Are the cross-culturally adapted versions of the Tampa Scale for Kinesiophobia 11-item valid, reliable, and responsive? A COSMIN-informed systematic review of measurement properties

Short running title: **Measurement properties of the TSK-11**

Vanessa Alpalhão[1,2] ; João R. Vaz[1]; Nuno Cordeiro[2,3]; Pedro Pezarat-Correia[4,5]

[1.] Egas Moniz Center for Interdisciplinary Research (CiiEM), Egas Moniz School of Health and Science, Monte da Caparica, Lisbon, Portugal.

[2.] Physical Therapy Department, Polytechnic Institute of Castelo Branco, Superior Health School, Castelo Branco, Portugal.

[3.] AGE.COMM—Interdisciplinary Research Unit—On Building Functional Ageing Communities, Polytechnic Institute of Castelo Branco, Castelo Branco, Portugal.

[4.] Neuromuscular Research Lab, Faculdade de Motricidade Humana, Universidade de Lisboa, Lisbon, Portugal.

[5.] CIPER, Faculdade de Motricidade Humana, Universidade de Lisboa, Lisbon, Portugal.

Corresponding author

Vanessa B. Alpalhão, Egas Moniz School of Health and Science, Campus Universitário, Quinta da Granja, 2829-511 Monte da Caparica, Almada, Portugal. Email: vanessaluisa@hotmail.com. Phone:+351 965844339.

The systematic review protocol was prospectively registered in PROSPERO with registration ID CRD42022376292.

ABSTRACT

Numerous cultural adaptations of the Tampa Scale of Kinesiophobia 11-item version (TSK-11) scale have emerged since the original version was introduced. We conducted a COSMIN-informed systematic review of measurement properties to identify the cross-cultural adaptation of the TSK-11 and report, critically appraise, and systematize its measurement properties. Six databases were searched for studies published since 2005. Studies reporting on the measurement properties of culturally adapted versions of the TSK-11, published in English, Portuguese, and Spanish were considered for inclusion. Results were synthesized by measurement properties and rated against the COSMIN criteria for good measurement properties. The quality of the evidence was assessed using the GRADE approach and presented in a summary of findings table. Twenty-three studies were included and cultural adaptations for 15 languages were identified: English, Chinese, Cantonese, Swedish, German, Dutch, Arabic, Turkish, Danish, Spanish, Japanese, Brazilian Portuguese, Marathi, Thai, and Persian. There is "high" certainty in the evidence for "sufficient" criterion validity (TSK-17, r=0.84) and "insufficient" measurement error (SDC range 5.6-6.16). "Moderate" certainty in the evidence for "sufficient" construct validity (87.8% of hypotheses confirmed), test-retest reliability ($ICC_{2,1}$= 0.747-0.87), and "low" certainty in evidence for "sufficient" responsiveness. The numerous sources of heterogeneity prevent conclusions from being drawn regarding structural validity. Measurement error, responsiveness, and structural validity of the TSK-11 require further investigation. Clinicians should complement the use of TSK-11 with other instruments. Future studies on the structural validity of the questionnaire should standardize the data analysis methods.

Perspective: This article presents the measurement properties of the cross-cultural adaptations of the TSK-11. Clinicians should be aware that cultural and clinical aspects may influence the structural validity of the questionnaire. Using the TSK-11 as a standalone instrument may omit relevant clinical progression in the patient's condition.

Keywords: Clinimetrics; psychometric; fear of movement; phobia; patient-reported outcome measure.

1. INTRODUCTION

Since it was first described as “an excessive, irrational and debilitating fear of physical movement and activity resulting from a feeling of vulnerability to painful injury or re-injury”,[1] kinesiophobia has been gaining increasing attention. Recognized as an intermediary factor in the transition from acute to chronic pain stages, its initial study was primarily focused on patients with Chronic Low Back Pain.[2] Since then, the scope of research on kinesiophobia has been extended to numerous clinical conditions.[3–7] Kinesiophobia may constitute a barrier to physical activity,[8,9] impacting not only disability and quality of life,[10] but also in adherence to rehabilitation programs[11].

Interventions directed at kinesiophobia have been developed with evidence in reducing disability, dysfunctional thoughts, catastrophizing, and pain and improving patients' quality of life.[12,13]

To measure kinesiophobia and assess its variation over time, the Tampa Scale for Kinesiophobia (TSK) was developed in 1991, under the presentation of a 17-item questionnaire.[14] Subsequently, several versions of the TSK have been proposed,[15–21] with the 11-item version (TSK-11) considered most advantageous due to its brevity and lower patient burden.[20] As a result the TSK-11 has become a widespread instrument across settings (hospitals,[22–30] clinics,[31–40] primary care centers,[41] and domicile[42–45]) and clinical populations (pain,[3,26,27,31,34,35,41,44–64] post-surgical,[22–24,26,65–83] neurological,[84–86] and oncological,[87,88]).

The TSK-11 consists of 11 items from the original TSK, grouped into two subscales: Somatic Focus (TSK-SF) and Activity Avoidance (TSK-AA).[20] Each item is scored on a 4-point Likert scale, with higher scores indicating higher levels of kinesiophobia.

Since its introduction in 2005,[20] the TSK-11 has undergone numerous adaptations and translations for use in culturally diverse populations[21,23,44,89–93]. In the context of measurement properties, "culturally different populations" was interpreted broadly,[94] concerning not only languages or ethnicities but also genders, ages, or clinical conditions.[94] The cross-cultural adaptation process is imperative when using an instrument in a different population, setting, and time.[95] This process involves more than linguistic translation as some items in the questionnaire may lose their meaning or be irrelevant in another culture or population.[96] Once the instrument has been translated and culturally adapted, its measurement properties must be rigorously evaluated. There is much emphasis on using validated instruments, as they increase the certainty of accurate measures.[97]

Scientific literature reports inconsistencies regarding the TSK-11 and its use in assessing kinesiophobia, concerning both the structural and construct validity of the questionnaire.[98–100] The TSK-11 structural validity seems to vary not only in clinically diverse populations but also in different stages of the same condition.[73] Concerning construct validity, it was suggested that the scale better reflects the construct of beliefs regarding movement than the construct of kinesiophobia itself.[99,100] Evidence also suggested that the TSK-11 may underestimate the fear of movement concerning specific activities.[101]

To conclude about the quality of the questionnaire, a review of its measurement properties is required. Recently, a review was carried out but it only considered studies conducted in populations with musculoskeletal pain.[102] Given the widespread use of TSK-11, it is imperative to evaluate its measurement properties in culturally and clinically diverse populations. This review aimed to assess the measurement properties of cross-culturally adapted versions of the TSK-11, identify methodological flaws, and

the need for further validation, which could improve the existing instrument and potential adaptations. It also aimed to identify the languages and clinical populations for which the questionnaire has been adapted, promoting the dissemination of its use. Specifically, it was guided by the following questions:

i) For which populations and/or contexts has the TSK-11 been adapted?

ii) Are the cross-culturally adapted versions of the TSK-11 valid, reliable, and responsive for the assessment of kinesiophobia?

2. METHODS

This systematic review followed the COSMIN methodology for conducting systematic reviews of measurement properties,[94] the JBI evidence synthesis template for systematic reviews of measurement properties,[103] and the PRISMA 2020 statement[104]. We also submitted registration materials prior to the literature search in PROSPERO with registration ID CRD42022376292.

Search strategy

A four-step strategy was used in the search for relevant studies [103]: 1) An initial search in PubMed was carried out in June 2022 using the key terms “TSK-11” and “Tampa scale”, and the sensitive search filter developed by the COSMIN to aid researchers in finding studies on measurements properties of measurement instruments.[105] Keywords used in the title/abstract and index terms were identified and subsequently considered in the development of a full search strategy for each database. 2) A database-specific search was performed on April 28th and 29th, 2023 on MEDLINE (PubMed), EMBASE, Scopus, Web of Science, APA PsycNet, and EBSCOhost (Supplementary File 1). The COSMIN filters were used in PubMed and EMBASE.[105] 3) An updated database search

was conducted on April 12th, 2024. 4) The reference lists of studies selected for critical appraisal were scanned to identify additional relevant studies.

Inclusion criteria

Population

Studies including participants of any age and sex, in any country or setting, and under any clinical condition were eligible for this systematic review.

Instrument

All studies reporting the measurement properties of cross-cultural adaptations of the TSK-11 were considered for inclusion.

Construct

The construct of interest in this review was kinesiophobia defined as "an excessive, irrational, and debilitating fear of physical movement and activity resulting from a feeling of vulnerability to painful injury or re-injury".[1]

Outcomes

The outcomes of interest were the measurement properties: validity, reliability, and responsiveness. Studies assessing one of these measurement properties in a cross-culturally adapted version of the TSK-11 were considered for inclusion. In a measurement context, validity represents the degree to which an instrument measures the construct it purports to measure.[94] Two validity domains were considered: construct validity (Hypotheses testing, structural validity, cross-cultural validity) and criterion validity. Reliability represents the extent to which the score for patients who have not changed condition is the same for repeat measurement under various circumstances: using different sets of items from the same instrument (internal consistency), over time

(test-retest), by different persons on the same occasion (inter-rater), or by the same person (raters or responders) on different occasions (intra-rater). In the context of reliability, measurement error (systematic and random error of a patient's score that is not attributed to true changes in the construct to be measured) was also considered.[94] Responsiveness, in turn, represents the ability of an instrument to detect changes over time in the construction to be measured, commensurate with the amount of changes that have occurred in that construction.

Types of studies

Peer-reviewed studies with publication dates from 2005 (introduction of the TSK-11) onwards,[20] published in English, Portuguese, or Spanish, with full-text availability. Studies evaluating the measurement properties of the original version,[20] in clinically distinct populations were included; however, studies that duplicated data from that version were excluded. Studies that only used the TSK-11 as an outcome measure instrument, qualitative research papers, systematic reviews, opinion texts, and editorials were also excluded.

Study selection

All potential records were uploaded to Zotero 6.0.8 (Roy Rosenzweig Centre for History and New Media, USA), and duplicates were removed before the title and abstract screening. Two reviewers independently screened all records by title/abstract against the inclusion criteria. An initial pilot test of 25 randomly selected evidence sources was performed to ensure clarity and consistency in the application of inclusion and exclusion criteria. Following the pilot test an agreement between reviewers of 92% was achieved (Cohen Kappa= 0.7024). A team meeting was held to discuss discrepancies and clarify eligibility criteria. All potentially relevant studies were

subsequently assessed using full-text analysis. Where disagreements arose, the reviewers met to discuss and resolve.

Assessment of methodological quality

Two reviewers (V.A. and N. C.) independently assessed the methodological quality of each study using the COSMIN risk-of-bias checklist.[106] A third reviewer discussed any inconsistencies or disagreements. The COSMIN risk-of-bias checklist includes 10 domains and assists in the assessment of each domain: PROM development, content validity, structural validity, internal consistency, cross-cultural validity/measurement invariance, reliability, measurement error, criterion validity, and hypotheses testing for construct validity and responsiveness. The measurement properties evaluated in a study determined which domains needed to be completed. Each domain contains several items, and each item can be reported as 'very good, adequate, doubtful, inadequate, or not applicable.' The lowest rating for any standard was used to determine the overall domain score.

In line with COSMIN's recommendation, before assessing the methodological quality of the studies on criterion validity, the review team determined that the TSK-17 would be considered a reasonable gold standard. Although a perfect gold standard seldom exists in practice, when developing a shorter version of a questionnaire the long version can be considered the gold standard.[96] Thus, only studies using the TSK-17 were considered for criterion validity.

Data extraction

Considering the COSMIN data extraction forms for psychometric properties,[106] a data extraction template was developed to extract relevant data from the included studies. The data extracted were directly derived from the objective and specific questions of the

review and included study identification (author(s), year of publication, title, country), instrument identification (TSK-11 version, language, mode of administration, studied structure), population characteristics (number of participants, age, gender, clinical condition, clinical condition duration and severity), context characteristics (country, settings), and data on measurement properties. Two reviewers (V.A. and N. C.) pilot-tested the data extraction template in five studies to check its suitability for the content and to ensure systematic extraction. To avoid missing relevant information and to minimize errors, the same two reviewers independently extracted data from all included studies with the support of a third independent reviewer in case of conflict.

Data extracted on measurement properties were presented in three domains, as suggested by the COSMIN methodology: validity, reliability, and responsiveness.[94] For validity, data were extracted for structural validity, hypotheses testing, cross-cultural validity/measurement invariance, and criterion validity. For reliability, we extracted data on internal consistency, test-retest reliability, intra- and inter-observer reliability when applicable, and measurement error. Relevant indicators such as the Intraclass Correlation Coefficient [ICC], Cronbach's alpha [α], Standard Error of Measurement [SEM] Minimum Detectable Change [MDC], Pearson's correlation coefficient [r], Minimal Clinical Important Difference [MCID] and Area Under the Curve [AUC] were extracted when available.

Appraisal of measurement properties

The results of each study were rated against the updated COSMIN criteria for good measurement properties (Supplementary File 2).[94] Each result was rated as sufficient (+), insufficient (–), or indeterminate (?). The quality rating result for each measurement

property was added directly to the table in line with the results extracted for that property.

To appraise the measurement properties of studies on hypotheses testing for construct validity based on the relationships proposed by theoretical models,[2,107] and results from previous studies,[108,109] the following hypotheses were formulated:
1) The expected correlations with instruments measuring related but dissimilar constructs would be moderate to strong ($|r| \geq 0.3$),[110] including constructs related to fear of pain, catastrophizing, Fear Avoidance Beliefs (FAB), escape/avoidance, anxiety, physical activity, confidence while performing movements, and pain acceptance.
2) The expected correlations with other constructs, such as pain intensity, pain interference, depression, disability, quality of life, health status, physical function, and fatigue, would be weak to moderate ($|r| \leq 0.3$).

The hypotheses were defined in the form of an interval for the correlation coefficient (r) because they were considered both for relationships with the TSK-11 total score in unidimensional models, as well as for subscale scores in versions of the TSK-11 with more than one factor.

Data synthesis

The results for hypotheses testing, criterion validity, test-retest reliability, measurement error, and responsiveness were qualitatively summarized. This qualitative summary considered, for example, the percentage of confirmed hypotheses for construct validity or the range of Internal Coefficient Correlation (ICC) for test-retest reliability. To reach an overall conclusion on the quality of the measurement properties, the summarized results per measurement property of studies with good methodological quality were rated against the COSMIN quality criteria for good measurement properties.[94] The overall rating for the summarized results was classified as sufficient (+), insufficient (–),

inconsistent (±), or indeterminate (?). When all results were consistent, the overall rating was sufficient (+) or insufficient (-). To rate the qualitatively summarized results as sufficient (or insufficient), 75% of the results should meet the criteria. The criteria for classifying the overall results of studies in construct validity as "sufficient" was that at least 75% of the results must be in line with the hypotheses defined. For inconsistent results, explanations for the inconsistency were explored. The summarized results per measurement property were reported in a Summary of Finding (SoF) table, accompanied by a rating of the summarized results (+ / – / ± / ?), the grading of the quality of evidence (high, moderate, low, and very low), and a narrative synthesis.[111] The many sources of heterogeneity in studies on structural validity and internal consistency prevent us from carrying out a qualitative summary. Their results were narratively summarized considering the studied factor structure and the methodological quality. The cross-cultural validity/measurement invariance results have also been narratively summarized as they cannot be pooled.

Assessing certainty in the findings

To grade the quality of summarized evidence, a modified Grading of Recommendations, Assessment, Development, and Evaluation (GRADE) approach was used, as recommended by the COSMIN. In accordance, the quality of the evidence was graded as “high,” “moderate,” “low,” or “very low.” The starting point in grading evidence was the assumption that the overall result was of high quality. The quality of evidence was subsequently downgraded by one or two levels when there was a risk of bias, (unexplained) inconsistency, imprecision, or indirect results. The quality of evidence was downgraded by three levels when the evidence was based on only one inadequate study. The GRADE approach uses five factors to determine evidence quality: risk of bias, inconsistency, indirectness, imprecision, and publication bias.

However, in this review, we did not consider publication bias. As suggested by Prinsen et al it is difficult to assess publication bias in this type of review because of the lack of registries for studies on the measurement properties.[94]

3. RESULTS

A total of 320 potential records were identified through the initial search (April 2023), and 53 full texts were screened against the inclusion criteria. By common agreement, two reviewers excluded 32 articles (three were conference papers, and 29 did not address the 11-item version of the TSK). A total of 21 articles were selected for inclusion in this phase [17,21,23,35,44,61,64,73,87,89,90,92,93,98,112–118]. In the update search (April 2024), an additional 39 potential records were identified. Seven were assessed in full text for eligibility, of which only 2 met the criteria.[91,119] In total, 23 articles were included in this review. A flowchart of the literature search and study selection process is presented in Figure 1.

Characteristics of included studies.

Table 1 summarizes the characteristics of the 23 included studies. Almost half (n=10; 43.5%) of the studies were published between 2007 and 2012.[21,61,73,87,92,98,112,114,115,117] The downward trend in scientific production on cross-cultural adaptations of the TSK-11 between 2013 and 2018 (n=5; 21.7%),[17,44,89,93,116] has been reversed in the last five years (n=8; 34.8%).[23,35,64,90,91,113,118,119]

The most frequent method of administration of the TSK-11 was through a self-completed questionnaire (n=20; 87.0%),[17,21,23,35,44,61,73,87,89–93,98,112–116,119] two (8.7%) studies administered it online,[21,93] and one (4.8%) study administered it as a computer-based questionnaire[17]. In 13% (n=3) of the studies, the TSK-11 was administered by

interview,[92,117,118] and in 2 studies (8.7%),[21,64] the mode of administration was not reported.

Cultural adaptations for 15 languages were identified: English,[61,73,112,114,115,116] Chinese,[23] Cantonese,[117] Swedish,[44,98] German,[17] Dutch,[21,87] Arabic,[90] Turkish,[89] Danish,[113] Spanish,[42] Japanese,[43] Brazilian Portuguese,[64,119] Marathi,[35] Thai,[118] and Persian[91]. More than a quarter of the studies (n=6; 26.1%) assessed the measurement properties of the English version of the TSK-11.[61,73,112,114–116] The Swedish,[44,98] Brazilian,[64,119] and Dutch versions,[87,108] were evaluated, each, in 8.7% (n = 2) of the included studies. The remaining language versions were assessed in only 1 study (4.3%).

The majority of the studies (n=14; 60.9%) assessed the measurement properties of the TSK-11 as a 2-factor model,[17,21,44,87,90–92,98,112,115–119] and 56.5% (n=13) assessed it as a unidimensional model[21,35,44,61,64,91,93,98,112–114,117,119]. A 3- factor model was considered in 8.7% (n=2) of the studies,[23,73] and only one (4.3%) study considered a 4-factor model[89] (Supplementary File 3).

The most frequently reported measurement properties were validity (n=22; 95.7%),[17,21,23,35,44,61,64,73,87,89–93,98,112,114–119] and reliability (n=19; 82.6%)[17,21,23,35,44,61,64,73,87,89–93,112,113,115,117,118]. Responsiveness was the least studied property,[61] with only one (4.3%) study reporting on it (Supplementary File 4).

Characteristics of participants and contexts

The characteristics of studied populations and contexts are summarized in Table 2. In the included studies, the sample sizes were quite uneven, ranging from 85 to 2825 participants.[21,89] The measurement properties of the TSK-11 were assessed across a wide range of ages, from 16 to 98 years.[44,113] One of the included studies evaluated the

measurement properties of the TSK-11 in a population with high-impact chronic pain and aged between 16 and 78.[113] However, the questionnaire's measurement properties remain to be investigated in adolescents. In terms of gender representation, 13 out of 23 studies (56.5%) showed equal representation for both sexes. Male representativity was higher in six (26.10%) studies,[64,87,92,114,116,118] and female in four (17.4%)[35,91,93,119]. More than half of the studies (n=13; 56.5%) assessed the measurement properties of the TSK-11 in people with chronic pain conditions.[21,35,44,61,64,90–92,98,113,115,117,119] Acute and post-surgical conditions were considered respectively, in one (4.3%),[92] and three (13.0%)[23,73,112] of the included studies. Specific populations, such as those with cardiac and oncological conditions, have been poorly studied (n=1, 4.3%).[87,89] Geographically, there was homogeneity in the distribution of studies conducted on the psychometric properties of the culturally adapted versions of the TSK-11. The American continent was the most representative (n=9; 39.1%),[61,64,73,91,112,114–116,119] followed by the European (n=7; 30.4%),[17,21,44,87,92,98,113] and Asian (n=6; 26.1%)[23,35,90,93,117,118]. Of the included studies, 5 (21.7%) were conducted without direct contact with participants, with data collected via online, mailed, or postal questionnaires.[21,44,93,98,114] Regarding the data collected in context, the most frequent settings were pain centers (n=5; 21.7%),[87,92,113,115,117] universities (n=4; 17.4%),[23,64,112,116] and rehabilitation center/physiotherapy departments (n=4; 17.4%)[35,73,91,119]. These were followed by hospitals (n=3; 13%),[61,90,118] and orthopedics (n=2; 8.7%)[17,117](Supplementary File 5).

Methodological quality

The structural validity of the TSK-11 was assessed in 19 studies, of which 64.7% (n=11) had "very good" methodological quality,[17,21,44,87,91,98,112,115,117–119] 36.8% (n=7) had "adequate" methodological quality,[23,73,90–92,114,116] and 10.5% (2) had "doubtful" methodological quality[89,92]. The main reason for "adequate" methodological quality was

the methodology used in the factorial analysis (exploratory),[23,73,90–92] The small sample size was the main reason for the "doubtful" methodological quality of 2 studies.[89,92]

The majority (50%; n=8) of studies on hypotheses testing had “adequate” methodological quality,[17,35,44,73,93,115,117,118] 37.5% (n=6) had “very good” methodological quality,[21,23,61,64,87,92] one (6,3%) study was classified as “doubtful”,[89] and another as “inadequate”[91]. The main methodological flaw in hypotheses testing studies was the lack of description of the psychometric properties of the comparator instrument(s) in the studied population.[35,89,91,93,115,117,118]

The methodological quality of the two studies on cross-cultural validity/measurement invariance was “doubtful”; indeed, it is not clear whether the samples were similar for all the relevant characteristics except the group variable.[87,108]

The one study on criterion validity of the TSK-11 presented “very good” methodological quality.[64]

Most of the studies (n=12; 70.6%) on the internal consistency of cross-cultural adaptation of the TSK-11 presented “very good” methodological quality,[17,23,44,73,87,90–92,108,115,117,118] three (17.6%) studies had “doubtful”,[35,64,93] and 2 (11.8%) studies “inadequate”[89,112] methodological quality. The three studies with “doubtful” methodological quality assessed TSK-11 internal consistency without evidence of structural validity in that specific culturally adapted version.[35,64,93] Two studies were evaluated as having “inadequate” methodological quality because the internal consistency coefficient was not calculated for each unidimensional subscale.[89,112]

Most studies (n=4; 50%) on test-retest reliability presented “adequate” methodological quality,[23,44,61,64] one (12.5%) study had “very good” methodological quality,[113] and in three (37.5%) studies the appropriateness of the time interval for the retest was considered "doubtful"[35,89,91]. Measurement error of the cross-culturally

adapted versions of the TSK-11 was assessed in four studies with “very good”,[113] “adequate”,[61,64] and “doubtful”[91] methodological quality.

Responsiveness of the TSK-11 was assessed in only one study of “adequate” methodological quality [61] (Supplementary File 6).

Results on TSK-11 validity

Construct Validity

Structural Validity

The results of the structural validity studies are presented in Table 3. The structural validity of the cross-cultural adaptation of the TSK-11 was assessed in 18 studies.[17,21,23,44,73,87,89–92,98,112,114–119] The unidimensionality and composition of the questionnaire proposed by Woby et al.[20] were validated in one study with "very good" methodological quality, which fulfills the criteria for good psychometric properties (+).[91] Indeed, the questionnaire presented a unidimensional structure in the Persian version in people suffering from Complex Regional Pain Syndrome (CRPS).[91] The unidimensionality of the English version of the TSK-11 has also been proven through Item Response Theory (IRT) in a community-based population with arthritis[114]; however, the items included in the questionnaire are different from those suggested by Woby et al.[20]. Both authors removed the reverse scoring items (4, 8, 12, and 16), but while Woby et al. removed items 9 and 14 while Mielenz et colleagues removed items 13 and 6.[20,114]

Although a 2-factor model of the TSK-11 has emerged in various studies,[17,21,44,87,90–92,98,112,115–119] there were some discrepancies regarding the items in its composition. Among the included studies, the 2-factor model was first proposed in 2007 by Roelofs et al., using the Dutch version of the TSK-11 among people with work-

related upper extremity injuries.[21] The study presents "very good" methodological quality and fulfills the criteria for good measurement properties (+). The proposed structure consists of the "Activity Avoidance" subscale (items 1,2,10,13,15,17) and the "Somatic Focus" subscale (items 3,6,7,9,11).[21] The 2-factor model described was corroborated through Confirmatory Factor Analyses (CFI) in three of the included studies, all of good methodological quality and which met the criteria for good psychometric properties (+).[17,87,117] In this way, the 2-factor model proposed by Roelofs et al.,[21] seems to better represent the construct of "kinesiophobia" in the following cross-cultural adaptation of the TSK-11: German version in people suffering from low back pain,[17] Dutch version in cancer survivors,[87] and Cantonese version in people with chronic pain from a sample of orthopedics[117]. The model didn´t, however, meet the criteria for good measurement properties in five studies with "very good" methodological quality: Brazilian version in CLBP patients,[119] English version in adults treated by spinal surgery,[112] Swedish version in patients with persistent musculoskeletal pain,[98] and in older adults with chronic pain,[44] and Cantonese version in a subsample of patients from a pain clinic[117]. Five other models consisting of 2-factors emerged from seven of the included studies,[90–92,115,116,118,119] of which five did not meet the criteria for good measurement properties[90–92,116,119]. The remaining two studies proposed a 2-factor model of the English version of the TSK-11 in people with heterogeneous chronic pain ("Activity Avoidance" items 5,6,9,11,12,13; "Somatic Focus" items 1,2,3,7,8) and a 2-factor model of the Thai version in older people with knee osteoarthritis ("Activity Avoidance" items 1,2,7,9,10,11; "Somatic Focus" items 3,4,5,6,8).[115,118]

A 3-factor structure was proposed for the Chinese version of the TSK-11 in patients who had undergone total knee arthroplasty and for the English version in patients in the early phase after anterior cruciate ligament reconstruction.[23,73] The two

studies presented adequate methodological quality but did not fulfill the criteria for good measurement properties (-).

The Turkish version of the TSK-11 in cardiac patients has been proposed as a 4-factor model, however, the methodological quality of the study is "doubtful" according to the COSMIN checklist.[89]

Hypotheses Testing

Overall, in 4909 patients, 87.8% (43/49) of the results were in accordance with our *a priori* hypothesis. With regard to our first hypothesis, all but one study,[92] found moderate to strong correlations between the TSK-11 scores and those of instruments that measure related but dissimilar constructs (catastrophizing, FAB, anxiety, and confidence while performing movement). Gómez-Pérez et al.,[92] reported weak correlations between catastrophizing (r=0.27), anxiety (r=0.29), FAB (r=0.23), and kinesiophobia assessed by the Spanish version of the TSK-11, in a sub-sample of patients with acute musculoskeletal pain. However, in a sub-sample of patients with chronic heterogeneous pain, moderate relationships were found with catastrophizing (r=0.49).[92] Our second hypothesis was also confirmed as most of the studies reported weak to moderate correlations between the TSK and pain intensity,[17,21,23,44,64,73,92,93,115,117,118] physical function (r=-0.11 to –0.339),[23,64,73,92,115] physical activity (r=-0.378),[44] disability,[17,21,44,115,117] depression (r= 0.206 to 0.38),[17,64,92,115,117] quality of life (r= 0.374),[89] health status (r=-0.333 to -0.570),[87] and fatigue (r=0.3)[87]. Nevertheless, the results of two studies with "adequate" and one study with "doubtful" methodological quality were not in accordance with our *a priori-defined* hypothesis.[35,91,93] Satpute et al.,[35] reported a strong correlation between the TSK-11 Marathi version score, pain intensity (r=0.6), and disability (r=0.7) in patients with chronic LBP. A strong correlation was also identified between the TSK-11

Japanese version score and health status (r=− 0.57),[93] and the TSK-11 Persian version score and pain intensity (r=0.61)[91] (Table 4).

Cross-Cultural Validity/Measurement Invariance

Two studies of "doubtful" methodological quality investigated the cross-cultural validity and measurement invariance of the TSK-11.[21,87] Roelofs et al.,[21] reported an invariant 2-factor model of the TSK-11 across pain diagnoses (upper extremity disorders, CLBP, fibromyalgia, osteoarthritis, and musculoskeletal pain) and nationalities (Dutch, Swedish, English, and French). The proposed model consists of two subscales, "Somatic Focus" which comprises items 3,5,6,7 and 11, and "Activity Avoidance" which comprises items 1,2,10,13,15, and 17 of the original scale. Velthuis et al.,[87] investigated the cross-cultural validity of the TSK-11 model proposed by Roelofs et al.,[21] among cancer survivors. The authors concluded that the TSK-11 "proved to be invariant across sex and cancer type". However, due to the analysis approach adopted both studies were classified as "indeterminate" according to the COSMIN criteria for good measurement properties.

Criterion validity

Of the included studies, only one (4.3%) study of "very good" methodological quality assessed the correlation between the TSK-11 and TSK-17 scores and was therefore considered a study on criterion validity.[64] The scores of the Brazilian Portuguese version of the TSK-11 were strongly correlated with that of the 17-item version (r=0.84) and meet the criteria for good psychometric properties (+).[64]

Reliability

Internal Consistency

Internal consistency of the cross-cultural adapted versions of the TSK-11 was assessed in 17 studies.[17,21,23,35,44,64,73,87,89–93,112,115,117,118] We were unable to draw any conclusions about five (21.7%) studies,[35,64,89,93,112] as three (13%) of them did not present evidence of the structural validity of the model,[35,64,93] and the other two (8.7%) presented a unique Cronbach's alpha value when the structural validity analysis suggested a two,[112] and a four-factor[89] model. The internal consistency of unidimensional models with at least low evidence of structural validity has only been assessed in the Swedish (α=0.87) and Persian (α=0.93) versions, which met the criteria for good measurement properties.[44,91]

The 2-factor model proposed by Roelofs et colleagues only met the COSMIN criteria for good internal consistency in the Swedish version in older adults with chronic pain and patients with musculoskeletal pain.[21,44] In different languages and populations, the model's internal consistency did not meet the criteria for good internal consistency in both subscales simultaneously: "Somatic Focus" and "Activity Avoidance" (SF, α=0.64/0.76; AA, α =0.62/0.74)[21,87]. The German, Thai, Cantonese, and Arabic versions of the TSK-11 also failed to meet the criteria for good internal consistency in the two sub-scales.[17,90,117,118] The English version showed good internal consistency in patients with heterogeneous chronic pain.[115] The Spanish version also showed good internal consistency in a chronic pain sample (TSK-AA, α =0.79; TSK-H α =0.70) and in an acute pain sample (TSK-AA, α =0.79; TSK-H, α =0.70).[92]

The English and Chinese versions of the TSK-11 were considered a 3-factor model in patients after TKA and ACL reconstruction.[23,73] Both studies had “very good” methodological quality and satisfied the COSMIN criteria for good internal consistency (Table 5).

Test-retest Reliability

The test-retest reliability data are presented in Table 6. All the cross-culturally adapted versions of the TSK-11 met the COSMIN criteria for good measurement properties of test-retest reliability (ICC ranging from 0.747 to 0.93),[23,35,44,61,64,91,113] except one study of "doubtful" methodological quality[89]. Overall, in a total sample of 433 participants from studies of "adequate" and "very good" methodological quality, the Intraclass Correlation Coefficient ranged between 0.747/0.87. For the Swedish version of the TSK-11, a strong agreement has been reported for the total scores of the instrument between the two evaluations ($ICC_{2.1}$=0.7479), however, the agreement for the individual items was relatively low (weighted k values ranging from 0.37 to 0.58).[44]

Measurement error

Four (17.4%) studies assessed the measurement error of the culturally adapted version of the TSK-11,[61,64,113] but one had "doubtful" methodological quality[91]. In the studies with good methodological quality, the reported Standard Error of Measurement (SEM) ranged from 2.10 to 2.65 points. The Small Detectable Change (SDC) was 5.82 points for the Danish version in patients with high-impact chronic pain,[113] 5.6 points for the English version in patients with heterogeneous chronic pain,[61] and 6.16 points for the Brazilian Portuguese version in patients with fibromyalgia[64] (Table 7). None of the studies met COSMIN's criteria for good measurement properties of measurement error. In all studies, the SDC was higher than the Minimal Important Change (MIC) of four points reported by Woby et al.[20] (Table 7).

Responsiveness

Responsiveness of the English version of the TSK-11 was assessed in patients with heterogeneous chronic pain after attending a 4-week interdisciplinary chronic pain

management program. The study presented “adequate” methodological quality and met the COSMIN criteria for good responsiveness (Area Under the ROC Curve =0.73, 95% CI [0.57–0.88]).[61]

Certainty in the findings

A summary of the findings is presented in Table 8. We have “moderate” confidence that there is sufficient evidence for the construct validity of the cross-culturally adapted version of the TSK-11 evaluated through hypotheses testing. More than 75% of the results were in line with the *a priori* hypotheses; however, some inconsistencies were observed. Regarding criterion validity, we are very confident that the true measurement property lies close to that of its estimate because the result came from a study with “very good” methodological quality with an adequate sample size. We also have “moderate” confidence of “sufficient” test-retest reliability among the culturally adapted versions of the TSK-11. The summarized results only considered studies of “adequate” and “very good” methodological quality, the total sample size was adequate, but still, there were some inconsistencies concerning agreement for the individual items of the questionnaire. We have “high” confidence in the evidence for the “insufficient” measurement error property of the studied TSK-11 versions. The results were consistent and came from studies with “adequate” and “very good” methodological quality, and the sample size was adequate. Regarding responsiveness, we have “low” confidence in the evidence for “sufficient” responsiveness of the cross-culturally adapted version of the TSK-11. The evidence came from only one study with “adequate” methodological quality, and we downgraded one more level for imprecision, as the sample size was < 100.

4. DISCUSSION

The present review aimed to identify the populations and contexts for which the TSK-11 has been culturally adapted, and to assess the measurement properties of the culturally adapted versions of the questionnaire. The TSK-11 has been culturally adapted into Swedish, Dutch, Arabic, Turkish, Chinese, Danish, Spanish, Japanese, German, Brazilian Portuguese, Marathi, Cantonese, German, Persian, and Thai.[17,21,35,44,64,89–93,113,117,117–119]Although for many years, TSK has mostly been used to assess kinesiophobia in populations with chronic pain, the study of the measurement properties of TSK-11 has been extended to clinically diverse populations, such as: patients after spinal surgery, TKA, and ACL reconstruction[23,73,112]; cancer survivors[87]; cardiac conditions, and patients with acute musculoskeletal pain[89,92]. However, evidence on the measurement properties of TSK-11 in clinically diverse populations is still sparse and there is a need for further investigation. Also, the suitability or need to adapt the questionnaire for use in specific age groups, such as adolescents, remains unstudied.

Regarding structural validity, a variety of models emerged from the included studies (1- factor, 2-factor, 3- factor, and 4- factor model). The diversity in cultural and clinical characteristics, the variety of statistical methodologies employed, and differences in the methodological quality of the studies prevented us from synthesizing results. Cultural and clinical aspects may impact the structure of the questionnaire as well as the relative weight of its items. In addition to culturally rooted values and beliefs, experiences of pain, perceptions, and attitudes towards it, as well as the particularities of clinical conditions, can influence the personal experience of fear of movement/(re)injury. For example, a unidimensional model was shown to fit in a community-based population with arthritis,[114] however, the model was different from those proposed by Woby et al.,[20] by removing items 6 and 13 instead of items 9 and 14. The item 6 state “My

accident has put my body at risk for the rest of my life"; as arthritis is a disease not caused by a traumatic episode, it is natural that the item proved to be irrelevant in this population. Evidence from two studies supports the idea that both cultural and clinical factors shape questionnaire structure: while a unidimensional model of the English version accurately reflects kinesiophobia in patients with arthritis,[114] in English patients with heterogeneous chronic pain,[115] the construct seems to be better explained by two lower-order constructs ("somatic focus" and "activity avoidance"). The results of a study of "very good" methodological quality seem to indicate that the factor structure of the TSK-11 may differ also according to the severity of the clinical conditions.[117] Indeed, in the pain clinic sample, patients had a higher level of severity than in the orthopedic sample.[117] An invariant 2-factor model of the TSK-11 across pain diagnoses and nationalities was reported but some caution is needed considering the study's methodological quality.[21] Also, the model did not fulfill the COSMIN criteria for good internal consistency for the two subscales simultaneously (Activity Avoidance and Somatic Focus) for most cultural adaptations, except for the Swedish version.[21] Indeed, although the TSK is not a condition-specific instrument, clinicians and researchers should be aware that its factor structure can vary across populations and contexts. Other authors have previously reported inconsistencies in the dimensionality of kinesiophobia.[98]

According to the COSMIN methodology, the assessment of methodological properties follows a hierarchy. As evidence for structural validity (or unidimensionality) is a prerequisite for the interpretation of internal consistency, structural validity assessment should be conducted prior to internal consistency analysis.[94] Indeed, the internal consistency statistic of a scale or sub-scale has no intrinsic meaning if there is no evidence of structural validity for it. Considering the evidence available in the

included studies, only the Swedish version (in older adults with chronic pain and patients with musculoskeletal pain), Spanish version (both in patients with acute pain and in patients with chronic pain), and English version (in patients with chronic heterogeneous pain) of the TSK-11fulfilled, simultaneously, the COSMIN criteria for good structural validity and internal consistency.[21,44,92,115]. In this regard, some methodological flows have been identified,[89,112] and future studies on measurement properties should calculate an internal consistency statistic for each unidimensional subscale in multidimensional models of the questionnaire.

As for the construct validity assessed through hypothesis testing, our *a priori* hypotheses were confirmed in 87.8% of cases. Nevertheless, the results of four studies were inconsistent.[35,91–93] Satpute et al.,[35] found a strong relationship between pain, disability, and TSK-11 scores in patients with CLBP; Kikuchi et al.,[93] reported a strong association between health status and kinesiophobia in patients with a whiplash neck injury or LBP due to motor vehicle accidents and, Farzard and colleagues,[91] confirmed a strong association between pain intensity and kinesiophobia in patients suffering from CRPS. Considering that the three studies reported data from Asian populations, we cannot rule out the possibility that cultural differences have an impact on the divergent results. Indeed, previous studies have suggested that individuals from an Asian ethnic background experience greater levels of perceived pain than those from a Caucasian ethnic background.[120] In addition, the clinical characteristics of the samples should be considered. In the case of patients suffering from LBP the level of disability was high (ODI mean score 37.4+-14.5).[35] In the other case, the patients had been in a motor vehicle accident, which may have had a stronger impact on their health status.[93] Also, not in accordance with our a priori hypotheses, a weak association was found between catastrophizing, anxiety, fear-avoidance beliefs and the scores of the Spanish version of

the TSK-11 in a sample with acute pain.[92] Nevertheless, in the same version of the questionnaire the associations were as expected.[92] These data support the hypotheses that the construct validity of the TSK-11 may be influenced by clinical aspects.

The Brazilian Portuguese version of the TSK-11 showed good criterion validity through a strong correlation with the TSK-17.[64] Regarding criterion validity, some methodological flaws have been identified concerning the definition of a gold standard. Future studies aiming to establish criterion validity for culturally adapted versions of the TSK-11 should consider using the TSK-17 as the gold standard.

The overall test-retest reliability of the culturally adapted version of the TSK-11 was acceptable (ICC range 0.747-0.87) but, the results of one study showed fair to moderate agreement for the individual items (k=0.37/0.58).[44] Although the time elapsed between administrations was within the recommended (2 weeks), we cannot exclude the possibility of changes in the condition impacting the results of individual items.

The COSMIN criteria for good measurement error were not met by any of the included studies. The reported SDC values were higher than the MIC in a CLBP population (4 points).[20] Researchers and clinicians should consider complementing kinesiophobia assessments with other instruments. The TSK-11 alone may not have the ability to detect clinically relevant changes in patient conditions.

This review also found "low" evidence for the responsiveness of the TSK-11. The evidence came from one study with "adequate" quality, but the sample size requirements were not met. Also, the change criterion used was not specific to the construct evaluated by the TSK (‘‘To what extent do you think you have accomplished your goals in the past 4 weeks?’).[61] Future research should consider assessing the MIC of the TSK-11 in clinically diverse populations and the responsiveness of the instrument, ideally with a criterion of change specific to kinesiophobia.

To our knowledge, this is the first systematic review of the measurement properties of culturally adapted versions of TSK-11. This study identified the cultural adaptations of the TSK-11, which could help promote its use in different cultures. The measurement properties of the culturally adapted versions were assessed according to the COSMIN criteria and methodological flaws in the conduct of the studies were identified. We hope our results will guide researchers in future studies and help clinicians with important considerations when applying the instrument.

Clinical Implications

Clinicians should be aware that clinical aspects may influence the structural validity of the TSK-11. Despite being culturally adapted to a language, when using the questionnaire its measurement properties in specific clinical populations need to be considered. In some populations, the construct of kinesiophobia may be unidimensional wereas in others it may be better represented by different lower-order constructs. The use of the TSK-11 to assess kinesiophobia must be complemented with other instruments. The known MIC is lower than the SDC reported in included studies and, there is “low” evidence for “sufficient” responsiveness of the questionnaire. Using the TSK-11 as a standalone instrument may omit relevant clinical progression in the patient’s condition.

Limitations

In this review, content validity was not assessed since most of the studies that aimed to validate the TSK-11 and assess its measurement properties took the previously culturally adapted 17-item version as their starting point. However, the results of this review should be interpreted taking this limitation into account. The potential omission of relevant sources of information due to publication and language bias should also be considered.

5. CONCLUSION

This systematic review identified cross-cultural adaptations of the TSK-11 and evaluated its measurement properties according to COSMIN criteria. Fifteen culturally and linguistically adapted versions of TSK-11 were identified: English, Chinese, Swedish, Dutch, German, Arabic, Turkish, Danish, Spanish, Japanese, Brazilian Portuguese, Marathi (Indian), Cantonese, Persian, and Thai. The majority of the studies were conducted in patients with chronic pain. Evidence on the measurement properties of TSK-11 in clinically diverse populations begins to emerge but is still sparse. The structural validity of the TSK-11 seems to vary according to populations and contexts. Still, the diversity in statistical methodologies used prevents conclusions from being drawn.

There is “high certainty” in the evidence for “sufficient” criterion validity and for “insufficient” measurement error. “Moderate” certainty in the evidence for “sufficient” construct validity and test-retest reliability and, “low” certainty for “sufficient” responsiveness. Although the TSK-11 appears to be a valid and reliable patient-reported outcome measure, measurement error, responsiveness and structural validity of the instrument should be further investigated.

Disclosures

This research did not receive any specific grants from funding agencies in the public, commercial, or not-for-profit sectors.

The authors have no conflicts of interest to declare.

Figure1. Flowchart of the search strategy and selection of studies.

Adapted from: Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71. https://doi.org/10.1136/bmj.n71

References

1. Kori SH. Kinesiophobia: a new view of chronic pain behavior. *Pain Manage*. 1990;3:35-43.

2. Vlaeyen JWS, Kole-Snijders AMJ, Boeren RGB, van Eek H. Fear of movement/(re)injury in chronic low back pain and its relation to behavioral performance. *Pain*. 1995;62(3):363-372. https://doi.org/10.1016/0304-3959(94)00279-N

3. Vincent HK, Seay AN, Montero C, Conrad BP, Hurley RW, Vincent KR. Kinesiophobia and fear-avoidance beliefs in overweight older adults with chronic low-back pain: relationship to walking endurance--part II. *Am J Phys Med Rehabil*. 2013;92(5):439-445. https://doi.org/10.1097/PHM.0b013e318287633c

4. Şahin HB, Kalaycıoğlu E, Şahin M. The effect of cardiac rehabilitation on kinesiophobia in patients with coronary artery disease. *Turk J Phys Med Rehabil*. 2021;67(2):203-210. https://doi.org/10.5606/tftrd.2021.5164

5. Tajdini H, Letafatkar A, Brewer BW, Hosseinzadeh M. Association between Kinesiophobia and Gait Asymmetry after ACL Reconstruction: Implications for Prevention of Reinjury. *Int J Environ Res Public Health*. 2021;18(6):3264. https://doi.org/10.3390/ijerph18063264

6. Zhang Q, Zhang J, Ran W, Yu S, Jin Y. Effectiveness of cognitive behavioral therapy on kinesiophobia and oral health-related quality of life in patients with temporomandibular disorders, study protocol for a randomized controlled trial. *Medicine*. 2020;99(47):e23295. https://doi.org/10.1097/MD.0000000000023295

7. Ozer AY, Karaca S, Senocak E, Oguz S, Polat MG. Does kinesiophobia limit physical activity and quality of life in asthmatic patients?. *Int J Rehabil Res*. 2022;45(3):230-236. https://doi.org/10.1097/MRR.0000000000000534

8. Selçuk MA, Karakoyun A. Is There a Relationship Between Kinesiophobia and Physical Activity Level in Patients with Knee Osteoarthritis? *Pain Med*. 2020;21(12):3458-3469. https://doi.org/10.1093/pm/pnaa180

9. Sütçü G, Ayvat E, Kilinç M. Effects of fatigue and kinesiophobia on functional capacity, physical activity and quality of life in Parkinson's disease. *Int J Rehabil Res*. 2021;44(1):65-68. https://doi.org/10.1097/MRR.0000000000000449

10. Luque-Suarez A, Martinez-Calderon J, Falla D. Role of kinesiophobia on pain, disability and quality of life in people suffering from chronic musculoskeletal pain: a systematic review. *Br J Sports Med*. 2019;53(9):554-559. https://doi.org/10.1136/bjsports-2017-098673

11. Parent-Roberge H, Marechal R, Fontvieille A, et al. Breaking Barriers: Could Exercise Supervision Attenuate Kinesiophobia in an Older Cancer Patient? *Adv Geriatr Med Res*.2019;1:e190011. https://doi.org/10.20900/agmr20190011

12. Monticone M, Ferrante S, Teli M, et al. Management of catastrophising and kinesiophobia improves rehabilitation after fusion for lumbar spondylolisthesis and stenosis. A randomised controlled trial. *Eur Spine J*. 2014;23(1):87-95. https://doi.org/10.1007/s00586-013-2889-z

13. Monticone M, Ambrosini E, Rocca B, Cazzaniga D, Liquori V, Foti C. Group-based task-oriented exercises aimed at managing kinesiophobia improved disability in chronic low back pain. *Eur J Pain*. 2016;20(4):541-551. https://doi.org/10.1002/ejp.756

14. Miller RP, Kori SH, Todd DD. The Tampa Scale: a Measure of Kinisophobia. *Clin J Pain*. 1991;7(1):51.

15. Cordeiro N, Pezarat-Correia P, Gil J, Cabri J. Portuguese language version of the Tampa Scale for Kinesiophobia [13 items]. *J Musculoskelet Pain*. 2013;21(1):58-63. https://doi.org/10.3109/10582452.2012.762966

16. Monticone M, Giorgi I, Baiardi P, Barbieri M, Rocca B, Bonezzi C. Development of the Italian version of the Tampa Scale of Kinesiophobia (TSK-I): cross-cultural adaptation, factor analysis, reliability, and validity. *Spine (Phila Pa 1976)*. 2010;35(12):1241-1246. https://doi.org/10.1097/BRS.0b013e3181bfcbf6

17. Rusu AC, Kreddig N, Hallner D, Hülsebusch J, Hasenbring MI. Fear of movement/(Re)injury in low back pain: confirmatory validation of a German version of the Tampa Scale for Kinesiophobia. *BMC Musculoskelet Disord*. 2014;15:280. https://doi.org/10.1186/1471-2474-15-280

18. Wei X, Xu X, Zhao Y, Hu W, Bai Y, Li M. The Chinese version of the Tampa Scale for Kinesiophobia was cross-culturally adapted and validated in patients with low back pain. *J Clin Epidemiol*. 2015;68(10):1205-1212. https://doi.org/10.1016/j.jclinepi.2015.07.003

19. Houben RMA, Leeuw M, Vlaeyen JWS, Goubert L, Picavet HSJ. Fear of movement/injury in the general population: factor structure and psychometric properties of an adapted version of the Tampa Scale for Kinesiophobia. *J Behav Med*. 2005;28(5):415-424. https://doi.org/10.1007/s10865-005-9011-x

20. Woby SR, Roach NK, Urmston M, Watson PJ. Psychometric properties of the TSK-11: a shortened version of the Tampa Scale for Kinesiophobia. *Pain*. 2005;117(1-2):137-144. https://doi.org/10.1016/j.pain.2005.05.029

21. Roelofs J, Sluiter JK, Frings-Dresen MH, et al. Fear of movement and (re)injury in chronic musculoskeletal pain: Evidence for an invariant two-factor model of the Tampa Scale for Kinesiophobia across pain diagnoses and Dutch, Swedish, and Canadian samples. *Pain*. 2007;131(1-2):181-190. https://doi.org/10.1016/j.pain.2007.01.008

22. Alswat MM, Khojah O, Alswat AM, et al. Returning to Sport After Anterior Cruciate Ligament Reconstruction in Physically Active Individuals. *Cureus*. 2020;12(9):e10466. https://doi.org/10.7759/cureus.10466

23. Cai L, Liu Y, Woby SR, Genoosha N, Cui M, Guo L. Cross-Cultural Adaptation, Reliability, and Validity of the Chinese Version of the Tampa Scale for Kinesiophobia-11 Among Patients Who Have Undergone Total Knee Arthroplasty. *J Arthroplasty*. 2019;34(6):1116-1121. https://doi.org/10.1016/j.arth.2019.01.076

24. DiBartola AC, Magnussen RA, Wiet M, et al. Predictors of poor pre-operative psychological status among patients with cartilage defects. *The Knee*. 2021;33:11-16. https://doi.org/10.1016/j.knee.2021.08.026

25. Fernández-de-Las-Peñas C, Herrero-Montes M, Cancela-Cilleruelo I, et al. Understanding Sensitization, Cognitive and Neuropathic Associated Mechanisms behind Post-COVID Pain: A Network Analysis. *Diagn Basel Switz*. 2022;12(7) :1538-1548. https://doi.org/10.3390/diagnostics12071538

26. Kim KS, An J, Kim JO, Lee MY, Lee BH. Effects of Pain Neuroscience Education Combined with Lumbar Stabilization Exercise on Strength and Pain in Patients with Chronic Low Back Pain: Randomized Controlled Trial. *J Pers Med*. 2022;12(2):303-316. https://doi.org/10.3390/jpm12020303

27. Martin-Vera D, Fernández-Carnero J, Rodríguez-Sanz D, et al. Median Nerve Neural Mobilization Adds No Additional Benefit When Combined with Cervical Lateral Glide in the Treatment of Neck Pain: A Randomized Clinical Trial. *J Clin Med*. 2021;10(21):5178-5194. https://doi.org/10.3390/jcm10215178

28. Reinoso-Cobo A, Ortega-Avila AB, Ramos-Petersen L, et al. Relationship between Kinesiophobia, Foot Pain and Foot Function, and Disease Activity in Patients with Rheumatoid Arthritis: A Cross-Sectional Study. *Med Kaunas Lith*. 2023;59(1):147-155. https://doi.org/10.3390/medicina59010147

29. Rhim HC, Lee SJ, Jeon JS, et al. Effectiveness of modeling videos on psychological responses of patients following anterior cruciate ligament reconstruction: A pilot randomized trial. *Medicine*. 2020;99(45):e23158. https://doi.org/10.1097/MD.0000000000023158

30. Selhorst M, Fernandez-Fernandez A, Schmitt L, Hoehn J. Adolescent psychological beliefs, but not parent beliefs, associated with pain and function in adolescents with patellofemoral pain. *Phys Ther Sport*. 2020;45:155-160. https://doi.org/10.1016/j.ptsp.2020.07.003

31. Barriga-Valenzuela E, Caballero-Sepúlveda C, Medina-Muñoz E, Núñez-Cortés R, Urrutia-Quezada R, Espinoza-Ordoñez C. The influence of cognitive factors in relation to the patients' treatment adherence for non-specific chronic low-back pain. A case series. *J Bodyw Mov Ther*. 2022;29:271-278. https://doi.org/10.1016/j.jbmt.2021.10.006

32. Hadadi M, Heghighat F, Hossein I. Comparison of Health-Related Quality of Life in Different Types of Chronic Ankle Instability. *Phys Med Rehabil Kurortmed*. 2020;30(04):230-236. https://doi.org/10.1055/a-1128-6424

33. Palomo-López P, Becerro-de-Bengoa-Vallejo R, Losa-Iglesias ME, et al. Kinesiophobia and Pain Intensity Are Increased by a Greater Hallux Valgus Deformity Degree-Kinesiophobia and Pain Intensity in Hallux Valgus. *Int J Environ Res Public Health*. 2020;17(2):626-638. https://doi.org/10.3390/ijerph17020626

34. San-Antolín M, Rodríguez-Sanz D, Vicente-Campos D, et al. Fear Avoidance Beliefs and Kinesiophobia Are Presented in Athletes who Suffer from Gastrocnemius Chronic Myofascial Pain. *Pain Med*. 2020;21(8):1626-1635. https://doi.org/10.1093/pm/pnz362

35. Satpute KH, Ranade PS, Hall TM. Development of the Marathi version of the Tampa scale of kinesiophobia 11: Cross-cultural adaptation, validity, and test-retest reliability in patients with low back pain. *Indian Spine J*. 2019;2(2):146-151. https://doi.org/10.4103/isj.isj_13_18

36. Woolnough LU, Lentini L, Sharififar S, Chen C, Vincent HK. The relationships of kinesiophobia and physical function and physical activity level in juvenile idiopathic arthritis. *Pediatr Rheumatol Online J*. 2022;20(1):73. https://doi.org/10.1186/s12969-022-00734-2

37. Goldberg P, Zeppieri G, Bialosky J, et al. Kinesiophobia and Its Association With Health-Related Quality of Life Across Injury Locations. *Arch Phys Med Rehabil*. 2018;99(1):43-48. https://doi.org/10.1016/j.apmr.2017.06.023

38. Zelle DM, Corpeleijn E, Klaassen G, Schutte E, Navis G, Bakker SJL. Fear of Movement and Low Self-Efficacy Are Important Barriers in Physical Activity after Renal Transplantation. *PloS One*. 2016;11(2):e0147609. https://doi.org/10.1371/journal.pone.0147609

39. Walton DM, Eilon-Avigdor Y, Wonderham M, Wilk P. Exploring the clinical course of neck pain in physical therapy: a longitudinal study. *Arch Phys Med Rehabil*. 2014;95(2):303-308. https://doi.org/10.1016/j.apmr.2013.09.004

40. Lentz TA, Sutton Z, Greenberg S, Bishop MD. Pain-related fear contributes to self-reported disability in patients with foot and ankle pathology. *Arch Phys Med Rehabil*. 2010;91(4):557-561. https://doi.org/10.1016/j.apmr.2009.12.010

41. Ferrer-Peña R, Vicente-de-Frutos G, Flandez-Santos D, Martín-Gómez C, Roncero-Jorge C, Calvo-Lobo C. Patient-reported outcomes measured with and without dizziness associated with non-specific chronic neck pain: implications for primary care. *PeerJ*. 2019;7:e7449. https://doi.org/10.7717/peerj.7449

42. Terradas-Monllor M, Navarro-Fernández G, Ruiz MA, et al. Postoperative Psychosocial Factors in Health Functioning and Health-Related Quality of Life After Knee Arthroplasty: A 6-Month Follow up Prospective Observational Study. *Pain Med*. 2021;22(9):1905-1915. https://doi.org/10.1093/pm/pnab025

43. Erden A, Guner S. Impact of exercise on quality of life, body awareness, kinesiophobia and the risk of falling among young older adults. *Cukurova Med J*. 2018;43(4):941-950. https://doi.org/10.17826/cumj.382689

44. Larsson C, Hansson EE, Sundquist K, Jakobsson U. Psychometric properties of the Tampa Scale of Kinesiophobia (TSK-11) among older people with chronic pain. *Physiother Theory Pract*. 2014;30(6):421-428. https://doi.org/10.3109/09593985.2013.877546

45. Larsson C, Ekvall Hansson E, Sundquist K, Jakobsson U. Kinesiophobia and its relation to pain characteristics and cognitive affective variables in older adults with chronic pain. *BMC Geriatr*. 2016;16:128. https://doi.org/10.1186/s12877-016-0302-6

46. Bisson EJ, Katz L, Vader K, Bossio JA, Wilson R, Duggan S. Pain-Related Fear in Adults Living With Chronic Pain: Development and Psychometric Validation of a Brief Form of the Tampa Scale of Kinesiophobia. *Arch Phys Med Rehabil*. 2022;103(5):875-881.e1. https://doi.org/10.1016/j.apmr.2021.09.001

47. Cresswell C, Galantino ML, Myezwa H. The prevalence of fear avoidance and pain catastrophising amongst patients with chronic neck pain. *South Afr J Physiother*. 2020;76(1):1326. https://doi.org/10.4102/sajp.v76i1.1326

48. Dagenais M, Brun C, Ohayon A, Mercier C. Virtual Reality in Fibromyalgia: Does Altering Visual Feedback Impact on Pain and Movement During Reaching? *Front Virtual Real*. 2021;2. https://doi.org/10.3389/frvir.2021.681034

49. Fernández-Carnero J, Beltrán-Alacreu H, Arribas-Romano A, et al. Prediction of Patient Satisfaction after Treatment of Chronic Neck Pain with Mulligan's Mobilization. *Life Basel Switz*. 2022;13(1):48-63. https://doi.org/10.3390/life13010048

50. Ferrer-Peña R, Moreno-López M, Calvo-Lobo C, López-de-Uralde-Villanueva I, Fernández-Carnero J. Relationship of Dynamic Balance Impairment with Pain-Related and Psychosocial Measures in Primary Care Patients with Chronic Greater Trochanteric Pain Syndrome. *Pain Med*. 2019;20(4):810-817. https://doi.org/10.1093/pm/pny160

51. Liew BXW, Valera-Calero JA, Varol U, et al. Distress and Sensitization as Main Mediators of Severity in Women with Fibromyalgia: A Structural Equation Model. *Biomedicines*. 2022;10(5):1188-1198. https://doi.org/10.3390/biomedicines10051188

52. Osumi M, Sumitani M, Otake Y, et al. Kinesiophobia modulates lumbar movements in people with chronic low back pain: a kinematic analysis of lumbar bending and returning movement [published correction appears in Eur Spine J. 2019 Jun 13;:]. *Eur Spine J*. 2019;28(7):1572-1578. https://doi.org/10.1007/s00586-019-06010-4

53. Vicente-Campos D, Sanchez-Jorge S, Terrón-Manrique P, et al. The Main Role of Diaphragm Muscle as a Mechanism of Hypopressive Abdominal Gymnastics to Improve Non-Specific Chronic Low Back Pain: A Randomized Controlled Trial. *J Clin Med*. 2021;10(21):4983-4996. https://doi.org//10.3390/jcm10214983

54. Beltran-Alacreu H, López-de-Uralde-Villanueva I, Fernández-Carnero J, La Touche R. Manual Therapy, Therapeutic Patient Education, and Therapeutic Exercise, an Effective Multimodal Treatment of Nonspecific Chronic Neck Pain: A Randomized Controlled Trial. *Am J Phys Med Rehabil*. 2015;94(10 Suppl 1):887-897. https://doi.org/10.1097/PHM.0000000000000293

55. Beltran-Alacreu H, López-de-Uralde-Villanueva I, Calvo-Lobo C, Fernández-Carnero J, La Touche R. Clinical features of patients with chronic non-specific neck pain per disability level: A novel observational study. *Rev Assoc Med Bras (1992)*. 2018;64(8):700-709. https://doi.org/10.1590/1806-9282.64.08.700

56. Garrigós-Pedrón M, La Touche R, Navarro-Desentre P, Gracia-Naya M, Segura-Ortí E. Effects of a Physical Therapy Protocol in Patients with Chronic Migraine and Temporomandibular Disorders: A Randomized, Single-Blinded, Clinical Trial. *J Oral Facial Pain Headache*. 2018;32(2):137-150. https://doi.org/10.11607/ofph.1912

57. La Touche R, Pardo-Montero J, Gil-Martínez A, et al. Craniofacial pain and disability inventory (CF-PDI): development and psychometric validation of a new questionnaire. *Pain Physician*. 2014;17(1):95-108.

58. Lopez-de-Uralde-Villanueva I, Beltran-Alacreu H, Fernandez-Carnero J, Kindelan-Calvo P, La Touche R. Widespread Pressure Pain Hyperalgesia in Chronic Nonspecific Neck Pain with Neuropathic Features: A Descriptive Cross-Sectional Study. *Pain Physician*. 2016;19(2):77-88.

59. Marcos-Martín F, González-Ferrero L, Martín-Alcocer N, Paris-Alemany A, La Touche R. Multimodal physiotherapy treatment based on a biobehavioral approach for patients with chronic cervico-craniofacial pain: a prospective case series. *Physiother Theory Pract*. 2018;34(9):671-681. https://doi.org/10.1080/09593985.2017.1423522

60. Gil-Martínez A, Navarro-Fernández G, Mangas-Guijarro MÁ, et al. Comparison Between Chronic Migraine and Temporomandibular Disorders in Pain-Related Disability and Fear-Avoidance Behaviors. *Pain Med*. 2017;18(11):2214-2223. https://doi.org/10.1093/pm/pnx127

61. Hapidou EG, O'Brien MA, Pierrynowski MR, de Las Heras E, Patel M, Patla T. Fear and Avoidance of Movement in People with Chronic Pain: Psychometric Properties of the 11-Item Tampa Scale for Kinesiophobia (TSK-11). *Physiother Can*. 2012;64(3):235-241. https://doi.org/10.3138/ptc.2011-10

62. Wong WS, McCracken LM, Fielding R. Factorial validity and reliability of the Chinese version of the Pain Vigilance and Awareness Questionnaire (ChPVAQ) in a sample of Chinese patients with chronic pain. *Pain Med*. 2011;12(7):1018-1025. https://doi.org/10.1111/j.1526-4637.2011.01169.x

63. Wong WS, McCracken LM, Fielding R. Factor structure and psychometric properties of the Chinese version of the 20-item Pain Anxiety Symptoms Scale (ChPASS-20). *J Pain Symptom Manage*. 2012;43(6):1131-1140. https://doi.org/10.1016/j.jpainsymman.2011.06.021

64. Santo Salvador E, Franco K, Miyamoto G, Franco Y, Cabral C. Analysis of the measurement properties of the Brazilian-Portuguese version of the Tampa Scale for Kinesiophobia-11 in patients with fibromyalgia. *Braz J Phys Ther*. 2021;25(2):168-174. https://doi.org/10.1016/j.bjpt.2020.05.004

65. Baez S, Collins K, Harkey M, et al. Kinesiophobia Is Associated with Peak Knee Abduction Angle during Jump Landing after ACL Reconstruction. *Med Sci Sports Exerc*. 2023;55(3):462-468. https://doi.org/10.1249/MSS.0000000000003075

66. Baez S, Cormier M, Andreatta R, Gribble P, Hoch JM. Implementation of In vivo exposure therapy to decrease injury-related fear in females with a history of ACL-Reconstruction: A pilot study. *Phys Ther Sport*. 2021;52:217-223. https://doi.org/10.1016/j.ptsp.2021.09.009

67. Browning RB, Clapp IM, Alter TD, Nwachukwu BU, Nho SJ. Pain Catastrophizing and Kinesiophobia Affect Return to Sport in Patients Undergoing Hip Arthroscopy for the Treatment of Femoroacetabular Impingement. *Arthrosc Sports Med Rehabil*. 2021;3(4):e1087-e1095. https://doi.org/10.1016/j.asmr.2021.03.014

68. Chmielewski TL, George SZ. Fear avoidance and self-efficacy at 4 weeks after ACL reconstruction are associated with early impairment resolution and readiness for advanced rehabilitation. *Knee Surg Sports Traumatol Arthrosc*. 2019;27(2):397-404. https://doi.org/10.1007/s00167-018-5048-6

69. Chmielewski TL, Jones D, Day T, Tillman SM, Lentz TA, George SZ. The association of pain and fear of movement/reinjury with function during anterior cruciate ligament reconstruction rehabilitation. *J Orthop Sports Phys Ther*. 2008;38(12):746-753. https://doi.org/10.2519/jospt.2008.2887

70. Chmielewski TL, Zeppieri GJ, Lentz TA, et al. Longitudinal changes in psychosocial factors and their association with knee pain and function after anterior cruciate ligament reconstruction. *Phys Ther*. 2011;91(9):1355-1366. https://doi.org/10.2522/ptj.20100277

71. Dudley RI, Lohman EB, Patterson CS, Knox KG, Gharibvand L. The relationship between kinesiophobia and biomechanics in anterior cruciate ligament reconstructed females. *Phys Ther Sport*. 2022;56:32-37. https://doi.org/10.1016/j.ptsp.2022.06.002

72. Feigenbaum LA, Baraga M, Kaplan LD, et al. Return to Sport Following Surgery for a Complicated Tibia and Fibula Fracture in a Collegiate Women's Soccer Player with a Low Level of Kinesiophobia. *Int J Sports Phys Ther*. 2015;10(1):95-103.

73. George SZ, Lentz TA, Zeppieri Jr G, Lee D, Chmielewski TL. Analysis of shortened versions of the Tampa Scale for Kinesiophobia and Pain Catastrophizing Scale for patients after anterior cruciate ligament reconstruction. *Clin J Pain*. 2012;28(1):73-80. https://doi.org/10.1097/AJP.0b013e31822363f4

74. Hartigan EH, Lynch AD, Logerstedt DS, Chmielewski TL, Snyder-Mackler L. Kinesiophobia after anterior cruciate ligament rupture and reconstruction: noncopers versus potential copers. *J Orthop Sports Phys Ther*. 2013;43(11):821-832. https://doi.org/10.2519/jospt.2013.4514

75. Hoch JM, Sinnott CW, Robinson KP, Perkins WO, Hartman JW. The Examination of Patient-Reported Outcomes and Postural Control Measures in Patients With and Without a History of ACL Reconstruction: A Case Control Study. *J Sport Rehabil*. 2018;27(2):170-176. https://doi.org/10.1123/jsr.2016-0105

76. Hsu CJ, George SZ, Chmielewski TL. Association of Quadriceps Strength and Psychosocial Factors With Single-Leg Hop Performance in Patients With Meniscectomy. *Orthop J Sports Med*. 2016;4(12):2325967116676078. https://doi.org/10.1177/2325967116676078

77. Hsu CJ, George SZ, Chmielewski TL. Fear-avoidance and self-efficacy psychosocial factors are altered after partial meniscectomy and associated with rehabilitation outcomes. *Int J Sports Phys Ther*. 2020;15(4):557-570.

78. Isaji Y, Yamada T, Oka T, Mori K, Aoyama N. Association between knee function and kinesiophobia 6 months after anterior cruciate ligament reconstruction. *J Phys Ther Sci*. 2023;35(2):128-132. https://doi.org/10.1589/jpts.35.128

79. Kuenze C, Lisee C, Triplett A, et al. Validation of a Survey to Characterize Barriers to Physical Activity After Anterior Cruciate Ligament Reconstruction. *J Athl Train*. 2023;58(10):841-848. https://doi.org/10.4085/1062-6050-0436.22

80. Lentz TA, Zeppieri GJ, George SZ, et al. Comparison of physical impairment, functional, and psychosocial measures based on fear of reinjury/lack of confidence and return-to-sport status after ACL reconstruction. *Am J Sports Med*. 2015;43(2):345-353. https://doi.org/10.1177/0363546514559707

81. Luc-Harkey BA, Franz JR, Losina E, Pietrosimone B. Association between kinesiophobia and walking gait characteristics in physically active individuals with anterior cruciate ligament reconstruction. *Gait Posture*. 2018;64:220-225. https://doi.org/10.1016/j.gaitpost.2018.06.029

82. Meierbachtol A, Obermeier M, Yungtum W, et al. Advanced training enhances readiness to return to sport after anterior cruciate ligament reconstruction. *J Orthop Res*. 2022;40(1):191-199. https://doi.org/10.1002/jor.25072

83. Müller U, Krüger-Franke M, Schmidt M, Rosemeyer B. Predictive parameters for return to pre-injury level of sport 6 months following anterior cruciate ligament reconstruction surgery. *Knee Surg Sports Traumatol Arthrosc*. 2015;23(12):36233631. https://doi.org/10.1007/s00167-014-3261-5

84. Finley MA, Euiler E. Association of musculoskeletal pain, fear-avoidance factors, and quality of life in active manual wheelchair users with SCI: A pilot study. *J Spinal Cord Med*. 2020;43(4):497-504. https://doi.org/10.1080/10790268.2019.1565717

85. Jiménez-Cebrián AM, Becerro-de-Bengoa-Vallejo R, Losa-Iglesias ME, et al. Kinesiophobia Levels in Patients with Parkinson's Disease: A Case-Control Investigation. *Int J Environ Res Public Health*. 2021;18(9):4791-4797. https://doi.org/10.3390/ijerph18094791

86. Ruiz-Sánchez FJ, Martins M do R, Soares S, et al. Kinesiophobia Levels in Patients with Multiple Sclerosis: A Case-Control Investigation. *Biology*. 2022;11(10):1428-1434. https://doi.org/10.3390/biology11101428

87. Velthuis M, Van den Bussche E, May A, Gijsen B, Nijs S, Vlaeyen J. Fear of movement in cancer survivors: validation of the Modified Tampa Scale of Kinesiophobia-Fatigue. *Psychooncology*. 2012;21(7):762-770. https://doi.org/10.1002/pon.1971

88. Wang YW, Qiu XX. Kinesiophobia and related factors in cancer patients with TIAPs during the long term: a cross-sectional survey. *Support Care Cancer*. 2022;30(6):4927-4934. https://doi.org/10.1007/s00520-022-06842-9

89. Acar S, Savci S, Keskinoğlu P, et al. Tampa scale of kinesiophobia for heart Turkish version study: Cross-cultural adaptation, exploratory factor analysis, and reliability. *J Pain Res*. 2016;9:445-451. https://doi.org/10.2147/JPR.S105766

90. Al-Shudifat A, Farah K, Hawamdeh ZM, Alqudah A, Juweid ME. Psychometric testing of a short form, 11-item Tampa Scale of Kinesiophobia-Arabic version: TSK-AV-11. *Medicine*. 2020;99(24):e20292. https://doi.org/10.1097/MD.0000000000020292

91. Farzad M, MacDermid J, Packham T. Fear of Movement in People With CRPS: A Psychometric Evaluation of the Tampa Scale-11 for Kinesiophobia. *Clin J Pain*. 2024;40(1):26-34. https://doi.org/10.1097/AJP.0000000000001169

92. Gomez-Perez L, Lopez-Martinez A, Ruiz-Parraga G. Psychometric Properties of the Spanish Version of the Tampa Scale for Kinesiophobia (TSK). *J Pain*. 2011;12(4):425-435. https://doi.org/10.1016/j.jpain.2010.08.004

93. Kikuchi N, Matsudaira K, Sawada T, Oka H. Psychometric properties of the Japanese version of the Tampa Scale for Kinesiophobia (TSK-J) in patients with whiplash neck injury pain and/or low back pain. *J Orthop Sci*. 2015;20(6):985-992. https://doi.org/10.1007/s00776-015-0751-3

94. Prinsen CAC, Mokkink LB, Bouter LM, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res*. 2018;27(5):1147-1157. https://doi.org/10.1007/s11136-018-1798-3

95. Herdman M, Fox-Rushby J, Badia X. A model of equivalence in the cultural adaptation of HRQoL instruments: the universalist approach. *Qual Life Res*.1998;7(4):323335. https://doi.org/10.1023/a:1024985930536

96. de W, Terwee CB, Mokkink LB, Knol DL. *Measurement in Medicine*. Cambridge University Press; 2011.

97. Kimberlin CL, Winterstein AG. Validity and reliability of measurement instruments used in research. *Am J Health Syst Pharm*. 2008;65(23):2276-2284. https://doi.org/10.2146/ajhp070364

98. Lundberg M, Styf J, Jansson B. On what patients does the Tampa Scale for Kinesiophobia fit? *Physiother Theory Pract*. 2009;25(7):495-506. https://doi.org/10.3109/09593980802662160

99. Bunzli S, Smith A, Watkins R, Schutze R, O'Sullivan P. What Do People Who Score Highly on the Tampa Scale of Kinesiophobia Really Believe? A Mixed Methods Investigation in People With Chronic Nonspecific Low Back Pain. *Clin J PAIN*. 2015;31(7):621-632. https://doi.org/10.1097/AJP.0000000000000143

100. Turk DC, Robinson JP, Sherman JJ, Burwinkle T, Swanson K. Assessing Fear in Patients with Cervical Pain: Development and Validation of the Pictorial Fear of Activity Scale – Cervical (PFActS-C). *Pain*. 2008;139(1):55-62. https://doi.org/10.1016/j.pain.2008.03.001

101.Tissot LPM, Evans DW, Kirby E, Liew BXW. Tampa Scale of Kinesiophobia may underestimate task-specific fear of movement in people with and without low back pain. *Pain Rep*. 2023;8(4):e1081. https://doi.org/10.1097/PR9.0000000000001081

102. Dupuis F, Cherif A, Batcho C, Massé-Alarie H, Roy JS. The Tampa Scale of Kinesiophobia: A Systematic Review of Its Psychometric Properties in People With Musculoskeletal Pain. *Clin J Pain*.2023;39(5):236-247. https://doi.org/10.1097/AJP.0000000000001104

103.Stephenson M, Riitano D, Wilson S, et al. Chapter 12: Systematic Reviews of Measurement Properties. In: Aromataris E, Munn Z, eds. *JBI Manual for Evidence Synthesis*. JBI; 2020:452-468. https://doi.org/10.46658/JBIMES-20-13

104.Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021;372:n71. https://doi.org/10.1136/bmj.n71

105. Terwee CB, Jansma EP, Riphagen II, de Vet HC. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. *Qual Life Res*.2009;18(8):1115-1123. https://doi.org/10.1007/s11136-009-9528-5

106. Mokkink LB, de Vet HCW, Prinsen CAC, et al. COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. *Qual Life Res*. 2018;27(5):1171-1179. https://doi.org/10.1007/s11136-017-1765-4

107.Vlaeyen JWS, Linton SJ. Fear-avoidance model of chronic musculoskeletal pain: 12 years on. *Pain*. 2012;153(6):1144-1147. https://doi.org/10.1016/j.pain.2011.12.009

108. Roelofs J, Goubert L, Peters ML, Vlaeyen JWS, Crombez G. The Tampa Scale for Kinesiophobia: further examination of psychometric properties in patients with chronic low back pain and fibromyalgia. *Eur J Pain Lond Engl*. 2004;8(5):495-502. https://doi.org/10.1016/j.ejpain.2003.11.016

109. Vlaeyen JW, Kole-Snijders AM, Rotteveel AM, Ruesink R, Heuts PH. The role of fear of movement/(re)injury in pain disability. *J Occup Rehabil*. 1995;5(4):235-252. https://doi.org/10.1007/BF02109988

110. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Routledge; 1988.

111. Schünemann H, Brożek J, Guyatt G, Oxman A, editors. *GRADE handbook for grading quality of evidence and strength of recommendations*. Updated October 2013. The GRADE Working Group, 2013. Available from guidelinedevelopment.org/handbook.https://gdt.gradepro.org/app/handbook/handbook.html

112. Archer KR, Phelps KD, Seebach CL, Song Y, Riley LH 3rd, Wegener ST. Comparative study of short forms of the Tampa Scale for Kinesiophobia: fear of movement in a surgical spine population. *Arch Phys Med Rehabil*. 2012;93(8):1460-1462. https://doi.org/10.1016/j.apmr.2012.03.024

113. Eiger B, Errebo M, Straszek C, Vaegter H. Less is more: reliability and measurement error for three versions of the Tampa Scale of Kinesiophobia (TSK-11, TSK-13, and TSK-17) in patients with high-impact chronic pain. *Scand J Pain*. 2023;23(1):217-224. https://doi.org/10.1515/sjpain-2021-0200

114. Mielenz TJ, Edwards MC, Callahan LF. First item response theory analysis on Tampa Scale for Kinesiophobia (fear of movement) in arthritis. *J Clin Epidemiol*. 2010;63(3):315-320. https://doi.org/10.1016/j.jclinepi.2009.04.011

115. Tkachuk G, Harris C. Psychometric Properties of the Tampa Scale for Kinesiophobia-11 (TSK-11). *J Pain*. 2012;13(10):970-977. https://doi.org/10.1016/j.jpain.2012.07.001

116. Walton D, Elliott J. A Higher-Order Analysis Supports Use of the 11-Item Version of the Tampa Scale for Kinesiophobia in People With Neck Pain. *Phys Ther*. 2013;93(1):60-68. https://doi.org/10.2522/ptj.20120255

117. Wong W, Kwok H, Luk K, et al. Fear of movement/(re)injury in Chinese patients with chronic pain: Factorial validity of the Chinese version of the Tampa Scale for Kinesiophobia. *J Rehabil Med*. 2010;42(7):620-629. https://doi.org/10.2340/16501977-0575

118. Youngcharoen P, Saraboon Y, Aree-Ue S, Chintapanyakun T, Kawinwonggowit V. Psychometric properties of the thai tampa scale of kinesiophobia among older people with Knee osteoarthritis. *J Med Assoc Thai*. 2021;104(8):1317-1325. https://doi.org/10.35755/jmedassocthai.2021.08.12901

119. Pontes-Silva A, Dibai A, de Jesus S, et al. The best structure of the Tampa Scale for Kinesiophobia for patients with chronic low back pain has two domains and nine items. *Clin Rehabil*. 2023;37(3):407-414. https://doi.org/10.1177/02692155221128829

120.Robertson O, Robinson SJ, Stephens R. Swearing as a response to pain: A cross-cultural comparison of British and Japanese participants. *Scand J Pain*. 2017;17:267-272. https://doi.org/10.1016/j.sjpain.2017.07.014

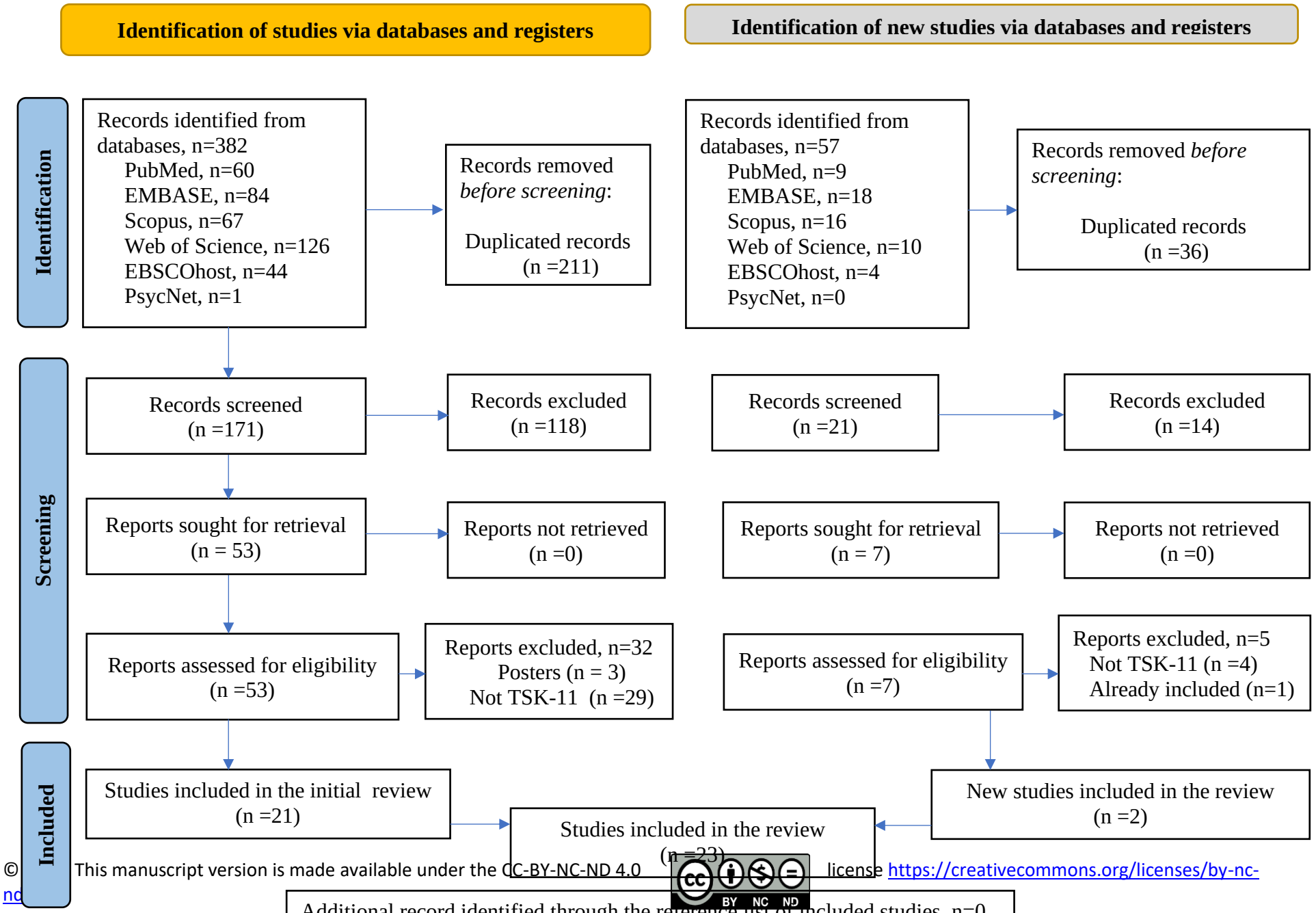

Table 1. Characteristics of included studies

Characteristics	Value [a]
Year of publication	
2007-2012	10 (43.5)
2013-2018	5 (21.7)
2019-2024	8 (34.8)
TSK-11 mode of administration	
Self-completed questionnaire	20 (87.0)
Online questionnaire	2 (8.7)
Computer-based	1 (4.3)
Interview	3 (13.0)
Phone interview	1 (4.3)
Not reported	2 (8.7)
TSK-11 language of administration	
English	6 (26.1)
Swedish	2 (8.7)
Dutch	2 (8.7)
Brazilian Portuguese	2 (8.7)
Arabic	1 (4.3)
Turkish	1 (4.3)
Chinese	1 (4.3)
Danish	1 (4.3)
Spanish	1 (4.3)
Japanese	1 (4.3)
Cantonese	1 (4.3)
German	1 (4.3)
Marathi	1 (4.3)
Thai	1 (4.3)
Persian	1 (4.3)
TSK-11 structure	
Unidimensional	13 (56.5)
2-factor model	14 (60.9)
3-factor model	2 (8.7)
4-factor model	1 (4.3)
Assessed measurement properties	
Validity	22 (95.7)
Construct Validity	21 (91.3)
Structural	19 (82.6)
Hypotheses testing	16 (69.6)
Cross-cultural/measurement invariance	2 (8.7)
Criterion	1 (4.3)
Reliability	19 (82.6)
Internal consistency	17 (73.9)
Test-retest	8 (34.8)
Measurement error	4 (17.4)
Responsiveness	1 (4.3)

[a] Values are reported as number (percentage) of studies unless otherwise indicated.

[b] On the basis of the 21 included studies.

Table 2. Characteristics of participants and context

Characteristics of participants	Value [a]
Number	[85; 2825]
Age	[16; 98]
Gender	
Female	[29%; 98.5%]
Male	[1.5%; 71%]
Clinical condition	
Chronic pain	13 (56.5)
Heterogeneous	5 (21.7)
CLBP	4 (17.4)
Fibromyalgia	2 (8.7)
High-impact chronic pain	1 (4.3)
CRPS	1 (4.3)
Older adults	1 (4.3)
Post-surgical	3 (13.0)
Spinal surgery	1 (4.3)
Total Knee Arthroplasty	1 (4.3)
Anterior Cruciate Ligament reconstruction	1 (4.3)
Acute musculoskeletal pain	1 (4.3)
Osteoarthritis	2 (8.7)
Knee osteoarthritis	1 (4.3)
Whiplash/low back pain	1 (4.3)
Low Back Pain	1 (4.3)
Neck pain	1 (4.3)
Cancer	1(4.3)
Cardiac	1 (4.3)
Work-related upper extremity disorder	1 (4.3)
Characteristics of contexts	
Continent/country	
America	9 (39.1)
Canada	4 (17.4)
USA	3 (13.0)
Brazil	2 (8.7)
Europa	7 (30.4)
Sweden	2 (8.7)
Netherland	2 (8.7)
Denmark	1 (4.3)
Spain	1 (4.3)
Germany	1 (4.3)
Asia	6 (26.1)
China	2 (8.7)
Thailand	1 (4.3)
Jordan	1 (4.3)
Japan	1 (4.3)
India	1 (4.3)
Euro-Asia	1 (4.3)
Turkey	1 (4.3)
Settings	
Online, mailed, or postal questionnaire	5 (21.7)
Pain centre	5 (21.7)
University	4 (17.4)
Physiotherapy /rehabilitation center	4 (17.4)
Hospital	3 (13.0)
Orthopaedics	2 (8.7)

[a] Values are reported as number (percentage) of studies, [range] unless otherwise indicated.

Table 3. Structural Validity of the cross-cultural adaptations of the TSK-11

TSK-11 version	Population	n	Method. quality	Result (rating)
English [112]	Adults treated by spinal surgery	137	Adequate	1-factor model CFI= - 0.05; RMSEA=0.197 **(-)** 2-factor model CFI= 0.86; RMSEA=0.073 **(-)**
English [73]	Patients after ACLR (Early post-operative phase)	105	Adequate	3-factor model with a cumulative response variance of 64.3% **(-)**
English [114]	Community-based with arthritis	275	Adequate	1-factor model CFI=0.98; RMSEA=0.06 **(+)**
English [115]	Patients with heterogeneous chronic pain	276	Very good	2-factor model CFI=0.95; RMSEA=0.06; SRMR=0.05 **(+)**
English [116]	Patients with neck pain	207	Adequate	2-factor model Item-trait interaction:χ^2=0.69, P=0.95; MFR=0.27 **(-)**
Chinese [23]	Patients after TKA	254	Adequate	The 3-factor model accounted for 65.2% of the explanatory variance **(-)**
Cantonese [117]	Patients with chronic pain	O:216	Very good	2-factor model CFI=0.930; RMSEA=0.025 **(+)**
		PC:109		2-factor model CFI=0.820; RMSEA=0.068 **(-)**
Spanish [92]	Heterogeneous chronic pain	125	Adequate	The 2-factor model accounted for 48.37% of the explanatory variance **(-)**
	Acute musculoskeletal pain	86	Doubtful	The 2-factor model explained 50.01% of the variance **(-)**
Swedish [44]	Older people with chronic pain	433	Very good	1-factor model CFI=0.923; RMSEA=0.124 **(-)** 2-factor model CFI=0.945; RMSEA=0.103 **(-)**
Swedish [98]	Patients with persistent musculoskeletal pain	539	Very good	1-factor model CFI=0.808; RMSEA=0.098 **(-)** 2-factor model CFI=0.855; RMSEA=0.086 **(-)**
TSK-F Dutch [87]	Cancer survivors	329	Very good	2-factor model CFI=0.96; RMSEA=0.039 **(+)**

Dutch [21]	Patients with work-related upper extremity disorders	1109	Very good	1-factor model CFI=0.85; RMSEA=0.071 **(-)** 2-factor model CFI=0.93, RMSEA=0.051 **(+)**
TSK-GV German [17]	Patients with Low Back Pain	191	Very good	2-factor model CFI=0.933; RMSEA=0.047 **(+)**
Brazilian [119]	Patients with CLBP	122	Very good	1-factor model CFI=0.906; RMSEA=0.083 **(-)** 2-factor model (Roelofs model) CFI=0.918; RMSEA=0.079 **(-)** 2-factor model (Al-Shudifat model) CFI=0.928; RMSEA=0.078 **(-)**
Thai [118]	Older people with knee OA	200	Very good	2-factor model CFI=0.871; RMSEA=0.075; SRMR=0.064 **(+)**
Persian [91]	Patients suffering from CRPS	142	Very good Adequate	1-factor model CFI=0.97; RMSEA>0.08 **(+)** 2-factor model Explained 64.1% of the total variance **(-)**
TSK Heart Turkish [89]	Cardiac patients	85	Doubtful	The 4-factor model cumulatively accounted for 64.9% of the total variance **(-)**
TSK-AV-11 Arabic [90]	CLBP patients	101	Adequate	The 2-factor model cumulatively accounted for 46.54% of the total variance **(-)**

ACLR= Anterior Cruciate Ligament Reconstruction; CFI= Comparative Fit Index; MFR=Mean Fit Residuals; CLBP= Chronic Low Back Pain; CRPS= Complex Regional Pain Syndrome; O= Orthopedic sample; OA= Osteoarthritis; PC= Pain clinic sample; RMSEA= Root Mean Square Error of Approximation; SRMR= Standardized Root Mean Residuals; TKA= Total Knee Arthroplasty.

Rating: + =sufficient; − = insufficient; ? = indeterminate

Table 4. Construct Validity of the cross-cultural adaptations of the TSK-11

Construct/Instrument	Correlation level	CC range	Results (Rating)
Catastrophizing PCS [6,8,15,16,19–21]	Weak [92] Moderate [92,118] Strong [6,16,19–21]	0.27/0.751	Not in line 1 hypo (1-) In line 7 hypo (7+)
Fear Avoidance Beliefs FABQ [23,92] FABQ-PA [17]	Weak [92] Moderate [17] Strong [23]	0.23/0.719	Not in line 1 hypo (1-) In line 2 hypo (2+)
Anxiety HADS-A [92,117]	Weak [92] Moderate [117]	0.29/0.30	Not in line 1 hypo (1-) In line 1 hypo (1+)
Confidence in movement VAS [22]	Strong [22]	0.60	In line 1 hypo (1+)
Pain-related acceptance CPAQ [115]	Moderate [115]	-0.46	In line 1 hypo (1+)
Pain intensity NRS [2,6,8,9,13,15,16,20,21] VAS [12,22] MPI-PS [115] CPG [117]	Weak [2,4,8,12,15,20] Moderate [6,7,9,13,21] Strong[16,22]	0.03/0.69	In line 11 hypo (11+) Not in line 1 hypo (2-)
Physical function SF-12, PH [23] IKDC [73] IFI [92] PSFS [20] Stair climbing [115] Sit-Stand [115]	Weak[4,8,20] Moderate[23,73]	-0.11/-0.339	In line 6 hypo (6+)
Depression HADS-D [92,117] BDI [13,20] BDI-II [115]	Weak[17,117] Moderate[4,8,20]	0.206/0.38	In line 5 hypo (5+)
Physical activity GAS [44]	Moderate[44]	-0.378	In line 1 hypo (1+)
Disability ODI [13,22] MPI-INT [115] CPG [117] ADL staircase [44] DASH functioning [21] DASH sports/hobby [12] DASH work [21]	Weak[7,9,12,13] Moderate[4,7,12] Strong[22]	0.18/0.72	In line 7 hypo (7+) Not in line 1 hypo (1-)
Fatigue FACT-F [87]	Moderate[87]	0.3	In line 1 hypo (1+)
Health Status EQ-5D [21] EORTC-QLQC30 [87]	Moderate[87] Strong[21]	-0.333/-0.570	Not in line 1 hypo (1-) In line 1 hypo (1+)
Quality of life NHP [89]	Moderate[89]	0.374	Not in line 1 hypo (1-)
Summary result * (overall rating)	**Total sample size n= 4909**		**| 43+ and 6- | Overall: +**

*Considering the results of studies with good methodological quality

ADL=Activities of Daily Living; BDI=Beck Depression Inventory; BDI-II=Beck Depression Inventory-2nd Edition; CC=Correlation Coefficient; CPAQ=Chronic Pain Acceptance Questionnaire; CPG=Chronic Pain Grade Questionnaire; DASH=Disability of the Arm, Shoulder and Hand; EQ-5D=EuroQol 5 Dimension; EORTC-QLQC30=EORTC Core Quality of Life Questionnaire; FABQ=Fear Avoidance Beliefs Questionnaire; FABQ-PA=Fear Avoidance Beliefs Questionnaire, Physical Activity subscale; FACT-F=Functional Assessment of Cancer Therapy-Fatigue subscale; GAS=Grimby's activity scale; HADS-A=Hospital Anxiety and Depression Scale, Anxiety subscale; HADS-D=Hospital Anxiety and Depression Scale, Depression subscale IFI=Impairment and Functioning Inventory; IKDC= International Knee Documentation Committee subjective form; MPI=Multidimensional Pain Inventory; MPI-INT= Multidimensional Pain Inventory-Interference; MPI-PS=Multidimensional Pain Inventory-Pain Severity; NHP=Nottingham Health Profile; NRS=Numerical Rating Scale; ODI=Oswestry Disability Index; PASS-DE=Pain Anxiety Symptom Scale; PCS=Pain Catastrophizing Scale; PF=Physical Function subscale of the 12-Item Short Form Health Survey; PSFS=Patient-Specific Functional Scale; SES=Self-Evaluation Scale; SF-12, PH=Short Form Health Survey 12 items, Physical Health subscale; VAS=Visual Analogue Scale.

Table 5. Internal Consistency of the cross-cultural adaptations of the TSK-11

Author(s), Language	Population	n	Meth. Qual	Result (rating)
1-Factor Model				
Larsson et al., Swedish [44]	Older people with chronic pain	433	Very good	TSK-11 α=0.87 **(+)**
Frazad et al., Persian [91]	Patients suffering from CRPS	142	Very good	TSK-11 α=0.93 **(+)**
2-Factors Model				
AL-Shudifat et al., Arabic [90]	CLBP patients	101	Very good	TSK-AV-AA, α =0.74 **(+)** TSK-AV-SF, α =0.68 **(-)**
Acar et al., Turkish [89]	Cardiac patients	85	Inadequate	α =0.759 **(?)**
Archer et al., English [112]	Spinal surgery	137	Inadequate	α =0.73 **(?)**
Gómez-Pérez et al., Spanish [92]	Chronic pain sample	125	Very good	TSK- AA α =0.79 **(+)** TSK-H α =0.70 **(+)**
Gómez-Pérez et al., Spanish [92]	Acute pain sample	86	Very good	TSK- AA α =0.79 **(+)** TSK-H α =0.70 **(+)**
Larsson et al., Swedish [44]	Older people with chronic pain	433	Very good	TSK-SF α=0.74 **(+)** TSK-AA α=0.85 **(+)**
Roelofs et al.,	Upper extremity disorders	1109	Very good	TSK-AA α =0.67 **(-)** TSK-SF α =0.76 **(+)**
Dutch [120]	CLBP	225		TSK-SF α =0.74 **(+)** TSK-AA α =0.68 **(-)**
	Fibromyalgia	391		TSK-SF α =0.70 **(+)** TSK-AA α =0.64 **(-)**
	Osteoarthritis	254		TSK-SF α =0.67 **(-)** TSK-AA α =0.67 **(-)**
English	Musculoskeletal pain	335		TSK-AA=0.67 **(-)** TSK-SF=0.68 **(-)**
French		175		TSK-AA=0.68 **(-)** TSK-SF=0.68 **(-)**
Swedish		336		TSK-AA=0.73 **(+)** TSK-SF=0.71 **(+)**
Rusu et al., German [17]	LBP	191	Very good	TSK-SF α =0.64 **(-)** TSK-SF α =0.63 **(-)**
Tkachuk et al., English [115]	Heterogeneous chronic pain	276	Very good	TSK-SF, α=0.71 **(+)** TSK-AA, α=0.75 **(+)**
Velthuis et al., Dutch [87]	Cancer survivors	658	Very good	TSK-F-SF, α=0.68 **(-)** TSK-F-AA, α=0.62 **(-)**
Wong et al., Cantonese [117]	Chronic pain	216	Very good	TSK-SF, α=0.56 **(-)** TSK-AA, α=0.60 **(-)**
Youngcharoen et al., Thai [118]	Older people with knee osteoarthritis	200	Very good	TSK-SF, α=0.61 **(-)** TSK-AA, α=0.69 **(-)**

Farzad et al., Persian [91]	Patients suffering from CRPS	142	Very good	TSK-FA, α=0.90 **(+)** TSK-MH, α=0.90 **(+)**
3-Factors Model				
George et al., English [73]	ACL reconstruction (early phase)	105	Very good	TSK-SOMFE α =0.78 **(+)** TSK-FOPE α =0.78 **(+)** TSK-FOIE α =0.76 **(+)**
Cai et al., Chinese [23]	Patients following TKA	254	Very good	TSK-AA α =0.73 **(+)** TSK-SF α =0.83 **(+)** TSK-AB α =0.72 **(+)**
Cross-cultural adaptations without evidence for structural validity				
Kikuchi et al., Japanese [93]	Whiplash/ LBP	956	Doubtful	TSK-J11 α =0.92 **(?)**
Santo Salvador et al., Brazilian Portuguese [64]	Fibromyalgia	130	Doubtful	TSK-11 α = 0.77 **(?)**
Saltpute et al., Marathi [35]	CLBP	100	Doubtful	TSK-11 α = 0.85 **(?)**

ACL= Anterior Cruciate Ligament; CLBP= Chronic Low Back Pain; LBP= Low Back Pain; TKA= Total Knee Arthroplasty; TSK-AA- Activity Avoidance subscale; TSK-AB= Avoidance Beliefs subscale; TSK-AV-AA= Tampa Scale Arabic Version, Activity Avoidance subscale; TSK-AV-SF= Tampa Scale Arabic Version, Somatic Focus subscale; TSK-FOIE = TSK subscale Fear of Injury in Early phase; TSK-FOPE = TSK subscale Fear Of Pain in Early phase; TSK-H= Harm subscale; TSK-J11- Tampa Scale for Kinesiophobia Japanese version, 11 item; TSK-SF= Somatic Focus subscale; TSK-SOMFE = TSK subscale somatic focus of Early phase after Anterior Cruciate Ligament Reconstruction.

Rating: += sufficient; − = insufficient; ? = indeterminate.

Table 6. Test-retest reliability of the cross-cultural adaptations of the TSK-11

TSK-11 version	n	Meth qual	Result (rating)
TSK Heart Turkish [89]	85	Doubtful	$ICC_{2,1}$=0.564 (-)
Marathi [35]	100	Doubtful	$ICC_{2,1}$ = 0.93 (+)
Persian [91]	142	Doubtful	$ICC_{2,1}$ = 0.93 (+)
Chinese [23]	20	Adequate	$ICC_{2,1}$=0.798 (+)
Danish [113]	77	Very good	$ICC_{2,1}$ =0.87 (+)
English [61]	18	Adequate	$ICC_{2,1}$= 0.81 (+)
Swedish [44]	264	Adequate	$ICC_{2,1}$= 0.747 (+) Weighted k 0.37 to 0.58 (-)
Brazilian Portuguese [64]	54	Adequate	$ICC_{2,1}$ =0.85 (+)
summary result * (overall rating)	**433**		**$ICC_{2,1}$ =0.747/0.87 (+)**

*Considering the results of studies with good methodological quality.

ICC= Intraclass Correlation Coefficient.

Rating: + = sufficient; − = insufficient; ? = indeterminate.

Table 7. Measurement Error of the cross-cultural adaptations of the TSK-11

TSK-11 Version	n	Meth qual	Result (rating)
Persian [91]	142	Doubtful	SEM= 4.3; SDC=11.7 **(-)**
Danish [113]	77	Very good	SEM=2.10; SDC=5.82 **(-)**
English [61]	18	Adequate	SEM=2.41; SDC= 5.6 **(-)**
Brazilian Portuguese [64]	54	Adequate	SEM =2.65; SDC= 6.16 **(-)**
summary result * (overall rating)	**149**		SEM= 2.10 - 2.65; SDC= 5.6 - 6.16 **(-)**

*Considering the results of studies with good methodological quality.

SDM= Small Detectable Change; SEM= Standard Error of Measurement.

Rating: +=sufficient; − = insufficient; ? = indeterminate.

Table 8. Summary of findings

Psychometric property	Nr. participants (studies)	Summarized results	Overall rating	Quality of evidence (GRADE)
Hypotheses testing for Construct validity	4909 (14)	43 out of 49 hypotheses confirmed	Sufficient	**Moderate**: as there were some inconsistencies.
Criterion validity	130 (1)	TSK-17/TSK-11, r=0.84	Sufficient	**High**: results from 1 "very good" study with adequate sample size.
Test-retest reliability	433 (5)	ICC range 0.747/0.87 Weighted k 0.37 to 0.58	Sufficient	**Moderate**: as there were some inconsistencies.
Measurement error	149 (3)	SDC range 5.6/6.16	Insufficient	**High**: results from "adequate" and "very good" studies; consistent results; adequate sample size
Responsiveness	63 (1)	AUC =0.79	Sufficient	**Low**: result from 1 "adequate" study; sample size<100

AUC= Area Under the Curve; ICC= Intraclass Correlation Coefficient; SDC= Smal Detectable Change.

Rating: sufficient (+); insufficient (–); inconsistent (±); indeterminate (?)