

Self-Report Measures of Physical Activity

Emma L. Healey,¹  Kelli D. Allen,² Kim Bennell,³ Jocelyn L. Bowden,⁴ Jonathan G. Quicke,¹ and Robert Smith⁵

INTRODUCTION

Many people with musculoskeletal (MSK) conditions can reduce their pain and improve their quality of life by being more physically active (1). Physical activity (PA) is internationally recommended as a core treatment for common MSK conditions, such as osteoarthritis (OA) (2–5). However, people with MSK conditions are often less active than those without such conditions (6–8), with less than half of adult patients with OA meeting PA guidelines (9).

PA levels can be measured by using objective methods, such as accelerometry and pedometers, or subjective methods, such as self-report measures. Use of self-report PA measures is a popular approach because they are easy to use and are of low cost (10). Two systematic reviews have evaluated the measurement properties of self-report PA measures in patients with OA (11,12), but to date, there is still no consensus regarding which self-report PA measure is the most suitable for use in MSK populations.

In this article, we aim to summarize and critically assess the most widely used self-report PA measures in studies of common MSK conditions. The authors selected measures based on the following criteria: 1) administered by self-report, 2) most commonly cited for use in MSK populations (eg, OA, low back pain [LBP], rheumatoid arthritis [RA], ankylosing spondylitis [AS], and fibromyalgia [FM]) in the last 5 years, and 3) has evidence of psychometric data in MSK populations. A two-stage computerized literature search by using Medline and Embase was performed. In the first search, medical subject headings for MSK conditions and MSK pain were used in conjunction with terms for exercise and PA to identify studies that used self-report PA measures. A second search was then conducted to retrieve studies that

evaluated measurement properties of the self-report PA measures identified from the first search.

The measures reviewed below include the Baecke Physical Activity Questionnaire (BPAQ), the International Physical Activity Questionnaire (IPAQ) (three versions), the Physical Activity Scale for the Elderly (PASE), and the Short Questionnaire to Assess Health-Enhancing Physical Activity (SQUASH). Details on the content and structure of each measure are presented (number of items, recall period, response options, presence of translations, and adaptations) (see Table 1). Information on cost and how to obtain the measures is also provided, as available. Important practical information is presented, including evidence for psychometric properties of each measure in common MSK populations (as available) (see Table 2). A critical appraisal of each measure is provided, and the review concludes with a summary and recommendations specific to the rheumatology community.

BAECKE PHYSICAL ACTIVITY QUESTIONNAIRE

Description

Purpose. The BPAQ (13) is a self-report PA measure developed, originally, for use in epidemiological studies to assess levels of PA in young adults.

Content or domains. Self-reported responses across three domains are used to assess PA levels in a typical week: 1) occupational physical activities (eight questions), 2) sport (four questions), and 3) leisure (four questions). Total PA is the sum of all indices.

The views expressed herein are those of the authors and do not necessarily represent those of the National Health Service, the National Institute for Health Research, or the Department of Health and Social Care.

Dr. Healey's work was partly supported by the National Institute for Health Research Applied Research Collaboration West Midlands. Dr. Allen's work was supported by the Center of Innovation for Health Services Research in Primary Care (grant CIN 13-410) at the Durham VA Health Care System and by the NIH-funded Core Center for Clinical Research (grant P30-AR-072580). Dr. Bennell's work was supported by the National Health and Medical Research Council and by Medibank Private for research into osteoarthritis. Dr. Quicke's work was partly supported by a National Institute for Health Research Clinical Research Network West Midlands Research Scholar Fellowship.

¹Emma L. Healey, PhD, Jonathan G. Quicke, PhD: Primary Care Centre Versus Arthritis and School of Primary, Community and Social Care, Keele

University, Keele, UK; ²Kelli D. Allen, PhD: The University of North Carolina at Chapel Hill, and Center of Innovation to Accelerate Discovery and Practice Transformation, Department of Veterans Affairs Healthcare System, Durham, North Carolina; ³Kim Bennell, PhD: Centre for Health Exercise and Sports Medicine, The University of Melbourne, Melbourne, Australia; ⁴Jocelyn L. Bowden, PhD: Institute of Bone and Joint Research, Kolling Institute, The University of Sydney, Sydney, Australia; ⁵Robert Smith, PhD: School of Nursing, The University of Hong Kong, Hong Kong.

No potential conflicts of interest relevant to this article were reported.

Address correspondence to Emma L. Healey, PhD, Keele University, School for Primary, Community and Social Care, Staffordshire, ST5 5BG, UK. E-mail: e.healey@keele.ac.uk.

Submitted for publication January 30, 2020; accepted in revised form April 7, 2020.

Number of items. Sixteen questions divided across three domains and summed for a total level of PA.

Response options/scale. Questions are scored on a five-point Likert scale ranging from “never” to “always” or “very often.”

Recall period for items. Typical week.

Cost to use. Free.

How to obtain. The questionnaire and scoring formula are available within the appendix of the original article (13).

Practical application

Method of administration. Self-completed.

Scoring. Each domain (occupation, sports, leisure) can receive a score from 1 to 5, which is achieved through a scoring formula for that domain (see below). Within domains, each question is also given a score from 1 to 5, with the exception of questions asking for main occupation and types of sports played. The total score is a sum of the scores for each domain and can range between 3 and 15, with a higher score corresponding to a higher PA level.

For the work domain, the respondent’s occupation is given a score of 1 (low activity), 3, (moderate activity), or 5 (high activity) based on work activity ratings from the Netherlands Nutrition Council. The overall work index is calculated by using the following formula (“Q” represents the question number): $[(6 - Q2) + [Q1 + Q3 + Q4 + Q5 + Q6 + Q7 + Q8)]/8$.

For the sport domain, the first question involves calculating a “simple sports score” based on the respondent’s sport activity (a calculation of their two most frequently played sports). First, it is determined what two sports the respondent plays most frequently. Sports are subdivided into three categories, each of which has a value corresponding to intensity (average energy expenditure). Second, it is determined how frequently the respondent engages in the sport in hours per week, with different numbers of hours given a different value. Finally, it is determined how many months in the year the respondent plays their most frequently played sports. A value is given for different numbers of months per year. Once these three values are determined, the sport score is determined by using the following formula: $([value\ for\ intensity\ of\ most\ frequent\ sport] \times [value\ for\ weekly\ time\ of\ most\ frequent\ sport] \times [value\ for\ yearly\ proportion\ of\ most\ frequent\ sport]) \times ([value\ for\ intensity\ of\ second\ sport] \times [value\ for\ weekly\ time\ of\ second\ sport] \times [value\ for\ yearly\ proportion\ of\ second\ sport])$. Finally, an overall sport index is calculated by using the following formula: $(simple\ sport\ score + Q10 + Q11 + Q12)/4$. For the leisure domain, the

overall leisure index is calculated by using the following formula: $[(6 - Q13) + [Q14 + Q15 + Q16)]/4$.

Score interpretation. Scores cannot be interpreted in relation to other metrics of PA.

Respondent time to complete. The BPAQ requires minimal burden; it is quick to complete because all items are multiple choice.

Administrative burden. Time to administer is short, and scoring can be done quickly by totaling scores for each index and summing all of these scores for a total score.

Translations/adaptations. There are no repositories of the BPAQ to identify all the translations, but the questionnaire has been used internationally in different populations with multiple languages, including Japanese (14), Dutch (15), Persian (16), French (17), Flemish (18), Greek (19), Korean (20), and Portuguese (21).

Psychometric information

Floor and ceiling effects. Not reported.

Reliability. *Internal consistency.* Not reported in any studies.

Test-retest. An intrasession intraclass correlation coefficient (ICC) of 0.77 (95% confidence interval [CI]: 0.65-0.84) was seen in patients with LBP (22). An ICC of 0.87 was seen in adult women with hip disorders (14).

Validity. *Content/face validity.* Not reported in any studies.

Criterion validity. Not reported in any studies.

Construct validity. There was significant and fair correlation in people with LBP ($\rho = 0.18$) between the BPAQ score and the number of steps and vector magnitude on an Actigraph wGT3X-BT accelerometer (22). In women with hip disorders, there was significant but low correlation ($\rho = 0.30-0.49$) across three measures of the BPAQ and a higher correlation between step counts and the total score on the BPAQ ($\rho = 0.49$) (14).

Responsiveness. Not reported in any studies.

Minimally important differences. Not reported in any studies.

Generalizability. The BPAQ was originally developed for epidemiological research and tested in a young Dutch male and female population (aged 32 years and younger). It was translated into multiple languages, but its application in older populations and in MSK populations is limited.

Use in clinical trials. No randomized controlled trials (RCTs) in MSK populations that used the BPAQ questionnaire could be identified; only one protocol was identified (23).

Critical appraisal of overall value to the rheumatology community

Strengths. Items are closed-response items, which allows for easy completion and scoring. Scores can be broken down into three indices, allowing for interpretations on how individuals are loading their PA. Total scores of the BPAQ appear to have adequate reliability (ICC > 0.7).

Caveats and cautions. The BPAQ is limited to two studies of evidence in measurement properties in MSK populations, and there are low correlations to objective measures. There is no evidence on sensitivity to changes/responsiveness. Scores cannot be interpreted in relation to recommended levels of PA.

Clinical usability. The BPAQ is quick and easy to administer and score. Individual scores across indices can be used to identify areas in which individuals are most and least active (occupation, sport, or leisure).

Research usability. No RCTs in MSK populations that used the BPAQ could be identified, and only one protocol was identified (23). The BPAQ was designed for use in epidemiological studies only. Several observational studies in MSK populations have used the BPAQ (24–27).

INTERNATIONAL PHYSICAL ACTIVITY QUESTIONNAIRE

Description

Purpose. The purpose of the IPAQ is to measure internationally comparable PA levels in adult populations (28). It was designed primarily for surveillance of PA at a population level and has been predominantly used in studies of people with OA. It is not recommended for use as an outcome measure in small-scale intervention studies.

Content or domains. The IPAQ is available for use with young and middle-aged individuals (ages 15–69 years) in long and short forms. There is also a short-form version validated for use in older adults (ages 65 years and older) (29).

The IPAQ Long Form (LF) covers five activity domains asked as separate sections, namely, PA undertaken related to 1) work, 2) transportation, 3) housework, 4) leisure-time activities, and 5) time spent sitting. The IPAQ-LF asks questions around the frequency (days) and duration (minutes) spent in each of these domains, with

a focus on three types of activity: vigorous intensity, moderate intensity, and time spent walking (28).

The IPAQ Short Form (SF) and IPAQ for the Elderly (E) ask about the four specific activity types (vigorous intensity, moderate intensity, time spent walking, and time spent sitting) undertaken during any work, transportation, housework, or leisure activity. Time spent sitting is asked as a separate question and used as an indicator of sedentary behavior (see scoring manual at www.ipaq.ki.se).

Number of items. The IPAQ-LF has 27 items; however, some can be skipped if the individual does not participate in any activities for a given domain (eg, no work-related vigorous or moderate activities undertaken).

The IPAQ-SF has six items, or seven if the individual reports any time spent walking.

The IPAQ-E covers the same items as the IPAQ-SF but is presented as only four questions.

Response options/scale. All forms ask people completing the survey to indicate the number of days per week and hours and minutes per day they spend doing PA within each domain. Individuals may indicate that they are not sure of the activity undertaken.

Recall period for items. Past week (last 7 days).

Cost to use. All versions are free to use and open to access (www.ipaq.ki.se).

How to obtain. The IPAQ forms are freely available in multiple languages, either in PDF or Word document format.

Practical application

Method of administration. The IPAQ-LF and IPAQ-SF are available in either a telephone-administered or self-completed format. The IPAQ-E is only available in a self-administrated format.

Scoring. An English version of the scoring protocol for the IPAQ-LF and IPAQ-SF is freely available. The scoring protocol for the IPAQ-SF can be used to score the IPAQ-E. Automatic scoring templates and reports are available for select languages on the website.

Walking, moderate, and vigorous scores are converted into metabolic equivalents (METs) in minutes per week by using the Ainsworth Compendium (30). An average MET score is based on time spent on each activity intensity and is then added together to make a total. This can then be converted into a categorical score of three categorical levels: low, moderate, and high levels of activity. Definitions of these categories are outlined in the IPAQ scoring protocol.

The IPAQ-LF provides separate domain-specific scores for vigorous-intensity activity, moderate-intensity activity, and walking within the four activity-related domains (excluding sitting). Computation of the total scores for this form involves summation of the frequency (days) and duration (minutes) for all activities in all domains. Domain-specific scores require summation of the scores of the different intensity activities within the specific domain. Activity-specific subscores can also be calculated by summing the scores of the specific types across the domains.

The IPAQ-SF provides separate scores for vigorous-intensity activity, moderate-intensity activity, and walking. However, domain-specific estimates cannot be calculated. Data for time spent sitting are not included in the summed score of PA but should be reported and used to categorize activity into low, moderate, or high levels (see the IPAQ-SF scoring protocol).

The IPAQ-E is scored similarly to the IPAQ-SF; however, Hurtig-Wennlöf et al (29) suggest caution should be used when converting to METs because of lower metabolic rates in older adults. Rather, minutes per day in each intensity can be reported.

Score interpretation. Scores are given in total METs in minutes per week, giving an estimate to energy expenditure in a week. The categorical score can be interpreted in terms of PA recommendations: low scores are below recommendations, moderate scores are meeting recommendations, and high scores are exceeding the recommendations.

Respondent time to complete. Time to complete has not been reported; however, because the IPAQ-SF and IPAQ-E have fewer than seven short-response items, time to complete is minimal. The IPAQ-LF is longer to administer, although it still takes a relatively short time to complete.

Administrative burden. Time to administer the questionnaires is very short because only a printout and a pen are required to complete it. Scoring can be completed by hand or by calculator or on a simple spreadsheet. No additional equipment or software is required.

Translations/adaptations. There are currently 24 translated versions available (www.ipaq.ki.se), although not for all form types. The website includes a guide for translating the IPAQ into languages not currently available.

Psychometric information

Floor and ceiling effects. Not reported in any studies.

IPAQ-LF reliability. *Internal consistency.* Not reported in any studies.

Test-retest reliability. An ICC of 0.65 was seen in patients with total hip replacement (THR) and/or total knee replacement (TKR) (31), an ICC of 0.83 was seen in patients with AS (32), an ICC of 0.77 was seen in women with FM (33), and an ICC of 0.37 was seen in patients with LBP (25). A standard error of the measurement (SEM) of 2668 METs in minutes per week and a minimal detectible change of 1115 METs in minutes per week was seen in patients with THR and/or TKR (31). In comparisons across 12 countries, the ICC ranged from 0.96 to 0.46 (28). The IPAQ-LF has shown weak reliability for sedentary behavior and moderate to vigorous activity in healthy older adult populations; therefore, care should be taken when using it to classify PA levels in older populations (34).

IPAQ-SF reliability. *Internal consistency.* Not reported in any studies.

Test-retest reliability. An ICC of 0.76 was seen in patients with THR, and an ICC of 0.87 was seen in patients with total knee replacement (TKR) (35). An ICC of 0.51 was seen in patients with THR and/or TKR (31). An SEM of 2487 METs in minutes per week and a minimal detectible change of 1039 METs in minutes per week was seen in patients with THR and/or TKR (31). In comparisons across 12 countries, the ICC ranged from 0.88 to 0.32 (28). An ICC of 0.64, an SEM of 3532 METs in minutes per week, and smallest detectable change (SDC) of 9791 METs in minutes per week was seen in patients with OA in a hip, knee, foot, and hand OA sample (36).

IPAQ-E reliability. Not reported in any studies.

IPAQ-LF validity. *Content/face validity.* Not reported in any studies.

Criterion validity. Correlations to Computer Science Application (CSA) accelerometers across 12 countries were equal to 0.33 (28).

Construct validity. Correlation to an ActiGraph GT1M accelerometer was equal to 0.43 in patients with THR and/or TKR (31), correlation to an ActiGraph GT1M was equal to 0.38 in patients with AS (32), and concordance correlation with a SenseWear Pro Armband was equal to 0.04 in women with FM (33) compared with an Actigraph GT3X for individual overestimates in RA (37). Correlation to Actigraph wGT3X-BT counts was equal to 0.33 in patients with LBP (25). Compared with an Actigraph, older adult self-report had moderate correlations for moderate to vigorous PA (MVPA) (0.43-0.56 and 0.70-0.26 for sedentary behavior), but they tended to underestimate both MVPA and sedentary behavior (38).

IPAQ-SF validity. *Content/face validity.* Not reported in any studies.

Criterion validity. Correlations to CSA accelerometers across 12 countries were equal to 0.30 (28).

Construct validity. Correlation to ActiGraph GT1M accelerometers was equal to 0.29 in patients with THR and/or TKR (31), correlation to the PASE was equal to 0.61 in patients with hip OA (39), and correlation to a SenseWear activity monitor was equal to 0.40 in patients with RA (40). Correlation to the PASE was equal to 0.56 in a hip, knee, foot, and hand OA sample (36).

IPAQ-E validity. *Content/face validity.* Not reported in any studies.

Criterion validity. Activity domains positively correlated to an ActiGraph GT1M at 0.28 to 0.47 in older adults (29). A main effect for each category (high, medium, and low) was observed with the highly sensitive serum C-reactive protein biomarker.

Construct validity. Not reported in any studies.

Responsiveness. The effect size was equal to -0.14 , the standard responsiveness measure was equal to -0.21 , and the responsiveness ratio was equal to 0.12 in six patients in a hip, knee, foot, and hand OA sample (36).

Minimally important differences. Not reported in any studies.

Generalizability. The IPAQ is designed as a population-based measure. It is generic and designed to be used across all adults aged 18 to 65 years (IPAQ-LF and IPAQ-SF) or 65 years and older (IPAQ-E) with or without clinical conditions (28). However, there are studies in certain populations that suggest it should be used cautiously with these groups.

Use in clinical trials. A number of RCTs have used the IPAQ-LF (41) and the IPAQ-SF (42–47) to assess and classify the level of PA in a study population.

Critical appraisal of overall value to the rheumatology community

Strengths. Scores for the IPAQ-SF and IPAQ-LF relate to weekly energy expenditure. Scores can be compared to recommended levels of PA or between different conditions. The IPAQ has been translated into different languages and is easy to administer and quick to complete. The forms are open access, readily accessible, and free to use. The IPAQ-SF is widely used to research different MSK conditions and has been used in a range of OA and rheumatology studies. It covers activities across multiple domains, including work, leisure, and home life. The IPAQ-LF has more evidence for reliability and construct validity in relation to objective measures compared with the IPAQ-SF or IPAQ-E. It was tested and developed in both high- and low-income countries (28).

Caveats and cautions. There is limited evidence of measurement properties in MSK populations for any of the forms, and there is no evidence in MSK populations for the IPAQ-E. All forms have low correlations to objective measures, and there is no evidence on sensitivity to changes/responsiveness. It is not recommended for use as an outcome measure for small intervention studies. Care should be taken when converting the IPAQ-E to METs (29). The IPAQ has not been validated for use in online studies.

Clinical usability. The questionnaires are quick and easy to administer and score, regardless of the form used. Individuals' scores can be related to evidence for levels of PA that can lead to health benefits. The IPAQ-E reports good acceptance by older users (29). It may not be appropriate for use with certain patient populations.

Research usability. The forms can be self-completed or administered via telephone. They have been translated in different languages and can be used in different countries/languages with direct comparisons. The inclusion of the sitting activity scores in the IPAQ-SF can provide data on inactivity and sedentary behavior in this population. The IPAQ-LF may be more applicable for research that requires more detailed assessments of PA.

Although the IPAQ questionnaire has been used in several RCTs, it was predominately designed for observational or population-based studies. Several observational studies in MSK populations have used the IPAQ-LF (48) and the IPAQ-SF (44,49–53).

PHYSICAL ACTIVITY SCALE FOR THE ELDERLY

Description

Purpose. The PASE was developed in the United States in a general older adult population and measures self-reported PA in older adults in the previous week (54).

Content or domains. The PASE contains three subdomains: leisure activities, household activities, and occupational work.

Number of items. The PASE has 12 items. The leisure activities domain contains five items (subdomains), the household activities domain contains six items, and the occupational work domain contains a single item.

Response options/scale. PASE scores are calculated by using both weights and frequency values for each of the 12 item-activity types. Respondents report activities by 1) providing categorical responses to the number of days per week and

average hours per day (leisure activity domain), 2) reporting if they have performed an activity or not (items in the household activity domain), and 3) reporting hours worked per week (occupational domain). The scale range is 0 to 400 or more (higher scores indicate higher PA level).

Recall period for items. In the last week.

Cost to use. There is a cost for the scoring manual and cost per use of the questionnaire (visit www.healthcore.com for more details).

How to obtain. The questionnaire and scoring protocol are available at www.healthcore.com.

Practical application

Method of administration. The PASE is self-completed or administered via telephone interview (recommended).

Scoring. Scoring involves totaling the scores from the three activity domains and rounding to the nearest integer. PASE scores are calculated by using both weights (intensity) and frequency values for each of the 12 item-activity types. However, each activity domain has a unique scoring method. The weighting of item activities was based on an algorithm derived by using accelerometry, activity diary, and global activity self-assessment (54) (see the scoring manual available at www.healthcore.com).

For leisure activities, individuals respond with categorical responses to the number of days per week (never, seldom, sometimes, or often) and average hours per day of activity (less than 1, 1-2, 2-4, 4 or more) within each item subdomain. A "PASE activity time to hours per day conversion table" is then used to convert these categorical data into hours per day. Different item activities are assigned different weight scores, which are then multiplied by the hours per day score for each item and totaled to give the domain subscore. For household activities, individuals provide a binary response to whether they have performed individual household activities in the last week, which is then weighted by each subdomain item and totaled for the domain subscore. Occupational hours worked is divided by seven and given a weight score for the occupational domain subscore.

Score interpretation. Higher PASE scores indicate higher levels of PA. The PASE estimates PA; however, its scores are not directly interpretable in meaningful PA units.

Respondent time to complete. Self-completed or interviewer-administered versions can be completed in 5 to 15 minutes.

Administrative burden. Administering time is 5 to 15 minutes. Time taken to score depends on the use of computer coding and is not provided in the literature.

Translations/adaptations. Originally developed in English in the United States, the PASE has been translated into Dutch (55), Norwegian (56), Japanese (57), Chinese (58), German (59), Malaysian (60), Turkish (61), Italian (62), and Persian (63). It has been adapted for Mexican-origin Hispanic patients in the southwestern United States by using an adapted scoring algorithm (64). It has also been adapted for Dutch populations by adding bicycling for transportation to the question about time spent walking (55).

Psychometric information

Floor and ceiling effects. Not specifically reported in any studies.

Reliability. *Internal consistency.* Not reported in any studies. *Test-retest.* An ICC of 0.77 was seen in patients with hip OA (39). An ICC of 0.77 was seen in men and an ICC of 0.58 was seen in women following TKR (65). An ICC of 0.77 was seen post-THR (59). An ICC of 0.68 was seen in a hip, knee, foot, and hand OA sample (36).

Measurement error. An SEM of 31 and SDC of 87 was seen in patients with hip OA (66). An SEM of 32% and 35% and an SDC of 89% and 97% was seen in men and women, respectively, following TKR (66). There was an SEM of 23.0% post-THR (59). An SEM of 46.7 and an SDC of 129.6 was seen in a hip, knee, foot, and hand OA sample (36).

Validity. *Content/face validity.* Not reported in any studies. *Criterion/convergent validity.* PASE scores significantly correlated in expected directions with performance in the 6-minute walk test, knee strength, knee pain frequency during transfer, and perceived difficulty with physical functioning in older adults with knee pain and physical disability (67).

Construct validity. Correlation of the total PASE score with accelerometer-based activity counts in patients with hip OA was 0.30 ($P = 0.089$) and ranged from 0.20 to 0.38 for the different PA categories (66). Correlation with accelerometer-based activity counts was 0.45 in men following TKR and 0.06 in women following TKR (65). Correlation with accelerometer-based activity counts was 0.27 in patients following THR (59).

Responsiveness. An effect size of -0.16 , a standard response measurement of -0.21 , and a response ratio of 0.09 was seen in a hip, knee, foot, and hand OA sample (36).

Minimally important differences. Not reported in any studies.

Generalizability. The PASE was developed in a US population but has undergone translation and validation in multiple countries (some of these have been in MSK populations) (39,59,65). One study (64) has questioned its generalizability (weighting of items) in Mexican elderly Americans.

Use in clinical trials. The measure has been used in multiple RCTs in OA populations (42,66,68–75).

Critical appraisal of overall value to the rheumatology community

Strengths. The PASE is designed specifically for older adults, is relatively quick to complete (5–15 minutes), is translated into multiple languages, and is often used in OA studies.

Caveats and cautions. The PASE has mixed reliability results, large measurement error, and poor responsiveness. It is better suited to older adult populations. It has weak correlations with objective measures of PA and is unable to discriminate between intensity of activity within individual subdomains (67). Cost is associated with use.

Clinical usability. The PASE is relatively quick and easy to administer, but scoring may be more time consuming and difficult in a consultation setting. Scores are not easily interpreted into meaningful units.

Research usability. The PASE is quick and easy to administer, so it could be used in large studies, including trials and observational studies (76–79). It is validated in older adult populations with joint pain (eg, OA populations). However, large measurement error and poor responsiveness properties suggest it is not useful in measuring change in PA.

SHORT QUESTIONNAIRE TO ASSESS HEALTH-ENHANCING PHYSICAL ACTIVITY

Description

Purpose. The SQUASH was developed in the Netherlands and measures the habitual activities in a normal week over the past months (80).

Content or domains. The SQUASH contains five subdomains: 1) commuting activities, 2) activity at work or school, 3) household activities, 4) leisure-time activities, and 5) sports.

Number of items. The SQUASH has up to 14 items, although not all need to be completed because each subdomain has a “not applicable” option as an item. The domains of commuting, work or school, and household activities each have two items, the leisure-time domain has four items, and the

sports domain can include up to four sports activities indicated by the respondent.

Response options/scale. Individuals respond with the number of days per week and average time per day (hours and minutes) spent on each activity within each subdomain.

Recall period for items. An average week over the past months.

Cost to use. Free.

How to obtain. The questionnaire and scoring protocol are published in the article by Wendel-Vos et al (80).

Practical application

Method of administration. Self-completed.

Scoring. Scoring is completed by taking a sum of the time spent active in each domain in total minutes and multiplying by the intensity score (80). Activities are divided into three intensity categories based on Ainsworth’s compendium of physical activities: 2.0 to less than 4.0 METs (light), 4.0 to less than 6.5 METs (moderate), and greater than or equal to 6.5 METs (vigorous) (30,81,82). Some studies have used different intensity categories for older adults: 2.0 to less than 3.0 METs (light), 3.0 to less than 5.0 METs (moderate), greater than or equal to 5.0 METs (vigorous) (83). Activities with an MET score below 2.0 are not counted.

Score interpretation. The SQUASH does not estimate energy expenditure but does estimate habitual activity in an average week for individuals. Some studies have summed the number of days per week for moderate and vigorous activity lasting at least 30 min/wk to evaluate adherence to American College of Sports Medicine and Dutch activity guidelines (84,85).

Respondent time to complete. Less than 5 minutes.

Administrative burden. Administering time is less than 5 minutes, but time taken to score could be relatively longer because intensity scores need to be assigned to activities, including open-ended sports questions.

Translations/adaptations. The SQUASH was originally developed in Dutch (80). An English version is available, but the process for adaptation/translation has not been published. The SQUASH has been translated into Turkish (86) and Japanese (87).

Psychometric information

Floor and ceiling effects. Not reported in any studies.

Reliability. *Internal consistency.* Not reported in any studies.

Test-retest. An ICC of 0.89 was found in patients with AS (32). Spearman's correlation of 0.57 was found in patients with THR (83).

Validity. *Content/face validity.* Not reported in any studies.

Criterion validity. Not reported in any studies.

Construct validity. Correlation (0.35) with accelerometer-based activity counts was found in patients with AS (32). Correlation with accelerometer parameters ranged from $r = 0.28$ to $r = 0.49$ in patients with knee OA (88). Correlations with accelerometer-based activity parameters were 0.20 to 0.67 in patients with THR (83).

Responsiveness. Not reported in any studies.

Minimally important differences. Not reported in any studies.

Generalizability. The SQUASH was developed in a Dutch population and has had limited use to date in other countries.

Use in clinical trials. The measure has been used in RCTs investigating knee OA (aqua cycling) (89) and RA (motivation and self-regulation for PA; combination therapies) (85,90).

Critical appraisal of overall value to the rheumatology community

Strengths. Scores can be related to time spent physically active, allowing individuals to be categorized in relation to recommended levels of PA. The SQUASH takes less than 5 minutes to complete. Respondents have the opportunity to report any sporting activities in open-ended questions.

Caveats and cautions. The SQUASH has limited use among individuals with rheumatic and MSK conditions. There are low correlations with objective measures of PA and mixed results on reliability. There is no evidence on sensitivity to changes/responsiveness, and the measure has limited use in RCTs. The SQUASH is time intensive to score.

Clinical usability. The SQUASH is quick and easy to administer, but scoring may be more time consuming and difficult in a consultation setting. Individual scores on weekly minutes of PA can be related to public health recommendations.

Research usability. The SQUASH is quick and easy to administer, so it could be used in large studies. However, some psychometric properties are not well established. Use of the measure in MSK conditions has involved mostly observational

studies, including measurement during daily activities in patients with knee OA (88) and multiple studies focused on establishing the psychometric properties of the scale in different patient groups (32,83,84,86,87). Use in RCTs has been limited, with studies among patients with knee OA (89) and RA (85,90).

CONCLUSIONS

To our knowledge, no self-report measure of PA has been developed specifically for use in populations with MSK conditions. It is therefore important for clinicians and researchers to understand what options are available and how commonly used measures reflect actual PA levels.

Four self-report PA measures were selected and reviewed based on their frequent use in MSK populations in the last 5 years and based on the identification of psychometric evidence properties in MSK populations. The authors acknowledge that other commonly used self-report measures, such as the Community Healthy Activities Model Program for Seniors (91), the Minnesota Leisure Time Physical Activity Questionnaire (92), and the Yale Physical Activity Survey (93), could potentially be suitable for use in MSK populations; however, they were not included in this review for two main reasons. Firstly, they are not commonly used in MSK research, and secondly, their psychometric evidence in MSK and older populations is lacking.

All of the measures included in this review were found to be quick and easy to complete. The majority of measures can also be scored relatively quickly (BPAQ, IPAQ forms, and PASE). The SQUASH, however, features open response items to allow respondents to report any sports or physical activities that may not naturally fall into other subdomains, which can make scoring more difficult.

Although it would be beneficial to use self-report PA measures to determine whether an individual or group is meeting current PA guidelines or to allow for clinically useful categorization of PA levels (eg, inactive, low activity, meeting recommendations), of all the measures included in this review, only the IPAQ and SQUASH allow for this. The impact of sedentary behavior on MSK conditions has gained increasing interest in recent years (94–96); however, of the measures included in this review, only the IPAQ forms assess sedentary time (sitting time). In addition, the SQUASH actually discounts low-level activities (less than 2.0 METs), which may be particularly important to some MSK populations because they may only be able to perform low-level activities. It is important that self-report PA measures record all levels of activity, otherwise they may be underestimating overall levels of PA.

The PASE and IPAQ-SF are currently the most commonly used measures in OA research, and the BPAQ appears to be most frequently used in LBP research. Overall, psychometric evidence of all the measures identified is lacking in MSK populations (see Table 2), and based on the evidence currently

available, none of the measures demonstrate adequate measurement properties in terms of all components of reliability and validity. More studies have examined the measurement properties of the IPAQ-LF, yet the IPAQ-SF appears to be a more frequently used measure, possibly because of the reduced number of items and time to complete. There was a clear lack of evidence in terms of responsiveness for the measures in this review. The evidence for the PASE suggests poor responsiveness (36), and others have questioned its ability to detect change in PA levels (97).

In conclusion, because the measures included in this review lack evidence of their psychometric properties and responsiveness to change in MSK populations, the authors suggest that caution be taken when using self-report PA measures. It is also important to note the wider limitations of all self-report measures, that is, potential for social desirability bias, recall bias, over- and underestimation of activities, and misclassification of activities (98,99). Therefore, when possible, the use of objective measures of PA (eg, accelerometry) should be considered. There is greater evidence of their validity and reliability (100), and they can objectively capture all dimensions of PA, including time spent sedentary, which is known to detrimentally affect the general health and functional status of MSK populations (96). Further research is needed to investigate the measurement properties of commonly used self-report PA measures in MSK populations to allow for informed recommendations and decisions on their use.

AUTHOR CONTRIBUTIONS

All authors drafted the article, revised it critically for important intellectual content, and approved the final version to be published.

REFERENCES

1. Versus Arthritis. Providing physical activity interventions for people with musculoskeletal conditions report. 2017. URL: <https://www.versusarthritis.org/policy/policy-reports/providing-physical-activity/>.
2. Fernandes L, Hagen KB, Bijlsma JW, Andreassen O, Christensen P, Conaghan PG, et al. EULAR recommendations for the non-pharmacological core management of hip and knee osteoarthritis. *Ann Rheum Dis* 2013;72:1125–35.
3. McAlindon TE, Bannuru RR, Sullivan MC, Arden N, Berenbaum F, Bierma-Zeinstra SM, et al. OARSI guidelines for the non-surgical management of knee osteoarthritis. *Osteoarthritis Cartilage* 2014;22:363–88.
4. National Institute for Health and Care Excellence. Osteoarthritis: care and management: clinical guideline [CG177]. 2014. URL: <https://www.nice.org.uk/guidance/cg177>.
5. Kolasinski SL, Neogi T, Hochberg MC, Oatis C, Guyatt G, Block J, et al. 2019 American College of Rheumatology/Arthritis Foundation guideline for the management of osteoarthritis of the hand, hip, and knee. *Arthritis Care Res (Hoboken)* 2020;72:149–62.
6. Hernández-Hernández V, Ferraz-Amaro I, Díaz-González F. Influence of disease activity on the physical activity of rheumatoid arthritis patients. *Rheumatology (Oxford)* 2014;53:722–31.
7. Swinnen TW, Scheers T, Lefevre J, Dankaerts W, Westhovens R, de Vlam K. Physical activity assessment in patients with axial spondyloarthritis compared to healthy controls: a technology-based approach. *PLoS One* 2014;9:e85309.
8. Herbolzheimer F, Schaap LA, Edwards MH, Maggi S, Otero Á, Timmermans EJ, et al. Physical activity patterns among older adults with and without knee osteoarthritis in six European countries. *Arthritis Care Res (Hoboken)* 2016;68:228–36.
9. Holden MA, Nicholls EE, Young J, Hay EM, Foster NE. Exercise and physical activity in older adults with knee pain: a mixed methods study. *Rheumatology (Oxford)* 2015;54:413–23.
10. Warren JM, Ekelund U, Besson H, Mezzani A, Geladas N, Vanhees L, et al. Assessment of physical activity: a review of methodologies with reference to epidemiological research: a report of the exercise physiology section of the European Association of Cardiovascular Prevention and Rehabilitation. *Eur J Cardiovasc Prev Rehabil* 2010;17:127–39.
11. Terwee C, Bouwmeester W, van Elsland SL, de Vet H, Dekker J. Instruments to assess physical activity in patients with osteoarthritis of the hip or knee: a systematic review of measurement properties. *Osteoarthritis Cartilage* 2011;19:620–33.
12. Smith RD, Dziedzic KS, Quicke JG, Holden MA, McHugh GA, Healey EL. Identification and evaluation of self-report physical activity instruments in adults with osteoarthritis: a systematic review. *Arthritis Care Res (Hoboken)* 2019;71:237–51.
13. Baecke JA, Burema J, Frijters JE. A short questionnaire for the measurement of habitual physical activity in epidemiological studies. *Am J Clin Nutr* 1982;36:936–42.
14. Ono R, Hirata S, Yamada M, Nishiyama T, Kurosaka M, Tamura Y. Reliability and validity of the Baecke Physical Activity Questionnaire in adult women with hip disorders. *BMC Musculoskelet Disord* 2007;8:61.
15. Vogels N, Westerterp KR, Posthumus DL, Rutters F, Westerterp-Plantenga MS. Daily physical activity counts vs structured activity counts in lean and overweight Dutch children. *Physiol Behav* 2007;92:611–6.
16. Sadeghisani M, Dehghan Manshadi F, Azimi H, Montazeri A. Validity and reliability of the Persian version of Baecke Habitual Physical Activity Questionnaire in healthy subjects. *Asian J Sports Med* 2016;7:e31778.
17. Van der Waerden J, Nakamura A, Pryor L, Charles MA, El-Khoury F, Dargent-Molina P, et al. Domain-specific physical activity and sedentary behavior during pregnancy and postpartum depression risk in the French EDEN and ELFE cohorts. *Prev Med* 2019;121:33–9.
18. Beunen GP, Philippaerts RM, Delvaux K, Thomis M, Claessens AL, Vanreusel B, et al. Adolescent physical performance and adult physical activity in Flemish males. *Am J Hum Biol* 2001;13:173–9.
19. Kaspiris A, Zaphiropoulou C, Vasilidi E. Range of variation of genu valgum and association with anthropometric characteristics and physical activity: comparison between children aged 3–9 years. *J Pediatr Orthop B* 2013;22:296–305.
20. Lee JY, Yun YH, Park EC, Seo HW, Lee JH, Shin HR, et al. Reliability and validity of the modified Korean version of Baecke Questionnaire on Physical Activity. *Korean J Epidemiol* 2004;26:20–31.
21. Bellafrente NT, Serafini RK, Chiarello PG. Relationship between total physical activity and physical activity domains with body composition and energy expenditure among Brazilian adults. *Am J Hum Biol* 2019;31:e23317.
22. Carvalho FA, Maher CG, Franco MR, Morelhão PK, Oliveira CB, Silva FG, et al. Fear of movement is not associated with objective and subjective physical activity levels in chronic nonspecific low back pain. *Arch Phys Med Rehabil* 2017;98:96–104.
23. Oliveira CB, Franco MR, Maher CG, Tiedemann A, Silva FG, Damato TM, et al. The efficacy of a multimodal physical activity

- intervention with supervised exercises, health coaching and an activity monitor on physical activity levels of patients with chronic, nonspecific low back pain (Physical Activity for Back Pain (PAY-BACK) trial): study protocol for a randomised controlled trial. *Trials* 2018;19:40.
24. Bento TP, Cornelio GP, Perrucini PO, Simeão SF, de Conti MH, de Vitta A. Low back pain in adolescents and association with sociodemographic factors, electronic devices, physical activity and mental health. *J Pediatr (Rio J)* 2019. E-pub ahead of print.
 25. Carvalho FA, Morelhão PK, Franco MR, Maher CG, Smeets RJ, Oliveira CB, et al. Reliability and validity of two multidimensional self-reported physical activity questionnaires in people with chronic low back pain. *Musculoskelet Sci Pract* 2017;27:65–70.
 26. Chiment RL, Scholtes SA, van Dillen LR. Activity characteristics and movement patterns in people with and people without low back pain who participate in rotation-related sports. *J Sport Rehabil* 2013;22:161–9.
 27. Handrakis JP, Friel K, Hoeffner F, Akinkunle O, Genova V, Isakov E, et al. Key characteristics of low back pain and disability in college-aged adults: a pilot study. *Arch Phys Med Rehabil* 2012;93:1217–24.
 28. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International Physical Activity Questionnaire: 12-country reliability and validity. *Med Sci Sports Exerc* 2003;35:1381–95.
 29. Hurtig-Wennlöf A, Hagströmer M, Olsson LA. The International Physical Activity Questionnaire modified for the elderly: aspects of validity and feasibility. *Public Health Nutr* 2010;13:1847–54.
 30. Ainsworth BE, Haskell WL, Whitt MC, Irwin ML, Swartz AM, Strath SJ, et al. Compendium of physical activities: an update of activity codes and MET intensities. *Med Sci Sports Exerc* 2000;32 Suppl:S498–504.
 31. Blikman T, Stevens M, Bulstra SK, van den Akker-Scheek I, Reininga IH. Reliability and validity of the Dutch version of the International Physical Activity Questionnaire in patients after total hip arthroplasty or total knee arthroplasty. *J Orthop Sports Phys Ther* 2013;43:650–9.
 32. Arends S, Hofman M, Kamsma YP, van der Veer E, Houtman PM, Kallenberg CG, et al. Daily physical activity in ankylosing spondylitis: validity and reliability of the IPAQ and SQUASH and the relation with clinical assessments. *Arthritis Res Ther* 2013;15:R99.
 33. Segura-Jimenez V, Munguia-Izquierdo D, Camiletti-Moiron D, Alvarez-Gallardo IC, Ortega FB, Ruiz JR, et al. Comparison of the International Physical Activity Questionnaire (IPAQ) with a multi-sensor armband accelerometer in women with fibromyalgia: the al-Ándalus project. *Clin Exp Rheumatol* 2013;31 Suppl 79:S94–101.
 34. Ryan DJ, Wullems JA, Stebbings GK, Morse CI, Stewart CE, Onambele-Pearson GL. Reliability and validity of the international physical activity questionnaire compared to calibrated accelerometer cut-off points in the quantification of sedentary behaviour and physical activity in older adults. *PLoS One* 2018;13:e0195712.
 35. Naal FD, Impellizzeri FM, Leunig M. Which is the best activity rating scale for patients undergoing total joint arthroplasty? [original article]. *Clin Orthop Relat Res* 2009;467:958–65.
 36. Smith RD. Self-reported physical activity levels: measurement and assessment in community dwelling adults with or at risk of osteoarthritis [thesis]. Keele (UK): Keele Univ.; 2017.
 37. Yu CA, Rouse PC, Veldhuijzen van Zanten JJ, Ntoumanis N, Kitas GD, Duda JL, et al. Subjective and objective levels of physical activity and their association with cardiorespiratory fitness in rheumatoid arthritis patients. *Arthritis Res Ther* 2015;17:59.
 38. Cleland C, Ferguson S, Ellis G, Hunter RF. Validity of the International Physical Activity Questionnaire (IPAQ) for assessing moderate-to-vigorous physical activity and sedentary behaviour of older adults in the United Kingdom. *BMC Med Res Methodol* 2018;18:176.
 39. Svege I, Kolle E, Risberg MA. Reliability and validity of the Physical Activity Scale for the Elderly (PASE) in patients with hip osteoarthritis. *BMC Musculoskelet Disord* 2012;13:26.
 40. Tierney M, Fraser A, Kennedy N. Criterion validity of the International Physical Activity Questionnaire Short Form (IPAQ-SF) for use in patients with rheumatoid arthritis: comparison with the SenseWear Armband. *Physiotherapy* 2015;101:193–7.
 41. Zacharia S, Taylor EL, Branscum PW, Cheney MK, Hoffer CW, Crowson M. Effects of a yoga intervention on adults with lower limb osteoarthritis: a randomized controlled trial. *Am J Health Stud* 2018;33:89–98.
 42. Dzedzic KS, Healey EL, Porcheret M, Afolabi EK, Lewis M, Morden A, et al. Implementing core NICE guidelines for osteoarthritis in primary care with a model consultation (MOSAICS): a cluster randomised controlled trial. *Osteoarthritis Cartilage* 2018;26:43–53.
 43. Eichler S, Rabe S, Salzwedel A, Müller S, Stoll J, Tilgner N, et al. Effectiveness of an interactive telerehabilitation system with home-based exercise training in patients after total hip or knee replacement: study protocol for a multicenter, superiority, no-blinded randomized controlled trial. *Trials* 2017;18:438.
 44. Gay C, Guiguet-Auclair C, Pereira B, Goldstein A, Bareyre L, Coste N. Efficacy of self-management exercise program with spa therapy for behavioral management of knee osteoarthritis: research protocol for a quasi-randomized controlled trial (GEET one). *BMC Complement Altern Med* 2018;18:279.
 45. Rodrigues da Silva JM, de Rezende MU, Spada TC, da Silva Francisco L, Sabine de Farias FE, Clemente da Silva CA, et al. Educational program promoting regular physical exercise improves functional capacity and daily living physical activity in subjects with knee osteoarthritis. *BMC Musculoskelet Disord* 2017;18:546.
 46. Sit RW, Chan KK, Yip BH, Zhang DD, Reeves KD, Chan YH, et al. Clinical effectiveness of patella mobilisation therapy versus a waiting list control for knee osteoarthritis: a protocol for a pragmatic randomised clinical trial. *BMJ Open* 2018;8:e019103.
 47. Vassão PG, de Souza MC, Silva BA, Junqueira RG, de Camargo MR, Dourado VZ, et al. Photobiomodulation via a cluster device associated with a physical exercise program in the level of pain and muscle strength in middle-aged and older women with knee osteoarthritis: a randomized placebo-controlled trial. *Lasers Med Sci* 2020;35:139–48.
 48. Figueiredo Neto EM, Queluz TT, Freire BF. Physical activity and its association with quality of life in patients with osteoarthritis. *Rev Bras Reumatol* 2011;51:544–9.
 49. Magnusson K, Hagen KB, Østerås N, Nordsletten L, Natvig B, Haugen IK. Diabetes is associated with increased hand pain in erosive hand osteoarthritis: data from a population-based study. *Arthritis Care Res (Hoboken)* 2015;67:187–95.
 50. Lee SY, Ro HJ, Chung SG, Kang SH, Seo KM, Kim DK. Low skeletal muscle mass in the lower limbs is independently associated to knee osteoarthritis. *PLoS One* 2016;11:e0166385.
 51. Shim HY, Park M, Kim HJ, Kyung HS, Shin JY. Physical activity status by pain severity in patients with knee osteoarthritis: a nationwide study in Korea. *BMC Musculoskelet Disord* 2018;19:380.
 52. Fu K, Makovey J, Metcalf B, Bennell KL, Zhang Y, Asher R, et al. Sleep quality and fatigue are associated with pain exacerbations of hip osteoarthritis: an internet-based case-crossover study. *J Rheumatol* 2019;46:1524–30.
 53. Kiliñç H, Karahan S, Atilla B, Kinikli Gİ. Can fear of movement, depression and functional performance be a predictor of physical activity level in patients with knee osteoarthritis? [original article]. *Arch Rheumatol* 2018;34:274–80.

54. Washburn RA, Smith KW, Jette AM, Janney CA. The Physical Activity Scale for the Elderly (PASE): development and evaluation. *J Clin Epidemiol* 1993;46:153–62.
55. Schult AJ, Schonten EG, Westerterp KR, Saris WH. Validity of the Physical Activity Scale for the Elderly (PASE): according to energy expenditure assessed by the doubly labeled water method. *J Clin Epidemiol* 1997;50:541–6.
56. Loland N. Reliability of the Physical Activity Scale For The Elderly (PASE). *Eur J Sport Sci* 2002;2:1–12.
57. Hagiwara A, Ito N, Sawai K, Kazuma K. Validity and reliability of the Physical Activity Scale for the Elderly (PASE) in Japanese elderly people. *Geriatr Gerontol Int* 2008;8:143–51.
58. Ngai SP, Cheung RT, Lam PL, Chiu JK, Fung EY. Validation and reliability of the Physical Activity Scale for the Elderly in Chinese population. *J Rehabil Med* 2012;44:462–5.
59. Casartelli NC, Bolszak S, Impellizzeri FM, Maffiuletti NA. Reproducibility and validity of the Physical Activity Scale For The Elderly (PASE) questionnaire in patients after total hip arthroplasty. *Phys Ther* 2015;95:86–94.
60. Ismail N, Hairi F, Choo WY, Hairi NN, Peramalah D, Bulgiba A. The Physical Activity Scale for the Elderly (PASE): validity and reliability among community-dwelling older adults in Malaysia. *Asia Pac J Public Health* 2015;27 Suppl:62S–72S.
61. Ayvat E, Kiliç M, Kirdi N. The Turkish version of the Physical Activity Scale For The Elderly (PASE): its cultural adaptation, validation, and reliability. *Turk J Med Sci* 2017;47:908–15.
62. Covotta A, Gagliardi M, Berardi A, Maggi G, Pierelli F, Mollica R, et al. Physical Activity Scale For The Elderly: translation, cultural adaptation, and validation of the Italian version. *Curr Gerontol Geriatr Res* 2018;2018:8294568.
63. Keikavoosi-Arani L, Salehi L. Cultural adaptation and psychometric adequacy of the Persian version of the Physical Activity Scale For The Elderly (P-PASE). *BMC Res Notes* 2019;12:555.
64. Siordia C. Alternative scoring for Physical Activity Scale for the Elderly (PASE). *Maturitas* 2012;72:379–82.
65. Bolszak S, Casartelli NC, Impellizzeri FM, Maffiuletti NA. Validity and reproducibility of the Physical Activity Scale for the Elderly (PASE) questionnaire for the measurement of the physical activity level in patients after total knee arthroplasty. *BMC Musculoskelet Disord* 2014;15:46.
66. Svege I, Nordsletten L, Fernandes L, Risberg MA. Exercise therapy may postpone total hip replacement surgery in patients with hip osteoarthritis: a long-term follow-up of a randomised trial. *Ann Rheum Dis* 2015;74:164–9.
67. Martin KA, Rejeski WJ, Miller ME, James MK, Ettinger WH Jr, Messier SP. Validation of the PASE in older adults with knee pain and physical disability. *Med Sci Sports Exerc* 1999;31:627–33.
68. Petrella RJ, Bartha C. Home based exercise therapy for older patients with knee osteoarthritis: a randomized clinical trial. *J Rheumatol* 2000;27:2215–21.
69. Bossen D, Veenhof C, Dekker J, de Bakker D. The usability and preliminary effectiveness of a web-based physical activity intervention in patients with knee and/or hip osteoarthritis. *BMC Med Inform Decis Mak* 2013;13:61.
70. Bennell KL, Ahamed Y, Jull G, Bryant C, Hunt MA, Forbes AB, et al. Physical therapist-delivered pain coping skills training and exercise for knee osteoarthritis: randomized controlled trial. *Arthritis Care Res (Hoboken)* 2016;68:590–602.
71. Bennell KL, Campbell PK, Egerton T, Metcalf B, Kasza J, Forbes A, et al. Telephone coaching to enhance a home-based physical activity program for knee osteoarthritis: a randomized clinical trial. *Arthritis Care Res (Hoboken)* 2017;69:84–94.
72. Bade M, Struessel T, Dayton M, Foran J, Kim R, Miner T, et al. Early high-intensity versus low-intensity rehabilitation after total knee arthroplasty: a randomized controlled trial. *Arthritis Care Res (Hoboken)* 2017;69:1360–8.
73. Hinman RS, Lawford BJ, Campbell PK, Briggs AM, Gale J, Bills C, et al. Telephone-delivered exercise advice and behavior change support by physical therapists for people with knee osteoarthritis: protocol for the telecare randomized controlled trial. *Phys Ther* 2017;97:524–36.
74. Quicke JG, Foster NE, Ogollah RO, Croft PR, Holden MA. Relationship between attitudes and beliefs and physical activity in older adults with knee pain: secondary analysis of a randomized controlled trial. *Arthritis Care Res (Hoboken)* 2017;69:1192–200.
75. Allen KD, Arbeeva L, Callahan LF, Golightly YM, Goode AP, Heiderscheit BC, et al. Physical therapy vs internet-based exercise training for patients with knee osteoarthritis: results of a randomized controlled trial. *Osteoarthritis Cartilage* 2018;26:383–96.
76. Dunlop DD, Song J, Semanik PA, Sharma L, Chang RW. Physical activity levels and functional performance in the osteoarthritis initiative: a graded relationship. *Arthritis Rheum* 2011;63:127–36.
77. Felson DT, Niu J, Yang T, Torner J, Lewis CE, Aliabadi P, et al. Physical activity, alignment and knee osteoarthritis: data from MOST and the OAI. *Osteoarthritis Cartilage* 2013;21:789–95.
78. Fransen M, Su S, Harmer A, Blyth FM, Naganathan V, Sambrook P, et al. A longitudinal study of knee pain in older men: Concord Health and Ageing in Men Project. *Age Ageing* 2014;43:206–12.
79. Bindawas SM, Vennu V. Longitudinal effects of physical inactivity and obesity on gait speed in older adults with frequent knee pain: data from the Osteoarthritis Initiative. *Int J Environ Res Public Health* 2015;12:1849–63.
80. Wendel-Vos GC, Schuit AJ, Saris WH, Kromhout D. Reproducibility and relative validity of the Short Questionnaire To Assess Health-Enhancing Physical Activity. *J Clin Epidemiol* 2003;56:1163–9.
81. Ainsworth BE, Haskell WL, Leon AS, Jacobs DR Jr, Montoye HJ, Sallis JF, et al. Compendium of Physical Activities: classification of energy costs of human physical activities. *Med Sci Sports Exerc* 1993;25:71–80.
82. Ainsworth BE, Haskell WL, Herrmann SD, Meckes N, Bassett DR Jr, Tudor-Locke C, et al. 2011 Compendium of Physical Activities: a second update of codes and MET values. *Med Sci Sports Exerc* 2011;43:1575–81.
83. Wagenmakers R, van den Akker-Scheek I, Groothoff JW, Zijlstra W, Bulstra SK, Kootstra JW, et al. Reliability and validity of the Short Questionnaire To Assess Health-Enhancing Physical Activity (SQUASH) in patients after total hip arthroplasty. *BMC Musculoskelet Disord* 2008;9:141.
84. De Hollander EL, Zwart L, de Vries SI, Wendel-Vos W. The SQUASH was a more valid tool than the OBiN for categorizing adults according to the Dutch physical activity and the combined guideline. *J Clin Epidemiol* 2012;65:73–81.
85. Konijn NP, van Tuyl LH, Boers M, den Uyl D, Ter Wee MM, Kerstens P, et al. Effective treatment for rapid improvement of both disease activity and self-reported physical activity in early rheumatoid arthritis. *Arthritis Care Res (Hoboken)* 2016;68:280–4.
86. Nicolaou M, Gademan MG, Snijder MB, Engelbert RH, Dijkshoorn H, Terwee CB, et al. Validation of the SQUASH Physical Activity Questionnaire in a multi-ethnic population: the HELIUS study. *PLoS One* 2016;11:e0161066.
87. Makabe S, Makimoto K, Kikkawa T, Uozumi H, Ohnuma M, Kawamata T. Reliability and validity of the Japanese version of the Short Questionnaire To Assess Health-Enhancing Physical Activity (SQUASH) scale in older adults. *J Phys Ther Sci* 2015;27:517–22.

88. Verlaan L, Bolink SA, van Laarhoven SN, Lipperts M, Heyligers IC, Grimm B, et al. Accelerometer-based physical activity monitoring in patients with knee osteoarthritis: objective and ambulatory assessment of actual physical activity during daily life circumstances. *Open Biomed Eng J* 2015;9:157–63.
89. Rewald S, Mesters I, Lenssen AF, Emans PJ, Wijnen W, de Bie RA. Effect of aqua-cycling on pain and physical functioning compared with usual care in patients with knee osteoarthritis: study protocol of a randomised controlled trial. *BMC Musculoskelet Disord* 2016;17:88.
90. Knittle K, de Gucht V, Hurkmans E, Peeters A, Ronday K, Maes S, et al. Targeting motivation and self-regulation to increase physical activity among patients with rheumatoid arthritis: a randomised controlled trial. *Clin Rheumatol* 2015;34:231–8.
91. Stewart AL, Mills KM, King AC, Haskell WL, Gillis D, Ritter PL. CHAMPS physical activity questionnaire for older adults: outcomes for interventions. *Med Sci Sports Exerc* 2001;33:1126–41.
92. Taylor HL, Jacobs DR Jr, Schucker B, Knudsen J, Leon AS, Debacker G. A questionnaire for the assessment of leisure time physical activities. *J Chronic Dis* 1978;31:741–55.
93. Di Pietro L, Caspersen CJ, Ostfeld AM, Nadel ER. A survey for assessing physical activity among older adults. *Med Sci Sports Exerc* 1993;25:628–42.
94. Demmelmaier I, Åsenlöf P, Bergman P, Nordgren B, Opava CH. Pain rather than self-reported sedentary time explains variation in perceived health and activity limitation in persons with rheumatoid arthritis: a cross sectional study in Sweden. *Rheumatol Int* 2017;37:923–30.
95. Pinto D, Song J, Lee J, Chang RW, Semanik PA, Ehrlich-Jones LS, et al. Association between sedentary time and quality of life from the Osteoarthritis Initiative: who might benefit most from treatment? [original article]. *Arch Phys Med Rehabil* 2017;98:2485–90.
96. Sliepen M, Mauricio E, Lipperts M, Grimm B, Rosenbaum D. Objective assessment of physical activity and sedentary behaviour in knee osteoarthritis patients: beyond daily steps and total sedentary time. *BMC Musculoskelet Disord* 2018;19:64.
97. Quicke JG, Foster NE, Croft PR, Ogollah RO, Holden MA. Change in physical activity level and clinical outcomes in older adults with knee pain: a secondary analysis from a randomised controlled trial. *BMC Musculoskelet Disord* 2018;19:59.
98. Adams SA, Matthews CE, Ebbeling CB, Moore CG, Cunningham JE, Fulton J, et al. The effect of social desirability and social approval on self-reports of physical activity. *Am J Epidemiol* 2005;161:389–98.
99. Prince SA, Adamo KB, Hamel ME, Hardt J, Connor Gorber S, Tremblay MA. A comparison of direct versus self-report measures for assessing physical activity in adults: a systematic review. *Int J Behav Nutr Phys Act* 2008;5:56.
100. Westerterp KR. Reliable assessment of physical activity in disease: an update on activity monitors. *Curr Opin Clin Nutr Metab Care* 2014;17:401–6.

Table 1. Practical applications for self-report physical activity measures*

Measure	Number of Items	Content/Domains	Method of Administration	Recall Period	Response Format	Range of Scores	Score Interpretation	Availability of Normative Data	Cross-Cultural Validation
BPAQ	16	Work-related, leisure time, and sport activities	Self-completed	Not clear	Multiple-choice answers	Not clear	No scoring interpretation	Not reported	Not reported
IPAQ-LF	27	Job-related activities, transportation-related activities, household-related activities, leisure time activities, and sitting activities	Self-completed, telephone administered	Last 7 d	Number of days, number of hours, and number of minutes spent active	NA	METs in minutes per week	NA	Multiple languages available with evidence for cross-cultural validity
IPAQ-SF	7	Moderate- and vigorous-intensity physical activity, walking and sitting activities	Self-completed, telephone administered	Last 7 d	Number of days, number of hours, and number of minutes spent active	NA	METs in minutes per week	NA	Multiple languages available with evidence for cross-cultural validity
IPAQ-E	4	Moderate- and vigorous-intensity physical activity, walking and sitting activities	Self-completed	Last 7 d	Number of days, number of hours, and number of minutes spent active	NA	METs in minutes per week, caution should be used when converting to METs for older adults	NA	Available in English and Swedish
PASE	12	Activities during leisure time and in occupational and household settings	Self-completed or telephone-administered	Last 7 d	Multiple choice	0 to ≥400	No scoring interpretation	Not reported	Some studies reported translations, but these are not freely available
SQUASH	14	Commuting activities, activity at work or school, household activities, leisure time activities, and sports	Self-completed	Usual week in the past months	Days per week, average time per day, intensity of activity	NA	Total time spent physically active over 2.0 METs in a week	Not reported	Original measure in Dutch; English, Turkish, and Japanese versions available, although cross-cultural validation was only reported in the Japanese version

* BPAQ = Baecke Physical Activity Questionnaire; IPAQ-E = International Physical Activity Questionnaire for the Elderly; IPAQ-LF = International Physical Activity Questionnaire Long Form; IPAQ-SF = International Physical Activity Questionnaire Short Form; MET = metabolic equivalent; NA = not applicable; PASE = Physical Activity Scale for the Elderly; SQUASH = Short Questionnaire to Assess Health-Enhancing Physical Activity; d = days.

Table 2. Psychometrics (musculoskeletal populations)*

Measure	Floor, Ceiling Effects	Reliability	Validity	Responsiveness	Minimally Important Differences	Generalizability	Used in RCTs	
BPAQ	0	ICC = 0.87 for women with hip disorders (14); ICC = 0.77 for LBP (22)	Correlation to accelerometers = 0.30-0.49 for women with hip disorders (14); correlation to accelerometers = 0.18 for LBP (22)	0	0	0	Ref. 23	
IPAQ-LF	0	ICC = 0.65, SEM = 2668 METs (minutes per week), and SDC = 1115 METs (minutes per week) for THR and/or TKR (31); ICC = 0.83 for AS (32); ICC = 0.77 for women with FM (33); ICC = 0.37 for LBP (22)	Correlation to accelerometer = 0.43 for THR and/or TKR (31); correlation to accelerometer = 0.38 for AS (32); concordance correlation with accelerometer = 0.04 for women with FM (33), compared with accelerometer with individual overestimate in IPAQ-LF for RA (38); correlation to accelerometer = 0.33 for LBP (22)	0	0	0	Ref. 42	
IPAQ-SF	0	ICC = 0.76 for THR; ICC = 0.87 for TKR (36); ICC = 0.51 for THR and/or TKR (31); SEM = 2487 METs (minutes per week); SDC = 1039 METs (minutes per week) for THR and/or TKR (31); ICC = 0.64, SEM = 3532, and SDC = 9791 in a sample with hip, knee, foot, and hand OA (37)	Correlation to accelerometer = 0.29 for THR and/or TKR (31); correlation to PASE = 0.61 for hip OA (40); correlation to accelerometer = 0.40 for RA (41); correlation to PASE = 0.56 (37)	ES = -0.14, SRM = -0.21, RR = 0.12 in a sample with hip, knee, foot, and hand OA (37)	0	0	Refs. 43-48	
IPAQ-E	0	0	0	0	0	0	0	
PASE	0	ICC = 0.77 for hip OA (40); ICC = 0.77 in men with TKR; ICC = 0.58 in women with TKR (66); ICC = 0.77 for THR (60); ICC = 0.68, SEM = 46.7, and SDC = 129.6 in a sample with hip, knee, foot, and hand OA (37); SEM = 31 and SDC = 87 for hip OA (67); SEM = 32%-35% and SDC = 89%-97% for TKR (67); SEM = 23% for THR (60)	Correlation with accelerometer = 0.30 for hip OA (67); correlation with accelerometer = 0.45 in men with TKR and 0.06 in women with TKR (66); correlation with accelerometer = 0.27 for THR (60)	ES = -0.16, SRM = 0.21, and RR = 0.09 in a sample with hip, knee, foot, and hand OA (37)	0	0	0	Refs. 40, 43, and 70-76
SQUASH	0	ICC = 0.89 for AS (32); Spearman's correlation = 0.57 for THR (84)	Correlation to accelerometer = 0.35 for AS (32); correlation to accelerometer ranged from $r = 0.28$ to $r = 0.49$ for knee OA (89); correlation to accelerometer = 0.20-0.67 for THA (84)	0	0	0	Refs. 86, 90, and 91	

* AS = ankylosing spondylitis; BPAQ = Baecke Physical Activity Questionnaire; ES = effect size; FM = fibromyalgia; ICC = intraclass correlation; IPAQ-E = International Physical Activity Questionnaire for the Elderly; IPAQ-LF = International Physical Activity Questionnaire Long Form; IPAQ-SF = International Physical Activity Questionnaire Short Form; LBP = low back pain; MET = metabolic equivalent; OA = osteoarthritis; PASE = Physical Activity Scale for the Elderly; RA = rheumatoid arthritis; RR = responsiveness ratio; SDC = smallest detectable change; SEM = standard error of the measurement; SQUASH = Short Questionnaire to Assess Health-Enhancing Physical Activity; SRM = standard responsiveness measure; THR = total hip replacement; TKR = total knee replacement; 0 = no existing evidence.