activity Questionnaire
PAQA	Physical Activity Questionnaire for Adolescents, locally modified
PAQ-A	Physical Activity Questionnaire for Adolescents, modified
PAQ-C	Physical Activity Questionnaire for Older Children
PDPAR	Previous Day Physical Activity Recall
QAPACE	Quantification de l'activité physique en altitude chez les enfants
SAPAC	Self-Administered Physical Activity Checklist
SAPAQ	Self-administered Physical Activity Questionnaire
SHAPES	School Health Action, Planning and Evaluation System
SNAP	Synchronised Nutrition and Activity Program
SWAPAQ	Swedish Adolescent Physical Activity Questionnaire
WHO HBSC	World Health Organization Health Behaviour in Schoolchildren questionnaire
YPAQ	Youth PA Questionnaire
YRBS	Youth Risk Behavior Survey

accelerometers ranging from 'very poor' to 0.77. Responsiveness was only studied in one questionnaire: the HAQ.[57] These results suggest that the response patterns of the HAQ are comparable to that of the Caltrac accelerometer or a diary.

In general, PA questionnaires for adolescents correlated better with accelerometer scores than PA questionnaires for children. This finding may be due to difficulties in recalling PA, in comprehensibility of the questions or the difference in the activity patterns of children and adolescents.

Few instruments have been evaluated in multiple studies (e.g. the PAQ-C,[17,43] CLASS[14,19] and the GAQ).[18,20] The reliability of the PAQ-C was good in one study,[17] and its validity was moderate in another.[43] Both studies[14,19] that investigated the reliability of the CLASS found it to be adequate, while validity relative to accelerometry was poor. For the GAQ, reliability was adequate in one of the two studies,[18] while validity relative to accelerometry was poor in both.[18,20] The Godin-Shephard questionnaire,

which was originally developed for adults, was modified for children in three studies.[23,48,55] However, all three studies evaluated a different version.

Since there were no questionnaires with both acceptable reliability and validity, we propose that the most promising questionnaires are improved and evaluated in multiple high-quality studies. Promising questionnaires for children are the PAQ-C,[17] GAQ,[18,20] CLASS,[14,19] the Physical Activity Questionnaire for Parents and Teachers,[40] the ACTIVITY[41] and the CPAR.[25] For adolescents, the QAPACE,[27] OPAQ,[28] SNAP,[39] PDPAR[50] and SAPAC[51] seem promising.

As with any systematic review, this review is limited by the quality of the included studies. Because of the large variation in study design, incomplete reporting of the studies and the limited methodological quality of the majority of the primary studies, it was not possible to apply our intended criteria of adequacy for the methodological quality and study results. Frequent methodological shortcomings of the studies were small sample sizes (25 studies with sample sizes of <50), inadequate time intervals between test and retest (frequently too long), not taking systematic differences into account in assessing reliability (i.e. using a correlation instead of ICCs in seven studies), only evaluating relative validity and not absolute validity (i.e. using correlations instead of measures for agreement in all but two studies[14,39]). In concordance with Sallis and Saelens[65] and Oliver et al.,[66] we also found that almost all studies only examined relative validity expressed as correlations. Correlations do not pick up systematic differences between two measures. Thus, two measures may have a strong and statistically significant correlation while the agreement between both measures may be low. In cases where measures have the same unit of measurement, it is preferable to calculate the absolute agreement by using, for instance, Bland Altman plots. This method has seldom been used in validation studies of PA questionnaires. Only four studies included in our review calculated Bland Altman plots.[16,27,46,59]

We only included studies that intentionally evaluated measurement properties of PA questionnaires. It is possible that more evidence is available in the literature that could be used to determine the validity or responsiveness of the questionnaires (e.g. in studies that examine the validity of other PA measures). Furthermore, we included only English-language publications and, therefore, we may have missed some publications on additional PA questionnaires in other languages. Questionnaires that received a negative rating are not necessarily bad questionnaires. It may also be that reliability has been inadequately studied or that the report of the study was incomplete.

Three measurement properties were not rated in our review: content validity, criterion validity and measurement error. Content validity refers to the degree to which the content of an instrument is an adequate reflection of the construct to be measured. Content validity was not rated because no studies examined or reported on content validity. Criterion validity refers to the degree to which the scores of an instrument are an adequate reflection of a 'gold standard'. There is no gold standard for the assessment of PA; thus, criterion validity could not be rated. Although the doubly labelled water technique or the respiratory chamber is considered a gold standard for the assessment of energy expenditure, these methods are not considered a gold standard for the assessment of PA. Measurement error is the systematic and random error of a subject's score that is not attributed to true changes in the construct to be measured. None of the included studies evaluated measurement error.

3.1 Reliability

A reliability study should have an adequate time interval between the two administrations. For questionnaires recalling the previous day or previous week, retests need to cover the same timeframe as the initial test.[65] Otherwise, lower ICCs may be the result of actual differences in the activity pattern between the recalled days. Recalls of 'usual' PA should be less sensitive to the time interval between tests. We acknowledge that the criteria relating to the appropriate time interval between test and retest are arbitrarily chosen.

3.2 Validity

A reasonable gold standard for measuring PA does not exist; thus, criterion validity cannot be assessed. Instead, the construct validity of instruments measuring PA can be evaluated. In construct validity or responsiveness studies, it is important to state an *a priori* hypotheses. When these hypotheses are not specified, the risk of bias is large because often only the positive results will be presented. This is an ongoing process. Furthermore, the construct of PA is a formative model, i.e. the items in the questionnaire measuring PA need not be highly correlated. Therefore, structural validity (the degree to which the scores of an instrument are an adequate reflection of the dimensionality of the construct to be measured), usually evaluating with a factor analyses, is also not applicable to PA questionnaires. Therefore, to evaluate validity of a PA questionnaire, one can only rely on content validity and construct validity.

3.3 Comparison Measures

The selection of an appropriate comparison measure against which to validate PA questionnaires is difficult. As such, in the included studies, many different criteria were used to validate PA questionnaires for youth: direct observation; accelerometers; heart rate monitors; pedometers; fitness tests; and other questionnaires. Each of these comparison measures has advantages and disadvantages, and it depends on the dimension of interest as to the most appropriate comparison measure. According to Sirard and Pate,[67] direct observation is the most practical and appropriate measure for PA. However, observation is a highly demanding method for the researcher, and the actual presence of an observer may also influence the behaviour of a subject. Other measures (e.g. doubly labelled water and indirect calorimetry) are not practical for use in large populations under free-living conditions. Moreover, these measures are only suitable for assessing energy expenditure.

The accelerometer is a commonly used tool against which to compare PA surveys. This is because of its ability to objectively detect amount, frequency and duration of movement,[68,69] and its predictive relationship with heart rate and energy expenditure in the laboratory.[70,71] Accelerometers have also shown their validity during free-living activities in youth.[72] However, accelerometers are better at detecting ambulatory activity (e.g. walking and jogging) than non-ambulatory activities (e.g. cycling), the lifting of heavy objects and surface incline or decline during locomotion such as stair walking.[70] Other limitations of accelerometry include errors associated with regression equations used to derive cut-off points for moderate- and vigorous-intensity activity.[73,74] There is no consensus about appropriate cut-off points for classifying accelerometer output into different intensity levels for youth; intensity cut-off points vary widely. Corder et al.[75] recommend moving away from the use of arbitrary count-based cut-off points towards a more universally comparable approach of using acceleration (metres/second) to summarize accelerometry data. In particular, activity patterns of young children may include more horizontal motion, such as rolling, crawling and climbing, highlighting the need for more sophisticated accelerometers capturing omnidirectional movement rather than only vertical accelerations. The epoch time used may influence the results. Most studies assessing PA in youths have set the epoch at 60 seconds.[34,76,77] However, typical for youth are short, intermittent bursts of PA with frequent rest periods of a longer duration.[9,78] The median duration of moderate- and high-intensity exercise appears to be only 6 and 3 seconds, respectively. As a result, moderate- and high-intensity exercise bouts may become inconsequential when summed over a 60-second epoch, which suggests the need to use a smaller epoch time. The new accelerometer/heart rate monitors show promising results and might be better at estimating PA levels than either measure alone.

Other measures seem less suitable for validation studies. Heart rate monitors have been validated against doubly labelled water and seem valid for classifying groups of individuals rather than estimating individual PA levels.[67] Heart rate is not only sensitive to emotional stress and body position, but also to body mass.[79]

In regards to pedometers, more high-quality research is needed to show their validity and reliability for assessing PA in children and adolescents.[80,81] Pedometers detect only total counts or steps, and cannot assess activity patterns or intensity. Validation against other questionnaires or diaries is problematic since both are dependent on self-report, and we cannot say which is superior. Physical fitness tests should not be used to validate PA questionnaires since these are two different constructs. Aerobic fitness is weakly associated with PA, especially in youth.[82,83] Moreover, changes in PA only influence physical fitness in the longer term.

3.4 Recommendations Regarding Future Studies

Terwee et al.,[10] provide general recommendations for evaluating the measurement properties of PA questionnaires (QAPAQ). For assessing validity of youth PA questionnaires relative to accelerometry we propose the following additional recommendations: a monitoring period of at least 6 days; using a smaller epoch time (e.g. 15 seconds); standard methods for analysis of accelerometer data; and, preferably, the use of accelerometers that capture omnidirectional movement rather than just vertical accelerations.

4. Conclusions

Considerably more high-quality research is needed to examine the validity and reliability of promising PA questionnaires for youth. Since there is no gold standard for assessing PA, validation against different measures such as direct observation combined with accelerometry should be considered. Furthermore, in validity or responsiveness studies, it is important to state *a priori* hypotheses. When these hypotheses are not specified, the risk of bias is large because often only positive results are presented. Standardized, quality criteria (such as QAPAQ) for studies examining measurement properties of PA questionnaires are important for the improvement of the methodological quality of future validity and reliability studies.

Acknowledgements

This review was financially supported by the Department of Public and Occupational Health, the Department of Epidemiology and Biostatistics, the EMGO Institute for Health and Care Research, VU University Medical Center and Body@Work, Research Center Physical Activity, Work and Health, TNO-VU University Medical Center, Amsterdam, the Netherlands. The authors have no conflicts of interest directly relevant to the contents of this article.

References

1. Ekblom B, Astrand PO. Role of physical activity on health in children and adolescents. Acta Paediatr 2000; 89 (7): 762-4
2. Hallal PC, Victora CG, Azevedo MR, et al. Adolescent physical activity and health: a systematic review. Sports Med 2006; 36 (12): 1019-30
3. Ward DS, Evenson KR, Vaughn A, et al. Accelerometer use in physical activity: best practices and research recommendations. Med Sci Sports Exerc 2005; 37 (11 Suppl.): S582-8
4. Trost S. Measurement of physical activity in children and adolescents. Am J Lifestyle Med 2007; 1 (4): 299-314
5. Salmon J, Booth ML, Phongsavan P, et al. Promoting physical activity participation among children and adolescents. Epidemiol Rev 2007; 29: 144-59
6. Montoye HJ, Kemper HCG, Saris WHM, et al. Measuring physical activity and energy expenditure. Champaign (IL): Human Kinetics, 1996
7. Going SB, Levin S, Harrell J, et al. Physical activity assessment in American Indian schoolchildren in the Pathways study. Am J Clin Nutr 1999; 69 (4 Suppl.): 788-95S
8. Sallis JF. Self-report measures of children's physical activity. J Sch Health 1991; 61 (5): 215-9
9. Baquet G, Stratton GVPEBS. Improving physical activity assessment in prepubertal children with high-frequency accelerometry monitoring: a methodological issue. Prev Med 2007; 44 (2): 143-7
10. Terwee CB, Mokkink LB, van Poppel MNM, et al. Qualitative attributes and measurement properties of physical activity questionnaires: a checklist. Sports Med 2010; 40 (7): 525-37
11. Scientific Advisory Committee of the Medical Outcomes Trust. Assessing health status and quality-of-life instruments: attributes and review criteria. Qual Life Res 2002; 11: 193-205
12. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 2007; 60 (1): 34-42
13. de Vet HCW. Observer reliability and agreement. In: Armitage P, Colton T, editors. Encyclopedia of biostatistics. Boston (MA): John Wiley & Sons Ltd., 1998: 3123-8
14. Telford A, Salmon J, Jolley D, et al. Reliability and validity of physical activity questionnaires for children: the Children's Leisure Activities Study Survey (CLASS). Pediatr Exerc Sci 2004; 16: 64-78
15. Janz KF, Broffitt B, Levy SM. Validation evidence for the Netherlands physical activity questionnaire for young

children: the Iowa Bone Development study. Res Q Exerc Sport 2005; 76 (3): 363-9

16. Corder K, van Sluijs EM, Wright A, et al. Is it possible to assess free-living physical activity and energy expenditure in young people by self-report? Am J Clin Nutr 2009; 89 (3): 862-70
17. Crocker PR, Bailey DA, Faulkner RA, et al. Measuring general levels of physical activity: preliminary evidence for the Physical Activity Questionnaire for Older Children. Med Sci Sports Exerc 1997; 29 (10): 1344-9
18. Treuth MS, Sherwood NE, Butte NF, et al. Validity and reliability of activity measures in African-American girls for GEMS. Med Sci Sports Exerc 2003; 35 (3): 532-9
19. Salmon J, Telford A, Crawford D. Assessment of physical activity among primary school aged children: the Children's Leisure Activities Study (CLASS). Australas Epidemiolog 2002; 9: 10-14
20. Treuth MS, Sherwood NE, Baranowski T, et al. Physical activity self-report and accelerometry measures from the Girls Health Enrichment Multi-site Studies. Prev Med 2004; 38 Suppl.: S43-9
21. Ching PLYH, Dietz WH. Reliability and validity of activity measures in preadolescent girls. Ped Exerc Sci 1995; 7: 389-99
22. Brown TD, Holland BV. Test-retest reliability of the self-assessed physical activity checklist. Percept Mot Skills 2004; 99 (3 Pt 2): 1099-102
23. Koo MM, Rohan TE. Comparison of four habitual physical activity questionnaires in girls aged 7–15 yr. Med Sci Sports Exerc 1999; 31 (3): 421-7
24. Janz KF, Witt J, Mahoney LT. The stability of children's physical activity as measured by accelerometry and self-report. Med Sci Sports Exerc 1995; 27 (9): 1326-32
25. McMurray RG, Harrell JS, Bradley CB, et al. Comparison of a computerized physical activity recall with a triaxial motion sensor in middle-school youth. Med Sci Sports Exerc 1998; 30 (8): 1238-45
26. Taren DL, Freeman MB, Brandenburg NA. Evaluation of dietary and activity questionnaires for elementary school children. Ann N Y Acad Sci 1993; 699: 298-300
27. Barbosa N, Sanchez CE, Vera JA, et al. A physical activity questionnaire: reproducibility and validity. J Sports Sci Med 2007; 6: 505-18
28. Lubans DR, Sylva K, Osborn Z. Convergent validity and test-retest reliability of the Oxford Physical Activity Questionnaire for secondary school students. Behav Change 2008; 25 (1): 23-34
29. Prochaska JJ, Sallis JF, Long B. A physical activity screening measure for use with adolescents in primary care. Arch Pediatr Adolesc Med 2001; 155 (5): 554-9
30. Rangul V, Holmen TL, Kurtze N, et al. Reliability and validity of two frequently used self-administered physical activity questionnaires in adolescents [abstract]. BMC Med Res Methodol 2008; 8: 47
31. Aaron DJ, Kriska AM, Dearwater SR, et al. Reproducibility and validity of an epidemiologic questionnaire to assess past year physical activity in adolescents. Am J Epidemiol 1995; 142 (2): 191-201
32. McMurray RG, Ring KB, Treuth MS, et al. Comparison of two approaches to structured physical activity surveys for adolescents. Med Sci Sports Exerc 2004; 36: 2135-43
33. Booth ML, Okely AD, Chey TN, et al. The reliability and validity of the Adolescent Physical Activity Recall Questionnaire. Med Sci Sports Exerc 2002; 34 (12): 1986-95
34. Treuth MS, Hou N, Young DR, et al. Validity and reliability of the Fels physical activity questionnaire for children. Med Sci Sports Exerc 2005; 37 (3): 488-95
35. Lachat CK, Verstraeten R, Khanh le NB, et al. Validity of two physical activity questionnaires (IPAQ and PAQA) for Vietnamese adolescents in rural and urban areas. Int J Behav Nutr Phys Act 2008; 5: 37
36. Booth ML, Okely AD, Chey T, et al. The reliability and validity of the physical activity questions in the WHO health behaviour in schoolchildren (HBSC) survey: a population study. Br J Sports Med 2001; 35 (4): 263-7
37. Burdette HL, Whitaker RC, Daniels SR. Parental report of outdoor playtime as a measure of physical activity in preschool-aged children. Arch Pediatr Adolesc Med 2004; 158 (4): 353-7
38. Nishikido N, Kashiwazaki H, Suzuki T. Preschool children's daily activities: direct observation, pedometry or questionnaire. J Hum Ergol (Tokyo) 1982; 11 (2): 214-8
39. Moore HJ, Ells LJ, McLure SA, et al. The development and evaluation of a novel computer program to assess previous-day dietary and physical activity behaviours in school children: the Synchronised Nutrition and Activity Program (SNAP). Br J Nutr 2008; 99 (6): 1266-74
40. Harro M. Validation of a questionnaire to assess physical activity of children ages 4–8 years. Res Q Exerc Sport 1997; 68 (4): 259-68
41. Tremblay MS, Inman JW, Willms JD. Preliminary evaluation of a video questionnaire to assess activity levels of children. Med Sci Sports Exerc 2001; 33 (12): 2139-44
42. Ridley K, Olds TS, Hill A. The Multimedia Activity Recall for Children and Adolescents (MARCA): development and evaluation. Int J Behav Nutr Phys Act 2006; 3: 10
43. Kowalski KC, Crocker PR, Faulkner RA. Validation of the physical activity questionnaire for older children. Pediatr Exerc Sci 1997; 9 (2): 174-86
44. Janz KF, Medema-Johnson HC, Letuchy EM, et al. Subjective and objective measures of physical activity in relationship to bone mineral content during late childhood: the Iowa Bone Development Study. Br J Sports Med 2008; 42 (8): 658-63
45. Sallis JF, Strikmiller PK, Harsha DW, et al. Validation of interviewer- and self-administered physical activity checklists for fifth grade students. Med Sci Sports Exerc 1996; 28 (7): 840-51
46. Basterfield L, Adamson AJ, Parkinson KN, et al. Surveillance of physical activity in the UK is flawed: validation of the Health Survey for England Physical Activity Questionnaire. Arch Dis Child 2008; 93 (12): 1054-8
47. Chen X, Sekine M, Hamanishi S, et al. Validation of a self-reported physical activity questionnaire for schoolchildren. J Epidemiol 2003; 13 (5): 278-87
48. Jurisson A, Jurimae T. The validity of the Godin-Shephard physical activity questionnaire in children. Biol Sport 1996; 13: 291-5
49. Moore JB, Hanes Jr JC, Barbeau P, et al. Validation of the Physical Activity Questionnaire for Older Children in children of different races. Pediatr Exerc Sci 2007; 19 (1): 6-19

50. Weston AT, Petosa R, Pate RR. Validation of an instrument for measurement of physical activity in youth. Med Sci Sports Exerc 1997; 29 (1): 138-43
51. Ekelund U, Neovius M, Linne Y, et al. The criterion validity of a last 7-day physical activity questionnaire (SAPAQ) for use in adolescents with a wide variation in body fat: the Stockholm Weight Development Study. Int J Obes (Lond) 2006; 30 (6): 1019-21
52. Wong SL, Leatherdale ST, Manske S. Reliability and validity of a school-based physical activity questionnaire. Med Sci Sports Exerc 2006; 38 (9): 1593-600
53. Janz KF, Lutuchy EM, Wenthe P, et al. Measuring activity in children and adolescents using self-report: PAQ-C and PAQ-A. Med Sci Sports Exerc 2008; 40 (4): 767-72
54. Kowalski KC, Crocker PR, Kowalski NP. Convergent validity of the physical activity questionnaire for adolescents. Pediatr Exerc Sci 1997; 9 (4): 342-52
55. Gao S, Schmitz K, Fulton J, et al. Reliability and validity of a brief tool to measure children's physical activity. J Phys Act Health 2006; 3 (4): 415-22
56. Philippaerts RM, Matton L, Wijndaele K, et al. Validity of a physical activity computer questionnaire in 12- to 18-year-old boys and girls. Int J Sports Med 2006; 27 (2): 131-6
57. Kimm SY, Glynn NW, Kriska AM, et al. Longitudinal changes in physical activity in a biracial cohort during adolescence. Med Sci Sports Exerc 2000; 32 (8): 1445-54
58. Verheul ACM, Prins AN, Kemper HCG, et al. Validation of a weight-bearing physical activity questionnaire in a study of bone density in girls and women. Pediatr Exerc Sci 1998; 10: 38-47
59. Hagstromer M, Bergman P, De BI, et al. Concurrent validity of a modified version of the International Physical Activity Questionnaire (IPAQ-A) in European adolescents: the HELENA study. Int J Obes (Lond) 2008; 32 Suppl. 5: S42-8
60. Shiely F, MacDonncha C. Meeting the international adolescent physical activity guidelines: a comparison of objectively measured and self-reported physical activity levels. Ir Med J 2009; 102 (1): 15-9
61. Biddle SJ, Mitchell J, Armstrong N. The assessment of physical activity in children: a comparison of continuous heart rate monitoring, self-report, and interview recall techniques. Br J Phys Educ 1991; (Research Suppl.): 5-8
62. Narring F, Cauderay M, Cavadini C, et al. Physical fitness and sport activity of children and adolescents: methodological aspects of a regional survey. Soz Praventivmed 1999; 44 (2): 44-54
63. Schmidt GJ, Walkuski JJ, Stensel DJ. The Singapore Youth Coronary Risk and Physical Activity Study. Med Sci Sports Exerc 1998; 30 (1): 105-13
64. Guyatt GH, Deyo RA, Charlson M, et al. Responsiveness and validity in health status measurement: a clarification. J Clin Epidemiol 1989; 42 (5): 403-8
65. Sallis JF, Saelens BE. Assessment of physical activity by self-report: status, limitations, and future directions. Res Q Exerc Sport 2000; 71 (2): S1-14
66. Oliver M, Schofield GM, Kolt GS. Physical activity in preschoolers: understanding prevalence and measurement issues. Sports Med 2007; 37 (12): 1045-70
67. Sirard JR, Pate RR. Physical activity assessment in children and adolescents. Sports Med 2001; 31 (6): 439-54
68. Freedson PS, Miller K. Objective monitoring of physical activity using motion sensors and heart rate. Res Q Exerc Sport 2000; 71 (2 Suppl.): S21-9
69. Bassett Jr DR, Ainsworth BE, Swartz AM, et al. Validity of four motion sensors in measuring moderate intensity physical activity. Med Sci Sports Exerc 2000; 32 (9 Suppl.): S471-80
70. Melanson Jr EL, Freedson PS. Validity of the Computer Science and Applications, Inc. (CSA) activity monitor. Med Sci Sports Exerc 1995; 27 (6): 934-40
71. Slootmaker SM, Chin A Paw MJ, Schuit AJ, et al. Concurrent validity of the PAM accelerometer relative to the MTI Actigraph using oxygen consumption as a reference. Scand J Med Sci Sports 2009; 19 (1): 36-43
72. Ekelund U, Sjostrom M, Yngve A, et al. Physical activity assessed by activity monitor and doubly labeled water in children. Med Sci Sports Exerc 2001; 33 (2): 275-81
73. Hendelman D, Miller K, Baggett C, et al. Validity of accelerometry for the assessment of moderate intensity physical activity in the field. Med Sci Sports Exerc 2000; 32 (9 Suppl.): S442-9
74. Welk GJ, Blair SN, Wood K, et al. A comparative evaluation of three accelerometry-based physical activity monitors. Med Sci Sports Exerc 2000; 32 (9 Suppl.): S489-97
75. Corder K, Ekelund U, Steele RM, et al. Assessment of physical activity in youth. J Appl Physiol 2008; 105 (3): 977-87
76. Hesketh K, Crawford D, Salmon J. Children's television viewing and objectively measured physical activity: associations with family circumstance. Int J Behav Nutr Phys Act 2006; 3: 36
77. Mattocks C, Ness A, Leary S, et al. Use of accelerometers in a large field-based study of children: protocols, design issues, and effects on precision. J Phys Act Health 2008; 5 Suppl. 1: S98-111
78. Bailey RC, Olson J, Pepper SL, et al. The level and tempo of children's physical activities: an observational study. Med Sci Sports Exerc 1995; 27 (7): 1033-41
79. Tang RB, Lee PC, Chen SJ, et al. Cardiopulmonary response in obese children using treadmill exercise testing. Zhonghua Yi Xue Za Zhi (Taipei) 2002; 65 (2): 79-82
80. Tudor-Locke C, Williams JE, Reis JP, et al. Utility of pedometers for assessing physical activity: construct validity. Sports Med 2004; 34 (5): 281-91
81. Tudor-Locke C, Williams JE, Reis JP, et al. Utility of pedometers for assessing physical activity: convergent validity. Sports Med 2002; 32 (12): 795-808
82. Raudsepp L, Jurimae T. Relationships between somatic variables, physical activity, fitness and fundamental motor skills in prepubertal boys. Biol Sport 1996; 13 (4): 279-89
83. Hands B, Larkin D, Parker H, et al. The relationship among physical activity, motor competence and health-related fitness in 14-year-old adolescents. Scand J Med Sci Sports 2009; 19 (5): 655-63

Correspondence: Dr *Mai J.M. Chin A Paw*, Department of Public and Occupational Health, EMGO Institute for Health and Care Research, VU University Medical Center, Van der Boechorststraat 7, 1081 BT Amsterdam, the Netherlands.
E-mail: m.chinapaw@vumc.nl