Validity, reliability, and readability of single-item and short physical activity questionnaires for use in surveillance: a systematic review

Authors:

Antonina Tcymbal[1*], Sven Messing[1], Rachel Mait[1], Roberto Galindo Perez[1], Taiyeba Akter[1], Ivo Rakovac[2], Peter Gelius[1], Karim Abu-Omar[1]

1 Department of Sport Science and Sport, Friedrich-Alexander-Universität Erlangen-Nürnberg, Gebbertstr. 123b, 91058 Erlangen, Germany

2 World Health Organization Regional Office for Europe, Marmorvej 51, DK-2100 Copenhagen, Denmark

* Corresponding author:

Dr. Antonina Tcymbal

antonina.tcymbal@fau.de

Abstract

Background: Accurate and fast measurement of physical activity is important for surveillance. Even though many physical activity questionnaires (PAQ) are currently used in research, it is unclear which of them is the most reliable, valid, and easy to use. This systematic review aimed to identify existing brief PAQs, describe and compare their measurement properties, and assess their level of readability.

Methods: We performed a systematic review based on the PRISMA statement. Literature searches were conducted in six scientific databases in March 2022. Articles were included if they evaluated validity and/or reliability of brief (i.e., with a maximum of three questions) physical activity or exercise questionnaires intended for healthy adults. Due to the heterogeneity of studies, data were summarized narratively. The level of readability was calculated according to the Flesch-Kincaid formula.

Results: In total, 34 articles published in English or Spanish were included, evaluating 31 distinct brief PAQs. The studies indicated moderate to good levels of reliability for the PAQs. However, the majority of results showed weak validity when validated against objective measurements and demonstrated weak to moderate validity when validated against other PAQs. Most of the assessed PAQs met the criterion of being "short," allowing respondents to complete them in less than one minute either by themselves or with an interviewer. However, only 17 questionnaires had a readability level that indicates that the PAQ is easy to understand for the majority of the population.

Conclusions: This review identified a variety of brief PAQs, but most of them were evaluated in only a single study. The methods used to assess measurement properties varied widely across studies, limiting the comparability between different

PAQs and making it challenging to identify a single tool as the most suitable. Furthermore, PAQs employed different concepts for measuring physical activity, necessitating consideration of measurement properties and assessment goals when selecting a specific tool. None of the evaluated brief PAQs allowed for the measurement of whether a person fulfills the main WHO physical activity recommendations. Future development or adaptation of PAQs should prioritize readability as an important factor to enhance their usability.

Keywords: physical activity, questionnaire, measurement properties, validity, reliability, readability, surveillance

Background

It has been demonstrated that regular physical activity (PA) can help improve physical and mental functions as well as reverse some effects of chronic diseases (1). Regularly engaging in 150 minutes of PA per week (2), alongside following a healthy diet and abstaining from smoking and alcohol consumption, is seen as being key to prevent non-communicable diseases.

However, accurate measurement of PA levels is important to determine the amount of activity needed to improve health and identify links with other health outcomes and behaviors (3). To this day, self-report questionnaires are the most common measures to collect PA data, often as part of surveillance systems such as the Behavioral Risk Factor Surveillance System (4), the WHO STEPwise approach to noncommunicable disease risk factor surveillance (5), and the European Health Interview Survey (6). While such self-report questionnaires have seen widespread use because of their

efficiency, they have also been shown to have limitations related to bias and data accuracy (7).

Many different self-report questionnaires have been developed over the years with the Global Physical Activity Questionnaire (GPAQ), the International Physical Activity Questionnaire (IPAQ; also available as a short form, IPAQ-SF), and the European Health Interview Survey Physical Activity Questionnaire (EHIS-PAQ) being the most widely utilized in global PA surveillance (8). All three questionnaires assess PA across different domains, asking respondents to report their PA in a typical week (GPAQ, EHIS-PAQ) or during the last 7 days (IPAQ-SF). All three are comparatively complex and range from seven items (IPAQ-SF) to 16 (GPAQ). Nevertheless, the measurement properties of these questionnaires are modest. For GPAQ (9), IPAQ-SF (10, 11), and the EHIS-PAQ (12), different studies have demonstrated reasonable reliability but comparatively low validity. In a recent review (8) of IPAQ-SF, GPAQ, and EHIS-PAQ, the questionnaires showed low-to-moderate validity against objective measures of PA such as accelerometers, and moderate-to-high validity against subjectively measured PA such as other questionnaires.

It is well known that questionnaires with many items can increase the response burden on respondents (13, 14), which has resulted in the development of several short self-report instruments. In contrast to the detailed questionnaires mentioned above, the purpose of short PA questionnaires is to simplify and speed up the procedure for assessing PA levels.

In addition, language-related difficulties are currently an important topic in public health. Patient education materials can increase patient compliance, but only if they are written in a language that is easy for the patient to understand (15). Regarding PA questionnaires, Altschuler et al. (16) found significant gaps between respondents'

interpretations of some PA questions and researchers' original assumptions about what those questions were intended to measure. One of the characteristics of language difficulty is the level of readability, which indicates how easily readers can understand the text. Research of texts used in healthcare consistently shows that materials intended for patients often require a high level of education and are too complicated for the average person (15, 17). In relation to physical activity questionnaires (PAQs), readability can influence the amount of time a person needs to understand the question and, if the text is too complicated, may potentially decrease the response rate and correctness of the answer.

A number of existing reviews have investigated the measurement properties of PA questionnaires. Van Poppel et al. (18) reviewed the validity and reliability methodology of 85 versions of PAQs with no clear consensus regarding the best questionnaire for PA measurement. Helmerhost et al. (19) studied reliability and objective criterion-related validity of 34 newly-developed and 96 existing PAQs. However, to our knowledge, a dedicated review of short and brief PA questionnaires measurement properties has not yet been performed. Therefore, the aim of this systematic review is to identify existing brief PAQs and to comparatively describe their measurement properties and level of readability.

Methods

This review follows the guidelines of the PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) statement (20) and the Consensus-based Standards for the Selection of Health Measurement Instruments (COSMIN) guideline for systematic reviews of patient-reported outcome measures (21).

Information sources and search strategy

A systematic search was performed in six databases (PubMed, Web of Science, Scopus, CINAHL, SportDiscus, PsycInfo) in March 2022. Additionally, reference lists from previous reviews of PA questionnaires and other relevant publications were screened to identify additional studies. A comprehensive search strategy was developed with a combination of keywords in the categories of measured construct (physical activity) and type of instruments (brief/short questionnaire).

The resulting search string was as follows:

("physical activit*" OR "physical inactivit*")

AND

(questionnaire OR measure OR evaluat* OR assess* OR surveillance OR monitor* OR screening)

AND

("single item" OR single-item OR "single question" OR "one item" OR one-item OR "one question" OR "brief questionnaire " OR "short questionnaire " OR "brief assessment" OR "short assessment" OR "two-item*" OR "two-item*" OR "two questions" OR "brief physical activity assessment" OR "Single response" OR "Single-response")

No restrictions were made regarding language or publication date.

Eligibility criteria

Articles were included in the review if they fulfilled the following eligibility criteria:

1. The article described a self-report PA or exercise questionnaire intended for healthy adults.
2. The evaluated questionnaire was brief and included a maximum of three questions.
3. The article investigated one or more measurement properties of the questionnaire.
4. The article was published in a peer-reviewed journal.

Articles were excluded based on the following exclusion criteria:

1. The article focused on a questionnaire that measured only physical inactivity, screen time, or sedentary behavior.
2. The evaluated questionnaire was intended solely for children and adolescents or people with specific conditions.
3. The article did not investigate the measurement properties of the questionnaire.

Study selection

Two reviewers independently screened and selected the relevant articles. First, all articles were screened based on titles and abstracts. If the title and/or abstract indicated that the study fulfilled the inclusion criteria, both reviewers screened the full text for eligibility. When necessary, supplementary files were also reviewed for additional information. Disagreements between the reviewers were discussed within the research team until a consensus was reached.

Records were managed using the Covidence systematic review software (Veritas Health Innovation, Melbourne, Australia; www.covidence.org) and EndNote X9 (Clarivate Analytics, Philadelphia, PA, USA).

Data extraction

Data of included studies were extracted and summarized by one reviewer and verified by a second reviewer to reduce bias and error. Discrepancies were discussed between reviewers to achieve consensus. Extracted information included: publication details (first author, year of publication, country), sample characteristics (number of participants, age category, special health conditions), the measurement tool(s) explored, who assessed PA levels of participants, assessed measurement properties, other measurement tool(s) used as a comparison, reliability test-retest interval, and the results of the study.

Risk of bias assessment

The methodological quality of the individual studies included in this review was assessed with 18 questions based on the Appraisal tool for Cross-Sectional Studies (AXIS) (22). In order to specifically assess the quality of the tools' validity and reliability measurement properties, some questions were modified based on the COSMIN risk of bias checklist (21) and on suggestions from previous systematic reviews examining PA assessment measures (23). This quality assessment evaluated articles based on study design, sample size, participant selection process, appropriate blinding, examiner experience, method of measurement, adequate data reporting, internal consistency, and six other categories (see Additional file 1). Risk of bias assessment was conducted independently by two reviewers, with any

discrepancies resolved through discussion in the research team. Studies were scored 1 if they satisfied the quality element, and 0 if they did not. The summary score (range: 0–18) indicates the risk of bias, with a higher score indicating higher quality and therefore a lower risk of bias.

Data synthesis and analysis of measurement properties

The primary objective of this systematic review was to identify existing short questionnaires suitable for assessing PA levels in surveillance and primary care settings, and to compare their measurement properties. To accomplish this goal, relevant information from the included studies was summarized separately for each questionnaire. In order to evaluate the reliability and validity of the identified questionnaires, a range of tests were employed in the included studies. Some studies reported results for a total questionnaire summary score, while others assessed reliability and validity for specific aspects, intensities, or domains of the questionnaire. Additionally, certain studies examined these measurement properties within subgroups of the test population. Due to the heterogeneity in the methods used and the lack of standardized reporting across studies, a quantitative meta-analysis was not feasible. Consequently, the information from the included studies was summarized narratively, highlighting the key findings for each questionnaire. This narrative synthesis allows for a comprehensive overview of the reliability and validity findings, highlighting strengths and limitations of each questionnaire.

Analysis of length and readability of questionnaires

To determine how quickly questionnaires could be answered and how easy it is to understand the questions, identified PAQs were analyzed for word count and readability level. The expected time the tool would take for self-administration (silent reading speed) and interviewer administration (spoken word speed) was calculated based on the respective questionnaire's word count and English reading speeds established by Brysbaert (24). The level of readability was calculated according to the Flesch-Kincaid formula, which was chosen because it is the most commonly used tool to calculate the readability level of written health information (25). The formula has two forms: the Flesch Reading-Ease-Score, and the Flesch–Kincaid Grade Level. The Flesch Reading-Ease-Score test produces a score from 0 to 100, and higher scores indicate material that is easier to read; lower scores mark passages that are more difficult to read. The Flesch–Kincaid Grade Level formula matches the text to the grade level achievement (number of years of education) required to understand the text. We applied both formulas for each questionnaire using an online calculator (26).

Results

Study selection process

The search across six databases resulted in 2,422 publications. After removing 1,255 duplicates, 1,167 articles were screened based on title and abstract. 65 studies were found eligible for full text assessment. One recently published article (27) and one article identified while screening the reference lists of relevant publications (28) satisfied all eligibility criteria and were therefore included in the analysis. Subsequently, 33 full text articles were excluded due to the length of the questionnaire (n=14), the publication type (n=14), a lack of self-reporting on PA and

exercise (n=2), and a study design that did not consider validity and reliability (n=3). In total, 34 studies were included for data extraction and methodological quality assessment. A summary of the search results is presented in Figure 1.

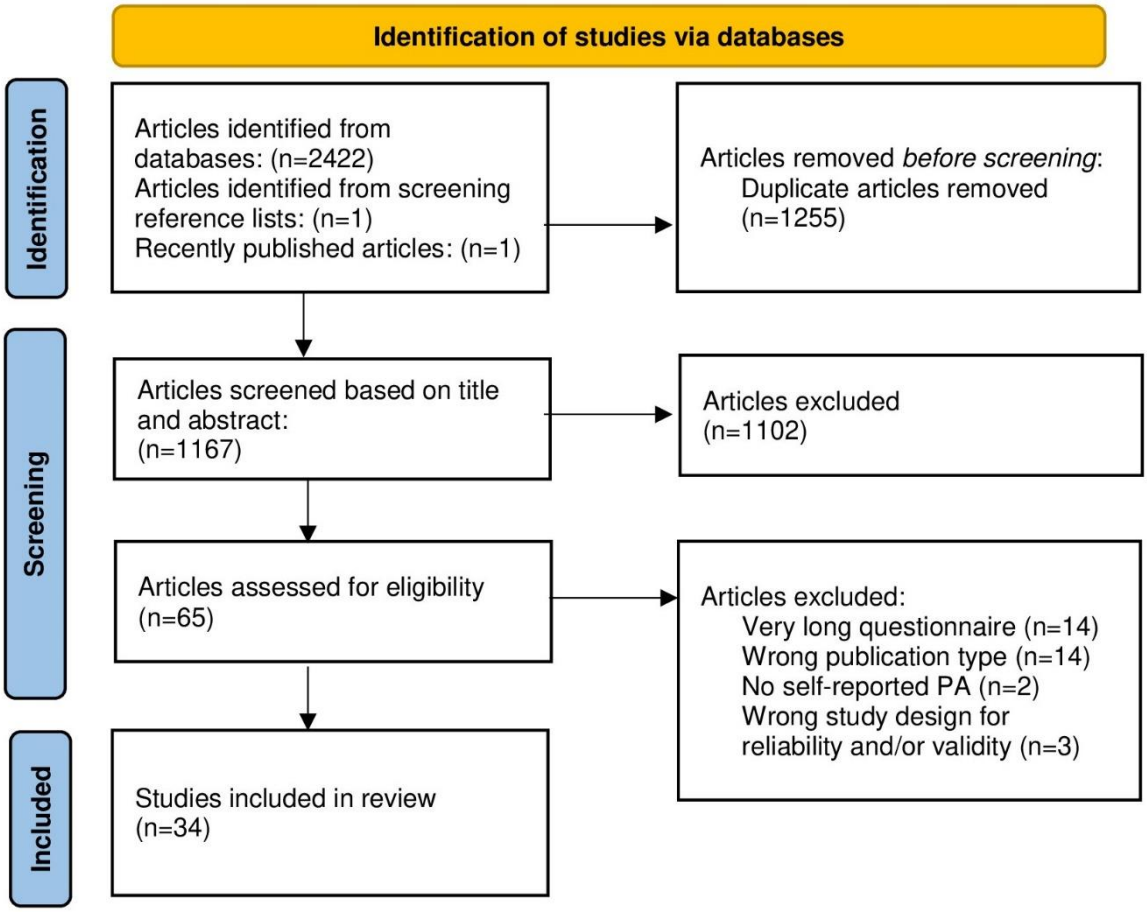

Figure 1. PRISMA flowchart of study eligibility (PA=physical activity)

Characteristics of included studies

A summary of the characteristics of the 34 included studies is presented in Table 1. The included studies were conducted in Western European countries (n=14), the USA (n=10), Australia (n=5), Canada (n= 2), New Zealand (n=2), and Japan (n=1).

[Table 1 about here]

A total of 114,143 adults were assessed using the 31 unique brief PAQs (with sample sizes ranging from 9 to 39,379). In 25 studies, the sample consisted only of healthy adults, and nine studies also included specific populations such as older adults (29) (30) (31), overweight/obese adults (32, 33), as well as patients with rheumatoid arthritis (34), coronary heart disease (35) (36), and chronic obstructive pulmonary disease (37).

Thirty-one studies documented who completed the PAQ. In 16 studies, respondents filled out the questionnaire by themselves, and in 15 studies the PA level was collected from an interviewer reading a questionnaire.

Questionnaires asked for the amount of PA (n=6), the number of days per week with a sufficient amount of PA (n=4), general exercise participation (n= 3), self-reported activity compared with peers (n=2), or for respondents to choose a categorical descriptor of PA levels ranging from “inactive” to “very active” (n=16).

Seven of the included studies investigated the measurement properties of the Single Item Physical Activity Measure (SIPAM) (27, 30, 38, 39, 40, 41, 42), five were related to the Brief Physical Activity Assessment Tool (BPAAT) (37, 43, 44, 45, 46), and two to the Stanford Leisure-Time Activity Categorical Item (L-Cat) (32, 33). The other 28 PAQs were assessed in a single study.

Risk of bias assessment

The average quality score of included studies was 11.75, with a range between 7 and 17 (out of 18). Thirteen studies received 14 or more points, which can be considered a high methodological quality. In most of the studies, the aims and objectives, the process of the measurement properties investigation, the statistical analysis and the results were sufficiently described, and the study design was appropriate. However,

the included papers reported poorly on whether the examiners had enough experience with the tool and whether they were blinded to participant characteristics, previous findings, or other observed findings. Only four studies justified their choice of sample size. Additionally, many studies failed to consistently report reasons for dropout and characteristics of non-responders. Quality scores of the individual studies are reported in Table 1.

Questionnaires' measurement properties

A summary of the reliability, validity, and diagnostic test accuracy data found in the identified studies is presented in Table 2. Studies commonly use a number of different statistical analyses to define absolute (agreement between the two measurement tools) or relative (the degree to which the two measurement tools rank individuals in the same order) validity and reliability. These types of statistical analysis include correlations (Pearson's; Spearman's; interclass), regression, kappa statistics, and area under the receiver operating curve (AUC-ROC).

[Table 2 about here]

Eleven studies measured the reliability of brief PAQs. All of them used a test-retest procedure to measure the consistency of the PAQs. Statistical methods and test-retest intervals varied widely between studies. Overall, studies showed moderate to good reliability levels of the PAQs.

The validity of PAQs was assessed in 33 studies. As a "gold standard" for validation, 13 studies used other PAQs, 15 studies validated brief PAQs against accelerometers or pedometers, and 11 studies compared results of the brief PAQs to other objective measurements, such as BMI (28, 32, 47, 48), VO_2 max (47, 49), or doubly labeled water (34). Validity coefficients were in general considerably lower than reliability

coefficients. The majority of results showed weak validity of brief PAQs against objective measurements and weak to moderate validity against other PAQs.

Length and complexity of the questionnaires

The texts of all questionnaires and information about their length and readability level are presented in Table 3. It should be noted that, while some PAQs were created and conducted in other languages, length and readability were calculated for the English versions.

[Table 3 about here]

Most of the assessed PAQs (n=26) can be read by in silence by the respondent or aloud by the interviewer in less than one minute. However, some questionnaires presented as “brief” by the authors of the identified studies contain a large amount of text and require more than one minute for reading: the Nordic Physical Activity Questionnaire (NPAQ) short, the Total Activity Measure (TAM), the Self-report Scale to Assess Habitual Physical Activity, the Stanford Leisure-Time Activity Categorical Item, and the Stanford Brief Activity Survey.

The calculation of the readability levels with the Flesch-Kincaid formula showed that only 17 out of 31 brief PAQs had readability levels of “easy to read” (n=9) or “plain English,” (n=8) and can be easily understood by the majority of the population (see Figure 2). Other questionnaires have long sentences and/or many complicated words (three syllables and more), making them difficult to read. Seven questionnaires scored as “difficult to read” and require college-level education; it is likely that people with lower education levels will find them hard to understand.

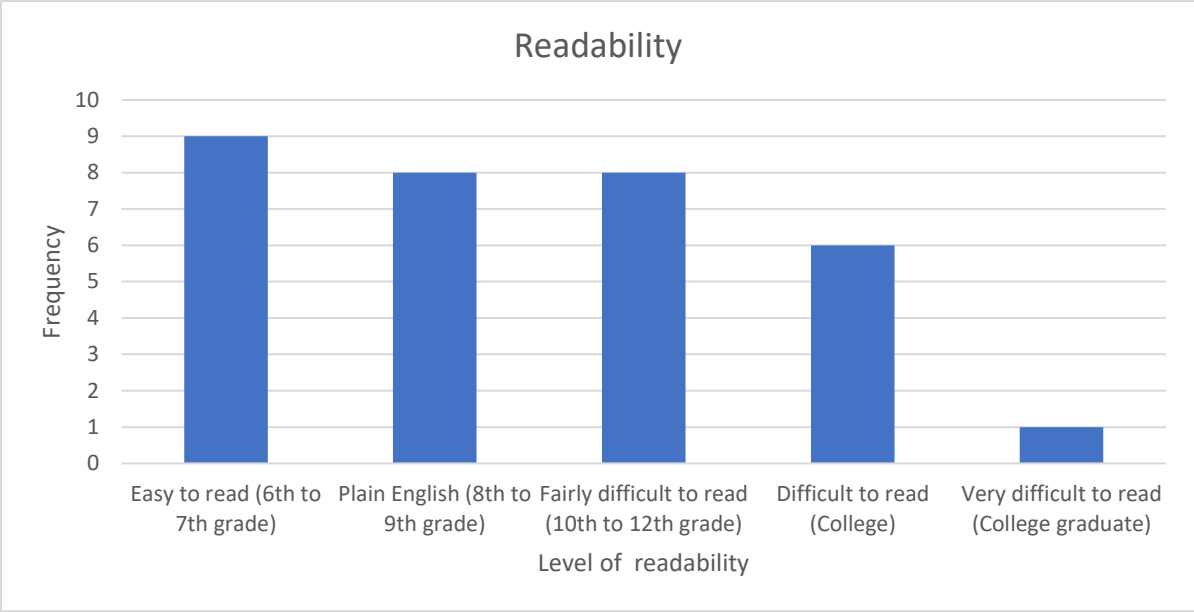

Figure 2. Readability levels for brief PA questionnaires (English language versions).

Comparison of brief PAQs

Data on both validity and reliability were only available for nine of the 31 assessed PAQs. These brief PAQs were chosen for an in-depth comparison. Their summarized measurement properties and linguistic characteristics are presented in Table 4.

Table 4. Comparison of the brief physical activity questionnaires by reliability, validity, length and readability.

Tool	Reliability	Validity		Word count	Readability
		Against objective measurement	Against other PAQ		
Single Item Physical Activity Measure (SIPAM)	Good (0,63-0,82)	Poor to Good (0,22-0,81)	Poor to Good (0,33-0,81)	63 – original 111 – parents version	Fairly difficult to read Parents version - Very difficult to read
Brief Physical Activity Assessment Tool (BPAAT)	Moderate to good (0,53-0,72)	Poor to moderate (0,18-0,43)	Moderate to good (0,45-0,64)	106	Difficult to read

Three-Question Assessment	Moderate to good (0,56-0,63)	Poor (0,24-0,31)	Moderate (0,39-0,43)	140	Fairly difficult to read
Nordic physical activity questionnaire (NPAQ) - short (open ended) *	Good (0,82)	Poor (0,33)	NM	237	Difficult to read
Total Activity Measure (Version 2)	Good (0,82)	Poor (0,36-0,38)	NM	241	Fairly difficult to read
Japan Collaborative Cohort (JACC) Questionnaire*	Moderate (0.39 - 0.56)	NM	Moderate (0.43-0.60)	90	Plain English
Relative PA Question (Compared to peers)	Moderate (0,56)	Poor to moderate (0,28-0,57)	NM	25	Easy to read
The Seven-level Single-Question Scale *	Good (0,7)	Poor (0,33)	NM	144	Fairly difficult to read
Absolute PA Question	Good (0,7)	Poor (0,1-0,33)	NM	34	Fairly difficult to read
A single-item 5-point rating of usual PA (Usual PA Scale)	Good (0,68-0,88)	NM	Good (0,66)	147	Plain English
The Stanford Leisure-Time Activity Categorical Item (L-CAT)	Good (0,64-0,8)	Poor (0,36-0,38)	NM	331	Plain English

* = the study used not the English version of the questionnaire, PAQ=physical activity questionnaire

The Single Item Physical Activity Measure (SIPAM) uses a single question to assess the number of days per week on which 30 minutes or more of PA are performed (excluding housework and work-related PA). For this PAQ, the largest number of studies on validity and reliability were found. The SIPAM showed good reliability levels in all studies, however, results on validity varied considerably between studies – from poor to good – against both objective measurements and other subjective PAQs. Zwolinsky et al. (2015) also measured SIPAM's ability to identify people that meet or do not meet WHO PA recommendations. The tool showed a low diagnostic capacity compared to the IPAQ. The agreement between the SIPAM and the IPAQ was low for identification of participants who meet WHO PA recommendations (kappa= 0.13, 95% CI 0.12 to 0.14) and moderate for the classification of inactive participants (kappa= 0.45, 95% CI 0.43 to 0.47) (42).

The Brief Physical Activity Assessment Tool (BPAAT) consists of two questions, one regarding the frequency and duration of vigorous PA and the other regarding moderate PA and walking performed in an individual's usual week. By combining the results of both questions (scores can range from 0 to 8), the subject can be classified as insufficiently (0–3 score) or sufficiently active (≥4 score). Results of reviewed studies showed that the BPAAT has moderate to good levels of reliability and validity in comparison with other PAQs, however comparison with accelerometers identified only poor to moderate validity.

Smith et al. (2005) evaluated the Three-Question Assessment variant of the BPAAT which has separate questions about moderate PA and walking. The results of the study did not find a considerable difference in validity and reliability between the two- and three-question versions. One study also reported that more physicians preferred the two-question version (BPAAT) as it is shorter and therefore easier to use (Smith, 2005).

The Nordic Physical Activity Questionnaire-short (NPAQ-short) includes one question about moderate to vigorous PA and a second question about vigorous PA. Danquah et al. (2018) compared open-ended and closed-ended questions with the open-ended version achieving better results. The open-ended version showed good reliability but performed similarly to other questionnaires, showing poor validity when compared against objective PA measures. The analyses showed that the questionnaire was one of the longest and was rated "difficult to read". The agreement with accelerometer data in identification of people that meet or do not meet the WHO PA recommendations was low (kappa=0.42).

The Total Activity Measures (TAM) includes three open-ended questions about strenuous, moderate, and mild PA. The revised second version, TAM2, asks about

the total time spent at each activity level over a 7-day period. The TAM2 showed good reliability when validated against objectively-measured PA. The word count was high at 241, and readability was rated as “fairly difficult”.

The Japan Collaborative Cohort (JACC) Questionnaire has three items. Two questions focus on leisure-time PA, i.e., time per week engaging in sport or PA (with options ranging from “little” to “at least 5 hours”), and frequency of engagement in sport over the past year (options from “seldom” to “at least twice a week”). The third question asks about daily walking patterns (options from “little” to “more than 1 hour”). The questionnaire showed moderate reliability after one year and moderate validity when compared against a more in-depth interview by a trained researcher. The questions are 90 words in total and were rated as “plain English.”

The Relative PA Question is the shortest of the compared PAQs and has an easy readability level. It allows respondents to compare their level of PA with other people of the same age and to choose from five categories ranging from “much more active” to “much less active.” The questionnaire has a moderate reliability level and showed poor validity against objective measures.

The Absolute PA Question asked participants to choose what best describes their activity level from three options: (1) vigorously active for at least 30 minutes, three times per week; (2) moderately active at least three times per week; or (3) seldom active, preferring sedentary activities. The questionnaire is relatively short, but as most of the other brief PAQs, it showed high reliability and poor validity against objective measures.

The Seven-Level Single-Question Scale (SR-PA L7) requires respondents to assign themselves to a level ranging from “I do not move more than is necessary in my daily routines/chores” and “I participate in competitive sports and maintain my fitness

through regular training." These seven items are supposed to categorize respondents as maintaining low, medium, or high PA. The SR-PA L7 showed poor validity when compared to objectively measured PA but good reliability. The scale is moderately long at 144 words and was rated as "fairly difficult to read."

A single-item 5-point rating of usual PA (Usual PA Scale) includes descriptions of three PA levels: highly active, moderately active, and inactive. Respondents need to read a description of each category and identify their usual PA level from the list. The questionnaire's readability level was ranked as "plain English," and it showed good reliability and validity levels. However, validation was done only against another PAQ which is not well-known and requires further research.

The Stanford Leisure-Time Activity Categorical Item (L-CAT) is a single-item questionnaire that consists of six descriptive PA categories ranging from inactive to very active. Each category consists of one or two statements describing common activity patterns over the past month, differing in frequency, intensity, duration, and types of activity. The categories are described in "plain English," but this made the questionnaire longer than the other questionnaires. The L-CAT showed high reliability but poor validity against pedometer and accelerometer validation.

Discussion

This review assessed the validity, reliability, and readability of brief PAQs for adults. To our knowledge, it is the first review focused specifically on brief PAQs and also the first to assess their readability. The sheer number of brief PAQs we identified (n=31) indicates a high research interest in such instruments. However, it also indicates a lack of harmonization when it comes to PA assessment using brief questionnaires.

Overall, most PAQs were reported to have poor or moderate validity. They showed higher validity levels against other self-reported tools than against objective PA measures. Although reliability is an important measurement property, it was assessed only in 11 studies. About half of all PAQs showed a moderate to good level of reliability. These results are in line with other reviews of PAQs (18, 19, 50, 51).

A significant difficulty in conducting this review was that the studies used different methods for validation, varying time-intervals between repeated measurements, and different statistical methods to analyze data. Complete data on validity and reliability were only available for nine PAQs, and only those could be compared in greater detail. All of these factors prevented a more detailed comparison of results across PAQs. In general, the methodological quality of the included studies was modest. The most common flaws were comparably small sample sizes, a lack of sample size justification, and a poor description of the validity and reliability assessment process. Additionally, most studies utilized convenience samples, making it impossible to assess if measurement properties would differ between adults with different levels of socioeconomic status and/or educational attainment.

Also, the included PAQs used different concepts of measuring PA, further complicating a direct comparison between instruments. For example, the SIPAM measures on how many days per week respondents perform 30 minutes or more of PA; the Absolute PAQ requires respondents to choose from several descriptions of different PA levels; and the Relative PA question asks them to compare their level of PA with peers. Some PAQs, such as the TAM, aim to assess the total volume of moderate-to-vigorous PA. Others focus on particular PA domains. And yet others assess only leisure-time PA. This variety can be partly explained by efforts to keep PAQs short at the price of excluding some PA dimensions (type, duration, intensity,

volume). However, it could also be interpreted as a lack of common understanding about which dimensions of PA should be assessed with brief PAQs. It should also be taken into account that none of the reviewed brief PAQs allowed for the measurement of whether a person fulfills the WHO physical activity recommendations (52).

All of the included studies were conducted in highly developed nations, and 23 of them took place in English-speaking countries. This most certainly biased the results, since terminology related to PA differs between languages, as do levels of literacy (53). Ultimately, to assess reliability and validity of PAQs, many more studies should be conducted in developing nations to obtain a more accurate assessment of their suitability for international surveillance systems. Certainly, in order to stimulate such research, funding opportunities need to be made available to research teams from such nations.

The results of this study point out that many PAQs have low readability levels. This is particularly disturbing when considering that most of these PAQs were tested in developed nations with comparatively high levels of literacy. Potentially, poor readability is related to the modest measurement properties that PAQs commonly come with. Questionnaires that are more readable or might even have been co-developed (54) with the intended population groups are likely to yield better measurement properties. This would also strengthen the case of integrating them into PA surveillance systems. However, at this point in time, there is a dearth of research relating PAQs and their measurement properties to readability. Advancing knowledge in this field would also benefit longer PA questionnaires (such as IPAQ and GPAQ) that are widely utilized in surveillance and potentially score low on readability as well.

This review comes with certain limitations. The search was limited to scientific databases, and only studies published in peer reviewed scientific journals were included. This can lead to a publication bias, as all other types of publications and gray literature were excluded. The varying measurement methods and conditions complicated the comparison of findings from different studies and limited data analysis to a narrative description of differences between tools. This led to more subjective results and greater difficulty in identifying the tools with the best measurement properties. It should be also taken into account that the Flesch-Kincaid readability formula was created for longer texts and not adapted for short questionnaires. Consequently, the readability levels of PAQs presented here should be interpreted with caution.

Conclusion

This systematic review sheds light on the validity, reliability, and readability of short PAQs. Results indicate that additional research is needed, notably regarding their reliability and validity, their readability, and their applicability to non-English speaking and/or developing countries.

Recent years have seen a shift from self-report questionnaires towards objective measures in PA surveillance. This shift has been partially motivated by the persistently modest measurement properties of PAQs, as well as technological advancements in the quality and affordability of accelerometers and other PA measurement tools. However, objective PA assessment comes with its own set of limitations (55), and self-report and objective measures have been described as measuring entirely different parameters. Right now, it seems unclear what the future for PA surveillance might hold and if brief PAQs will continue to play a role.

In this regard, the key merit of short PAQs is that they require substantially less time to complete than the established PA surveillance questionnaires (GPAQ, IPAQ, and EHIS-PAQ). This makes them highly appealing for use in surveillance systems. In order for them to be included in existing systems, more studies would need to investigate how they compare to longer PAQs. Potentially, if their validity and reliability are found to be at a comparable level, we might see surveillance systems gradually shift from longer towards briefer PAQs.

List of abbreviations:

BPAAT: Brief Physical Activity Assessment Tool

EHIS-PAQ: European Health Interview Survey Physical Activity Questionnaire

GPAQ: Global Physical Activity Questionnaire

IPAQ: International Physical Activity Questionnaire

JACC: Japan Collaborative Cohort Questionnaire

L-CAT: The Stanford Leisure-Time Activity Categorical Item

NPAQ-short: The Nordic Physical Activity Questionnaire-short

PA: physical activity

PAQ: physical activity questionnaire

SIPAM: Single Item Physical Activity Measure

SR-PA L7: The Seven-level Single-Question Scale for Self-Reported Leisure Time Physical Activity

TAM: Total Activity Measure

Declarations

Ethics approval and consent to participate: Not applicable.

Consent for publication: Not applicable.

Availability of data and materials: Data sharing is not applicable to this article as no datasets were generated or analysed during the current study.

Competing interests: The authors declare that they have no competing interests.

Funding: This research was conducted as part of projects funded by the German Federal Ministry of Health (ZMI5-2522WHO001, ZMI5-2523WHO001). The ministry was neither involved in writing this manuscript nor in the decision to submit the article for publication. Furthermore, we acknowledge financial support by Deutsche Forschungsgemeinschaft and Friedrich-Alexander-Universität Erlangen-Nürnberg within the funding programme "Open Access Publication Funding".

Authors' contributions: AT, KA-O and IR conceptualized the study. AT conducted the literature search. AT, RA, RGP, and TA performed abstract/title screening, full-text selection, data extraction, and quality assessment. AT with support from KA-O, SM and PG prepared a draft of the manuscript. All authors critically reviewed the draft and approved the final manuscript.

Acknowledgements: Not applicable.

References:

1. McPhee JS, French DP, Jackson D, Nazroo J, Pendleton N, Degens H. Physical activity in older age: perspectives for healthy ageing and frailty. Biogerontology. 2016;17(3):567-80.
2. World Health Organization. Guidelines on physical activity and sedentary behaviour. Geneva: World Health Organization; 2020.
3. Welk G. Physical activity assessments for health-related research: Human Kinetics; 2002.
4. Centers for Disease Control and Prevention. Behavioral Risk Factor Surveillance System Survey Questionnaire. Atlanta, Georgia: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention; 2021.
5. World Health Organization. WHO STEPS Surveillance Manual: The WHO STEPwise approach to noncommunicable disease risk factor surveillance. 2020.
6. Finger JD, Tafforeau J, Gisle L, Oja L, Ziese T, Thelen J, et al. Development of the European Health Interview Survey - Physical Activity Questionnaire (EHIS-PAQ) to monitor physical activity in the European Union. Arch Public Health. 2015;73:59.
7. Sallis JF, Saelens BE. Assessment of physical activity by self-report: status, limitations, and future directions. Research quarterly for exercise and sport. 2000;71:1-14.
8. Sember V, Meh K, Sorić M, Starc G, Rocha P, Jurak G. Validity and Reliability of International Physical Activity Questionnaires for Adults across EU Countries: Systematic Review and Meta Analysis. International journal of environmental research and public health. 2020;17(19):7161.
9. Bull FC, Maslin TS, Armstrong T. Global physical activity questionnaire (GPAQ): nine country reliability and validity study. J Phys Act Health. 2009;6(6):790-804.

10. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. *Medicine and science in sports and exercise*. 2003;35(8):1381–95.
11. Lee PH, J. MD, Lam TH, Stewart SM. Validity of the International Physical Activity Questionnaire Short Form (IPAQ-SF): a systematic review. The international journal of behavioral nutrition and physical activity. 2011;8:115.
12. Baumeister SE, Ricci C, Kohler S, Fischer B, Topfer C, Finger JD, et al. Physical activity surveillance in the European Union: reliability and validity of the European Health Interview Survey-Physical Activity Questionnaire (EHIS-PAQ). The international journal of behavioral nutrition and physical activity. 2016;13:61.
13. Ware J, Jr., M. K, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Medical Care. 1996;34(3):220-33.
14. Rolstad S, Adler J, Rydén A. Response burden and questionnaire length: is shorter better? A review and meta-analysis. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2011;14(8):1101-8.
15. Grabeel KL, Russomanno J, Oelschlegel S, Tester E, Heidel RE. Computerized versus hand-scored health literacy tools: a comparison of Simple Measure of Gobbledygook (SMOG) and Flesch-Kincaid in printed patient education materials. J Med Libr Assoc. 2018;106(1):38-45.
16. Altschuler A, Picchi T, Nelson M, Rogers JD, Hart J, Sternfeld B. Physical activity questionnaire comprehension: lessons from cognitive interviews. Med Sci Sports Exerc. 2009;41(2):336-43.
17. Wallace LS, Lennon ES. American Academy of Family Physicians patient education materials: Can patients read them? Family Medicine. 2004;36(8):571-4.
18. van Poppel MN, Chinapaw MJ, Mokkink LB, van Mechelen W, Terwee CB. Physical activity questionnaires for adults: a systematic review of measurement properties. Sports Medicine (Auckland, N Z). 2010;40(7):565-600.
19. Helmerhorst HJH, S. B, Warren J, Besson H, Ekelund U. A systematic review of reliability and objective criterion-related validity of physical activity questionnaires. The international journal of behavioral nutrition and physical activity. 2012;9:103.
20. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71.
21. Prinsen CAC, Mokkink LB, Bouter LM, Alonso J, Patrick DL, de Vet HCW, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. Quality of Life Research : An International Journal of Quality of Life Aspects of Treatment, Care, and Rehabilitation. 2018;27(5):1147-57.
22. Downes MJ, Brennan ML, Williams HC, Dean RS. Development of a critical appraisal tool to assess the quality of cross-sectional studies (AXIS). BMJ Open. 2016;6(12):e011458.
23. Williams MA, McCarthy CJ, Chorti A, Cooke MW, Gates S. A systematic review of reliability and validity studies of methods for measuring active and passive cervical range of motion. Journal of Manipulative and Physiological Therapeutics. 2010;33(2):138-55.
24. Brysbaert M. How many words do we read per minute? A review and meta-analysis of reading rate. Journal of Memory and Language. 2019;109.
25. Wang LW, Miller MJ, Schmitt MR, Wen FK. Assessing readability formula differences with written health information materials: application, results, and recommendations. Res Social Adm Pharm. 2013;9(5):503-16.

26. charactercalculator.com. Flesch Reading Ease 2023 [Available from: https://charactercalculator.com/flesch-reading-ease/.
27. Bauman AE, Richards JA. Understanding of the Single-Item Physical Activity Question for Population Surveillance. Journal of Physical Activity and Health. 2022;19(10):681-6.
28. Schechtman KB, Barzilai B, Rost K, B. FJE. Measuring physical activity with a single question. American journal of public health. 1991;81(6):771-3.
29. Gill DP, Jones GR, Zou G, Speechley M. Using a single question to assess physical activity in older adults: A reliability and validity study. BMC Medical Research Methodology. 2012;12(20).
30. Macdonald HM, Nettlefold L, Bauman A, Sims-Gould J, McKay HA. Pragmatic Evaluation of Older Adults' Physical Activity in Scale-Up Studies: Is the Single-Item Measure a Reasonable Option? Journal of Aging and Physical Activity. 2022;30(1):25-32.
31. Portegijs E, Sipilä S, Viljanen A, Rantakokko M, Rantanen T. Validity of a single question to assess habitual physical activity of community-dwelling older people. Scandinavian Journal of Medicine & Science in Sports. 2017;27(11):1423-30.
32. Kiernan M, Schoffman DE, Lee K, Brown SD, Fair JM, Perri MG, et al. The stanford leisure-time activity categorical item (L-Cat): A single categorical item sensitive to physical activity changes in overweight/obese women. International Journal of Obesity. 2013;37(12):1597-602.
33. Ross KM, Leahey TM, Kiernan M. Validation of the Stanford Leisure-Time Activity Categorical Item (L-Cat) using armband activity monitor data. *Obesity science & practice*. 2018;4(3):276-82.
34. Johansson G, Westerterp KR. Assessment of the physical activity level with two questions: Validation with doubly labeled water. International Journal of Obesity. 2008;32(6):1031-3.
35. Blomqvist A, Bäck M, Klompstra L, Strömberg A, Jaarsma T. Utility of single-item questions to assess physical inactivity in patients with chronic heart failure. ESC Heart Fail. 2020;7(4):1467-76.
36. Orrell A, Doherty P, Miles J, Lewin R. Development and validation of a very brief questionnaire measure of physical activity in adults with coronary heart disease. European Journal of Cardiovascular Prevention and Rehabilitation. 2007;14(5):615-23.
37. Cruz J, Jácome C, Oliveira A, Paixão C, Rebelo P, Flora S, et al. Construct validity of the brief physical activity assessment tool for clinical use in COPD. The Clinical Respiratory Journal. 2021;15(5):530-9.
38. Hamilton K, White KM, Cuddihy T. Using a single-item physical activity measure to describe and validate parents' physical activity patterns. Research Quarterly for Exercise and Sport. 2012;83(2):340-5.
39. Milton K, Bull FC, Bauman A. Reliability and validity testing of a single-item physical activity measure. British Journal of Sports Medicine. 2011;45(3):203-8.
40. Milton K, Clemes S, Bull F. Can a single question provide an accurate measure of physical activity? British Journal of Sports Medicine. 2013;47(1):44-8.
41. O'Halloran P, Kingsley M, Nicholson M, Staley K, Randle E, Wright A, et al. Responsiveness of the single item measure to detect change in physical activity. PLoS ONE. 2020;15(6).
42. Zwolinsky S, McKenna J, Pringle A, Widdop P, Griffiths C. Physical activity assessment for public health: Efficacious use of the single-item measure. Public Health. 2015;129(12):1630-6.

43. Marshall AL, Smith BJ, Bauman AE, Kaur S. Reliability and validity of a brief physical activity assessment for use by family doctors. British Journal of Sports Medicine. 2005;39(5):294-7.
44. Puig Ribera A, Peña Chimenis Ò, Romaguera Bosch M, Duran Bellido E, Heras Tebar A, Solà Gonfaus M, et al. How to identify physical inactivity in Primary Care: Validation of the Catalan and Spanish versions of 2 short questionnaires. Atencion Primaria. 2012;44(8):485-93.
45. Puig-Ribera A, Martín-Cantera C, Puigdomenech E, Real J, Romaguera M, Magdalena-Belio JF, et al. Screening physical activity in family practice: Validity of the Spanish version of a brief physical activity questionnaire. PLoS ONE. 2015;10(9).
46. Smith BJ, Marshall AL, Huang N. Screening for physical activity in family practice: Evaluation of two brief assessment tools. American Journal of Preventive Medicine. 2005;29(4):256-64.
47. Gionet NJ, Godin G. Self-Reported Exercise Behavior of Employees - A Validity Study. Journal of Occupational and Environmental Medicine. 1989;31(12):969-73.
48. Li S, Carlson E, Holm K. Validation of a single-item measure of usual physical activity. Perceptual and Motor Skills. 2000;91(2):593-602.
49. Graff-Iversen S, Anderssen SA, Holme IM, Jenum AK, Raastad T. Two short questionnaires on leisure-time physical activity compared with serum lipids, anthropometric measurements and aerobic power in a suburban population from Oslo, Norway. European journal of epidemiology. 2008;23(3):167-74.
50. Forsén L, Loland NW, Vuillemin A, Chinapaw MJ, van Poppel MN, Mokkink LB, et al. Self-administered physical activity questionnaires for the elderly: a systematic review of measurement properties. Sports Medicine. 2010;40(7):601-23.
51. Smith TO, McKenna MC, Salter C, Hardeman W, Richardson K, Hillsdon M, et al. A systematic review of the physical activity assessment tools used in primary care. Family Practice. 2017;34(4):384-91.
52. Organization WH. Guidelines on physical activity and sedentary behaviour. Geneva: World Health Organization; 2020.
53. Levin-Zamir D, Leung AYM, Dodson S, Rowlands G. Health Literacy in Selected Populations: Individuals, Families, and Communities from the International and Cultural Perspective. Studies in Health Technology and Informatics. 2017;240:392-414.
54. Marsilio M, Fusco F, Gheduzzi E, Guglielmetti C. Co-Production Performance Evaluation in Healthcare. A Systematic Review of Methods, Tools and Metrics. International Journal of Environmental Research and Public Health. 2021;18(7).
55. Hankinson A. Association of activity and chronic disease risk factors: utility and limitations of objectively measured physical activity data. Journal of the American Dietetic Association. 2008;108(6):945-7.
56. Ball TJ, Joy EA, Goh TL, Hannon JC, Gren LH, Shaw JM. Validity of two brief primary care physical activity questionnaires with accelerometry in clinic staff. Primary Health Care Research & Development. 2015;16(1):100-8.
57. Cruz J, Jácome C, Oliveira A, Paixão C, Rebelo P, Flora S, et al. Construct validity of the brief physical activity assessment tool for clinical use in COPD. Clin Respir J. 2021;15(5):530-9.
58. Danquah IH, Petersen CB, Skov SS, Tolstrup JS. Validation of the NPAQ-short - A brief questionnaire to monitor physical activity and compliance with the WHO recommendations. BMC Public Health. 2018;18(1).

59. Hyvärinen M, Sipilä S, Kulmala J, Hakonen H, Tammelin TH, Kujala UM, et al. Validity and reliability of a single question for leisure-time physical activity assessment in middle-aged women. Journal of Aging and Physical Activity. 2020;28(2):231-41.
60. Iwai N, Hisamichi S, Hayakawa N, Inaba Y, Nagaoka T, Sugimori H, et al. Validity and reliability of single-item questions about physical activity. Journal of Epidemiology. 2001;11(5):211-8.
61. Jackson AW, Morrow JR, Jr., Bowles HR, Fitzgerald SJ, Blair SN. Construct validity evidence for single-response items to estimate physical activity levels in large sample studies. Research Quarterly for Exercise and Sport. 2007;78(2):24-31.
62. Moreno-Llamas A, García-Mayor J, De la Cruz-Sánchez E. Concurrent and convergent validity of a single, brief question for physical activity assessment. International Journal of Environmental Research and Public Health. 2020;17(6):1989.
63. Rose SB, Elley CR, Lawton BA, Dowell AC. A single question reliably identifies physically inactive women in primary care. The New Zealand Medical Journal. 2008;121(1268):U2897.
64. Taylor-Piliae RE, Norton LC, Haskell WL, Mahbouda MH, Fair JM, Iribarren C, et al. Validation of a new brief physical activity survey among men and women aged 60-69 years. American Journal of Epidemiology. 2006;164(6):598-606.
65. Webster S, Khan A, Nitz JC. A brief questionnaire is able to measure population physical activity levels accurately: A comparative validation study. Journal of Clinical Gerontology and Geriatrics. 2011;2(3):83-7.
66. Weiss TW, Slater CH, Green LW, Kennedy VC, Albright DL, Wun CC. The validity of single-item, self-assessment questions as measures of adult physical activity. Journal of Clinical Epidemiology. 1990;43(11):1123-9.
67. Yore MM, Bowles HR, Ainsworth BE, Macera CA, Kohl HW, III. Single versus multiple item questions on occupational physical activity. Journal of Physical Activity and Health. 2006;3(1):102-11.
68. Bauman AE, Richards JA. Understanding of the Single-Item Physical Activity Question for Population Surveillance. Journal of Physical Activity and Health. 2022.

Table 1. Studies overview.

First author, year	Country	Questionnaire	Population assessed	Person conducting assessment (self or interview)	Measurement properties	Measures to compare validity	Reliability test/retest interval	Quality assessment score
Ball, 2015 (56)	USA	PAVS and SNAP	45 adults (clinic staff)	Interviewer	Validity	Accelerometer	NM	13
Bauman, 2022 (27)	New Zealand	SI-days (Bauman, 2022 Question)	13887 adults	Interviewer	Validity	IPAQ, single question hours of PA per week	NM	11
Blomqvist, 2020 (35)	Sweden	SR-1 and SR-2	106 adults with chronic heart failure	Responder	Validity, Sensitivity, Specificity	Accelerometer	NM	12
Cruz, 2021 (57)	Portugal	BPAAT	110 adults with COPD	Responder	Validity, Sensitivity, Specificity	Accelerometer	NM	13
Danquah, 2018 (58)	Denmark	NPAQ-short open-ended and NPAQ-short closed-ended	122 adults	Responder	Validity, Reliability, Sensitivity, Specificity	Accelerometer	2 weeks	15
Gill 2012 (29)	USA	Absolute PA Question and Relative PA Question	159 older adults (mean age 80 years +- 3.9)	Interviewer	Validity, Reliability	Clinical assessment of balance, gait, fear of falling, etc.	7 days	14
Gionet, 1989 (47)	Canada	GCSSI and Elaborative Exercise Questionnaire	551 adults	Interviewer	Validity	VO2max, BMI, ME	NM	11
Graff-Iversen, 2008 (49)	Norway	Gothenburg instrument, CONOR Instrument	1497 adults	Interviewer	Validity	IPAQ-L, VO2max, BMI, Triglycerides, HDL-Cholesterol, Waist-to-hip ratio	3 Years	10
Hamilton, 2012 (38)	Australia	SIPAM - Parent Version	458 adults pedometer n=30	Interviewer	Validity	Pedometer, 7-day PAR	NM	12

Hyvärinen, 2020 (59)	Finland	SR-PA L7	795 adult women	Self-administered	Validity, Reliability	Accelerometer	120 days	14
Iwai, 2001 (60)	Japan	JACC Questionnaire	1,730 adults for validity and 1,075 adults for reliability	Interviewer	Validity, Reliability	Interview similar to the Minnesota LTPA questionnaire	1 Year	13
Jackson, 2007 (61)	USA	PA5, PA8	687 Adults	NR	Validity	Clinical assessment, maximal TT and MHPW	NM	10
Johansson, 2008 (34)	Netherlands	WLPAQ	9 adults with rheumatoid arthritis	Self-administered	Validity	DLW	NM	7
Kiernan, 2013 (32)	USA	L-CAT	267 overweight/obese women	Self-administered	Validity, Reliability	Pedometer and BMI	2-6 Weeks	11
Li, 2000 (48)	USA	Usual PA Scale	188 adult women	Self-reported	Validity	Sturgeon 1993 PAQ, BMI, psychosomatic symptoms, perimenopause-related quality of life.	14 days	12
Macdonald, 2022 (30)	Canada	SIPAM	205 older adults (60 year and older)	Self-reported	Validity	CHAMPS questionnaire	NM	13
Marshall, 2004 (43)	Australia	BPAAT	75 adults	Interview	Validity, Reliability	CSA Accelerometer	7 days	13
Milton, 2011 (39)	UK	SIPAM, SIPAM - month	480 adults	Self-reported	Validity, Reliability	SIPAM - GPAQ, SIPAM - month - APS	7 days	14
Milton, 2013 (40)	UK	SIPAM	66 adults	Self-reported	Validity, Specificity, Sensitivity	Accelerometer	NM	12
Moreno-Llamas, 2020 (62)	Spain	Eurobarometer Survey Question	39,379 adults	Self-reported	Validity	IPAQ	NM	7
O'Halloran, 2020 (41)	Australia	SIPAM	120 adults	NR	Reliability Validity,	Accelerometer	NM	14

					specificity, sensitivity			
Orrell, 2007 (36)	UK	TAM1 TAM2	73 adults with CHD	Questionnaires	Validity, Reliability	Accelerometer	7 days	15
Portegijs, 2017 (31)	Finland	Self-report scale to assess habitual physical activity	848 adults aged 75-90 years (174 for accelerometer data)	Interviewer	Validity	Accelerometer 15-item life space assessment Walking difficulty question Short Physical Performance Battery		12
Puig Ribera, 2012 (44)	Spain	Catalonian and Spanish versions of BPAAT	105 adults	Interview	Validity, Reliability	IPAQ-SF	14-28 Days	15
Puig-Ribera, 2015 (45)	Spain	Spanish BPAAT	1184 adults	Interviewer	Validity	7-day Physical Activity Recall (7DPAR) Accelerometer	NM	16
Rose, 2008 (63)	New Zealand	Single-item screening question	1171 women	Interviewer	Validity, Sensitivity, Specificity	New Zealand Physical Activity Questionnaire - Long Form (NZPAQ-LF)	NM	11
Ross, 2018 (33)	USA	L-CAT	76 adults with obesity	Responder (Online)	Validity, Sensitivity, Specificity	Accelerometer	NM	15
Schechtman, 1991 (28)	USA	St. Louis Working Hearts Program single exercise question	1004 adults	NR	Validity	BMI, HDL Cholesterol, oxygen capacity	NM	7
Smith, 2005 (46)	Australia	BPAAT 3QA	509 adults	Interviewer	Validity, Reliability	Active Australia Questionnaire, Accelerometer	3 days	17
Taylor-Piliae, 2006 (64)	USA	SBAS	1,010 adults	Self-administered	Validity	Stanford Seven Day Physical Activity Recall, cardiovascular disease risk biomarkers	NM	14

Webster, 2011 (65)	Australia	The six-point scale	100 adults	Self-administered	Validity	IPAQ HAP Questionnaire	NM	14
Weiss, 1990 (66)	USA	Job Related Activity Main Daily Activity Compared to Peers Question	33,630 adults	Interview	Validity	Continuous energy expenditure measure	NM	10
Yore, 2006 (67)	USA	BRFSS	5,847 adults	Interview	Validity	OPAQ	NM	12
Zwolinsky, 2015 (42)	England	SIPAM	7650 adults	Self-reported	% Agreement, Sensitivity, Specificity	IPAQ	NM	14

NM=not measured, NR=not reported, PAVS=Physical Activity and Vital Signs, SNAP=Speedy Nutrition and Physical Activity Assessment, SI-Days=single-item day question, IPAQ= International Physical Activity Questionnaire, SR-1=Self-report on activity 1, SR-2=Self-report on activity 2, BPAAT= Brief Physical Activity Assessment Tool, COPD= chronic obstructive pulmonary disorder, NPAQ=Nordic physical activity questionnaire, PA=physical activity, GCSSI=Godin and coworkers simple self-administered instrument, BMI=body mass index, ME=muscular endurance, CONOR=Cohort of Norway, IPAQ-L=International Physical Activity Questionnaire – Long Form, HDL=high density lipoproteins, SIPAM=Single Item Physical Activity Measure, PAR=physical activity recall, SR-PA L7=The Seven-level Single-Question Scale for Self-Reported Leisure Time Physical Activity, JACC=Japan Collaborative Cohort, LTPA=leisure time physical activity, PA5=single-response from five descriptors, PA8=single-response from eight descriptors, WLPAQ=Work Leisure Physical Activity Questionnaire, DLW=Doubly-labeled water, L-CAT=The Stanford Leisure-Time Activity Categorical Item, PAQ=physical activity questionnaire, CHAMPS=Community Health Activities Model Program for Seniors physical activity self-report questionnaire, CSA=Computer Science Application, GPAQ=Global Physical Activity Questionnaire, APS=Active People Survey, TAM=Total Activity Measure, CHD=Coronary Heart Disease, 7DPAR=7-day physical activity recall, NZPAQ-LF=New Zealand physical activity questionnaire-Long form, SBAS=Stanford Brief Activity Survey, HAP=Human activity profile, BRFSS=Behavioral Risk Factor Surveillance System, OPAQ=Occupational Physical Activity Questionnaire

Table 2. Summary of measurement properties.

Questionnaire	Reliability	Validity	Sensitivity	Specificity	Identification of the population meeting WHO criteria for PA
Number days of sufficient amount of PA per week					
Physical Activity and Vital Signs (PAVS) (56)	NM	Total weekly min MVPA Bout PA rP=0.50 Non-Bout rP=0.33 No. of days >30min MVPA Bout rP=0.52 Non-bout rP=0.30	NM	NM	kappa=0.46 (95%CI: 0.04-0.89)
Single Item Physical Activity Measure (SIPAM)	kappa=0.63 (95%CI: 0.54-0.72) Rho=0.72 (39)	Rho=0.53 (39)			
		Total MVPA Rho=0.46 MVBA Bouts Rho=0.57 (40)	Total MVPA: 48% BoutsMVPA: 83% (40)	Total MVPA: 88% BoutsMVPA: 56% (40)	Identifying meeting the guidelines kappa=0.13 (95%CI: 0.12-0.14) Classification of inactive participants kappa=0.45 (95%CI: 0.43-0.47) (40)
			Identifying meeting the guidelines: 18.7% Classification of inactive participants: 74.2% (42)	Identifying meeting the guidelines: 97.2% Classification of inactive participants: 79.7% (42)	Identifying meeting the guidelines: kappa=0.13 (95%CI: 0.12-0.14) Classification of inactive participants kappa=0.45 (95%CI: 0.43-0.47) (42)
		MVPAtotal Rho=0.22 MVPAbouts Rho=0.51	MVPAtotal= 63.1% (95%CI: 50.2-74.7%)	MVPAtotal=52.0% (95%CI: 31.3-72.2%)	

		MVPAtotalchange Rho=0.36 MVPAboutschange Rho=0.40 (41)	MVPAbouts=69.2% (95%CI: 56.6-80.1%) (41)	MVPAbouts=48.0% (95%CI: 21.2-68.7%) (41)	
		Rho=0.33-0.55 (30)			
		AUC-ROC: against question hours per week 0.82 (95%CI: 0.81-0.83); against IPAQ total 0.71 (95%CI: 0.7-0.72); against IPAQ LTPA 0.75 (95%CI: 0.74-0.76) (68)			
		Version for parents: 7-day PAR rP=0.51 Pedometer rP=0.81 (38)			
	Month version: kappaw=0.76 (95%CI 0.69-0.82) Rho=0.82 (39)	Month version: Rho=0.33-0.48 (39)			
Brief Physical Activity Assessment Tool (BPAAT)	Rho=0.61 (95%CI: 0.53-0.69) kappa=0.58 (95%CI: 0.47-0.69) (46)	AAQ Reference Rho=0.54 (95%CI: 0.44-0.63) kappa=46.7% (95%CI: 35.6%-57.9%) Accelerometer Rho=0.39 (95%CI: 0.28-0.49) kappa=0.18 (95%CI: 0.39-0.33) (46)			
	kappa=0.53 (95%CI: 0.33-0.72) (43)	kappa =0.40 (95%CI: 0.12-0.69) (43)			
		rho=0.394-0.435 (37)	50-52% (37)	84-91% (37)	
		Spanish version: Against accelerometer: MPA rP=0.22 (95%CI: 0.156-0.272) VPA rP=0.28	Spanish version: 0.75 (95%CI: 0.713-0.773) (45)	Spanish version: 0.74 (95%CI: 0.702-0.790) (45)	

		(95%CI:0.165-0.391) Agreement with 7DPAR Kappa=0.454 (95%CI: 0.405-0.505) AUC-ROC: 0.8 (95%CI: 0.78-0.83) (45)			
	Spanish version: kappa=0.70 (95%CI: 0.53-0.82) (44)	Spanish version: kappa=0.64 (95%CI: 0.50-0.81) (44)			
	Catalan version: kappa=0.72 (95%CI: 0.55-0.83) (44)	Catalan version: kappa=0.58 (95%CI: 0.43-0.77) (44)			
Three-Question Assessment (3QA) (46)	Rho=0.63 (95%CI: 0.53-0.70) kappa=0.56 (95%CI: 0.44-0.67)	AAQ Reference Rho=0.43 (95%CI: 0.32-0.53) kappa=38.7% (95%CI: 26.4%-51.1%) Accelerometer Rho=0.31 (95%CI: 0.18-0.43) kappa=24.3% (95%CI: 11.6%-36.9%)			
Amount of PA					
Elaborative Exercise Questionnaire (47)		rP=0.18 and lower			
Self-report on Activity 2 (SR-2) (35)		rho=0.21	60% (Youden method)	54% (Youden method)	
JACC Questionnaire (60)	kappaw=0.39 - 0.56	Rho=0.43-0.60			
Nordic Physical Activity Questionnaire (NPAQ-short) (58)	Open ended: MVPA rho=0.82 VPA rho=0.80	MVPA rho=0.33 VPA rho=0.32	90%	50%	kappa=0.42
	Closed-ended: MVPA kappa=0.66 VPA kappa=0.59	MVPA kappa=0.17 VPA kappa=0.21	81%	55%	kappa=0.34

Total Activity Measure (TAM) (36)	TAM 1: Rho=0.73 (95%CI: 0.56-0.83)	rP=0.257			
	TAM 2: Rho=0.82 (95%CI: 0.71-0.88)	rP=0.36-0.38			
Cohort of Norway (CONOR) Instrument (49)		For objective measures (VO2max, BMI, Triglycerides, HDL-Cholesterol, Waist to hip ratio) Rho=-0.14-0.13. Statistically significant increase of MET values IPAQ-L in more active groups.			
General exercise participation/physical activity involvement questions					
Self-report on Activity 1 (SR-1) (35)		rho=0.24	30%	92%	
Single-item screening question to identify physical inactivity (63)		kappa=0.46-0.56	76.7% (95%CI: 73.5-79.7)	81.1% (95%CI: 77.2-84.4)	
St. Louis Working Hearts Program single exercise question (28)		Exercisers had statistically significant lower BMI, greater oxygen capacity in both genders, and woman had a higher HDL cholesterol.			
Self-reported level of PA compared with peers					
Compared to Peers Question (66)		Rho=0.28 Rho=0.29 (with elaboration question)			

Relative PA Question (29)	kappaw=0.56 (95%CI: 0.30-0.82)	Rho from 0.28 (95%CI: 0.15-0.44) to 0.57 (95%CI: 0.38-0.78)			
Categorical descriptions of PA levels					
The Seven-level Single-Question Scale for Self-Reported Leisure Time Physical Activity (SR-PA L7) (59)	Rho=0.707 Three categories Rho=0.622	light PA Rho=0.105, MVPA Rho=0.318, total counts Rho=0.333 Three categories light PA Rho=0.109 MVPA Rho=0.337 total counts Rho=0.333			
Self-report scale to assess habitual physical activity (31)		step count Rho=0.49 moderate PA Rho=0.49 Low-Intensity PA Rho=0.40			

PA categories (61)		PA5: Partial correlations with controlling for age and gender with TT r=0.57; with MHPW, r=0.44. Participants who according to questionnaire met the recommended PA levels had a lower percentage of body fat and cholesterol. PA8: Partial correlations with controlling for age and gender with TT r=0.53; with MHPW, r=0.43. Participants who according to questionnaire met the recommended PA levels had a lower percentage of body fat and cholesterol.			
Absolute PA Question (29)	kappaw=0.75 (95%CI: 0.60-0.91)	Rho from 0.10 (95%CI: 0.00-0.26) to 0.33 (95%CI: 0.19-0.49)			
Job Related Activity (66)		Rho=-0.02			
Main Daily Activity (66)		Rho=0.22			
Godin and coworkers simple		rP=0.4 and lower			

self-administered instrument (GCSSI) (47)					
Usual PA Scale (48)	Rho=0.88 Internal consistency reliability (Correction for attenuation method Rho=0.68)	Against the PA questionnaire Rho=0.66.			
The Stanford Leisure-Time Activity Categorical Item (L-CAT)		Against Steps Rho=0.38-0.41, weekly minutes MVPA Rho=0.38-0.40, Weekly minutes of MVPA bouts Rho=0.38-0.4 (33)	baseline: 30.4% 6mo: 54.2% (33)	baseline: 88.7% 6mo: 74.5% (33)	
	Rho=0.80, Kappaw=0.64 (95%CI: 0.54-0.73) (32)	Linear regression for BMI: An increase in L-Cat category was associated with a lower BMI at baseline of ~0.5 BMI unit. (B= -0.4 (95%CI: -0.8 to -0.1; B= -0.14; P=0.02). Linear regression for Pedometer steps at 6 months: An increase in one L-Cat category was associated with 1059 more steps (95%CI: 712-1407; B= 0.38; P<0.001). (32)			
Work Leisure Physical Activity Questionnaire (WLPAQ) (34)		Bland-Altman plot showed small difference between EE measured by questionnaire and with DLW. Highest r value with the mean values=0.89.			
Speedy Nutrition and Physical Activity Assessment (SNAP) (56)		Total weekly min MVPA Bout PA Rho=0.32 Non-Bout Rho=0.41 No. of days >30min MVPA Bout Rho=0.31 Non-bout Rho=0.49			kappa=0.12(95%CI: 0.04-0.28)

Stanford Brief Activity Survey (SBAS) (64)		Higher active categories in SBAS have statistically significant higher minutes per week and EE per day (Stanford Seven-day Physical activity recall) and lower cardiovascular disease risk biomarkers.			
The six-point scale (65)		IPAQ kappa=0.46 (95%CI: 0.26-0.66) HAP kappa=0.57 (95%CI: 0.41-0.73) IPAQ Rho=0.49 HAP Rho=0.69			
Occupational Physical Activity Question from the Behavioral Risk Factor Surveillance System (BRFSS) (67)		kappa=0.56(95%CI: 0.54-0.58)			
Eurobarometer Survey Question (62)		AUC-ROC: 0.751 (95%CI: 0.743-0.759) to 0.947 (95%CI: 0.945-0.95) (higher for higher PA level)			
Gothenburg Instrument (49)		For objective measures (VO2max, BMI, Triglycerides, HDL-Cholesterol, Waist to hip ratio) Rho=-0.19-0.16. Statistically significant increase of MET values IPAQ-L in more active groups.			

PA=physical activity, NM=not measured, MVPA=moderate to vigorous physical activity, CI=confidence interval, AUC-ROC=Area under the receiver operating characteristic curve, PAR=physical activity recall, BPAAT=brief physical activity assessment tool, AAQ=active Australia questionnaire, PAVS=physical activity and vital signs, SIPAM=single-item physical activity questionnaire, MPA=moderate physical activity, VPA=vigorous physical activity, 7DPAR=7-day physical activity recall, TAM=total activity measure, BMI=body mass index, IPAQ-L=International physical activity questionnaire long form, HDL=high density lipoprotein, MET=metabolic equivalent, CONOR=Cohort of Norway, SR-1=Self-report on activity 1, SR-PA L7=the Seven-level Single-Question Scale for Self-Reported Leisure Time Physical Activity, PA-5= single-response from five descriptors, PA-8= single-response from eight descriptors, MHPW=MET hours of reported PA per Week, TT=treadmill test, GCSSI=Godin and coworkers simple self-administered instrument, L-

CAT=the Stanford Leisure-Time Activity Categorical Item, WLPAQ=Work Leisure Physical Activity Questionnaire, DLW=doubly-labeled water, EE=energy expenditure, IPAQ=International Physical activity questionnaire, HAP=human activity profile, BRFSS=Behavioral Risk Factor Surveillance System

Table 3. Length and readability levels of short physical activity questionnaires.

Tool	Text	Word Count	Flesch Reading Score*	Flesch-Kincaid Grade Level*	Self-administered Reading Time (seconds)	Interviewer Speaking Time (seconds)
Number days of sufficient amount of PA per week						
Physical Activity and Vital Signs (PAVS)	1 How many days in the past week have you performed PA where your heart beats faster and your breathing is harder than normal for 30 minutes or more? (in 3 times 10 minute bouts, or 1 time 30 minutes bout) *0 1 2 3 4 5 6 7 (circle one)* 2 How many days in a typical week have you performed activity such as this? *0 1 2 3 4 5 6 7 (circle one)*	76	80,22	6th grade (easy to read)	19	24
Single-Item Physical Activity Measure (SIPAM)	In the past week, on how many days have you done a total of 30 minutes or more of physical activity, which was enough to raise your breathing rate? This may include sport, exercise, and brisk walking or cycling for recreation or to get to and from places, but should not include housework or physical activity that may be part of your job.	63	55,35	10th to 12th grade (fairly difficult to read)	16	21
	Month variation: In the past month, on how many days have you done a total of 30 minutes or more of physical activity, which was enough to raise your breathing rate? This may include sport, exercise, and brisk walking or cycling for recreation or to get to and from places, but should not include housework or physical activity that may be part of your job.	63	55,35	10th to 12th grade (fairly difficult to read)	16	21

	Bauman variation: In the past week, on how many days have you done a total of 30 minutes or more of physical activity, which was enough to raise your breathing rate? This may include sport, traditional games, kapa haka, exercise and brisk walking or cycling for recreation or to get to and from places, but should not include housework or physical activity that may be part of your job.	67	49,09	College (difficult to read)	17	22
	Parents version: On how many days in the course of the past week (past 7 days) have you engaged in at least 15-30 minutes of at least a moderate-intensity PA? (any activity that makes you breathe somewhat harder than normal, e.g., brisk walking, playing energetic games with children, sweeping, cleaning gutters, regular paced swimming or cycling, low-impact aerobics) to activities that make you breathe much harder than normal (e.g., heavy manual labour, moving heavy furniture, playing strenuous games with the children, high impact aerobics, running, fast bicycling) of which could be built up during the day with a variety of activities of at least 10 minutes in length, or done in one session.	111	18,65	College graduate (very difficult to read)	28	36
Brief Physical Activity Assessment Tool (BPAAT)	1 How many times a week do you usually do 20 minutes or more of vigorous-intensity physical activity that makes you sweat or puff and pant? (e.g., heavy lifting, digging, jogging, aerobics, or fast bicycling). *3 or more times a week.* *1 to 2 times a week.* *None.* 2 How many times a week do you usually do 30 minutes or more of moderate-intensity physical activity or walking that increases your heart rate or makes you breathe harder than normal? (e.g., carrying light loads, bicycling at a regular pace, or playing doubles tennis). *5 or more times a week.* *3–4 times a week.* *1–2 times a week.* *None.*	106	49,06	College (difficult to read)	27	35
Three-Question Assessment (3QA)	1 How many times a week do you usually do 20 minutes or more of vigorous-intensity physical activity that makes you sweat or puff and pant? (e.g., heavy lifting, digging, jogging, aerobics, or fast bicycling)? *3 or more times a week* *1 to 2 times a week* *none*	140	52,87	10th to 12th grade (fairly difficult to read)	35	46

	2 How many times a week do you usually do 30 minutes or more walking? (e.g., walking from place to place for exercise, leisure or recreation) *5 or more times a week* *3–4 times a week* *1–2 times a week* *none* 3 How many times a week do you usually do 30 minutes or more of moderate-intensity physical activity that increases your heart rate or makes you breathe harder than normal? (e.g., carrying light loads, bicycling at a regular pace, or doubles tennis) *5 or more times a week* *3–4 times a week* *1–2 times a week* *none*					
Amount of PA						
Elaborative Exercise Questionnaire	Considering a 7-day period (a week), how many times on average do you do the following kinds of exercise for more than 15 minutes during your free time (write in each circle the appropriate number). Strenuous Exercise (heart beats rapidly): Running, jogging, hockey, football, soccer, squash, basketball, cross-country skiing, judo, roller skating, vigorous swimming, vigorous long-distance bicycling. Mark the average number of times that you perform one or more of the above activities during your typical week. Moderate Exercise (not exhausting): Fast walking, baseball, tennis, easy bicycling, volleyball, badminton, easy swimming, alpine skiing, popular and folk dancing. Mark the average number of times that you perform one or more of these activities during your typical week. Mild Exercise (minimal effort): Yoga, archery, fishing from a river bank, bowling, horseshoes, golf, snowmobiling, easy walking. Mark the average number of times that you perform one or more of these activities during your typical week.	152	34,80	College (difficult to read)	38	50

Self-Report on Activity 2 (SR-2)	Over the past week (even if it's not a typical week), how much time did you exercise or were you physically active (e.g. strength training, walking, swimming, gardening or other type of training)? *No time.* *Less than 30 minutes per week.* *30-60 minutes per week.* *1-3 hours per week.* *More than 3 hours per week.*	55	45,16	College (difficult to read)	14	18
Japan Collaborative Cohort (JACC) Questionnaire	How much time per week on average do you spend engaging in sports or physical exercise? *1) At least 5 hours.* *2) 3-4 hours.* *3) 1-2 hours.* *4) Little.* How much time per day on average do you spend walking indoors or outside? *1) More than 1 hour.* *2) 30 minutes - 1 hour.* *3) about 30 minutes.* *4) Little.* How often did you engage in sports or physical exercise over the past year or two? *1) Seldom.* *2) Sometimes.* *3) About once a week.* *4) At least twice a week.*	90	69,89	8th & 9th grade (plain English)	23	30
Nordic Physical Activity Questionnaire – Short Version (NPAQ-short)	*Close ended:* Physical activities in your free time. The following questions concern how physically active you are in your free time and during transport (including your commute to and from work/school/classes). On a typical week, how much time do you spend in total on moderate and vigorous physical activities where your heartbeat increases and you breathe faster (e.g. brisk walking, cycling as a means of transport or as exercise, heavy gardening, running or recreational sports). Only include activities that lasted at least 10 minutes at a time.	237	47,74	College (difficult to read)	60	78

	Less than ½ an hour (less than 30 minutes). *½ an hour – 1 ½ hours (30-90 minutes).* *1 ½ - 2 ½ hours (90-150 minutes).* *2 ½ - 5 hours (150-300 minutes).* *More than 5 hours (150-300 minutes).* How much of the time that you spend on physical activities in a typical week, which you indicated above, do you spend in total on vigorous physical activities? This includes activities that get your heart racing, make you sweat and leave you so short of breath that speaking becomes difficult (e.g. swimming, running, cycling at high speeds, cardio training, weight-lifting or team sports such as football). Only include activities that lasted at least 10 minutes at a time. *Less than ½ an hour (less than 30 minutes).* *½ an hour – 1 hours (30-60 minutes).* *1 - 1 ½ hours (60-90 minutes).* *1 ½ - 2 ½ hours (90-150 minutes).* *More than 2 ½ hours (more than 150 minutes).*					
	Open ended: Physical activities in your free time. The following questions concern how physically active you are in your free time and during transport (including your commute to and from work/school/classes). On a typical week, how much time do you spend in total on moderate and vigorous physical activities where your heartbeat increases and you breathe faster (e.g. brisk walking, cycling as a means of transport or as exercise, heavy gardening, running or recreational sports). Only include activities that lasted at least 10 minutes at a time. *Hours pr week.* *Minutes pr week.* How much of the time that you spend on physical activities in a typical week, which you indicated above, do you spend in total on vigorous physical activities? This includes activities that get your heart racing, make you sweat and leave you so short of breath that speaking becomes difficult (e.g. swimming, running, cycling at high speeds, cardio training, weight-lifting or team sports such as football). Only include activities that lasted at least 10 minutes at a time. *Hours pr week.* *Minutes pr week.*	175	47,74	College (difficult to read)	44	57

Total Activity Measure 1 (TAM1)	We'd like to know how active you are. 1 In an average week how many times do you do 'Strenuous' activities? Strenuous activity means any activity that take as much effort as – jogging or running, digging in heavy ground, aerobics classes, vigorous swimming, heavy DIY jobs like sawing thick or long pieces of wood, football, circuit training, squash... *'Strenuous activity' times a week.* 1b If you do a 'strenuous' activity, how many minutes do you do it for on average? 'Strenuous activity' minutes each time. 2 In an average week how many times do you do 'moderate' activities? Moderate activity means any activity that take as much effort as – brisk walking, housework, carrying a light shopping bag on level ground, mowing the lawn, general DIY like painting and decorating, easy swimming, easy cycling, ballroom dancing... *'Moderate activity' times a week.* 2b If you do a moderate activity, how many minutes do you do it for on average? *'Moderate activity' minutes each time.* 3 In an average week how many times do you do any mild activities? Mild activity means any activity that take as much effort as – easy walking, very light housework, browsing in shops, slow dancing, bowls, river fishing, golf, hand weeding in the garden... *'Mild activity' times a week.* 3b If you do a 'mild' activity, how many minutes do you do it for on average? *'Mild activity' minutes each time.* Please check that you have put a number in all 6 boxes, even if it is a 0. Thank you for your help.	259	56,69	10th to 12th grade (fairly difficult to read)	63	82
	TAM 2 We'd like to know how active you are. 1 In an average week how many times do you do 'Strenuous' activities? Strenuous activity means any activity that take as much effort as – jogging or running, digging in heavy ground, aerobics classes, vigorous swimming, heavy DIY jobs like sawing thick or long pieces of wood, football, circuit training, squash... *'Strenuous activity' times a week.*	244	55,31	10th to 12th grade (fairly difficult to read)	59	77

	1b For how many minutes in total do you do strenuous activities? *'Strenuous activity' total minutes.* 2 In an average week how many times do you do 'moderate' activities? Moderate activity means any activity that take as much effort as – brisk walking, housework, carrying a light shopping bag on level ground, mowing the lawn, general DIY like painting and decorating, easy swimming, easy cycling, ballroom dancing... *'Moderate activity' times a week.* 2b For how many minutes in total in an average week do you 'mild' activities? 3 In an average week, how many times do you do any mild activities? Mild activity means any activity that take as much effort as – easy walking, very light housework, browsing in shops, slow dancing, bowls, river fishing, golf, hand weeding in the garden... *'Mild activity' times a week.* 3b For how many minutes in total in an average week do you do mild activities? *'Mild activity' total minutes.* Please check that you have put a number in all 6 boxes, even if it is a 0. Thank you for your help.					
Cohort of Norway (CONOR) Instrument	How has your physical activity during leisure time been over the last year? Think of your weekly average for the year. Time spent going to or from work counts as leisure time. Hours per week: None/Less than 1/1-2/3 or more. Light activity (not sweating or out of breath). Hard physical activity (sweating/out of breath). Please note physical activity during the past year in your spare time. What does describe you better? If activity varies between summer and wintertime, note a mean value. (Tick one only) Reading, watching TV or any other sedentary activity? Walking, cycling, or other activity, other for at least 4 hours a week? (Count also walking back and forth from work) Light sports, heavy gardening? (At least 4 hours per week).	132	67,44	8th & 9th grade (Plain English)	33	43

	Hard exercise, competitive sports? Regularly and several times a week.					
General exercise participation/physical activity involvement questions						
Self-Report on Activity 1 (SR-1)	I exercise regularly. *1 2 3 4 5*	8	N/A	N/A	2	3
Single-item screening question to identify physical inactivity	As a rule, do you do at least half an hour of moderate or vigorous exercise (such as walking or a sport) on five or more days of the week? *Yes.* *No.*	32	72,05	7th grade (fairly easy to read)	8	10
St. Louis Working Hearts Program single exercise question	Do you currently participate in any regular activity or program (either on your own or in a formal class) designed to improve or maintain your physical fitness? *Yes.* *No.*	29	32,16	College (difficult to read)	6	8
Self-reported level of PA compared with peers						
Compared to Peers Question	Would you say that you are physically more active, less active, or about as active as other persons your age? *Elaboration:* Is that a lot more or a little more/a lot less or a little less active?	36	73,42	7th grade (fairly easy to read	9,08	11,80
Relative PA Question	Compared to other people your own age, do you think you are... - *Much more active* - *More active* - *About as active* - *Less active* - *Much less active*	25	88,91	6th grade (easy to read)	6,30	8,20

Categorical descriptions of PA levels						
The Seven-level Single-Question Scale for Self-Reported Leisure Time Physical Activity (SR-PA L7)	Which of the following descriptions best corresponds to your physical activity at the moment? I do not move more than is necessary in my daily routines/chores. I go for casual walks and engage in light outdoor recreation 1-2 times a week. I go for casual walks and engage in light outdoor recreation several times a week. I engage, 1-2 times a week, in brisk physical activity (e.g. yard work, walking, and cycling) to the point of perspiring and some degree of breathlessness. Several times a week (3-5), I engage in brisk physical activity (e.g. yard work, walking, and cycling) to the point of perspiring and some degree of breathlessness. I do keep-fit exercises several times a week in a way that causes rather strong shortness of breath and sweating during the activity. I participate in competitive sports and maintain my fitness through regular training.	144	54,03	10th to 12th grade (fairly difficult to read)	36	47
Self-report scale to assess habitual physical activity	If you think about the past 6 months, which of the following descriptions best pictures your physical activity? Mostly resting, hardly any activity. Mostly sitting. You are usually doing things in a seated position, reading and watching TV. Your only physical activities relate to activities of daily living (grooming, dressing). Light physical activity. You are doing light housework (e.g., preparing food, dusting) or light gardening or going for a walk two to three times a week. Moderate physical activity about 3 hours a week. You are doing common housework (e.g., vacuum cleaning/sweeping floors, lawnmowing) or going for longer walks (at least 2 km) or cycling. Moderate physical activity at least 4 hours a week or heavier physical activity up to 4 hours a week (daily more than 30 minutes). You are doing moderate physical activities (see before) for at least 4 hours or physically exercise 1–2 hours a week or doing heavy gardening/housework or home maintenance involving some breathlessness and sweating.	203	51,79	10th to 12th grade (fairly difficult to read)	51	67

	You are engaging in active sports several times a week, which makes you heavily sweat and breathless during the exercise or you are doing heavy gardening or leisure time activities (at least 3 hours a week). You are participating in competitive sports.					
PA categories	Single response from eight descriptors (PA8): 1 I do not exercise/walk regularly now and I do not intend to start in the near future. 2 I do not exercise/walk regularly but I have been thinking of starting. 3 I am trying to start to exercise or walk or I exercise/walk infrequently. 4 I am doing vigorous physical activity less than three times per week or moderate physical activity less than five times per week. 5 I have been doing moderate physical activity that accumulates to at least 30 minutes per day at least 5 days per week for 1-6 months. 6 I have been doing moderate physical activity that accumulates to at least 30 minutes per day at least 5 days per week for 7 or more months. 7 I have been doing vigorous physical activity at least 20 minutes a day 3-5 days per week for 1-6 months. 8 I have been doing vigorous physical activity at least 20 minutes a day 3-5 days per week for 7 or more months.	166	70,09	7th grade (fairly easy to read)	40	52
	Single response from five descriptors (PA5): 1 I don't exercise or walk regularly now and I don't plan to start in the near future. 2 I don't exercise or walk regularly now but I've been thinking about starting. 3 I'm doing moderate physical activity fewer than five times a week or vigorous activity fewer than three times a week. 4 I've been doing moderate physical activity 5 or more days a week, or vigorous activity at least 3 days a week, for the last 1 to 6 months. 5 I've been doing moderate physical activity 5 or more days a week or vigorous activity at least 3 days a week, for 7 months or longer.	109	68,29	8th & 9th grade (plain English)	27	36
Absolute PA Question	What best describes your activity level? Vigorously active for at least 30 minutes, 3 times per week. Moderately active for at least 3 times per week. Seldom active, preferring sedentary activities.	34	57,06	10th to 12th grade (fairly difficult to read)	9	11

Job Related Activity	How much hard physical work is required on your job? Would you say a great deal, a moderate amount, a little or none?	23	84,81	6th grade (easy to read)	6	8
Main Daily Activity	How much hard physical work is required in your main daily activity? Would you say a great deal, a moderate amount, a little or none?	25	72,32	7th grade (fairly easy to read)	6	8
Godin and coworkers simple self-administered instrument (GCSSI)	How often did you participate in one or more physical activities of 20 to 30 minutes duration per session during your leisure time within the past 6 months? *1 Not at all.* *2 Less than once a month.* *3 About once a month.* *4 About two or three times a month.* *5 About one to two times a week.* *6 Three or more times a week.*	66	54,54	10th to 12th grade (fairly difficult to read)	17	22
The Stanford Leisure-Time Activity Categorical Item (L-CAT)	During the past month, which statement best describes the kinds of physical activity you usually performed during your FREE TIME (i.e., recreational or leisure time)? 1 I did not engage in much physical activity. I mostly did things like watching television, reading, playing cards, or playing computer games. Only occasionally, no more than once or twice a month, did I do anything more active such as getting outdoors for a walk or playing tennis. 2 Once or twice a week, I engaged in light activities such as getting outdoors on the weekends for a walk. Or I did some light chores around the house such as sweeping floors or vacuuming. 3 About three times a week, I did some moderate activity such as brisk walking, swimming, or riding a bike for about 15–20 minutes each time. Or about once a week, I did some moderately difficult chores such as raking, washing windows, or mowing the lawn for about 45–60 minutes. Or about once a week, I played sports such as doubles tennis or basketball for about 45–60 minutes. 4 Almost daily, that is five or more times a week, I did some moderate activity such as brisk walking, swimming, or riding a bike for 30 minutes or more each time. Or about once a week, I did some moderately difficult chores or played team sports for 2 hours or more.	331	62,20	8th & 9th grade (plain English)	83	109

	5 About three times a week, I engaged in a regular program of physical fitness involving some kind of heavy or vigorous physical activity such as running or riding hard on a bicycle for 30 minutes or more each time. Or I did chores such as heavy gardening or played active sports such as handball or singles tennis for 60 minutes or more each time. 6 Almost daily, that is, five or more times a week, I engaged in a regular program of physical fitness involving some kind of heavy or vigorous physical activity for 30 minutes or more each time.					
	Version 2.2: During the past month, which statement best describes the kinds of physical activity you usually did? Do not include the time you spent working at a job. Please read all six statements before selecting one. 1 I did not do much physical activity. I mostly did things like watching television, reading, playing cards, or playing computer games. Only occasionally, no more than once or twice a month, did I do anything more active such as going for a walk or playing tennis. 2 Once or twice a week, I did light activities such as getting outdoors on the weekends for an easy walk or stroll. Or once or twice a week, I did chores around the house such as sweeping floors or vacuuming. 3 About three times a week, I did moderate activities such as brisk walking, swimming, or riding a bike for about 15–20 minutes each time. Or about once a week, I did moderately difficult chores such as raking or mowing the lawn for about 45–60 minutes. Or about once a week, I played sports such as softball, basketball, or soccer for about 45–60 minutes. 4 Almost daily, that is five or more times a week, I did moderate activities such as brisk walking, swimming, or riding a bike for 30 minutes or more each time. Or about once a week, I did moderately difficult chores or played sports for 2 hours or more. 5 About three times a week, I did vigorous activities such as running or riding hard on a bike for 30 minutes or more each time. 6 Almost daily, that is, five or more times a week, I did vigorous activities such as running or riding hard on a bike for 30 minutes or more each time.	293	72,38	7th grade (fairly easy to read)	74	96

Usual PA Scale	My job requires very hard physical labor (such as digging or loading heavy objects) at least four hours a day. Or I do vigorous activities (jogging, cycling, swimming, etc.) at least three times per week for 30-10 min. or more. Or I do at least one hour of moderate activity such as brisk walking at least four days a week. My job requires that I walk, lift, carry or do other moderately hard work for several hours per day (day care worker, stock clerk, or busboy/waitress). Or I spend much of my leisure time doing moderate activities (dancing, gardening, walking or housework). My job requires that I sit at a desk most of the day. And much of my leisure time is spent in sedentary activities (watching TV, reading, etc.). And I seldom work up a sweat and I cannot walk fast without having to stop to catch my breath.	150	69,36	8th & 9th grade (plain English)	38	49
Work Leisure Physical Activity Questionnaire (WLPAQ)	Describe your physical activity at work (even work at home, sick leave at home and studying, for instance in a university). 1 Very light, e.g., sitting at the computer most of the day or sitting at a desk. 2 Light, e.g., light industrial work, sales or office work that comprises light activities. 3 Moderate, e.g., cleaning, staffing at kitchen or delivering mail on foot or by bicycle. 4 Heavy, e.g., heavy industrial work, construction work or farming. Describe your physical activity at leisure time. If the activities vary between summer and winter, try to give a mean estimate. 1 Very light: almost no activity at all. 2 Light, e.g., walking, nonstrenuous cycling or gardening approximately once a week. 3 Moderate: regular activity at least once a week, e.g., walking, bicycling, or gardening or walking to work 10–30 minutes per day. 4 Active: regular activities more than once a week, e.g., intense walking or bicycling or sports. 5 Very active: strenuous activities several times a week.	165	56,49	10th to 12th grade (fairly difficult to read)	42	54

Speedy Nutrition and Physical Activity Assessment (SNAP)	How Active are You? Examples of activity are Walking Housework Work in the yard or garden Dancing Jobs that require walking, lifting or other hard work Exercise Are you active for 30 minutes on 5 days of the week? 1 No and I have no plans to be more active. 2 No, but I have been thinking about being more active. 3 Sometimes I am active for 30 minutes but not all the time. 4 Yes, I am active for 30 minutes on 5 days of the week.	89	88,12	6th grade (easy to read)	22	29
Stanford Brief Activity Survey (SBAS)	*Appendix Table 1* A If you have no job or regular work, check box A and go on to Appendix table 2. B I spent most of the day sitting or standing. When I was at work, I did such things as writing, typing, talking on the telephone, assembling small parts, or operating a machine that takes very little exertion or strength. If I drove a car or truck while at work, I did not lift or carry anything for more that a few minutes each day. C I spent most of the day walking or using my hands and arms in work that required moderate exertion. When I was at work, I did such things as delivering mail, patrolling on guard duty, doing mechanical work on automobiles or other large machines, house painting, or operating a machine that requires some moderate-activity work of me. If I drove a truck or lift, my job required me to lift and carry things frequently. D I spent most of the day lifting or carrying heavy objects or moving most of my body in some other way. When I was at work, I did such things as stacking cargo or inventory, handling parts or materials, or doing work like that of a carpenter who builds structures or a gardener who does most of the work without machines. E I spent most of the day doing hard physical labor. When I was at work, I did such things as digging or chopping with heavy tools or carrying heavy loads (bricks, for example) to the place where they were to be used. If I drove a truck or operated equipment, my job also required me to do hard	600	60,80	8th & 9th grade (plain English)	151	197

	physical work most of the day with only short breaks. *Appendix Table 2* F Most of my leisure time was spent without very much physical activity. I mostly did things like watching television, reading, or playing cards. If I did anything else, it was likely to be light chores around the house or yard or some easy-going game like bowling or catch. Only occasionally, no more than once or twice a month, did I do anything more vigorous, like jogging, playing tennis, or active gardening. G Weekdays, when I got home from work, I did few active things, but most weekends I was able to get outdoors for some light exercise—going for walks, playing a round of golf (without motorized carts), or doing some active chores around the house. H Three times per week, on average, I engaged in some moderate activity, such as brisk walking or slow jogging, swimming, or riding a bike for 15–20 minutes or more, or I spent 45 minutes to an hour or more doing moderately difficult chores, such as raking or washing windows, mowing the lawn or vacuuming, or playing games such a doubles tennis or basketball. I During my leisure time over the past year, I engaged in a regular program of physical fitness involving some kind of heavy physical activity at least three times per week. Examples of heavy physical activity are jogging, running, or riding fast on a bicycle for 30 minutes or more; heavy gardening or other chores for an hour or more; active games or sports such as handball or tennis for an hour or more; or a regular program involving calisthenics and jogging or the equivalent for 30 minute or more. J Over the past year, I engaged in a regular program of physical fitness along the lines described in the last paragraph (I), but I did it almost daily—five or more times per week.					
The six-point scale	1 Moving only for necessary chores. 2 Walking or other outdoor activities one or two times per week. 3 Walking or other outdoor activities several times per week. 4 Exercising one or two times per week to the point of perspiring and heavy breathing. 5 Exercising several times per week to the point of perspiring and heavy breathing. 6 Keep-fit heavy exercise or competitive sport several times per week.	69	64	8th & 9th grade (plain English)	17	23

Occupational Physical Activity Question	When you are at work, which of the following best describes what you do? *Mostly sitting or standing.* *Mostly walking.* *Mostly heavy labor or physically demanding work.* *Do not work.* *Don't know/not sure.* *Refused.*	35	89,90	6th grade (easy to read)	9	11
Eurobarometer Survey Question	How often do you exercise, play sport, or engage in other physical activity, such as cycling from one place to another, dancing, gardening, etc.? *Never.* *Occasionally.* *Few times a week.* *Almost daily.*	32	41,48	College (difficult to read)	8	10
Gothenburg Instrument	Please note physical activity during the past year in your spare time. What does describe you better. If activity varies between summer and wintertime, note a mean value. (Tick one only) 1 Reading, watching TV or any other sedentary activity? 2 Walking, cycling, or other activity, other for at least 4 hours a week? (Count also walking back and forth from work.) 3 Light sports, heavy gardening? (At least 4 hours per week.) 4 Hard exercise, competitive sports? Regularly and several times a week	80	63,36	8th & 9th grade (plain English)	20	26

* = Only the main text of the questionnaires was used to calculate the level of readability. The text written in italics was excluded from the analysis.
PA = Physical Activity