Title: Measures of Patient Activation and Self-Efficacy

Authors: Clayon B. Hamilton[1,2] PhD, and Linda C. Li[1,2] PT, PhD

[1] Department of Physical Therapy, University of British Columbia, Canada

[2] Arthritis Research Canada, Canada

Address correspondence to Linda C. Li, PhD PT

Arthritis Research Canada

5591 No. 3 Road, Richmond, BC V6X 2C7 Canada

Email: lli@arthritisresearch.ca

Introduction

People with rheumatic diseases, as do those with other chronic health conditions, benefit from being confident about their ability and taking an active role to manage their health and health care process. An important concept related to self-management of chronic conditions is, patient activation, which describes an individual's knowledge, skill, and confidence for managing their health and healthcare.(1) Another similar and important concept is self-efficacy, which refers specifically one's belief or confidence in their ability to perform certain actions even when there are unpredictable and stressful aspects to the situation.(2, 3)

This review covers some well recognized measures of patient activation and self-efficacy, which could benefit the clinical care and study of people with arthritis. In this paper, we reviewed four self-report measures: Patient Activation Scale (PAM), Arthritis Self-Efficacy Scale-8 Item (ASES-8), PROMIS® Measures of Self-Efficacy for Managing Chronic Conditions (PROMIS-SE), and PROMIS® Measures of General Self-Efficacy (PROMIS General Self-Efficacy). To select the measures, we searched Google Scholar for papers citing the development and first validation paper of each measure. When the number of citations were 200 or more, we searched the Medline bibliographic database, informed by the COSMIN PubMed filter, to identify validation studies for each selected measure.(4) Furthermore, we focus only on measures used in adult populations and available in English, and placed greater attention to studies published within the last 5 years. This review is a follow-up to the 2011 review, published in this journal, on measure of self-efficacy.(5) We extended that review by adding three measures that are conceptually similar and overlapping to those previously reported: PAM, PROMIS-SE, and PROMIS General Self-Efficacy. We selected only the ASES-8 from among the measures previous reported, recognizing its increased use and the increase availability of information on its measurement properties.

1 Patient Activation Measure (PAM)

Description

Purpose. The PAM was designed to measure a person's perceived knowledge, skills and confidence in taking an active role in managing their health and health care.(1) The original

version containing 22 items was developed in English and published in 2004 and followed by a 13-item version (PAM-13) published in 2005.(1) This overview focuses on the PAM-13, which has been most often used and reported on.

Content or domains. A literature review, expert consensus, and patient focus groups developed the concept of patient activation.(1, 6) The concept was made into questionnaire items using existing instruments and by creating new items.(1, 6) The items of PAM-13 are divided across four developmental and hierarchical stages of activation: Believes active role is important (2 items), Confidence and knowledge to take action (6 items), Take action (3 items), and Staying the course (2 items).(6)

Number of items. The PAM-13 has 13 items.(1, 6)

Response options/scale. Four-point response scale indicating degree of agreement: 'disagree strongly', 'disagree', 'agree' and 'agree strongly', and a 'not applicable' (N/A) option. The response scale has no neutral anchor.

Recall period for items. Not applicable. Each item elicits respondents' current view. Any recall period would be determined by the case.

Cost to use. The cost varies by type of license, whether it is for commercial or research use. Commercial licenses costs are determined based on the size of the target population and operational requirements, while non-commercial research meeting specified requirements are offered a special license when they agree to share all of their de-identified data.

How to obtain. Property and copyright are held by Insignia Health (https://www.insigniahealth.com/products/pam-survey). The license and form for each language version is available from Insignia Health.

Practical application

Method of administration. Self-administered using paper or computer, and interviewer-administered by telephone or in-person.(1, 7-10)

Scoring. To calculate the total PAM score, the raw score is divided by the number of items answered (excepting non-applicable items) and multiplied by total number of items in the PAM-13.(6) Items with no response or with a response of 'not applicable' are scored as 'missing'. The PAM total scores range from 13 to 52. The raw summative scores of the PAM-13 can then converted to an interval-level activation scores using standardized scores generated

through Rasch analysis and available to license holders.(6) These standardized scores may be available as a conversion table.

Score interpretation. Activation scores vary from 0 to 100, with higher score indicating greater activation. Activation scores have cut-points across four level, which aligned to the four sequential stages of patient activation. The cut-points are reported in the literature as follows but might vary with updates from Insignia Health (https://www.insigniahealth.com/products/pam-survey). Level 1 (disengaged and overwhelmed) for 0 to 47, indicates patients are not understanding an active role is important. Level 2 (becoming aware but still struggling) for 47.1 to 55.1, indicates patients are lacking knowledge and confidence to take action. Level 3 (taking action) for 55.2 to 67.0, indicates patients are beginning to take action. Level 4 (maintaining behaviors and pushing further) for 67.1 to 100, indicates patients are maintaining healthy behavior over time.(11, 12) Insignia Health has normative data that it uses to refine the PAM and provide normative scoring.

Respondent time to complete. Not specified but estimated to be at maximum 5 to 10 minutes for the PAM-13. This is because the English version has a three-sentence introduction and averages 15.5 words per item. The long length of the items would contribute to long reading times and possibly long cognitive processing times.

Administrative burden. Depending on the mode of administration and scope of use, training might be required to administer the PAM-13, such as when used in health systems or for large survey using computer-assisted interviewing.(9) Standardized administration is recommended, but this instruction is not specified on the questionnaire. In clinical settings, both visual inspection and scoring have been recommended, which may require training. Scores need to be converted from raw to interval-level total scores and to the levels of activation available from the developers.

Translations/adaptations. PAM-13 has at least 50 language translations. Insignia Health has reported over 35 validated translations. At least 11 translations have published validation studies: Dutch, Norwegian, Danish, German, Hebrew, Italian, Korean, Spanish, European Spanish, Singapore English, and Turkish.(10, 13-21) The Dutch version was slightly rephased after cognitive testing.(13)

A version was published for people with mental health conditions (PAM-MH).(7) There is a caregiver and a parent version in multiple language translations.(22) The original 22-item

PAM has been adapted to Brazilian.(23) There is also a 10-item version, which is reported on the Insignia Health website to have similar measurement properties to PAM-13.

Psychometric information

Floor and ceiling effects. No numeric results were identified for these effects at the total scale level. At the individual item level, all items of the Italian version have displayed low floor effect and a moderate to large ceiling effect.(24) The Danish version has displayed ceiling effect greater than 15% for all items.(25) All but three items showed ceiling effect for the Norwegian version, while two items showed floor effects.(8) Visual inspection of results from Rasch analysis suggested ceiling effect for the overall total score for rural settings, which gives cautions for interpreting change of total scores over time.(26)

Reliability. *Internal consistency*. Cronbach's α was high for all version and vary from 0.80 to 0.90.(8, 9, 13-16, 18, 19, 21, 24-30)

Test-retest reliability. Intraclass correlation coefficient was reported as 0.76, 0.85, and 0.98 across three separate studies.(8, 16, 29)

Validity. *Content/face.* PAM-13 has demonstrated satisfactory content and face validity. The original set of 75 items where created through a three-round face-to-face cognitive testing with 20 individuals with chronic conditions.(1) The items were generally well accepted across the studies. Validation studies report low percentage of missing items, generally below 3.5%.(13, 20) Not applicable response was moderate, >10%, for a small number of items.(13) Young people in digital consulting found 11 of 13 item to be relevant, while four of seven clinicians found all items to be relevant.(8) Comprehension issues and cultural influences led to variation in responses and adaptation of the English version in Singapore.(17) For the Turkish version, difficulty in understanding "medical treatment" led to changes in item 7.(16) The Dutch had no major alteration only slight rephasing of some items after pretesting through a focus group and interviews of people with chronic illness or disability.(25)

Criterion. No information is available specifically for the PAM-13. When its predecessor, the 22-item PAM, was being developed, three blinded judges independently categorized the transcripts of interviews from high and low scoring respondents correctly 83% of the time (25 of 30 classifications). The respondents' preliminary PAM scores and three independent judges' assessment had strong agreement, with Cohen's kappa of 0.80, 0.90, and 0.90.(1)

Construct. Statistically significant Pearson's and Spearman's correlation coefficient of PAM-13 total scores to other instruments follows.

For correlations with measures of mood:
BDI-II (Beck depression inventory): 0.43 ref(20)
PHQ-9 (Patient Health Questionnaire): 0.35 ref(10), -0.16 ref(17)
Trait Hope Scale: 0.73 ref(29)

For correlations with measures of health-related quality of life:
HUI-3 (Health Utilities Index Mark 3): 0.32 ref(24)
SF-12 (12-Item Short Form Health Survey): 0.39 ref(10)
SF-36 MCS (Mental Health Composite Scale): 0.35 ref(24)
SF-36 PCS (Physical Health Composite Scale): 0.20 and 0.14 ref(24, 31)
MS QOL (Leeds multiple sclerosis quality of life): 0.42 ref(20)

For correlations with measures of self-efficacy, locus or control and adherence:
GSE (General Self-Efficacy scale): 0.43 ref(19)
LOTR (Health Locus of Control Scale) measure of optimism: 0.75 ref(29)
MMAS (Morisky Medication Adherence Scale): 0.04 not significant ref(32)
PEI (Perceived Expectancies Index): 0.44 ref(32)
Self-efficacy Scale: 0.47 ref(10)
SSE (Stanford Self-Efficacy for Managing Chronic Disease 6-item scale): 0.47 ref(14)

For correlations with measures of other health concepts:
S-TOHFLA (Short Test of Functional Health Literacy in Adults): 0.20 ref(33)
SBSQ-D (Set of Brief Questions) for health literacy: 0.28 ref(13)
CAHPS (Consumer Assessment of Healthcare Providers and Systems): 0.007-0.125 ref(26)
SLIQ (Simple Lifestyle Indicator Questionnaire):0.29 ref(24)

Studies determined the PAM-13 has a unidimensional construct or single dominant factor (namely, patient activation) which explained 34.5% to 96% of variance in total scores.(8, 15, 16,

18, 19, 21, 26, 30) One study found the PAM-13 has a unidimensional construct only when item 1 was removed.(20) Another study found item 3 and 7 had noticeable misfit for forming total scores with the other items of the PAM-13.(24)

The difficulty order of items was different from American English version for several other language versions, which raises concerns for interpretation of activation scores across the stages of activation.(18, 24, 31) Evidence of differential item function for demographic and clinical variables was reported in some studies.(18, 21, 25, 26, 34) Studies also reported no significant difference in scores across categories of demographic variables.(30)

Responsiveness. The responsiveness of PAM-13 in people with arthritis is unclear. Effect size of an educational intervention among out-patients with mental health conditions was 0.83.(8)

Minimal important differences. No evidence found.(35)

Generalizability. The PAM-13 was developed to be used in any adult patient population, and now has a parent version for children. It has been used across entire health care systems.(36) There are no special populations for which the PAM-13 is not intended for use.

Use in clinical trials. Insignia Health list 566 research studies including clinical trials that have used PAM. The PAM-13 is extensively used as a primary or secondary outcome measure in randomized controlled trials.(12)

Critical appraisal of overall value to the rheumatology community

Strengths. The PAM-13 has been extensively tested and used in a variety of healthcare settings including entire health systems.(36) Studies have repeatedly confirmed it is a reliable and valid measure of a unidimensional construct. The construct appears to capture the developmental nature of the patient activation phenomenon.(15) It is one of the most widely used questionnaire for capturing how actively engaged people are in their own health and health care. The PAM-13 has been translated and tested in multiple languages and is adequately precise for assessing individual patients. Furthermore, it has established scoring guidelines and the activation/total scores can be calculated when responses to items are missing.(37)

Caveats and cautions. Using a sample that included 66.4% of participants with arthrosis, all items of the German version of the PAM-13 fit the construct of the PAM-13 scale.(18) In a sample of 270 Korean patients with osteoarthritis, item 1 of PAM-13 was not consistent with the

construct of the total scale.(21) Similarly, a study of patients with neurological conditions reported that scaling problems (misfit of item 3 and 7) may yield measurement error and biases for those with low levels of activation.(24) The use of items and scaling tailored to specific diagnostic groups may be warranted.(24, 31) This may include adding items that would reflect higher levels of activation.(31)

The use of the 'levels of activation' in clinical settings should be done with caution because the extent of patient activation represented by each item varied across a number of studies.(18, 21, 24, 30) When the PAM-13 is used as a research tool (e.g., evaluate interventions) or in situations where careful direction or additional help is not possible, the results of the PAM-13 need to be interpreted with caution.(38)

Clinical usability. PAM-13 may be used flexibly as an effective measure by using both a visual scan and the PAM scores or levels of activation.(6, 20) A visual scan might be effective in clinical encounters.(1) The PAM-13 is easy to score and provides score interpretation across four levels of activation.(6)

The PAM-13 may be useful for tailoring targeted and individualized interventions to improve patient engagement in healthcare and to train clinicians in adapting their communication to the level of a patient's activation.(14, 15) In such cases, PAM-13 may be used along with clinical interviews.(14) It can be used both as the basis for clinical interventions to improved patient engagement and as a measure of the effectiveness of interventions.(8, 37)

Research usability. PAM-13 is a psychometrically sound, widely use and accepted measure of patient activation. The Insignia Health website indicates the PAM has been documented in over 400 peer-reviewed published studies worldwide as being able to measure activation and predict a broad range of health-related behaviors and outcomes.

2 Arthritis Self-Efficacy Scale-8 Item (ASES-8)

Description

Purpose. The original ASES, developed in English with 20 items, was published in 1989 as the first arthritis-specific measure for assessing "patients' beliefs on the extent to which they could perform specific tasks or behaviors to cope with the consequences of arthritis".(39) It is recognized as the most widely used arthritis-specific measure of self-efficacy.(40, 41) There is an 11-item version that has been rarely reported on.(42, 43) An 8-item version called ASES-8

was published in Spanish in 1995.(42) The 2011 review,(5) in this journal, on the ASES-8 was published before the first validation study of the English version of ASES-8 was published in 2014.(44)

Content or domains. The ASES-8 has two domains: self-efficacy for pain and self-efficacy for other symptoms. It includes items from the ASES pain subscale (two items) and other symptoms subscale (four items), and two new items related to preventing pain and fatigue from interfering with things one wants to do.(5)

Number of items. Eight items with no subscales.

Response options/scale. Ten-point rating scale on a patient's belief in their certainty to do specific tasks. The anchors are 1 (very uncertain) to 10 (very certain), with each item having the question stem "How certain that you can. …"

Recall period for items. Not applicable. Each item elicits respondents' current view.

Cost to use. Free for use without permission.

How to obtain. English, French, and Spanish versions are available at the website of Self-Management Resource Center, URL: (https://www.selfmanagementresource.com/resources/evaluation-tools/).

Practical application

Method of administration. Either interviewer administered or self-administered by patient using paper.(5, 44-46)

Scoring. Scoring for the English, French, and Spanish versions are available on the Self-Management Resource Center website, including instructions for handling missing items and double responses to a single item by an individual. No computer is necessary for scoring. Finals scores are the mean ratings of the eight items.

Score interpretation. Final scores vary from 1 to 10, with higher scores meaning greater self-efficacy or confidence. No cut-points values or population norms are available.

Respondent time to complete. Time to complete was within 6 minutes for the Arabic version.(47) This is close to the previous estimate of <5 minutes.(5)

Administrative burden. It take less than 10 minutes to administer and is easy to score.(41) No training, software, or special equipment is needed, and scoring is simple calculation of a mean score.

Translations/adaptations. ASES-8 was originally developed for Spanish-speaking people,(42) and the English version was available for many years before any publication on its measurement properties.(5) The Spanish version refers only to arthritis in each item, while the English version of the ASES-8 refers to both arthritis and fibromyalgia.(44) A German version published in 2003 was developed from the Spanish version with people with arthritis and fibromyalgia.(48) An Arabic version published in 2020, Brazilian Portuguese version in 2019, Chinese version in 2017, and Yoruba (Nigerian indigenous language) version in 2016, were each translated from the English version.(45, 47-50) A modified English version (ASES-AS) was developed in 2011 for ankylosing spondylitis in a UK population by making minor changes in phraseology including changing 'rheumatoid arthritis and fibromyalgia' to 'Ankylosing Spondylitis'.(46) A French version of the ASES-8 is available on the Self-Management Resource Center website, but has not been tested. It is promising that other validated language translations of the ASES-8 could become available, given the ASES has already been validated in Swedish and more recently in French in 2019.(51-53)

Psychometric information

Floor and ceiling effects. Studies on the Arabic and German versions report no floor or ceiling effect over the standard 15% threshold.(47, 48) The Arabic version demonstrated no floor or ceiling effect among 67 patients in Palestine with rheumatoid arthritis.(47) Across the individual items of the Arabic version, the percent floor effect varied between 1.5 and 11.9% and the percent ceiling varied between 1.5 and 10.4%.(47)

Reliability. *Internal consistency.* Overall, the Cronbach's α of ASES-8 has been higher than 0.85 across its different language and adapted versions. The Spanish ASES-8 had a Cronbach's α that vary between 0.92 and 0.96.(42) The English ASES-8 had Cronbach's α of 0.89 among 401 people with self-reported doctor-diagnosed arthritis, and Cronbach's α varied from 0.87 to 0.94 for subgroups.(44) The Arabic ASES-8 had Cronbach's α that varied slightly between 0.86 to 0.88 when any single items were deleted.(47) The Brazilian ASES-8 had Cronbach's α of 0.985 among 30 patients with rheumatoid arthritis for at least one year.(45) Chinese ASES-8 had Cronbach's alpha of 0.942 among 136 patients with rheumatoid arthritis and varied slightly (0.929–0.940) when any single item was deleted.(49) The German ASES had Cronbach's α of 0.90 among 201 patients (148 with fibromyalgia and 53 with rheumatoid

arthritis).(48) Cronbach's α was not reported for the Yoruba version.(50) The ASES-AS Cronbach's alpha was of 0.93.(46)

Test-retest. Overall, the 2-week test-retest reliability was ≥ 0.69. The 10-days to 2-week test–retest was 0.69 for the Spanish version.(42) The Chinese version's 2-week test-retest reliability was 0.98.(49) The 2-week test-retest reliability of ASES-AS varied from 0.72 to 0.82.(46) The German version had an 8-week test–retest reliability of 0.51.(48)

Validity. *Criterion*. No evidence found.

Content/face. The Spanish ASES-8 was developed from a 13-item version of the ASES (the 11 item ASES plus two new items) through a rigorous process of forward and back translation with eight culturally-diverse bilingual Spanish-speakers, review by a bilingual committee, pretesting with arthritis patients, a scaling study, and a replication study.(42) The items were reduced to eight because of low test-retest correlations, ambiguity and redundancy.(42) The German version was developed from the Spanish ASES-8, but details of the content validation process was not published.(48) The Arabic, Brazilian Portuguese, Chinese, and Yoruba language versions of ASES-8 all used a forward and back translation from the English version.(45, 47-50) The Brazilian Portuguese version was tested for cultural equivalence on 30 patients with rheumatoid arthritis and was shown to be understandable and culturally adapted.(45) After cognitive interviews with five patient research partners, the Arabic ASES-8 required no changes because it was easy to understand and administer, and had no conceptual and cultural difference to the English version.(47)

Construct. No validation data were published on the Spanish ASES-8. The German translation of the ASES-8 demonstrated significantly positive correlations with conceptual measures of function (r = 0.20), coping techniques (r = 0.35), internal locus of control (r = 0.33), optimism (r = 0.39), and general self-efficacy (r = 0.40) among the 148 people with fibromyalgia, and negative correlations with depression (r = −0.53). (48)

The English ASES-8 scores from 401 people with arthritis were significantly negatively correlated with arthritis symptoms (r = -0.39 to -0.41), depressive symptoms (r = -0.43), self-reported disability(r = -0.43), and positively correlated with health-related quality of life (r = 0.41), self-rated health (r = 0.32), functional performance measures (r = 0.18 to 0.25), and total physical activity (r = 0.15).(44) An identical pattern of associations was seen for all subgroups, although some of the correlations did not reach statistical significance.(44) The ASES-AS scores

of 601 people had significant negative correlations with pain (r = -0.62), depression (r = -0.56), anxiety (r = -0.42) and positive correlations with ankylosing spondylosis-specific quality of life (r = 0.58 to 0.70) and general quality of life (r = 0.56 to 0.63).(46)

The Chinese ASES-8 score of 134 people with RA were significantly negatively correlated with pain (r = -0.487), depression (r = -0.583), anxiety (r = -0.656), and positively correlated with general functional status (r = 0.561), and fatigue (r = 0.660).(49)

The Brazilian ASES-8 scores of 32 people aged 18 to 60 year with RA were significantly negatively correlated with pain (r = -0.278), functional capacity (r = -0.437), depression (r = -0.562) and positively correlated with domains of quality of life (r = 0.321 to 0.558).(45)

The Arabic ASES-8 scores of 67 people aged 29 to 77 years with RA had significant negative correlations with pain (r = -0.05 to -0.16), disability (r = -0.57), anxiety (r = -0.50), and depression (r = -0.46 to -0.51).(47)

Responsiveness. The responsiveness of the Spanish or English ASES-8 have not reported, even though they have reported change in arthritis intervention studies.(54) The German ASES documented medium size change (effect size = 0.31) in a sample of 43 people with fibromyalgia in a clinical setting.(48) ASES-AS reported a moderate level of responsiveness to change in people reporting improvement in health, with a mean score of 0.71 and 0.43 for AS-specific health and general health improvement, respectively.(46)

Minimally important differences. No evidence found.

Generalizability. Validation studies support the use of ASES-8 with adults who have different forms of arthritis (particularly rheumatoid arthritis, osteoarthritis, and ankylosing spondylosis) and fibromyalgia. There are no studies using representative samples of populations with arthritis or fibromyalgia in the development, refinement, and evaluation of ASES-8.

Use in clinical trials. A 2019 meta-analysis identified 47 studies with sample sizes of minimum 50 participants that used ASES or ASES-8 to assess self-efficacy of adults with rheumatoid arthritis or knee osteoarthritis and reported on the scores cross-sectional association with three key arthritis outcomes: impairment, pain severity and emotional distress.(45) Five of those studies used the ASES-8. Among the 48 samples identified in the 47 studies, self-efficacy scores were significantly associated with the three arthritis outcomes. ASES-8 had significantly weaker associations than ASES with impairment but had statistically similar associations with

pain severity. No intervention studies were found looking at the association of ASES-8 with emotional distress.

Critical appraisal of overall value to the rheumatology community

Strengths. The ASES-8 is one of the most commonly used scales for measuring self-efficacy among adults with arthritis or fibromyalgia. Being disease-specific allows it to capture key elements of self-efficacy relevant to arthritis and fibromyalgia. Its low number of items makes it appealing for use in both clinical and research settings. It has been translated, culturally adapted, and validated in a few languages, which allows for its use among diverse language groups. The validation studies across version of the ASES-8 demonstrates it high internal consistency, good reliability, and significant correlations in a theoretical way to important measures of health outcome.

Caveats and cautions. Some studies that investigated the measurement properties of the ASES-8 use relatively small sizes, such as the validation of the Arabic version among 67 participants of which 79% were female. Some studies were limited in the healthcare settings from which participants were recruited and the representativeness of the sample for arthritis populations. There is a dearth of information about how to interpret the scores from the ASES-8, as there are no published clinical cut points, values of clinical meaningful change, and scale bias (or differential item functioning). The ASES-8 does not assume interval-level or standardized scoring.

Clinical usability. Clinical settings would appreciate the low administrative and respondent burden of ASES-8, but an inability to easily interpret the scores of patients is a substantive barrier to its usefulness.

Research usability. Clinical research on people with arthritis and fibromyalgia would benefit from the low administrative and respondent burden of ASES-8, especially in scenarios where multiple measures are administered to patients. The availability of several validated language versions allows ASES-8 to be use with diverse language groups. It has demonstrated good internal consistency and validity. Its test-retest reliability is inconsistent and several of its measurement properties related to interpretability should be investigated and confirmed before it can be implemented as a primary outcome measure in randomized controlled trials.

3 PROMIS® Measures of Self-Efficacy for Managing Chronic Conditions (PROMIS-SE)

Description

Purpose. The PROMIS-SE measures were designed to evaluate the confidence of patients in performing important tasks and behaviors for managing chronic diseases.(55) They are not disease-specific but are generic measures of self-efficacy that enable comparisons across people with different chronic diseases. The first validation study was published in 2016 for the English version measures from one of five domains,(55) followed by a publication of English version measures for all of the five domains in 2017.(56)

Content or domains. PROMIS-SE measures were developed as item banks and short forms of five domains to measure one's belief about their ability to carry out behaviors to reach their health goals.(56) The five domains of self-efficacy for managing chronic conditions include self-efficacy for managing: Daily Activities, Symptoms, Medications and Treatments, Emotions, and Social Interactions.(56) These measures were developed using a standard guide by the PROMIS, which included defining 'self-efficacy for managing chronic conditions' as "an individual's confidence in his/her ability to successfully perform specific tasks or behaviors related to one's health in a variety of situations", identifying the domains, reviewing and confirming the domains and content areas through 12 patient focus groups, identifying and selecting items by the research team, and cognitive testing with 30 patients.(56)

Number of items. Each domain of PROMIS-SE has three test forms: an item bank for Computerized Adaptive Test (CAT), and a stable eight-item and four-item short form. The item bank for Daily Activities, Symptoms, Medications and Treatments, Emotions, and Social Interactions each contains 35, 25, 26, 23, and 28 items respectively.(56) Respondents are required to answer between four and 12 items in order to obtain a CAT score for each domain.

Response options/scale. Five-point rating scale for levels of confidence: 1 (I am not at all confident); 2 (I am a little confident); 3 (I am somewhat confident); 4 (I am quite confident); and 5 (I am very confident).

Recall period for items. There is no recall period. Captures one's current level of confidence for performing activities and behaviors.

Cost to use. Protected by copyright, and all versions in English and Spanish are free to use without permission. Translations in other languages may require permission for access and subjected to a distribution fee.

How to obtain. Available along with the user manual at PROMIS website: http://www.healthmeasures.net/explore-measurement-systems/promis. (p1)

Practical application

Method of administration. Self-administered by paper, computer, and an app. CAT is not available on paper.

Scoring. The final scores for each domain are T_{clin}-scores. These are T-scores (centered at 50 with a standard deviation of 10) based on the distribution of responses from a US population sample with chronic diseases rather than the general population.(56) CAT directly provides users with T_{clin}-scores. An individual's T_{clin}-scores for short forms can be obtained simply by adding their responses and using the corresponding T_{clin}-scores and standard error in a conversion scoring table available on the PROMIS website: http://www.healthmeasures.net/images/PROMIS/manuals/PROMIS_Self_Efficacy_Managing_Chronic_Conditions_Scoring_Manual.pdf. Each item must have a response in order to obtain valid scores from the scoring tables. When a respondent skipped a question, the HealthMeasures Scoring Service (https://www.assessmentcenter.net/ac_scoringservice) can be used to generate the T_{clin}-scores. The use of the HealthMeasures Scoring Service produces the most accurate scores because it accounts for the item-level calibrations and response patterns for the concept being measured. The manual provides instructions on dealing with multiple responses to one item by an individual.

Score interpretation. Higher scores indicate higher levels of self-efficacy for each domain. The T_{clin}-scores of each domain were calibrated on a different T_{clin}-score distribution, which resulted in different T_{clin}-score ranges across the domains.(56, 57) The maximum T_{clin}-scores across the domains correspond to different portions of the population.(56, 57) T_{clin}-scores can be compared to see how far away a person lies from the average score for a domain. T_{clin}-scores of greater than 50 means persons reported higher levels of the domain-specific self-efficacy than the average in the general chronic disease population. The T_{clin}-scores should not

be compared across domains because the average scores and standard deviations of scores can vary across them.(56, 57)

Respondent time to complete. Not reported. The time would vary based on the number of items. The maximum number of items is 12, the response scales are the same across all items, and the instructions and items are not wordy, which leads us to estimate a maximum of five minutes to complete each measure. It is likely that 4-item measures would require less than three minutes.

Administrative burden. The short forms require no training, software, or special equipment to administer. Scoring time is required for calculating the total score by adding the responses and then converting the total to a T_{clin}-score. When using the CAT, specialize software and training are may be required.

Translations/adaptations. The measures are available in English and French on the PROMIS website. There are no reported validation studies on the French version.

Psychometric information

Floor and ceiling effects. No evidence was found for the final scores/T_{clin}-score of each domain. The 35 items of the Daily Activities domain were reported in the first published study to be mostly rated in the highest rating category and demonstrated ceiling effects (range 33.1–86.0 %).(55)

Reliability. *Internal consistency*. The Cronbach's α of domain measures varied from 0.96 to 0.97 for the full item banks, 0.90 to 0.95 for the 8-item short forms, and 0.85 to 0.92 for the 4-item short forms.(56) All short forms correlated highly with the full item banks (r >0.90 for 8-item short forms, r >.0.85 for 4-item short forms).(56) All measures demonstrated good item-total correlations (r >0.55 8-item short forms, r >0.85 for 4-item short forms).(56)

Test-retest. No evidence found.

Validity. *Content/face*. Limited information stated that "development of definition, identification of domains, patient focus groups, and qualitative item review based on cognitive interviews" of the domains and item banks followed standard guidelines set by PROMIS.(54) The final items were reviewed by the research team and the cognitive interviews involved 30 patients.(39)

Criterion. No evidence found.

Construct. Among 1087 participants, 33.3 percent with arthritis or rheumatism, statistically significant Pearson's correlation coefficients were observed for PROMIS-SE domains' final scores with other instruments. There were positive correlations between each domain's item back with another measure of self-efficacy: Daily Activities (r = 0.65), Symptoms (r = 0.76), Emotions (r = 0.69), Medications and Treatments (r = 0.57), and Social Interactions (r = 0.62). There were positive correlations between each domain and a measure of physical function, which was 0.78 for Daily Activities and varied between 0.38 to 0.48 for the other domains. There were positive correlations across all the domains with a measure of global mental health (r = 0.52 to 0.70). The domains' scores were negative correlations with a measure of fatigue (r = -0.39 to -0.54) and of depression (r = -0.48 to -0.69). The domains' scores were negatively correlations with disease severity (r = -0.16 to -0.43).

One study confirmed that the PROMIS-SE has a multidimensional construct made up of the five domains that are correlation and, therefore, are inter-dependent domains.(56) No items demonstrated differential item functioning when tested for age, sex, race, or data source (general medical conditions versus chronic neurologic conditions).(56) The measures were able to characterize patient acceptable symptom state (PASS) status with moderate accuracy (70%) for a large portion of patients (70%) among sample of 94 patients who attended a rural primary care clinic.(58)

Responsiveness. No evidence found.

Minimally important differences. No evidence found.

Generalizability. While the PROMIS-SE measures are generic, their development and validation limit their use to clinical populations with chronic conditions.(56) They are not disease-specific, and respondents must be adults (aged 18 years or older) with at least one chronic condition.

Use in clinical trials. We identified a pilot RCT and an RCT with arthritis populations that used the PROMIS-SE. Both used the PROMIS-SE as a primary outcome measure.(59, 60) The pilot study found people with rheumatoid arthritis had improve self-management behavior (i.e. self-efficacy) after using a mobile app with hand optical imaging capabilities compared to a control group.(59) In the RCT, assessment with the 8-item short form of each domain found an internet-based self-management program with not superior to a patient-focused educational book for improving self-efficacy of people with systemic sclerosis.(60) Domains of the PROMIS-SE

have been reported as variable of interest in ongoing RCT for people with cancer, and a published clinical trial for people with multiple sclerosis.(61-63)

Critical appraisal of overall value to the rheumatology community

Strengths. PROMIS-SE measures are suitable for use in a variety of chronic disease populations. The domains allow for focus on separate aspects of self-efficacy. The measures can be administered using CAT for more accurate estimates and the total scores are standardized to a reference population. One study of patients with a number of chronic neurologic conditions found that PROMIS-SE measures are better predictors of mental health, disability, and quality of life than disease severity or diagnosis. Patients final scores are interpretable against the reference sample using standardized scores.(64) The PROMIS-SE measures may allow for interpretation of self-efficacy scores across disease and demographic groups.

Caveats and cautions. The PROMIS-SE measures need further psychometric testing to establish additional measurement properties including clinical meaningful difference and clinical cut points. There are no validation studies on any translations for their use in other language groups.

Clinical usability. Self-efficacy is foundational to patient's decision to take control of their health by performing specific healthy behaviors to manage their illness or symptoms to achieve more favourable outcomes. Self-efficacy for managing chronic conditions as measured by the PROMIS-SE measures is a major determinant of health outcomes and would be valuable in assessing self-efficacy for identifying people at risk for poor health outcomes in order to provide them with targeted treatment and therapy.(64) The PROMIS-SE measures, particularly the short forms, have very attractive levels of respondent and respondent burden for domain-specific self-efficacy assessment in clinical settings.

Research usability. The PROMIS-SE measures enables the exploration of interventions and tools to improve self-management within and across chronic health conditions. It allows for evaluating interventions designed to improve self-management and increasing patients' confidence in carrying out healthy behaviors.(59) Further research is needed on the interpretability of the PROMIS-SE domains, such as clinical cut-point/thresholds scores and values clinical meaningful change.

4 PROMIS® Measures of General Self-Efficacy

Description

Purpose. The PROMIS General Efficacy measures were "designed to assess a person's belief in his/her [their] capacity to manage daily stressors and have control over meaningful events."(65) They were first published in English in 2019.(65)

Content or domains. They were developed by fitting the 10 items of the NIH Toolbox® Self-Efficacy Item Bank with levels of "confidence" for their response scale options. The NIH Toolbox Self-Efficacy Item Bank is a 10-item that was derived from the widely used General Self-Efficacy Scale that has been translated in over 30 languages.(66) The NIH Tool Self-Efficacy Item Bank was calibrated with 1111 participants aged 18 years and older from the US general population. The NIH Tool Self-Efficacy Item Bank fits a one-factor model and has demonstrated high internal consistency (Cronbach $\alpha = 0.93$) and no differential item functioning.(67) There are two test forms of the PROMIS General Self-Efficacy measures: an item bank and a short form.

Number of items. The item bank has 10 items and the short form has four items.

Response options/scale. Five-point rating scale for levels of confidence: 1 (I am not at all confident); 2 (I am a little confident); 3 (I am somewhat confident); 4 (I am quite confident); and 5 (I am very confident).

Recall period for items. There is no recall period. Used to capture current level of confidence performing activities and behaviors.

Cost to use. Protected by copyright, and both test forms are available in English and Spanish to use for free without permission. Translations in other languages may require permission for access and subjected to a distribution fee.

How to obtain. Available along with the user manual at PROMIS website: http://www.healthmeasures.net/explore-measurement-systems/promis.

Practical application

Method of administration. Self-administered by paper, computer, and an app. CAT is not available on paper.

Scoring. The final scores for each test form are T-scores (centered at 50 with a standard deviation of 10) based on the distribution of responses from representative sample the general population. CAT directly provides users with T-scores.(65) The T-scores for short forms can be determined by adding an individual's responses and using the corresponding T-scores and standard error in a conversion scoring table available on the PROMIS website: http://www.healthmeasures.net/images/PROMIS/manuals/PROMIS_Self_Efficacy_Managing_Chronic_Conditions_Scoring_Manual.pdf. Each items must have a response in order to obtain a valid score using the scoring tables. When a participant has skipped a question, the HealthMeasures Scoring Service (https://www.assessmentcenter.net/ac_scoringservice) can be used to generate the T-scores. As with the PROMIS-SE measures, the use of the HealthMeasures Scoring Service produces the most accurate scores because it accounts for the item-level calibrations and response pattern for the concept being measured. The manual provides instructions on dealing with multiple responses to one item from by an individual.

Score interpretation. Higher scores indicate higher levels of global self-efficacy. T-scores of greater than 50 means persons reported higher levels of self-efficacy than on average in the general population.(65)

Respondent time to complete. Completion time would vary based on the number of items. The maximum number of items is 10, the response scales are the same across all items, and the instructions and items are not wordy, which leads us to estimate a maximum of five minutes to complete.

Administrative burden. The short forms require no training, software, or special equipment to administer. Scoring time is required for calculating the total score and converting it to a T-score. When using the CAT, specialize software and training may be required.

Translations/adaptations. The measures are available in English, French and Spanish, but only the English versions has published results of its testing.

Psychometric information

Floor and ceiling effects. No evidence found.

Reliability. *Internal consistency*. Cronbach $\alpha = 0.94$ for the 10-item bank and 0.88 for the 4-item short form.(65)

Test-retest. No evidence found.

Validity. *Content/face*. Limited information provided on cognitive testing of the items among 20 patients with cancer. Majority of the items were easy to understand and logical. Two of the items were found to be difficult to respond to, but kept in the item bank because they are in the legacy general self-efficacy measure.(65)

Criterion. No evidence found.

Construct. Statistically significant Pearson's correlation coefficients were between 0.52 and 0.58 with the Life Orientation Test-Revised (LOT-R) and Generalized Expectancy for Success Scale (GESS).(65)

The item bank was confirmed to represent a single dominant factor that is unidimensional and accounts for 69% of the explained variance when assessing self-efficacy. The measures displayed no differential item functioning for age, gender, education, and race.(65)

Responsiveness. No evidence found.

Minimally important differences. No evidence found.

Generalizability. It is generic measure of self-efficacy and has been calibrated with a representative sample of adults (aged 18 years or older) from a general US population.(65)

Use in clinical trials. No evidence found.

Critical appraisal of overall value to the rheumatology community

Strengths. Unlike other existing measures of general self-efficacy, the PROMIS General Self-Efficacy measures used a large diverse sample from the general population to calibrate the items along the same metric and are confirmed to be free of DIF.(65) This enable the items to be used for CAT, which minimizes respondents time to complete the item bank measure without losing measurement precision. Furthermore, the T-scores facilitates easy interpretation of scores.

Caveats and cautions. The PROMIS General Self-Efficacy measures are new with no independent publications on their use or validity. Their measurement properties have not been assessed for use in clinical settings. No values have been published for cut-points for clinical utility and thresholds for clinical meaningful change.

Clinical usability. The use of these measures in clinical settings have not yet been explore. These measures could be used to guide clinicians in implementing strategies to improve self-efficacy in clients with arthritis conditions with reference to general population. This could lead to better health outcomes for those clients.

Research usability. The PROMIS General Self-Efficacy enables the exploration of interventions and tools to improve self-management in healthy individuals and individuals with a variety of acute and chronic illnesses. While the item bank has demonstrated excellent internal consistency and good construct validity, it is not clear how reliable are clinically meaningful it's scores are.

Conclusion

The measures reported in this review each have their benefits or challenges for use when assessing people with rheumatic diseases in clinical and research settings. The PAM-13 is the only measure reported for measuring patient activation, a concept that goes beyond confidence in one's ability as captured by tools of self-efficacy. PAM-13 extends beyond confidence to capture one's skills and knowledge for self-management of their health condition. The ASES-13 is the only disease-specific measure, while the other measures are generic for diagnosis. Both the domain-specific and general PROMIS self-efficacy measures are relatively new, and while demonstrating strong measurement properties for internal consistency and construct validity, have not yet demonstrated test-retest reliability.

All of the measures have reasonable respondent and administrative burden. The PAM-13 is best situated for use because of the readily available qualitative interpretation of its final scores, with cautions for interpretation across language versions. It could also benefit from arthritis-specific calibration and interpretation of it scores but has been extensively used even among people with arthritis. The PROMIS measures while having standardized scores that would facilitate easy interpretation, has not yet established clinical thresholds and qualitative interpretation of those scores. The ASES-13 is least poised for interpretation of its score, and is the only measure not benefiting from calibration for interval-level scoring. Interval-level scoring provides more confidence in interpreting the final scores obtained for individual patients.

References

1. Hibbard JH, Stockard J, Mahoney ER, Tusler M. Development of the Patient Activation Measure (PAM): conceptualizing and measuring activation in patients and consumers. Health services research. 2004;39(4p1):1005-26.

2. Bandura A. Human agency in social cognitive theory. Am Psychol. 1989;44(9):1175-84.
3. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev. 1977;84(2):191-215.
4. 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. Quality of Life Research. 2009;18(8):1115-23.
5. Brady TJ. Measures of self-efficacy: Arthritis Self-Efficacy Scale (ASES), Arthritis Self-Efficacy Scale-8 Item (ASES-8), Children's Arthritis Self-Efficacy Scale (CASE), Chronic Disease Self-Efficacy Scale (CDSES), Parent's Arthritis Self-Efficacy Scale (PASE), and Rheumatoid Arthritis Self-Efficacy Scale (RASE). Arthritis Care Res (Hoboken). 2011;63 Suppl 11:S473-85.
6. Hibbard JH, Mahoney ER, Stockard J, Tusler M. Development and testing of a short form of the patient activation measure. Health services research. 2005;40(6 Pt 1):1918-30.
7. Green CA, Perrin NA, Polen MR, Leo MC, Hibbard JH, Tusler M. Development of the Patient Activation Measure for mental health. Adm Policy Ment Health. 2010;37(4):327-33.
8. Moljord IE, Lara-Cabrera ML, Perestelo-Pérez L, Rivero-Santana A, Eriksen L, Linaker OM. Psychometric properties of the Patient Activation Measure-13 among out-patients waiting for mental health treatment: A validation study in Norway. Patient Educ Couns. 2015;98(11):1410-7.
9. Skolasky RL, Green AF, Scharfstein D, Boult C, Reider L, Wegener ST. Psychometric properties of the patient activation measure among multimorbid older adults. Health Serv Res. 2011;46(2):457-78.
10. Magnezi R, Glasser S. Psychometric properties of the hebrew translation of the patient activation measure (PAM-13). PLoS One. 2014;9(11):e113391.
11. Roberts NJ, Kidd L, Dougall N, Patel IS, McNarry S, Nixon C. Measuring patient activation: The utility of the Patient Activation Measure within a UK context-Results from four exemplar studies and potential future applications. Patient Educ Couns. 2016;99(10):1739-46.
12. Kinney RL, Lemon SC, Person SD, Pagoto SL, Saczynski JS. The association between patient activation and medication adherence, hospitalization, and emergency room utilization in patients with chronic illnesses: a systematic review. Patient education and counseling. 2015;98(5):545-52.

13. Rademakers J, Nijman J, van der Hoek L, Heijmans M, Rijken M. Measuring patient activation in The Netherlands: translation and validation of the American short form Patient Activation Measure (PAM13). BMC Public Health. 2012;12:577.
14. Ngooi BX, Packer TL, Kephart G, Warner G, Koh KW, Wong RC, et al. Validation of the Patient Activation Measure (PAM-13) among adults with cardiac conditions in Singapore. Qual Life Res. 2017;26(4):1071-80.
15. Graffigna G, Barello S, Bonanomi A, Lozza E, Hibbard J. Measuring patient activation in Italy: Translation, adaptation and validation of the Italian version of the patient activation measure 13 (PAM13-I). BMC Med Inform Decis Mak. 2015;15:109.
16. Kosar C, Besen DB. Adaptation of a patient activatıon measure (PAM) into Turkish: reliability and validity test. African health sciences. 2019;19(1):1811-20.
17. Ngooi BX, Packer TL, Warner G, Kephart G, Koh KWL, Wong RCC, et al. How adults with cardiac conditions in Singapore understand the Patient Activation Measure (PAM-13) items: a cognitive interviewing study. Disabil Rehabil. 2018;40(5):587-96.
18. Zill JM, Dwinger S, Kriston L, Rohenkohl A, Härter M, Dirmaier J. Psychometric evaluation of the German version of the Patient Activation Measure (PAM13). BMC Public Health. 2013;13:1027.
19. Brenk-Franz K, Hibbard JH, Herrmann WJ, Freund T, Szecsenyi J, Djalali S, et al. Validation of the German version of the patient activation measure 13 (PAM13-D) in an international multicentre study of primary care patients. PLoS One. 2013;8(9):e74786.
20. Stepleman L, Rutter MC, Hibbard J, Johns L, Wright D, Hughes M. Validation of the patient activation measure in a multiple sclerosis clinic sample and implications for care. Disabil Rehabil. 2010;32(19):1558-67.
21. Ahn Y-H, Yi C-H, Ham O-K, Kim B-J. Psychometric properties of the Korean version of the “Patient Activation Measure 13”(PAM13-K) in patients with osteoarthritis. Evaluation & the health professions. 2015;38(2):255-64.
22. DeCamp LR, Leifheit K, Shah H, Valenzuela-Araujo D, Sloand E, Polk S, et al. Cross-cultural validation of the parent-patient activation measure in low income Spanish- and English-speaking parents. Patient Educ Couns. 2016;99(12):2055-62.

23. Cunha CM, Nepomuceno E, Manzato RO, Cunha DCPT, Silva DD, Dantas RAS. Cultural adaptation and validation of the Brazilian Version of the Patient Activation Measure-22 items. Rev Bras Enferm. 2018;71(4):1891-8.
24. Packer TL, Kephart G, Ghahari S, Audulv Å, Versnel J, Warner G. The Patient Activation Measure: a validation study in a neurological population. Qual Life Res. 2015;24(7):1587-96.
25. Maindal HT, Sokolowski I, Vedsted P. Translation, adaptation and validation of the American short form Patient Activation Measure (PAM13) in a Danish version. BMC Public Health. 2009;9:209.
26. Hung M, Carter M, Hayden C, Dzierzon R, Morales J, Snow L, et al. Psychometric assessment of the patient activation measure short form (PAM-13) in rural settings. Qual Life Res. 2013;22(3):521-9.
27. Fowles JB, Terry P, Xi M, Hibbard J, Bloom CT, Harvey L. Measuring self-management of patients' and employees' health: further validation of the Patient Activation Measure (PAM) based on its relation to employee characteristics. Patient education and counseling. 2009;77(1):116-22.
28. Prey JE, Qian M, Restaino S, Hibbard J, Bakken S, Schnall R, et al. Reliability and validity of the patient activation measure in hospitalized patients. Patient Educ Couns. 2016;99(12):2026-33.
29. Skolasky RL, Mackenzie EJ, Riley LH, Wegener ST. Psychometric properties of the Patient Activation Measure among individuals presenting for elective lumbar spine surgery. Qual Life Res. 2009;18(10):1357-66.
30. Rademakers J, Maindal HT, Steinsbekk A, Gensichen J, Brenk-Franz K, Hendriks M. Patient activation in Europe: an international comparison of psychometric properties and patients' scores on the short form Patient Activation Measure (PAM-13). BMC Health Serv Res. 2016;16(1):570-.
31. Moreno-Chico C, González-de Paz L, Monforte-Royo C, Arrighi E, Navarro-Rubio MD, Gallart Fernández-Puebla A. Adaptation to European Spanish and psychometric properties of the Patient Activation Measure 13 in patients with chronic diseases. Fam Pract. 2017;34(5):627-34.

32. Munson GW, Wallston KA, Dittus RS, Speroff T, Roumie CL. Activation and perceived expectancies: correlations with health outcomes among veterans with inflammatory bowel disease. J Gen Intern Med. 2009;24(7):809-15.
33. Lubetkin EI, Lu WH, Gold MR. Levels and correlates of patient activation in health center settings: building strategies for improving health outcomes. J Health Care Poor Underserved. 2010;21(3):796-808.
34. Gong HS, Park JW, Shin YH, Kim K, Cho KJ, Baek GH. Use of a decision aid did not decrease decisional conflict in patients with carpal tunnel syndrome. BMC Musculoskelet Disord. 2017;18(1):118-.
35. Cooper V, Clatworthy J, Harding R, Whetham J, Em EC. Measuring empowerment among people living with HIV: a systematic review of available measures and their properties. AIDS Care. 2019;31(7):798-802.
36. Greene J, Hibbard JH, Sacks R, Overton V, Parrotta CD. When patient activation levels change, health outcomes and costs change, too. Health Aff (Millwood). 2015;34(3):431-7.
37. Linden A. Estimating Measurement Error of the Patient Activation Measure for Respondents with Partially Missing Data. Biomed Res Int. 2015;2015:270168.
38. Gao J, Arden M, Hoo ZH, Wildman M. Understanding patient activation and adherence to nebuliser treatment in adults with cystic fibrosis: responses to the UK version of PAM-13 and a think aloud study. BMC Health Serv Res. 2019;19(1):420-.
39. Lorig K, Chastain RL, Ung E, Shoor S, Holman HR. Development and evaluation of a scale to measure perceived self-efficacy in people with arthritis. Arthritis Rheum. 1989;32(1):37-44.
40. Brand E, Nyland J, Henzman C, McGinnis M. Arthritis self-efficacy scale scores in knee osteoarthritis: a systematic review and meta-analysis comparing arthritis self-management education with or without exercise. J Orthop Sports Phys Ther. 2013;43(12):895-910.
41. Miles CL, Pincus T, Carnes D, Taylor SJ, Underwood M. Measuring pain self-efficacy. Clin J Pain. 2011;27(5):461-70.
42. González VM, Stewart A, Ritter PL, Lorig K. Translation and validation of arthritis outcome measures into Spanish. Arthritis Rheum. 1995;38(10):1429-46.
43. Garratt AM, Klokkerud M, Løchting I, Hagen KB. Rasch analysis of the Norwegian version of the Arthritis Self-Efficacy Scale (ASES). Scand J Rheumatol. 2017;46(1):33-9.

44. Wilcox S, Schoffman DE, Dowda M, Sharpe PA. Psychometric properties of the 8-item english arthritis self-efficacy scale in a diverse sample. Arthritis. 2014;2014:385256.
45. Silva R, Silva FC, Meireles SM, Natour J. Translation to Brazilian Portuguese, cultural adaptation and psychometric properties of 8-item ArthritisSelf-Efficacy Scale (ASES-8). Sao Paulo Med J. 2019;137(1):6-12.
46. Sandhu J, Packham JC, Healey EL, Jordan KP, Garratt AM, Haywood KL. Evaluation of a modified arthritis self-efficacy scale for an ankylosing spondylitis UK population. Clin Exp Rheumatol. 2011;29(2):223-30.
47. Arab Alkabeya H, Daibes J, Hughes AM, Adams J. The Arabic Arthritis Self-Efficacy Scale-8 (ASES-8): a valid and reliable measure of evaluating self-efficacy in Palestinian patients with rheumatoid arthritis. Disabil Rehabil. 2020:1-7.
48. Mueller A, Hartmann M, Mueller K, Eich W. Validation of the arthritis self-efficacy short-form scale in German fibromyalgia patients. Eur J Pain. 2003;7(2):163-71.
49. Gao L, Zhang XC, Li MM, Yuan JQ, Cui XJ, Shi BX. Psychometric properties of the Chinese version of Arthritis Self-Efficacy Scale-8 (ASES-8) in a rheumatoid arthritis population. Rheumatol Int. 2017;37(5):751-6.
50. Adegoke BOA, Odole AC, Adekunle-Balogun AT, Umar I. Validation of a Yoruba version of the arthritis self-efficacy scale. Afr J Med Med Sci. 2016;45(2):189-95.
51. Lomi C, Nordholm LA. Validation of a Swedish version of the Arthritis Self-efficacy Scale. Scand J Rheumatol. 1992;21(5):231-7.
52. Lomi C, Burckhardt C, Nordholm L, Bjelle A, Ekdahl C. Evaluation of a Swedish version of the arthritis self-efficacy scale in people with fibromyalgia. Scand J Rheumatol. 1995;24(5):282-7.
53. Bareyre L, Gay C, Coste N, Bonnin A, Pereira B, Coudeyre E. French validation of the Arthritis Self-Efficacy Scale and further psychometric properties exploration among 168 people with osteoarthritis. Clin Rehabil. 2019;33(3):546-56.
54. Jackson T, Xu T, Jia X. Arthritis self-efficacy beliefs and functioning among osteoarthritis and rheumatoid arthritis patients: a meta-analytic review. Rheumatology (Oxford). 2020;59(5):948-58.

55. Hong I, Velozo CA, Li CY, Romero S, Gruber-Baldini AL, Shulman LM. Assessment of the psychometrics of a PROMIS item bank: self-efficacy for managing daily activities. Qual Life Res. 2016;25(9):2221-32.
56. Gruber-Baldini AL, Velozo C, Romero S, Shulman LM. Validation of the PROMIS(®) measures of self-efficacy for managing chronic conditions. Qual Life Res. 2017;26(7):1915-24.
57. Lee MJ, Romero S, Velozo CA, Gruber-Baldini AL, Shulman LM. Multidimensionality of the PROMIS self-efficacy measure for managing chronic conditions. Qual Life Res. 2019;28(6):1595-603.
58. Houck J, Kang D, Cuddeford T, Rahkola S. Ability of Patient-Reported Outcomes to Characterize Patient Acceptable Symptom State (PASS) After Attending a Primary Care Physical Therapist and Medical Doctor Collaborative Service: A Cross-Sectional Study. Arch Phys Med Rehabil. 2019;100(1):60-6.
59. Mollard E, Michaud K. A Mobile App With Optical Imaging for the Self-Management of Hand Rheumatoid Arthritis: Pilot Study. JMIR Mhealth Uhealth. 2018;6(10):e12221.
60. Khanna D, Serrano J, Berrocal VJ, Silver RM, Cuencas P, Newbill SL, et al. Randomized Controlled Trial to Evaluate an Internet-Based Self-Management Program in Systemic Sclerosis. Arthritis Care Res (Hoboken). 2019;71(3):435-47.
61. Yanez B, Baik SH, Oswald LB, Buitrago D, Buscemi J, Iacobelli F, et al. An Electronic Health Intervention for Latina Women Undergoing Breast Cancer Treatment (My Guide for Breast Cancer Treatment): Protocol for a Randomized Controlled Trial. JMIR Res Protoc. 2019;8(12):e14339.
62. Kratz AL, Alschuler KN, Ehde DM, von Geldern G, Little R, Kulkarni S, et al. A randomized pragmatic trial of telephone-delivered cognitive behavioral-therapy, modafinil, and combination therapy of both for fatigue in multiple sclerosis: The design of the "COMBO-MS" trial. Contemp Clin Trials. 2019;84:105821.
63. Benedict C, Ford JS, Schapira L, Simon P, Spiegel D, Diefenbach M. Family-building decision aid and planning tool for young adult women after cancer treatment: protocol for preliminary testing of a web-based decision support intervention in a single-arm pilot study. BMJ Open. 2019;9(12):e033630.

64. Shulman LM, Velozo C, Romero S, Gruber-Baldini AL. Comparative study of PROMIS(®) self-efficacy for managing chronic conditions across chronic neurologic disorders. Qual Life Res. 2019;28(7):1893-901.
65. Salsman JM, Schalet BD, Merluzzi TV, Park CL, Hahn EA, Snyder MA, et al. Calibration and initial validation of a general self-efficacy item bank and short form for the NIH PROMIS(®). Qual Life Res. 2019;28(9):2513-23.
66. Schwarzer R, Jerusalem M, Weinman J, Wright S, Johnston M. Measures in health psychology: A user's portfolio. Causal and control beliefs. Windsor, UK: Nfer-Nelson. 1995:35-7.
67. Kupst MJ, Butt Z, Stoney CM, Griffith JW, Salsman JM, Folkman S, et al. Assessment of stress and self-efficacy for the NIH Toolbox for Neurological and Behavioral Function. Anxiety Stress Coping. 2015;28(5):531-44.