Musculoskeletal
Review

OPEN

PAIN REPORTS

Measurement properties of the Pain Self-Efficacy Questionnaire in populations with musculoskeletal disorders: a systematic review

Marc-Olivier Dubé[a,b,c,*], Pierre Langevin[a,b,c], Jean-Sébastien Roy[a,b]

Abstract

A higher level of pain self-efficacy has been suggested as a predictor of a better outcome in patients with musculoskeletal disorders. The Pain Self-Efficacy Questionnaire (PSEQ) is one of the most frequently used patient-reported outcome measures for pain self-efficacy. The purpose of this study was to conduct a systematic review that would identify, appraise, and synthetize the psychometric properties of the PSEQ. Embase, MEDLINE, and CINAHL databases were searched for publications reporting on psychometric properties of the PSEQ in populations with musculoskeletal disorders. After applying selection criteria on identified citations, 28 studies (9853 participants) were included. The methodological quality as measured with the COSMIN risk of bias tool varied from *adequate* to *very good* for most measurement properties. The results showed a weighted mean intraclass correlation coefficient of 0.86 (range: 0.75–0.93) for test–retest reliability for the original 10-item PSEQ and the minimal detectable change at 95% confidence interval was 11.52 out of 60 points. Effect size and standardized response mean values were 0.53 and 0.63, respectively, whereas the minimal clinically important difference ranged from 5.5 to 8.5 in patients with chronic low back pain. Internal consistency (Cronbach alpha) ranged from 0.79 to 0.95. The results also showed that the PSEQ has low to moderate correlations with measures of quality of life, disability, pain, pain interference, anxiety, depression, and catastrophizing. Finally, the PSEQ has been adapted and validated in 14 languages. Overall, the results demonstrate that the PSEQ has excellent validity, reliability, and responsiveness. Further high-quality studies are needed to determine responsiveness in populations other than chronic low back pain.

Keywords: Pain self-efficacy, Pain, Musculoskeletal, PSEQ, Measurement

1. Introduction

One in 3 people live with chronic musculoskeletal (MSK) disorders.[12] Musculoskeletal disorders can lead to decreased participation and quality of life as well as increased all-cause mortality.[11] Many factors have been studied to better understand the persistence of MSK pain over time; pain self-efficacy being one of them.[28] Self-efficacy was originally defined as an individual's confidence or belief in their capacity to achieve goals or perform activities.[4] More specifically in the MSK field, pain self-efficacy is defined as the confidence that one has to perform their activities and achieve their goals despite the presence of symptoms or pain.[35] A higher level of pain self-efficacy has been suggested as a predictor of a better prognosis in those with MSK pain[28] and has been associated with less disability, pain, fatigue, and emotional distress.[27]

Several scales have been developed to measure pain-related self-efficacy, namely, the Arthritis Self-Efficacy Scale (ASES; 20 items), the Chronic Disease Self-Efficacy Scale (33 items), the Pain Self-Efficacy Questionnaire (PSEQ; 10 items), the Chronic Pain Self-Efficacy Scale (22 items), and the Self-Efficacy Scale (8 items).[32,47] The number of items in these scales varies from 8 to 33 items, leading to a variable response burden (completion times).[47] Among all self-efficacy questionnaires, the PSEQ is the most frequently used in clinical setting for MSK disorders.[15,36,47] According to a Delphi study, it is the preferred clinician tool to assess pain-related self-efficacy.[39] It is a short self-administered questionnaire that has been extensively studied with various MSK disorders[10,20,26] and translated in many languages.[2,25,37,38]

Sponsorships or competing interests that may be relevant to content are disclosed at the end of this article.

[a] *Department of Rehabilitation, Faculty of Medicine, Université Laval, Pavillon Ferdinand-Vandry, Quebec City, QC, Canada,* [b] *Centre for Interdisciplinary Research in Rehabilitation and Social Integration, Quebec Rehabilitation Institute (CIRRIS), Quebec City, QC, Canada,* [c] *Physio Interactive, Québec, QC, Canada*

**Corresponding author. Address: Centre for Interdisciplinary Research in Rehabilitation and Social Integration, Quebec Rehabilitation Institute, 525, Boulevard Wilfrid Hamel, Quebec City, QC G1M 2S8, Canada. Tel.: 418-907-5858. E-mail address: marc-olivier.dube.1@ulaval.ca (M.-O. Dubé).*

PR9 6 (2021) e972

http://dx.doi.org/10.1097/PR9.0000000000000972

Only 1 systematic review has been published on the psychometric properties of assessment tools for pain self-efficacy.[47] This systematic review included all available pain self-efficacy scales for a population with low back pain (LBP). The main findings were that the PSEQ and Chronic Pain Self-Efficacy Scale were the most commonly studied scales and that internal consistency was reported in all original included studies but other measurement properties such as reliability, validity, and responsiveness were poorly reported. In that systematic review, methodology for data extraction was not described and authors did not perform a quality assessment of included studies to evaluate the risk of bias of included studies. This lack of methodological evaluation and description in the review process limits the ability to conclude on the psychometric properties of the evaluated outcome measures.

Currently, no evidence synthesis or systematic review on the measurement properties of the PSEQ, the most widely used pain self-efficacy questionnaire, has been published for MSK disorders. To use this questionnaire in research and clinical contexts, it is important to review and assess the quality of evidence evaluating its measurement properties. The purpose of this study was to conduct a systematic review that would identify, appraise, and synthetize the psychometric properties of the PSEQ.

2. Methods

2.1. Description of the questionnaire

The PSEQ is a 10-item questionnaire, originally developed in English, aimed to assess the confidence of people with persistent pain to achieve different activities despite their pain.[35] Each item is rated by selecting a number on a 7-point numerical scale (scores from 0 to 6), where 0 means "not at all confident" and 6 means "completely confident." It can be used with different populations of people with persistent pain.[19,20,35] Items cover functions such as work, social activities, household chores, and coping with pain without medication. A total score is calculated by summing the scores for each of the 10 items, yielding a total score ranging from 0 to 60. Higher scores reflect stronger self-efficacy beliefs.[35] Shorter versions (PSEQ-2 and PSEQ-4) including 2 and 4 items, respectively, have also been developed.[1,8,10,31] Refer to **Table 1** for a list of the PSEQ Items.

Table 1
PSEQ-10 items.

Items
1. I can enjoy things, despite the pain.
2. I can do most of the household chores (tidying up, washing dishes, etc), despite the pain.
3. I can socialize with my friends or family members as often as I used to do, despite the pain.
4. I can cope with my pain in most situations.
5. I can do some form of work, despite the pain. ("work" includes housework, paid, and unpaid work).
6. I can still do many of the things I enjoy doing, such as hobbies or leisure activity, despite pain.
7. I can cope with my pain without medications.
8. I can still accomplish most of my goals in life, despite the pain.
9. I can live in a normal lifestyle, despite the pain.
10. I can gradually become more active, despite the pain.

PSEQ, Pain Self-Efficacy Questionnaire.

2.2. Literature search and study identification

A database search using Embase, MEDLINE, and CINAHL was performed January 29, 2021, by a librarian from *Université Laval*. The following keywords were used to search databases for eligible studies: Pain Self-Efficacy Questionnaire, PSEQ, PSEQ-10, PSEQ-2, PSEQ-4, Reliability, Reproducibility, Validity, Validation, Responsiveness, Translation, Minimal detectable change, MDC, minimal clinical important difference, MCID, Factor analysis, Rasch, Internal coherence, Internal consistency, Psychometric properties, and Measurement properties. Keywords related to the population of interest were not included for the databases search because the PSEQ was originally developed to be used in a population with chronic pain.[35] Complete search strategy, MeSH, and keywords are available in Appendix 1. Hand searches of retrieved study reference lists were also conducted.

2.3. Study selection

Covidence software was used for the selection phase of the study. The titles and abstracts of each article were first independently reviewed by 2 of the authors, and an article was accepted for a full review if it met the following inclusion criteria: (1) original studies that aim to report at least one of the validity and/or reliability indices of the PSEQ-10 or one of its shortened versions, (2) written in English or French, and (3) included adult population with any MSK disorder. Articles were excluded if any of the following were retrieved: (1) studies reporting the psychometric properties of the PSEQ in nonmusculoskeletal disorders and (2) systematic reviews, narrative reviews, scoping reviews, meta-analyses, study protocols, or conference proceedings. A consensus between 2 authors was needed to include an article.

2.4. Methodological quality appraisal scoring

A pair of raters independently reviewed each article that met the inclusion criteria. The COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) quality appraisal tool was used to appraise the quality of measurement properties described in included studies.[33] The COSMIN quality appraisal tool recommends rating the quality of each measurement property reported in each included study. Measurement properties were rated, when applicable, according to the following categories: (1) patient-reported outcome measure development, (2) content validity, (3) structural validity, (4) internal consistency, (5) cross-cultural validity or measurement invariance, (6) reliability, (7) measurement error, (8) criterion validity, (9) construct validity, and (10) responsiveness. Each category contains a series of items related to the specific measurement properties. Each item was rated as "very good," "adequate," "doubtful," or "inadequate" by assessing the information reported in the study being evaluated.[33] The lowest score among all items was then used to rate the quality of each measurement property evaluated for every individual study.[44] All the authors first met for a calibration review, where they independently reviewed 3 articles and then discussed each item to clarify the meaning and interpretation of critical appraisal items. Then, the pair of raters independently evaluated an assigned subset of articles. A consensus meeting was held to produce a consensus statement on each rated property of every included study. The Gwet AC2 coefficient (quadratic weights) was used to evaluate the

preconsensus agreement on the COSMIN scores between the 2 reviewers.[49]

2.5. Data extraction and analysis

A standardized data extraction form used in previous systematic reviews[9,40] was used to extract data from included studies. Information on the patient population of each study was extracted, including the number of participants, medical conditions, age, sex, and intervention provided when applicable. The following psychometric properties were also extracted: reliability (test–retest intraclass correlation coefficient [ICC], SEM, minimal detectable change [MDC], and internal consistency), validity (content, construct, factorial, known-group, and floor/ceiling effects), responsiveness (effect size [ES], standardized response mean [SRM], and minimal clinically important difference [MCID]), time to administer, and cross-cultural adaptation. One reviewer extracted the data, and a second reviewer verified the transcription to reduce risk of error. Overall relative and absolute reliability (ICC, SEM, and MDC) and responsiveness (ES and SRM) were pooled and determined by calculating a weighted average over all studies (weighted by sample size). Intraclass correlation coefficient were considered excellent if higher than 0.81, good from 0.61 to 0.80, moderate from 0.41 to 0.60, fair from 0.21 to 0.40 and poor if less than 0.20.[41] Effect size and SRM were considered large if higher than 0.8, moderate between 0.50 and 0.80, and small between 0.20 and 0.50.[18] Pearson or Spearman correlations were categorized as high if higher than 0.70, moderate between 0.50 and 0.70, low between 0.26 and 0.49, and very low between 0.01 and 0.25.[34] For construct validity, the following 3 hypotheses were formulated, and data were extracted to confirm or infirm those hypotheses. First, correlation would be positive and moderate to high ($r > 0.50$) between PSEQ-10 and shortened versions of the PSEQ. Second, positive and moderate correlations ($0.50 < r < 0.70$) between self-efficacy and quality of life were expected. Third, negative low to moderate correlations ($-0.26 < r < -0.70$) between self-efficacy and pain, pain interference, anxiety, depression, catastrophizing, and disability were expected. The second and third hypotheses are in line with the original development study of the PSEQ-10.[35]

3. Results

3.1. Literature search and study selection

The PRISMA study selection flowchart is presented in **Figure 1**. The literature search revealed a total of 347 citations. After removal of duplicates, the inclusion and exclusion criteria were applied by screening titles and abstracts of 111 studies. Thirty-four full-text studies were then screened for eligibility. Of these 34 studies, 5 were excluded because of their targeted outcome measure (2 studies), patient population (2 studies), or study design (1 study). Another study was excluded because it was a conference proceeding (abstract only). Therefore, 28 studies[1–3,7,8,10,13–17,19–23,25,26,29,31,35–38,45,46,48,50] were included, in which a total of 9853 participants were involved.

3.2. Characteristics of included studies

The detailed characteristics of included studies can be found in **Table 2**. While 22 included studies evaluated several psychometric properties of the PSEQ-10, 6 studies[1,10,16,31,36,50] evaluated the psychometric properties of either one or both of its short forms, namely, the PSEQ-2 (evaluated in the 6 studies) and the PSEQ-4 (evaluated in 3 of these 6 studies). Populations evaluated in included studies comprised chronic LBP,[14,17,20,21,29,50] fibromyalgia,[37] neck pain,[15] upper limb MSK pain,[8,10,23] knee osteoarthritis,[26] and any chronic MSK disorder.[1–3,8,10,13,16,19,22,25,31,35,36,38,45,46,48]

3.3. COSMIN risk of bias assessment of included studies

Table 3 presents the quality of properties evaluated by the COSMIN risk of bias tool. The risk of bias assessment demonstrated a good overall preconsensus agreement between the 2 reviewers (Gwet AC2 coefficient = 0.62; $P < 0.00001$; observed agreement = 84.96%; agreement by chance = 60.39%). The overall quality of the PSEQ development study[35] and the content validity were rated as doubtful mainly because the methods used to assess the comprehensibility and comprehensiveness of the PSEQ were poorly described. Overall, structural validity, internal consistency, cross-cultural validity, construct validity, and responsiveness designs were of *very good* quality in most of the included studies (75 of 79 [94.9%] measurement properties rated). For test–retest reliability and measurement error, failure to document (1) if patients were stable between 2 measurement time points, (2) the time interval between the 2 tests, or (3) similarity of the 2 measurements conditions resulted in 3 of 13 studies (23.1%) obtaining an *inadequate* score, 4 of 13 studies (30.8%) obtaining a *doubtful* score, and 4 of 13 studies (30.8%) obtaining an *adequate* score.

3.4. Validity

3.4.1. Content

The COSMIN quality of the PSEQ development study (Box 1) and content validity (Box 2) were judged as *doubtful*, mainly because there was not enough description of the qualitative process used to develop the PSEQ, neither from a patient nor from a clinician perspective. Nicholas[35] developed the original English version of the PSEQ, namely, the PSEQ-10 (development study published in 2007). The 10 items were selected from other questionnaires and from the authors' clinical experience. The content validity was then tested on 85 patients using a qualitative testing analysis of a patient's comprehensibility of the questionnaire. A confirmatory factor analysis then showed that the questionnaire was unidimensional with a corrected item-total correlation that varied from 0.67 to 0.84 when used in a population with chronic LBP.[35] The 1-dimensionality of the PSEQ-10 was confirmed in 16 other studies in populations with chronic MSK pain, neck pain, and upper limb MSK pain.[2,3,13–15,17,20,23,25,26,37,38,45,46,48,50] Regarding the COSMIN quality score for structural validity, 12 studies were rated as *very good*,[2,3,13–15,17,20,26,37,38,45,48,50] 2 as *adequate,*[23,25] and 1 as *inadequate*.[46] All items of the questionnaire performed well overall on factorial analysis except for item 7 of the questionnaire, "I can cope with my pain without medication," which performed constantly worse on item fit analysis than other items but was deemed clinically relevant by authors.[17,19,20,35,38] Two versions of the PSEQ-2 were developed. One version consisted of items 8 and 9[8] and another version consisted of items 5 and 9 (refer to **Table 1** for specific item wording).[36] The structural validity of both versions of the PSEQ-2 was confirmed (1 dimension) in 3 studies with a moderate to high corrected item- total correlation.[1,7,31] Finally, the PSEQ-4 consisted of items 4, 6, 8, and 9, and its structural validity was also confirmed (1 dimension).[31]

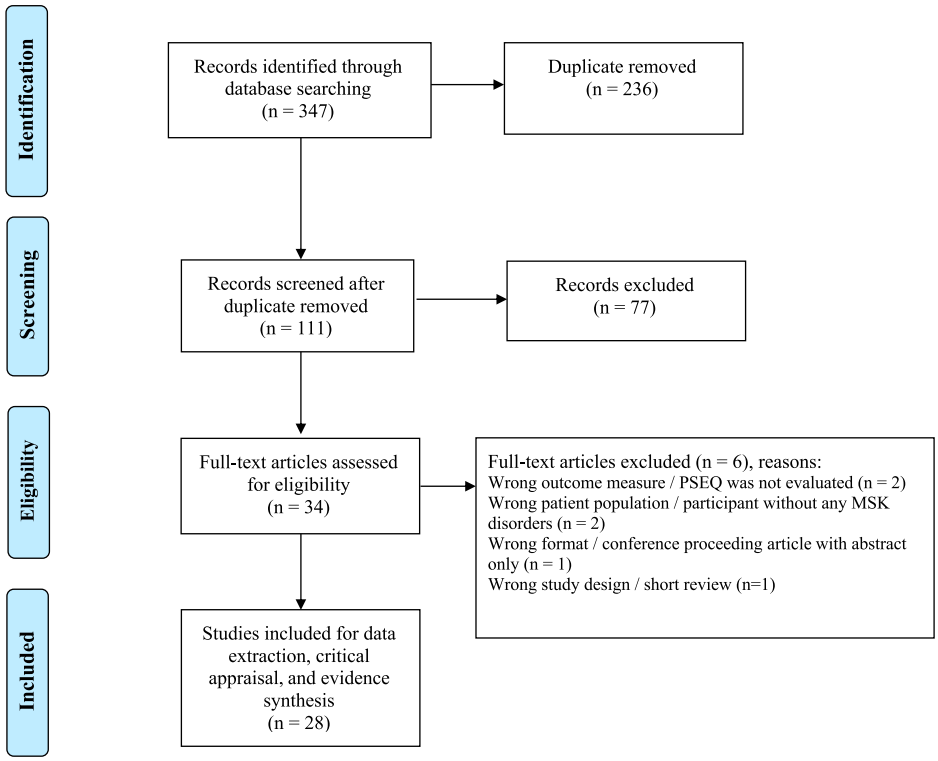

Figure 1. PRISMA flow diagram.

3.5. Floor and ceiling effects

Floor and ceiling effects were examined in several language versions of the PSEQ-10. The English version showed no floor effect and a ceiling effect that ranged from 0% to 20%.[10,20,23] The Italian,[17] Danish,[37] and Yoruba[21] versions of the PSEQ-10 did not show ceiling or floor effects. The Amharic version did not show a floor effect but showed a minimal ceiling effect of 4.8%.[14] The PSEQ-2 (version with items 8 and 9) showed a minimal floor effect of 1.6% but a ceiling effect of 41.7% which means that PSEQ-2 does not allow to discriminate among those who have a higher level of pain self-efficacy.[10]

3.6. Construct

Table 4 summarizes findings for construct validity. Convergent and divergent validity of the PSEQ were assessed by calculating correlations (Pearson (*r*) or Spearman (Rho)) between PSEQ total scores and total scores of several other patient-reported outcome measures in 19 included studies. COSMIN risk of bias for construct validity was rated as *very good* for all 19 included studies that evaluated this measurement property. Quality of life evaluated by the SF-12 or SF-36 questionnaires showed low to moderate positive statistically significant correlations with the PSEQ-10 for chronic LBP or MSK pain (7 studies; 1277 participants; correlation range: 0.38–0.57).[2,3,14,22,25,48,50] For divergent validity, there was a low to moderate negative statistically significant correlation between function or disability *(higher score = lower level of function/disability)* (10 studies; 2630 participants; correlation range: −0.71 to −0.29),[3,7,10,15,17,31,36,38,48,50] pain interference (5 studies; 1526 participants; correlation range: −0.67 to −0.40),[2,36,38,45,50] anxiety (5 studies; 805 participants; correlation range: −0.63 to −0.32),[1,22,25,45,50] or depression (9 studies; 2276 participants; correlation range: −0.68 to −0.32)[1,3,17,22,25,31,36,45,50] and the PSEQ-10 or shortened forms of the PSEQ for chronic LBP, upper limb pain, or MSK pain. Correlations between the PSEQ and catastrophizing variables mainly measured by the Pain Catastrophizing Scale varied from low to high and were statistically significant among studies that included various populations with chronic MSK pain (10 studies, 3495 participants, correlation range: −0.74 to −0.36).[1,2,13,15,17,25,31,35,36,45] Correlation levels between pain intensity and PSEQ varied from very low to low (12 studies; 2087 participants, correlation range: −0.50 to −0.05)[1,7,10,15,17,21,22,25,31,38,45,48] with the Chinese[25] and Yoruba[21] versions that did not meet statistical significance. Thus, our original hypotheses formulated in the methodology section were confirmed. Finally, the hypothesis that the short version was correlated with the PSEQ-10 was also confirmed with a high statistically significant positive correlation between the PSEQ-10 and the PSEQ-2 for chronic MSK pain and for upper limb MSK pain (4 studies; 1318 participants; correlation range: 0.76–0.94).[1,8,36,50]

3.7. Reliability

3.7.1. Internal consistency

Internal consistency was evaluated in most studies using Cronbach alpha. All studies were given a *very good* score on the COSMIN quality assessment tool for internal consistency. Twenty studies (5295 participants)[1–3,8,10,13,14,17,21–23,25,31,35–38,45,48,50] calculated Cronbach alpha for the PSEQ-10, with values ranging from 0.79 to 0.95. Six studies (1354 participants)[1,8,10,31,36,50] calculated Cronbach alpha for the PSEQ-2, with values ranging from 0.76 to 0.91. Two studies (430 participants)[1,31] calculated Cronbach alpha for the PSEQ-4, with values ranging from 0.87 to 0.90. **Table 5** presents ranges of Cronbach alphas. A score of 0.95 or higher may indicate a level of redundancy within the items of the questionnaire which was observed in 3 studies.[10,45,50]

Table 2

Characteristics of included studies.

Study ID	Questionnaire	Population				Properties evaluated
		Disorders	Total, n	Mean age (SD)	Gender (% F)	
Adachi et al., 2014[2]	PSEQ-10/Japanese	Chronic MSK pain	176	64.33 (15.12)	54.5%	Test–retest reliability Internal consistency Factorial validity Construct validity Predictive validity Cross-cultural/language translation
Adachi et al., 2019[1]	PSEQ-10, PSEQ-2, and PSEQ-4/Japanese	Chronic MSK pain	150	54.75 (15.73)	63%	Internal consistency Factorial validity Construct validity
Asghari et al., 2009[3]	PSEQ-10/Persian	Chronic MSK pain	517	40.6 (14.1)	39.5%	Test–retest reliability Internal consistency Factorial validity Construct validity Cross-cultural/language translation
Bot et al., 2013[7]	PSEQ-10/Web based	Upper limb MSK pain	99	49 (18)	46%	Construct validity
Bot et al., 2014[8]	PSEQ-10 and PSEQ-2	Upper limb MSK pain	316	46 (16)	47%	Internal consistency Content validity Construct validity
Briet et al., 2014[10]	PSEQ-10 and PSEQ-2	Upper limb MSK pain	249	47 (16)	54%	Test–retest reliability Measurement error Internal consistency Floor/ceiling effect Construct validity
Castarlenas et al., 2020[13]	PSEQ-10/Catalan	Chronic MSK pain	227	17.87 (3.08)	70%	Internal consistency Factorial validity Construct validity Cross-cultural/language translation
Chala et al., 2021[14]	PSEQ-10/Amharic	Chronic low back pain	240	40.93 (13.5)	59.2%	Test–retest reliability Measurement error Internal consistency Floor/ceiling effect Factorial validity Construct validity Cross-cultural/language translation
Chiarotto et al., 2014[17]	PSEQ-10/Italian	Chronic low back pain	165	49.9 (12.4)	64.8%	Test–retest reliability Measurement error Internal consistency Floor/ceiling effect Factorial validity Construct validity Cross-cultural/language translation

(*continued on next page*)

Table 2 (*continued*)

Characteristics of included studies.

Study ID	Questionnaire	Population				Properties evaluated
		Disorders	Total, n	Mean age (SD)	Gender (% F)	
Chiarotto et al., 2016[16]	PSEQ-10, PSEQ-2, and PSEQ-4/Italian	Chronic low back pain	104	48.85 (12.23)	68%	Responsiveness
Chiarotto et al., 2018[15]	PSEQ-10/Italian	Subacute and chronic neck pain	161	44.55 (15.81)	65%	Factorial validity Internal consistency Construct validity Responsiveness
Costa et al., 2016[19]	PSEQ-10	Chronic MSK pain	1511	48.9 (16.1)	57%	Content validity Factorial validity
DiPietro et al., 2014[20]	PSEQ-10	Subacute and chronic low back pain	610	48.1 (15.1)	50.7%	Internal consistency Floor/ceiling effect Factorial validity Construct validity
Fatoye et al., 2021[21]	PSEQ-10/Yoruba	Chronic low back pain	131	51.92 (16.1)	50.4%	Test–retest reliability Measurement error Internal consistency Floor/ceiling effect Construct validity Cross-cultural/language translation
Ferreira-Vallente et al., 2011[22]	PSEQ-10/Portuguese (European)	Chronic MSK pain	174	59.18 (16.11)	59.2%	Measurement error Internal consistency Construct validity Cross-cultural/language translation
Kortlever et al., 2015[23]	PSEQ-10	Upper limb MSK pain	134	52 *(Median);* 35–64 *(IQR)*	57%	Internal consistency Floor/ceiling effect Factorial validity Construct validity
Lim et al., 2007[25]	PSEQ-10/Chinese	Chronic MSK pain	120	40 (9.6)	58.3%	Test–retest reliability Internal consistency Factorial validity Construct validity Cross-cultural/language translation
Lincoln et al., 2017[26]	PSEQ-10	Knee osteoarthritis	192	67 (10)	53%	Content validity
Maughan et al., 2010[29]	PSEQ-10	Chronic low back pain	48	52	67%	Test–retest reliability Measurement error Responsiveness
McWilliams et al., 2015[31]	PSEQ-10 and PSEQ-4	Chronic MSK pain	280	47.9 (10.1)	66.1%	Internal consistency Construct validity
Nicholas et al., 2007 [35]	PSEQ-10	Chronic MSK pain	1554	41	NR	Test–retest reliability Internal consistency Factorial validity Content validity Construct validity

(*continued on next page*)

Table 2 (*continued*)

Characteristics of included studies.

Study ID	Questionnaire	Population				Properties evaluated
		Disorders	Total, n	Mean age (SD)	Gender (% F)	
Nicholas et al., 2015[36]	PSEQ-10; PSEQ-2	Chronic MSK pain	1558	48.8 (16.1)	51.5%	Test–retest reliability Internal consistency Construct validity Responsiveness
Rasmussen et al., 2015[37]	PSEQ-10/Danish	Fibromyalgia	99	46.5	94%	Test–retest reliability Internal consistency Floor/ceiling effect Content validity Factorial validity Construct validity Cross-cultural/language translation
Sarda et al., 2007[38]	PSEQ-10/Portuguese (Brazilian)	Chronic MSK pain	311	48.9 (14.06)	74%	Internal consistency Factorial validity Construct validity Cross-cultural/language translation
Tuck et al., 2020[45]	PSEQ-10/Mongolian	Chronic MSK pain	142	53.96 (15.50)	61%	Internal consistency Construct validity Cross-cultural/language translation
Ugwuanyi et al., 2020[46]	PSEQ-10/Nigerian	Chronic MSK pain	256	NR	NR	Test–retest reliability Internal consistency Factorial validity
Vong et al., 2009[48]	PSEQ-10/Chinese	Chronic MSK pain	120	41.9 (12.21)	67.5%	Internal consistency Factorial validity Construct validity
Yang et al., 2019[50]	PSEQ-10 and PSEQ-2/Chinese (Mainland)	Chronic low back pain	219	53.4 (11.5)	42.9%	Test–retest reliability Internal consistency Factorial validity Construct validity Cross-cultural/language translation

Table 3

COSMIN risk of bias of included studies.

Study	1a.	1b.	2.a	2.b	2.c	2.d	2.e	3.	4.	5.	6.	7.	8.	9a	9b	10a	10b	10c	10d
Nicholas et al., 2007[35]	D	D	—	—	D	—	—	—	V	—	I	—	—	V	—	—	—	—	—
Lim et al., 2007[25]	—	—	—	—	—	—	—	A	V	V	D	—	—	V	—	—	—	—	—
Sarda et al., 2007[38]	—	—	—	—	—	—	—	V	V	—	—	—	—	V	—	—	—	—	—
Asghari et al., 2009[3]	—	—	—	—	—	—	—	V	V	V	V	—	—	V	—	—	—	—	—
Vong et al., 2009[48]	—	—	—	—	—	—	—	V	V	—	—	—	—	V	—	—	—	—	—
Ferreira-Vallente et al., 2011[22]	—	—	—	—	—	—	—	—	V	V	—	—	—	V	—	—	—	—	—
Bot et al., 2013[7]	—	—	—	—	—	—	—	—	—	—	—	—	—	V	—	—	—	—	—
Bot et al., 2014[8]	—	—	—	—	—	—	—	—	V	—	—	—	—	—	—	—	—	—	—
Adachi et al., 2014[2]	—	—	—	—	—	—	—	—	V	V	D	—	V	V	—	—	—	—	—
Briet et al., 2014[10]	—	—	—	—	—	—	—	—	V	—	I	I	—	V	—	—	—	—	—
Chiarotto et al., 2014[17]	—	—	—	—	—	—	—	V	V	V	A	A	—	V	—	—	—	—	—
Di Pietro et al., 2014[20]	—	—	—	—	—	—	—	—	V	—	—	—	—	V	—	—	—	—	—
McWilliams et al., 2015[31]	—	—	—	—	—	—	—	—	V	—	—	—	—	V	—	—	—	—	V
Nicholas et al., 2015[36]	—	—	—	—	—	—	—	—	V	—	A	—	I	V	—	—	—	—	I
Kortlever et al., 2015[23]	—	—	—	—	—	—	—	A	V	—	—	—	—	V	—	—	—	—	—
Rasmussen et al., 2015[37]	—	—	—	—	—	—	—	V	V	V	V	—	—	—	—	—	—	—	—
Chiarotto et al., 2016[16]	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	V	V	—	V
Costa et al., 2016[19]	—	—	—	—	—	—	—	V	—	—	—	—	—	—	—	—	—	—	—
Lincoln et al., 2017[26]	—	—	—	—	—	—	—	V	—	—	—	—	—	—	—	—	—	—	—
Chiarotto et al., 2018[15]	—	—	—	—	—	—	—	V	V	—	—	—	—	V	—	—	V	V	V
Adachi et al., 2019[1]	—	—	—	—	—	—	—	—	V	—	—	—	I	V	—	—	—	—	—
Maughan et al., 2010[29]	—	—	—	—	—	—	—	—	—	—	D	D	—	—	—	—	—	—	V
Yang et al., 2019[50]	—	—	—	—	—	—	—	V	V	V	A	—	I	V	—	—	—	—	—
Castarlenas et al., 2020[13]	—	—	—	—	—	—	—	V	V	V	—	—	—	V	—	—	—	—	—
Tuck et al., 2020[45]	—	—	—	—	—	—	—	V	V	V	—	—	—	V	—	—	—	—	—
Uguwani et al., 2020[46]	—	—	—	—	—	—	—	I	V	—	I	—	—	—	—	—	—	—	—
Chala et al., 2021[14]	—	—	—	—	—	—	—	V	V	V	A	—	—	V	—	—	—	—	—
Fatoye et al., 2021[21]	—	—	—	—	—	—	—	—	V	V	D	D	—	V	—	—	—	—	—

3.8. Test–retest reliability

Test–retest reliability was evaluated using ICC in 11 studies (1016 participants).[2,3,10,14,17,21,25,29,36,37,50] COSMIN risk of bias of reliability was rated as *very good* for 2 included studies,[3,37] *adequate* for 4 studies,[14,17,36,50] *doubtful* for 4 studies,[2,21,25,29] and *inadequate* for 1 study.[10] **Table 5** presents weighted means and ranges of ICCs. Intraclass correlation coefficients for the PSEQ-10 ranged from 0.75 to 0.93 with an overall weighted mean of 0.86 (10 studies; 908 participants), which represents an excellent reliability overall. Intraclass correlation coefficients for the PSEQ-2 ranged from 0.79 to 0.88 with an overall weighted mean of 0.85 (3 studies; 422 participants), which represents an excellent reliability. The reliability of the PSEQ-4 has not been evaluated yet.

3.9. Absolute reliability

Four studies (381 participants)[10,17,21,29] evaluated the absolute reliability of the PSEQ-10 by calculating SEM or MDC. COSMIN risk of bias of measurement error was rated as *adequate* for 1 included study,[17] *doubtful* for 2 studies,[21,29] and *inadequate* for 1 study.[10] SEM, which represents the standard deviation of the PSEQ-10 score, ranged from 1.23 to 5.66 (weighted mean of 3.37 [3 studies; 356 participants).[10,17,21] MDC_{95}, which represents the estimate of the smallest amount of change that can be detected between 2 measurements representing a real difference between the same 2 measurements, ranged from 3.41 to 15.69 (weighted mean of 11.52 [3 studies; 273 participants]).[17,21,29]

3.10. Responsiveness

Responsiveness was evaluated in 4 studies[15,16,29,36] using ES, SRM, and MCID. COSMIN risk of bias for responsiveness was rated as *very good* for 3 included studies[15,16,29] and *inadequate* for 1 study.[36] Nicholas et al.[36] reported a large effect size (0.92) and SRM (0.94) for the PSEQ-2 in a population with chronic pain (140 participants) that had just participated in a cognitive intervention that aimed, in part, to modify self-efficacy. When analyzing results from those who improved (increase of 2 or more on the PSEQ-2 score), the ES and SRM increased to 1.54 and 2.22, respectively. In subjects who were classified as not improved (increase of less than 2 on the PSEQ-2 score), the ES and SRM decreased to 0.15 and 0.30, respectively. Chiarotto et al.[16] evaluated responsiveness in a sample (n = 104) of patients with chronic LBP, regardless of whether they improved or not, who had received a multimodal physical therapy intervention partially designed to modify self-efficacy. They

Table 4

Summary of findings for construct validity.

Convergent validity with:	Correlation (high, moderate, low, and very low)	Person (r) or Spearman (Rho) correlation coefficient, significance, PSEQ version (other than PSEQ-10) when applicable, and population studied.
Original vs short version		
PSEQ-10 vs PSEQ-2	High	$r = 0.94$, $P < 0.001$ for chronic MSK pain[50]
	High	$r = 0.93$, $P < 0.001$ for chronic MSK pain[36]
	High	$r = 0.92$, $P < 0.001$ for chronic MSK pain[1]
	High	$r = 0.76$, $P < 0.001$ for chronic upper limb MSK pain[8]
Quality of life		
SF-12 physical	Moderate	$r = 0.52$, $P \leq 0.001$ for chronic MSK pain[3]
SF-12 physical	Moderate	$r = 0.51$, $P < 0.01$ for chronic MSK pain[22]
SF-12 mental	Low	$r = 0.46$, $P < 0.01$ for chronic MSK pain[22]
SF-12 mental	Low	$r = 0.42$, $P \leq 0.001$ for chronic MSK pain[3]
SF-36 physical	Low	$r = 0.43$, $P < 0.001$ for chronic LBP[50]
SF-36 physical	Low	$r = 0.41$, $P < 0.001$ for chronic MSK pain[2]
SF-36 physical	Low	$r = 0.40$, $P < 0.001$ for chronic MSK pain[25]
SF-36 physical	Low	Rho = 0.38, $P < 0.01$ for chronic and acute LBP[14]
SF-36 physical	Low	$r = 0.38$, $P < 0.001$ for PSEQ-2 on chronic LBP[50]
SF-36 mental	Moderate	$r = 0.57$, $P < 0.001$ for chronic LBP[50]
SF-36 mental	Low	$r = 0.47$, $P < 0.001$ for PSEQ-2 on chronic LBP[50]
SF-36 mental	Low	$r = 0.46$, $P < 0.001$ for chronic MSK pain[48]
SF-36 mental	Low	$r = 0.40$, $P < 0.001$ for chronic MSK pain[2]
Pain intensity		
VAS	Moderate	$r = -0.50$, $P < 0.001$ for chronic upper limb MSK pain[7]
VAS	Low	Rho = −0.46, $P < 0.01$ for chronic MSK pain[45]
VAS	Low	$r = -0.45$, $P < 0.001$ for chronic upper limb MSK pain[10]
VAS	Low	$r = -0.40$, $P < 0.001$ for chronic MSK pain[17]
VAS	Low	$r = -0.38$, $P < 0.001$ for PSEQ-2 on chronic upper limb MSK pain[10]
VAS	Low	$r = -0.36$, $P < 0.001$ for chronic MSK pain[1]
VAS	Low	$r = -0.36$, $P < 0.001$ for chronic MSK pain[49]
VAS	Low	$r = -0.33$, $P < 0.001$ for PSEQ-2 on chronic upper limb MSK pain[7]
VAS	Low	$r = -0.29$, $P < 0.001$ for chronic MSK pain[31]
VAS	Low	$r = -0.28$, $P < 0.01$ for chronic MSK pain[22]
VAS	Very low	$r = -0.25$, $P \leq 0.001$ for chronic MSK pain[38]
VAS	Very low	$r = -0.24$, $P < 0.001$ for PSEQ-4 on chronic MSK pain[31]
VAS	Very low	$r = -0.18$, $P < 0.001$ for PSEQ-2 on chronic MSK pain[31]
VAS	Very low	$r = -0.121$, $P = 0.189$ for chronic MSK pain[25]
VAS	Very low	$r = -0.05$, $P = 0.59$ for chronic MSK pain[21]
NRS	Very low	Rho = −0.14 $P < 0.05$ for chronic neck pain[15]
Pain interference		
BPI	Moderate	$r = -0.67$, $P < 0.001$ for chronic MSK pain[36]
BPI	Moderate	$r = -0.64$, $P < 0.001$ for PSEQ-2 on chronic MSK pain[36]
BPI	Moderate	Rho = −0.60, $P < 0.01$ for chronic MSK pain[45]
BPI	Moderate	$r = -0.58$, $P \leq 0.001$ for chronic MSK pain[38]
BPI	Low	$r = -0.43$, $P < 0.001$ for chronic MSK pain[2]
BPI	Low	$r = -0.42$, $P < 0.001$ for PSEQ-2 on chronic LBP[50]
MPI	Low	$r = -0.40$, $P < 0.001$ for chronic LBP[50]
Function/disability		
QuickDASh	High	$r = -0.71$, $P < 0.001$ for chronic upper limb MSK pain[10]
QuickDASH	High	$r = -0.71$, $P < 0.001$ for PSEQ-2 on chronic upper limb MSK pain[10]
QuickDASH	Moderate	$r = -0.67$, $P < 0.001$ for chronic upper limb MSK pain[7]
QuickDASH	Low	$r = -0.44$, $P < 0.001$ for PSEQ-2 on chronic upper limb MSK pain[7]
RMDQ	Moderate	$r = -0.66$, $P < 0.001$ for chronic MSK pain[17]
RMDQ	Moderate	$r = -0.65$, $P < 0.001$ for chronic MSK pain[49]
RMDQ	Moderate	$r = -0.58$, $P < 0.001$ for chronic MSK pain[38]
RMDQ	Moderate	$r = -0.55$, $P < 0.001$ for chronic MSK pain[36]
RMDQ	Moderate	$r = -0.54$, $P < 0.001$ for PSEQ-2 on chronic MSK pain[36]
RMDQ	Low	$r = -0.43$, $P < 0.001$ for PSEQ-2 on chronic LBP[50]
RMDQ	Low	$r = -0.41$, $P < 0.001$ for chronic LBP[50]
RMDQ	Low	$r = -0.40$, $P < 0.001$ for chronic MSK pain[3]
RMDQ	Low	$r = -0.35$, $P < 0.001$ for PSEQ-4 on chronic MSK pain[31]
RMDQ	Low	$r = -0.29$, $P < 0.001$ for PSEQ-2 on chronic MSK pain[31]
AVD-IASP-S	Low	$r = -0.41$, $P < 0.001$ for chronic MSK pain[31]
NDI	Low	Rho = −0.38 $P < 0.05$ for chronic neck pain[15]
Anxiety		
HADS-A	Moderate	$r = -0.63$, $P < 0.001$ for PSEQ-2 on chronic LBP[50]
HADS-A	Moderate	$r = -0.59$, $P < 0.001$ for chronic LBP[50]
HADS-A	Moderate	$r = -0.56$, $P < 0.001$ for chronic MSK pain[25]
HADS-A	Moderate	$r = -0.55$, $P < 0.001$ for chronic MSK pain[1]
HADS-A	Low	$r = -0.39$, $P < 0.01$ for chronic MSK pain[22]
DASS-A	Low	Rho = −0.32, $P < 0.01$ for chronic MSK pain[45]

(*continued on next page*)

Table 4 (*continued*)

Summary of findings for construct validity.

Convergent validity with:	Correlation (high, moderate, low, and very low)	Person (r) or Spearman (Rho) correlation coefficient, significance, PSEQ version (other than PSEQ-10) when applicable, and population studied.
Depression		
HADS-D	Moderate	$r = -0.68$, $P < 0.001$ for chronic MSK pain[1]
HADS-D	Moderate	$r = -0.67$, $P < 0.001$ for chronic LBP[50]
HADS-D	Moderate	$r = -0.66$, $P < 0.001$ for chronic MSK pain[25]
HADS-D	Moderate	$r = -0.51$, $P < 0.01$ for chronic MSK pain[22]
HADS-D	Low	$r = -0.48$, $P < 0.001$ for chronic MSK pain[3]
HADS-D	Low	$r = -0.49$, $P < 0.001$ for PSEQ-2 on chronic LBP[50]
HADS-D	Low	Rho $= -0.45$, $P < 0.01$ for chronic MSK pain[45]
HADS-D	Low	$r = -0.37$, $P < 0.001$ for chronic MSK pain[17]
DASS-D	Moderate	$r = -0.59$, $P < 0.001$ for chronic MSK pain[36]
DASS-D	Moderate	$r = -0.51$, $P < 0.001$ for PSEQ-2 on chronic MSK pain[36]
DASS-D	Low	Rho $= -0.40$, $P < 0.01$ for chronic MSK pain[45]
BDI	Low	$r = -0.45$, $P < 0.001$ for chronic MSK pain[31]
BDI	Low	$r = -0.42$, $P < 0.001$ for PSEQ-4 on chronic MSK pain[31]
BDI	Low	$r = -0.32$, $P < 0.001$ for PSEQ-2 on chronic MSK pain[31]
Catastrophizing		
PCS	Moderate	$r = -0.68$, $P < 0.001$ for chronic MSK pain[1]
PCS	Moderate	Rho $= -0.59$, $P < 0.05$ for chronic neck pain[15]
PCS	Low	$r = -0.49$, $P < 0.001$ for chronic MSK pain[2]
PCS	Low	$r = -0.48$, $P < 0.001$ for chronic MSK pain[13]
PCS	Low	Rho $= -0.43$, $P < 0.01$ for chronic MSK pain[45]
PCS	Low	$r = -0.42$, $P < 0.001$ for chronic MSK pain[17]
PCS	Low	$r = -0.41$, $P < 0.001$ for chronic MSK pain[25]
PCS	Low	$r = -0.39$, $P < 0.001$ for chronic MSK pain[31]
PCS	Low	$r = -0,38$, $P < 0.001$ for PSEQ-2 on chronic MSK pain[31]
PCS	Low	$r = -0.36$, $P < 0.001$ for PSEQ-4 on chronic MSK pain[31]
PBQ	High	$r = -0.74$, $P < 0.001$ for chronic MSK pain[35]
PRSS-C	Moderate	$r = -0.50$, $P < 0.001$ for PSEQ-2 on chronic MSK pain[36]
PRSS-C	Moderate	$r = -0.55$, $P < 0.001$ for chronic MSK pain[36]

AVD-IASP-S, adapted version of the disability subscale of the International Association for the Study of Pain scale; BPI, Brief Pain Inventory; BDI, Beck Depression Inventory; DASS-D, Depression Anxiety Stress Scale-Depression; DASS-A, Depression Anxiety Stress Scale-Anxiety; HADS-A, Hospital Anxiety and Depression Scale-Anxiety subscale; HADS-D, Hospital Anxiety and Depression Scale-Depression subscale; MPI, Multidimensional Pain Inventory; PCS, Pain Catastrophizing Scale; PRSS-C, Pain Response Self-Statements Catastrophizing Scale; RMDQ, Roland–Maurice Disability Questionnaire; VAS, visual analogue scale.

reported an ES and SRM of 0.53 and 0.63 for the PSEQ-10, 0.46 and 0.50 for the PSEQ-2, and 0.46 and 0.54 for the PSEQ-4. These values represent a small to moderate change. Finally, Chiarotto et al.[15] reported ES and SRM values of 0.73 and 1.15 in a sample of patients (146 participants) with neck pain disorders after 10 sessions of multimodal physical therapy over 5 weeks. When analyzed in subgroups, responsiveness was higher in patients with idiopathic neck pain (ES: 1.36 and SRM: 1.21) compared with patients with whiplash-associated disorder (ES: 0.57 and SRM: 1.09). Finally, reported MCID values ranged from

Table 5

Summary of findings for internal consistency and reliability.

	n studies	*n* participants	Population	Range	Weighted mean
Internal consistency (Cronbach alpha) PSEQ-10	20	5295	Chronic MSK pain, upper limb MSK pain, chronic LBP, subacute and chronic neck pain, and fibromyalgia	0.79–0.95	*N/A*
Internal consistency (Cronbach alpha) PSEQ-4	2	430	Chronic MSK pain	0.87–0.90	*N/A*
Internal consistency (Cronbach alpha) PSEQ-2	6	1354	Chronic MSK pain, upper limb MSK pain, and chronic LBP	0.76–0.91	*N/A*
Test–retest reliability (ICC)					
PSEQ-10	10	908	Chronic MSK pain, upper limb MSK pain, chronic LBP, and fibromyalgia	0.75–0.93	0.86
PSEQ-2	3	422	Chronic MSK pain, upper limb MSK pain, and chronic LBP	0.79–0.88	0.85
Absolute reliability (SEM)					
PSEQ-10	3	356	Chronic MSK pain and chronic LBP	1.23–5.66	3.37
Absolute reliability (MDC95)					
PSEQ-10	3	273	Chronic MSK pain, and chronic LBP	3.41–15.69	11.52

ICC, intraclass correlation coefficient; MDC_{95}, minimal detectable change at 95% confidence interval; MSK, musculoskeletal; LBP, low back pain.

5.5 to 8.5 points of 60 for the PSEQ-10 in a population of chronic LBP (2 studies; 151 participants; sensitivity (Sn) = 0.648; specificity (Sp) = 0.875; area under the curve (AUC) = 0.73 to 0.79),[16,29] whereas 1 study reported a MCID of 1.5 of 12 for the PSEQ-2 (1 study; 103 participants; Sn = 0.507; Sp = 0.844; AUC = 0.75) and 1.5 of 24 for the PSEQ-4 (1 study; 103 participants; Sn = 0.803; Sp = 0.687; AUC = 0.81) for chronic LBP.[16] To determine MCID, participants were dichotomized based on patients' self-reported improvement on the global perception of change scales, which were either general[16,29] or specific to self-efficacy.[16] Patients classifying themselves on the global perception of change scales as "totally improved," "much improved," and "rather improved" were considered improved; those who described themselves as "slightly improved," "unchanged," and "slightly worsened" were considered as unchanged.

3.11. Language and cultural translation

Several studies assessed the psychometric properties of the translated and adapted versions of the PSEQ-10 or one of its shortened versions in languages such as Amharic,[14] Chinese-*Hong Kong*[25], Chinese-*Mainland*,[50] Catalan,[13] Danish,[37] Italian,[17] Japanese,[2] Mongolian,[45] Persian,[3] Portuguese-*Brazilian*,[38] Portuguese-*European*,[22] and Yoruba.[21] They all applied rigorous methods to complete the translation and cross-cultural adaptation process.[6] Fatoye et al.[21] modified item 7 "I can cope with my pain without medication" by adding a word related to charms to the Yoruba version to better reflect their traditional medical culture. Chala et al.[14] also adapted 2 items (2 and 10) to the local culture and context for the Amharic version. Other examples of tasks were added to item 2 "I can do most of the household chores (tidying up, washing dishes, etc), despite the pain" to better reflect household chores accomplished by men (making a bed, picking things up, cooking a meal, splitting firewood, etc). Finally, the word "physically" was added to item 10 "I can gradually become more active, despite pain" because the Amharic translation of "active" is broad and can result in other interpretations of its meaning such as ones related to the level of consciousness. Mean and standard deviation scores of the PSEQ-10 obtained from those cross-cultural adaptations are presented in **Table 6**.

3.12. Administration burden

Two studies (299 participants) reported the time to complete the PSEQ-10.[17,23] Kortlever et al.[23] reported that participants took an average of 78 seconds (range: 24–316 seconds; IQR: 60–101 seconds) to complete the PSEQ-10 (original English version), whereas Chiarotto et al.[17] reported a median time of completion of 180 seconds (IQR: 120–300 seconds) for the PSEQ-10/Italian.

4. Discussion

This study summarized evidence from 28 studies including 9853 participants with various MSK disorders who completed the PSEQ-10 or one of its shortened versions (PSEQ-2 and/or PSEQ-4). The PSEQ is a unidimensional measurement tool to assess pain-related self-efficacy in populations with chronic pain, including MSK disorders. It has been adapted and validated in 14 languages which make it easier to use in several countries. Hypotheses on construct validity were confirmed in most included studies showing that the PSEQ is positively correlated with quality of life measures and negatively correlated with measures of disability, pain, pain interference, anxiety, depression, and catastrophizing. These divergent and convergent construct validity results show that patients with higher pain self-efficacy have a better quality of life and lower levels of physical and mental symptoms.

Table 6

Mean and standard deviation scores of the PSEQ-10 obtained from cross-cultural adaptations.

Version	Population	*n*	Mean score (SD)
Amharic[14]	Chronic low back pain	240	43.8 (12.0)
Catalan[13]	Chronic MSK pain	227	21.28 (5.94)
Chinese-*Hong Kong*[25]	Chronic MSK pain	120	28.5 (13.3)
Chinese-*Mainland*[50]	Chronic MSK pain	219	32.74 (13.18)
Danish[37]	Fibromyalgia	99	27.32 (5.12)
English[35]	Chronic MSK pain	1306	23.0 (12.7)
Italian[17]	Chronic low back pain	165	36.12 (12.90)
Japanese[2]	Chronic MSK pain	176	33.06 (13.51)
Mongolian[45]	Chronic MSK pain	142	36.7 (15.3)
Nigerian[46]	Chronic MSK pain	256	*Not available*
Persian[3]	Chronic MSK pain	517	36.7 (14.3)
Portuguese-*Brazilian*[38]	Chronic MSK pain	311	34.8 (14.8)
Portuguese-*European*[22]	Chronic MSK pain	174	40.83 (11.31)
Yoruba[21]	Chronic low back pain	131	*Not available*

MSK, musculoskeletal; PSEQ, Pain Self-Efficacy Questionnaire.

Boxes of the COSMIN risk of bias tool

1. Development
   - 1a. Design
   - 1b. Cognitive interview study or other pilot test
2. Content validity
   - 2a. Asking patient about relevance
   - 2b. Asking patient about comprehensiveness
   - 2c. Asking patient about comprehensibility
   - 2d. Asking professionals about relevance
   - 2e. Asking professionals about relevance
3. Structural validity
4. Internal consistency
5. Cross-cultural Validity
6. Reliability
7. Measurement error
8. Criterion validity
9. Construct validity
   - 9a. Comparison with other outcome measurement instruments (convergent validity)
   - 9b. Comparison between subgroups (discriminative or known-groups validity)
10. Responsiveness
    - 10a. Criterion approach (I.E. comparison to a gold standard)
    - 10b. Construct approach (ie, hypotheses testing: comparison with other outcome measurement instruments)
    - 10c. Construct approach: (I.E. hypotheses testing: comparison between subgroups)
    - 10d. Construct approach: (I.E. hypotheses testing: before and after intervention)

V=very good, A = adequate, D = doubtful, I=inadequate, - = not applicable.

Internal consistency is excellent overall for all versions of the PSEQ. A high Cronbach alpha score (≥0.95) indicates that certain items may be redundant.[42] Given that the PSEQ-10 Cronbach alpha was 0.95 for 3 included studies, it is possible that a certain level of redundancy is present within the questionnaire. However, because this high score was only found in 3 studies of 20 and that no study suggested redundancy nor the removal of an item based on factor loading analysis, the internal consistency was considered excellent. Another key point highlighted by authors was that item 7 of the questionnaire, "I can cope with my pain without medication," constantly had the lowest factor loading and correlation with other items.[17,19,20,25,35,38] One possible explanation for this could be that this item does not strictly measure the construct of self-efficacy or is the item that measures it the least.[20] Another reason could be that patients can cope well with their pain but rely heavily on their medication to do so. Test–retest reliability and MDC_{95} also showed excellent overall values. Clinicians can consider MDC_{95} to be 11.5 of 60 PSEQ-10 points (representing 19% of the total score) in the population with chronic LBP. A MDC higher than the MCID means that the MDC needs to be considered as the relevant value to determine whether a change has truly occurred between 2 measurements in time and whether the change is clinically relevant.[24]

Floor and ceiling effects were extracted and analysed. Floor effects refer to the capacity of a test to discriminate among patients with lower scores on the test. On the other hand, the ceiling effect refers to the capacity of a test to discriminate among patients with higher scores on the test.[43] If 15% or fewer participants achieve the lowest or highest available score, the effect is considered to be acceptable.[30] The PSEQ-10 and its short versions did not demonstrate a floor effect in any study, signifying that the measure is able to differentiate among people who have low pain self-efficacy. The PSEQ-10 ceiling effect was acceptable; however, the PSEQ-2 (version with items 8 and 9) ceiling effect was high, implying that the short form is less able to differentiate among patients with a higher level of pain self-efficacy. This suggests that the PSEQ-2 should not be used in place of the original version.

Clinicians can consider the MCID of the PSEQ-10 to range between 5.5 and 8.5 of 60 PSEQ-10 points (representing between 9% and 14% of the total score). MCIDs of the PSEQ-2 and PSEQ-4 were determined to be 1.5 of 12 points (12.5% of the total score) and 1.5 of 24 points (6.25% of the total score), respectively, in the population with chronic LBP. However, because few studies have evaluated it, the responsiveness needs to be further investigated, especially in populations other than chronic LBP. A few included studies reported the responsiveness of the PSEQ-10 by describing ES and SRM values.[15,16,36] Because it is complex to assess changes in a construct such as self-efficacy, it is difficult to calculate responsiveness values in people who improved their pain self-efficacy. Only 1 study[16] used the global perception of change scale specific to the construct of self-efficacy. All other studies[15,29,36] reporting on responsiveness used the global perception of change scale rated on the impression of general improvement in patients with chronic low back or neck pain which can lead to a bias for the change calculation of pain self-efficacy–specific construct. Observed ES and SRM values demonstrated that the PSEQ-10, PSEQ-4, and PSEQ-2 are responsive enough to be used in a clinical or research context to evaluate changes after an intervention on chronic MSK disorders, chronic LBP, and chronic neck pain, including whiplash-associated disorders.

4.1. Comparison with other reviews

To the best of our knowledge, this is the first systematic review, with a complete reported methodological quality appraisal process, assessing PSEQ psychometric properties in the population with chronic MSK pain. However, other reviews have reported PSEQ properties, and their results lean in the same direction.

Miles et al.[32] published a systematic review in 2011 where they looked at different patient-reported outcome measures for self-efficacy in a population with chronic pain. However, they did not mention the use of a methodological quality appraisal tool. Among all studies included, 14 of them reported values for the PSEQ. Authors of the systematic review concluded that the PSEQ had good content and construct validity (correlations with measures of anxiety, depression, pain ratings, and unhelpful coping strategies) which is similar to the current findings. They also mentioned that the PSEQ was easy to score and quick to administer (<10 minutes). Finally, they suggested that further research should focus on the assessment of responsiveness and test–retest reliability because there was not enough solid evidence to report on those properties. Since their publication, 10 studies on reliability and 3 on responsiveness have been published and were included in the current review, considerably increasing the body of evidence on these psychometric properties, especially for reliability.

In 2018, Banerjee et al.[5] published a systematic review of the psychometric properties of self-management questionnaires used in clinical trials with patients with chronic pain. They reported risk of bias of the methodology of those trials, but not for the assessment of each psychometric property. Surprisingly, they included only 7 studies using the PSEQ compared with 28 in the current review. Based on those 7 studies, the authors concluded that the PSEQ had good content and construct validity as well as good internal consistency which are results that align with our findings.

Finally, a 2020 systematic review by Vergeld et al.,[47] which included 47 studies (14 of them reporting on the PSEQ-10), assessed psychometric properties of the PSEQ-10 when used with a population with LBP. They reported excellent internal consistency, with Cronbach alpha ranging from 0.88 to 0.94 and excellent test–retest reliability with ICCs ranging from 0.82 to 0.92. These values are similar to the findings of the present review. However, they did not provide information on study quality assessment and data extraction. Similarly to Miles et al.,[32] they also suggested that future studies look at PSEQ responsiveness because there were not enough data to conclude on this property.

4.2. Clinical relevance

The results from this study show that the PSEQ is suitable to be used in clinical practice to measure patients' level of pain self-efficacy and its evolution in time. The MDC of 11.5 of the PSEQ-10 can be used as the MCID to assess change over a period ranging from 4 to 12 weeks. The PSEQ-10 is relatively short to complete, and it has been translated in different languages making it a useful tool in both clinical and research settings in several countries. Clinician can use the PSEQ-2 for a screening purpose as suggested by Bot et al. and Nicholas et al.[8,36] However, given that there are 2 versions of the PSEQ-2 which can be confusing, that the full version PSEQ-10 is short to complete, that the ceiling effect of the PSEQ-2 is high, and that the PSEQ-4 has not been studied enough, it is our

opinion that short forms are less valuable in clinic and research settings.

4.3. Study limitations

Although this systematic review was conducted using a rigorous method and a previously validated and comprehensive study quality assessment tool, some limitations can be identified. First, because we limited our literature search to 3 databases, we may have missed some potential articles. We are, however, quite confident that the databases used allowed us to perform an extensive search of the literature, especially because reference lists of all retrieved studies and previous reviews were searched for further relevant studies. Second, we only searched for articles written in French or English, which could have limited the scope of our results and prevented us from including relevant articles written in other languages. However, no study was excluded in the selection phase based on language.

4.4. Study strengths

This study also had several strengths because it constitutes an extensive and up-to-date review of the literature regarding the psychometric properties of the PSEQ when used in populations with MSK disorders. It was conducted using validated guidelines and represents an improvement in methodological quality compared with previous reviews where methodology for data extraction was not described and authors did not perform a quality assessment of included studies to evaluate their risk of bias.

4.5. Suggestion for future research

Based on our results, future research evaluating this questionnaire's psychometric properties could improve reliability testing methods because this category had the lowest COSMIN score. Such studies should make sure to document the similarity between the 2 measurement conditions and the time interval between administrations. In addition, they should document whether patients were stable or not during the interval period in-between measurements, using tools such as the Global Rating of Change scales. Finally, further studies are needed to establish MDC and MCID values for populations other than chronic LBP because there are currently values available for this population alone.

5. Conclusion

The PSEQ is a robust and frequently used questionnaire assessing pain self-efficacy in chronic MSK disorders. Methods used for most of the measurement properties evaluated were of *adequate* to *very good* quality (76/92; 90%) according to the COSMIN risk of bias tool. The PSEQ has good content validity and structural validity which shows that every version of the PSEQ is a unidimensional questionnaire. Construct validity shows low to moderate correlations with measures of quality of life, disability, pain, pain interference, anxiety, depression, and catastrophizing. Concurrently, a high correlation between the PSEQ-10 and the 2 versions of the PSEQ-2 was found. Test–retest reliability and internal consistency are excellent. Pooled estimates of SEM (3.37) and MDC (11.52) were calculated and deemed good. MCID ranged from 5.5 to 8.5 of 60 in patients with chronic LBP which are lower than the MDC meaning that the MDC can be used to measure change by clinicians and researchers. The results of this study suggest that clinicians can use the questionnaire with confidence to measure the pain self-efficacy state of patients and their change in time. Further high-quality studies are needed to determine MCID in populations other than chronic LBP.

Disclosures

The authors have no conflicts of interest to declare.

Acknowledgments

The authors want to acknowledge Charles Batcho for his guidance and advice throughout the production process. They also acknowledge Marianne Roos for her contribution to the English language review of the manuscript. M-O. Dubé and P. Langevin are supported by Training Scholarships from the Fonds de Recherche Québec-Santé (FRQ-S). J.-S. Roy is supported by salary awards from the FRQ-S. Review was not registered, and there is no available protocol for the present review. The data that support the findings of this study are available from the corresponding author, JSR, on reasonable request.
Authors Contribution: M.-O. Dubé and P. Langevin conceived and designed the review. They conducted the literature search, extracted the data, and wrote the manuscript. J.-S. Roy commented on the draft of the manuscript. All authors have seen and approved the final version.

Article history:
Received 10 May 2021
Received in revised form 13 September 2021
Accepted 29 September 2021

References

[1] Adachi T, Enomoto K, Yamada K, Inoue D, Nakanishi M, Takahashi N, Nishigami T, Shibata M. Evaluating the psychometric properties of two-item and four-item short forms of the Japanese Pain Self-Efficacy Questionnaire: a cross-sectional study. J Anesth 2019;33:58–66.
[2] Adachi T, Nakae A, Maruo T, Shi K, Shibata M, Maeda L, Saitoh Y, Sasaki J. Validation of the Japanese version of the pain self-efficacy questionnaire in Japanese patients with chronic pain. Pain Med 2014; 15:1405–17.
[3] Asghari A, Nicholas MK. An investigation of pain self-efficacy beliefs in Iranian chronic pain patients: a preliminary validation of a translated English-language scale. Pain Med 2009;10:619–32.
[4] Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Adv Behav Res Ther 1978;1:139–61.
[5] Banerjee A, Hendrick P, Bhattacharjee P, Blake H. A systematic review of outcome measures utilised to assess self-management in clinical trials in patients with chronic pain. Patient Educ Couns 2018;101:767–78.
[6] Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine 2000; 25:3186–91.
[7] Bot AGJ, Menendez ME, Neuhaus V, Mudgal CS, Ring D. The comparison of paper- and web-based questionnaires in patients with hand and upper extremity illness. Hand 2013;8:210–14.
[8] Bot AGJ, Nota SPFT, Ring D. The creation of an abbreviated version of the PSEQ: the PSEQ-2. Psychosomatics 2014;55:381–5.
[9] Bouffard J, Bertrand-Charette M, Roy JS. Psychometric properties of the musculoskeletal function assessment and the short musculoskeletal function assessment: a systematic review. Clin Rehabil 2016;30:393–409.
[10] Briet JP, Bot AGJ, Hageman MGJS, Menendez ME, Mudgal CS, Ring DC. The pain self-efficacy questionnaire: validation of an abbreviated two-item questionnaire. Psychosomatics 2014;55:578–85.
[11] Briggs AM, Cross MJ, Hoy DG, Sànchez-Riera L, Blyth FM, Woolf AD, March L. Musculoskeletal health conditions represent a global threat to healthy aging: a report for the 2015 world health organization world report on ageing and health. Gerontologist 2016;56(suppl 2):S243–255.
[12] Briggs AM, Woolf AD, Dreinhöfer K, Homb N, Hoy DG, Kopansky-Giles D, Åkesson K, March L. Reducing the global burden of musculoskeletal conditions. Bull World Health Organ 2018;96:366–8.

[13] Castarlenas E, Solé E, Galán S, Racine M, Jensen MP, Miró J. Construct validity and internal consistency of the Catalan version of the pain self-efficacy questionnaire in young people with chronic pain. Eval Health Prof 2020;43:213–21.
[14] Chala MB, Donnelly C, Wondie Y, Ghahari S, Miller J. Cross-cultural translation, adaptation, and validation of the Amharic version pain self-efficacy questionnaire in people with low back pain in Ethiopia. BMC Musculoskelet Disord 2021;22. doi: 10.1186/s12891-021-03985-4.
[15] Chiarotto A, Falla D, Polli A, Monticone M. Validity and responsiveness of the pain self-efficacy questionnaire in patients with neck pain disorders. J Orthop Sports Phys Ther 2017;48:204–16.
[16] Chiarotto A, Vanti C, Cedraschi C, Ferrari S, de Lima E Sà Resende F, Ostelo RW, Pillastrini P. Responsiveness and minimal important change of the pain self-efficacy questionnaire and short forms in patients with chronic low back pain. J Pain 2016;17:707–18.
[17] Chiarotto A, Vanti C, Ostelo RW, Ferrari S, Tedesco G, Rocca B, Pillastrini P, Monticone M. The pain self-efficacy questionnaire: cross-cultural adaptation into Italian and assessment of its measurement properties. Pain Pract 2015;15:738–47.
[18] Cohen J. Statistical power analysis for the behavioral sciences. New York, NY: Academic Press, 2013.
[19] Costa DSJ, Asghari A, Nicholas MK. Item response theory analysis of the pain self-efficacy questionnaire. Scand J Pain 2017;14:113–17.
[20] Di Pietro F, Catley MJ, McAuley JH, Parkitny L, Maher CG, Costa LdaCM, Macedo LG, Williams CM, Moseley GL. Rasch analysis supports the use of the pain self-efficacy questionnaire. Phys Ther 2014;94:91–100.
[21] Fatoye F, Mbada CE, Oladayo TO, Idowu OA, Oyewole OO, Fatoye C, Oke KI. Validation of the Yoruba Version of the Pain Self-Efficacy Questionnaire in Patients with Chronic Low Back Pain. Spine (Phila Pa 1976) 2021;46:E528–E533.
[22] Ferreira-Valente MA, Pais-Ribeiro JL, Jensen MP. Psychometric properties of the Portuguese version of the pain self-efficacy questionnaire. Acta Reumatol Port 2011;36:260–7.
[23] Kortlever JTP, Janssen SJ, van Berckel MMG, Ring D, Vranceanu AM. What is the most useful questionnaire for measurement of coping strategies in response to nociception?. Clin Orthop Relat Res 2015;473:3511–18.
[24] Leopold SS, Porcher R. Editorial: the minimum clinically important difference—the least we can do. Clin Orthop Relat Res 2017;475:929–32.
[25] Lim HS, Chen PP, Wong TCM, Gin T, Wong E, Chan ISF, Chu J. Validation of the Chinese version of pain self-efficacy questionnaire. Anesth Analg 2007;104:918–23.
[26] Lincoln N, Moreton B, Turner K, Walsh D. The measurement of psychological constructs in people with osteoarthritis of the knee: a psychometric evaluation. Disabil Rehabil 2017;39:372–84.
[27] Martinez-Calderon J, Flores-Cortes M, Morales-Asencio JM, Fernandez-Sanchez M, Luque-Suarez A. Which interventions enhance pain self-efficacy in people with chronic musculoskeletal pain? A systematic review with meta-analysis of randomized controlled trials, including over 12 000 participants. J Orthop Sports Phys Ther 2020;50:418–30.
[28] Martinez-Calderon J, Zamora-Campos C, Navarro-Ledesma S, Luque-Suarez A. The role of self-efficacy on the prognosis of chronic musculoskeletal pain: a systematic review. J Pain 2018;19:10–34.
[29] Maughan EF, Lewis JS. Outcome measures in chronic low back pain. Eur Spine J 2010;19:1484–94.
[30] McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: are available health status surveys adequate? Qual Life Res 1995;4:293–307.
[31] McWilliams LA, Kowal J, Wilson KG. Development and evaluation of short forms of the pain catastrophizing scale and the pain self-efficacy questionnaire. Eur J Pain 2015;19:1342–9.
[32] Miles CL, Pincus T, Carnes D, Taylor SJC, Underwood M. Measuring pain self-efficacy. Clin J Pain 2011;27:461–70.
[33] Mokkink LB, de Vet HCW, Prinsen CaC, Patrick DL, Alonso J, Bouter LM, Terwee CB. COSMIN risk of bias checklist for systematic reviews of patient-reported outcome measures. Qual Life Res 2018;27:1171–9.
[34] Munro BH. Statistical methods for health care research. 5th ed. Philadelphia: Lippincott Williams & Wilkins, 2005. Available at: http://catdir.loc.gov/catdir/toc/ecip0416/2004007098.html. Accessed March 1, 2021.Refstyled
[35] Nicholas MK. The pain self-efficacy questionnaire: taking pain into account. Eur J Pain 2007;11:153–63.
[36] Nicholas MK, McGuire BE, Asghari A. A 2-item short form of the Pain Self-efficacy Questionnaire: development and psychometric evaluation of PSEQ-2. J Pain 2015;16:153–63.
[37] Rasmussen MU, Rydahl-Hansen S, Amris K, Samsøe BD, Mortensen EL. The adaptation of a Danish version of the Pain Self-Efficacy Questionnaire: reliability and construct validity in a population of patients with fibromyalgia in Denmark. Scand J Caring Sci 2016;30:202–10.
[38] Sardá J, Nicholas MK, Pimenta CAM, Asghari A. Pain-related self-efficacy beliefs in a Brazilian chronic pain patient sample: a psychometric analysis. Stress and Health 2007;23:185–90.
[39] Sleijser-Koehorst MLS, Bijker L, Cuijpers P, Scholten-Peeters GGM, Coppieters MW. Preferred self-administered questionnaires to assess fear of movement, coping, self-efficacy, and catastrophizing in patients with musculoskeletal pain-A modified Delphi study. PAIN 2019;160:600–6.
[40] St-Pierre C, Desmeules F, Dionne CE, Frémont P, MacDermid JC, Roy JS. Psychometric properties of self-reported questionnaires for the evaluation of symptoms and functional limitations in individuals with rotator cuff disorders: a systematic review. Disabil Rehabil 2016;38:103–22.
[41] Streiner DL. Learning how to differ: agreement and reliability statistics in psychiatry. Can J Psychiatry 1995;40:60–6.
[42] Streiner DL. Starting at the beginning: an introduction to coefficient alpha and internal consistency. J Pers Assess 2003;80:99–103.
[43] Terwee CB, Bot SDM, de Boer MR, van der Windt DAWM, Knol DL, Dekker J, Bouter LM, de Vet HCW. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 2007;60:34–42.
[44] Terwee CB, Mokkink LB, Knol DL, Ostelo RWJG, Bouter LM, de Vet HCW. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. Qual Life Res 2012;21:651–7.
[45] Tuck NL, Khuvtsagaan B, Rashid U, Aamir T, Goucke R, Regjii B, Dorjbal E-A, Lundeg G, Bean DJ. The adaptation of pain assessment tools from high-income to low- and middle-income countries: psychometric properties of a set of chronic pain questionnaires in Mongolian and New Zealand patient samples. Pain Med 2021;22:948–60.
[46] Ugwuanyi CS, Okeke CIO. Psychometric properties of the pain self-efficacy questionnaire using Nigerian university students with chronic pain. IJPHRD 2020;11:1530–5.
[47] Vergeld V, Utesch T. Pain-related self-efficacy among people with back pain: a systematic review of assessment tools. Clin J Pain 2020;36:480–94.
[48] Vong SKS, Cheing GLY, Chan CCH, Chan F, Leung ASL. Measurement structure of the Pain Self-Efficacy Questionnaire in a sample of Chinese patients with chronic pain. Clin Rehabil 2009;23:1034–43.
[49] Wongpakaran N, Wongpakaran T, Wedding D, Gwet KL. A comparison of Cohen's Kappa and Gwet's AC1 when calculating inter-rater reliability coefficients: a study conducted with personality disorder samples. BMC Med Res Methodol 2013;13:61.
[50] Yang Y, Yang M, Bai J, Zhao J, Chen K, Zhou X, Wei X, Li M. Validation of simplified Chinese version of the pain self-efficacy questionnaire (SC-PSEQ) and SC-PSEQ-2 for patients with nonspecific low back pain in mainland China. Spine 2019;44:E1219–26.