

REVIEWS

Selecting Patient-Reported Outcome Measures to Contribute to Primary Care Performance Measurement: a Mixed Methods Approach

San Keller, PhD¹, Sydney Dy, MD^{2,3}, Renee Wilson, MS², Vadim Dukhanin, MD², Claire Snyder, PhD^{2,3}, and Albert Wu, MD^{2,3}

¹American Institutes for Research in the Behavioral and Social Sciences, Chapel Hill, NC, USA; ²Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA; ³Johns Hopkins School of Medicine, Baltimore, MD, USA.

New models of primary care include patient-reported outcome measures (PROMs) to promote patient-centered care. PROMs provide information on patient functional status and well-being, can be used to enhance care quality, and are proposed for use in assessing performance. Our objective was to identify a short list of candidate PROMs for use in primary care practice and to serve as a basis for performance measures (PMs). We used qualitative and quantitative methods to identify relevant patient-reported outcome (PRO) domains for use in performance measurement (PRO-PM) and their associated PROMs. We collected data from key informant groups: patients ($n=13$; one-on-one and group interviews; concept saturation analysis), clinical thought leaders ($n=9$; group discussions; thematic analysis), primary care practices representatives ($n=37$; six focus groups; thematic analysis), and primary care payer representatives ($n=10$; 12-question survey; frequencies of responses). We merged the key informant group information with findings from environmental literature scans. We conducted a targeted evidence review of measurement properties for candidate PROMs. We used a scoping review and key informant groups to identify PROM evaluation criteria, which were linked to the National Quality Forum measure evaluation criteria. We developed a de novo schema to score candidate PROMs against our criteria. We identified four PRO domains and 10 candidate PROMs: 3 for depressive symptoms, 2 for physical function, 3 for self-efficacy, 2 for ability to participate. Five PROMs met $\geq 70\%$ of the evidence criteria for three PRO domains: PHQ-9 or PROMIS Depression (depression), PF-10 or PROMIS-PF (physical functioning), and PROMIS Self-Efficacy for Managing Treatments and Medications (self-efficacy). The PROMIS Ability to Participate in Social Roles and Activities met 68% of our criteria and might be considered for inclusion. Existing evidence and key informant data identified 5 candidate PROMs to use in primary care. These instruments can be used to develop PRO-PMs.

KEY WORDS: primary care; patient-reported outcome measures; performance measurement; patient-centered care; PROM; PRO-PM; multiple chronic conditions; MCCs.

J Gen Intern Med 35(9):2687-97
DOI: 10.1007/s11606-020-05811-4
© Society of General Internal Medicine 2020

INTRODUCTION

The goal of primary care is to provide continuous, comprehensive, and coordinated care.^{1, 2} However, the addition of a population health perspective, along with increased specialization of care, can impede achievement of those goals.³ The number of clinical issues addressed during primary care visits has grown,⁴ along with the complexity of diagnostic testing and prescribing, increasing the need to coordinate with multiple providers.⁵ To reflect changes in what primary care providers (PCPs) do, new models of healthcare delivery have been implemented.^{6, 7} The Centers for Medicare and Medicaid Services (CMS) Innovation Center was created to test new models of healthcare delivery to improve care quality while lowering costs.⁸ The Innovation Center's advanced delivery model for primary care, Comprehensive Primary Care Plus (CPC+), was launched in 2017.

The goal of the CPC+ model is to promote coordinated, patient-centered care. To do so, the model employs innovative methods to measure and improve access and quality, including expanded use of the electronic health record (EHR) and patient-reported outcome measures (PROMs).⁹ The digitization of healthcare has provided opportunities to improve the patient-centeredness and quality of care¹⁰⁻¹², including incorporating PROM data into the medical record by collecting these data directly from patients electronically and merging the data into EHRs.^{13, 14}

PROMs provide reliable information coming directly from patients about what they are able to do (i.e., their functioning) and how they feel (i.e., their symptoms). Interest in incorporating PROMs into clinical practice is based on growing evidence that this information can help clinicians and patients to improve patient outcomes by supplementing information provided by traditional clinical measures.¹⁵⁻¹⁹ PROMs also

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s11606-020-05811-4>) contains supplementary material, which is available to authorized users.

Received October 22, 2019

Accepted March 17, 2020

Published online June 3, 2020



have the potential to serve as a measure of healthcare quality, such as the amount of improvement in functioning, or reduction in symptoms, that occurs over a period of time.^{20, 21} Moreover, deployment of patient-reported outcome performance measures (PRO-PMs) could enable a shift to providing the goal-oriented care especially needed for people with multiple chronic conditions (MCCs).²²

Recognizing the potential value of incorporating PROMs into an enhanced primary care delivery model, the Innovation Center sought to identify a short list of PROMs to consider for use in the CPC+ model. The research we detail here was designed to help the Innovation Center determine the one or two best PROMs for use in primary care to enhance patient care and for developing PRO-PMs to evaluate performance. We had four research objectives:

1. Identify high-priority, patient-reported outcome (PRO) domains that would provide useful information to guide PCPs, with the emphasis on health-related domains important to patients with MCCs
2. Identify existing PROMs for those domains
3. Identify criteria to evaluate and compare the candidate PROMs
4. Select, using the criteria, the one or two best PROMs for each PRO domain

METHODS

Overview

The sheer volume of possible PRO domains (for example, the Patient-Reported Outcome Measurement Information System [PROMIS] alone includes 97 PRO domains) and possible PROMs to assess them required a comprehensive, yet result-tailored, approach. We based our recommendations for PRO domains and PROMs on a synthesis of qualitative and quantitative data gathered from key informant groups and an environmental scan (Tables 1 and 2).

The purpose of the key informant groups was to identify the priority topics that stakeholders in primary care quality assessment thought should be addressed, factors they would advise us to consider when selecting a PROM, and their experiences with specific PROMs. The purpose of the environmental scan was to identify recommendations in the published and gray literature regarding (1) PRO domains for assessment in primary care, (2) PROMs to assess those PRO domains, and (3) criteria to apply to identify the best PROMs. After identifying a small subset of PRO domains and ten candidate PROMs, we conducted a targeted review of the evidence base to summarize the measurement properties of each candidate PROM. The relationship between our data sources and our research objectives is summarized in online appendix 1 and detailed below.

Key Informant Data Collection Methods

We tailored our methods for collecting key informant data to the preferences of each type of key informant group. The methods used to analyze the data were tailored to the nature of the data. Table 1 summarizes the characteristics of each key informant group and the characteristics of the data collection and analysis methods for each.

Methods for Targeted Environmental Scan, Evaluation of Yield, and Data Synthesis

We briefly describe below how we tailored targeted environmental scan strategies to each research objective. Details about the search strategies are available in online appendix 2.

Objective 1: Identify High-Priority PRO Domains for Primary Care. Before we could select candidate PROMs, we had to identify a subset of high-priority PRO domains that should be measured. To do so, we searched professional society websites, selected patient group websites, and published and gray literature describing primary care patient and clinician priorities.^{23–29} We also conducted searches in the Google Scholar using published search strategies developed for a similar purpose.³⁰ Next, we took a census of the PRO domains that emerged from these searches and identified those that were common across the sources. Finally, we compared this subset of domains to those identified via key informant data collection to select the subset of domains common to all or most sources.

Objective 2: Identify Candidate PROMs for PRO Domains. Once a subset of four high-priority PRO domains was selected, we developed a targeted PubMed-search strategy for each, conducted additional searches in EMBASE and PsycINFO, and searched reference lists of articles. For three of four PRO domains (ability to participate, depressive symptoms, and physical functioning), our initial searches for PROMs yielded far too many articles to review. For those domains, we searched for review articles published in the past 5 years. We did not find a published, comprehensive review of multiple PROMs for self-efficacy; however, we were able to obtain an unpublished review from the researcher (Dr. L. Shulman) who developed the self-efficacy measures for PROMIS. We supplemented this review with a search of the PROMIS measure database, as the Shulman review did not contain any PROMs appropriate for primary care. We excluded from consideration PROMs that would not be appropriate for patient self-report to assess primary care performance in the USA. Specific reasons for excluding PROMs from candidacy are provided below.

PROMs Excluded

-
- For which there was no English language translation
 - That had never been used in the USA
 - For which modes of administration did not include self-report
 - That were specific to a particular condition
-

We merged information on PROM candidates with information identified via key informant data collection.

Table 1 Key Informant Group Data Collection and Analysis Methods

Methods detail	Patients, n = 13	Clinical thought leaders, n = 9	CPC+ practice representatives, n = 37	CPC+ payer representatives, n = 10
Role or expertise of participants	<ul style="list-style-type: none"> • Patients with multiple chronic conditions • Patient advocates living with or caring for someone with multiple chronic conditions 	<ul style="list-style-type: none"> • Internal and geriatric medicine • Nursing • Primary care practice transformation • Healthcare quality improvement • PROM development and evaluation • Health policy 	<ul style="list-style-type: none"> • Administrator • Practitioner • Quality improvement staff 	<ul style="list-style-type: none"> • Regional payer representatives • National payer representatives
Recruitment considerations	Variation in geographic location, gender, age, race, and diagnosis	Variation in expertise	Variation in geographic region	Mix of regional and national payers Mix of experience with CPC+ model
Data collection method	<ul style="list-style-type: none"> • Semi-structured, one-on-one interviews • Semi-structured TEP group discussion 	Semi-structured TEP group discussion	Six semi-structured focus groups	12-question survey with closed and open-ended questions
Data analysis	<ul style="list-style-type: none"> • Interviews: Conducted a concept saturation analysis²⁵ • Group discussion: Extracted recommended criteria with illustrative quotes 	Created exhaustive list of PRO domains and PROM selection criteria that emerged	Summarized themes about PROM selection criteria and identified PROMs being used in narrative form	Calculated the frequencies of responses to questions and listed all unique responses

Objective 3: Identify Criteria to Evaluate PROMs. To identify PROM selection criteria, we conducted environmental scans^{31, 32} based on an initial set of 15 guidelines from standard setting

bodies detailing evaluation criteria for PROMs and quality measures. The criteria included in each document along with citations to the documents is presented in online appendix 4.

Table 2 Most Frequently Mentioned PRO Domains Across Key Informant Group Input and Environmental Scan of the Literature

PRO domain*	Patients (n = 13)	Clinical thought leaders (n = 9)	CPC+ practice representatives† (n = 37)	CPC+ payer representatives (n = 9)	Literature	Number of sources
Depression	✓	✓	✓	✓	✓	5
Physical function,‡	✓	✓	✓	✓	✓	5
ADLs	✓	✓	✓	✓	✓	4
Pain	✓	✓	✓	✓	✓	4
Self-efficacy§	✓	✓	✓	✓	✓	4
Ability to participate	✓	✓	✓	✓	✓	4
Cognitive ability	✓	✓	✓	✓	✓	3
Sleep	✓	✓	✓	✓	✓	3
Fatigue	✓	✓	✓	✓	✓	2
General anxiety	✓	✓	✓	✓	✓	2
Overall mental health	✓	✓	✓	✓	✓	2
Overall physical health	✓	✓	✓	✓	✓	2
Anxiety about health	✓	✓	✓	✓	✓	1
Musculoskeletal	✓	✓	✓	✓	✓	1
Overall general health	✓	✓	✓	✓	✓	1
Overall quality of life	✓	✓	✓	✓	✓	1
Health confidence¶	✓	✓	✓	✓	✓	1
Sexual functioning	✓	✓	✓	✓	✓	1
Social functioning	✓	✓	✓	✓	✓	1
Social isolation	✓	✓	✓	✓	✓	1
Symptom burden	✓	✓	✓	✓	✓	1

*Domains are listed alphabetically within frequency category

†These were PRO domains that practices were currently measuring with PROMs; practices were not queried about and did not spontaneously comment on which PRO domains should be measured. Practices also mentioned using drug abuse screening measures that were patient reported but are not used as outcome measures

‡Includes the ability to execute specific actions, perform tasks, and engage in simple and more demanding activities

§Self-efficacy for managing one's health/chronic condition

¶Ability to participate in social roles

*Perceived wellness/health confidence

We then searched the reference lists of that initial set to identify additional sources. Next, we conducted a content analysis of the yield of this review and the qualitative data obtained from key informants to identify a consensus set of evaluation criteria. Finally, we compared the resulting set to the measure evaluation criteria put forth by the National Quality Forum.

Objective 4: Select the Best PROMs for Primary Care. We developed a schema in which we assigned a PROM one point for each criterion for which there was supporting evidence. When the evidence for a criterion was mixed, the PROM received a half point. When there was no evidence or when the only evidence found was unfavorable to the PROM, the PROM received no point for that criterion. We did not use previously published methods to “score” the appropriateness of PROMs, because they lacked criteria specific to the application of the PROMs in primary care or performance measurement.^{33, 34}

RESULTS

Objective 1: PRO Domains

The row headings in Table 2 provide a list of all the PRO domains that were identified in the thematic analysis of interview and focus group results, confirmed in the survey of payers, or included in gray or published literature (see online appendix 3 for additional detail). The last column in Table 2 shows the degree of consensus, across different sources, on which PROs may be important to patient care. The PRO domains with the greatest support across stakeholders and literature were as follows:

- Ability to participate in social roles
- Depression
- Pain
- Physical function
- Self-efficacy for managing one’s health/chronic condition

We proceeded with identifying PROMs for each of these PRO domains except pain due to the current controversy over the potential unintended consequence of a mandate to assess pain.³⁵

Objective 2: Candidate PROMs

Initially, we identified a total of 503 PROMs for the four PRO domains. We eliminated approximately 85% of these ($n = 429$) because they did not meet selection criteria. We reviewed 74 PROMs in greater depth. Of these, we eliminated 64 because the PROM questions did not match the PRO domain or because using the PROM would not be feasible (that is, the PROM was too long or too difficult to obtain). Our PROM screening procedure is illustrated in Fig. 1.

After applying all the exclusions documented in Fig. 1, there were 10 PROMs for which we searched the primary literature to obtain additional evidence:

Ten PROMs Selected for Further Analysis

Ability to participate

- Keele Assessment of Participation (KAP)
- PROMIS Ability To Participate in Social Roles and Activities Short Forms

Depressive symptoms

- Patient Health Questionnaire-9 (PHQ-9)
- Center for Epidemiologic Studies Depression Scale—Revised (CESD-R)
- Patient-Reported Outcome Measurement Information System (PROMIS) Depression Short Forms (SFs) and Computer Adaptive Tests (CATs)

Physical function

- PF-10 (Short Form-36’s physical functioning 10-item subscale)
- PROMIS Physical Function Short Forms and CAT

Self-efficacy

- Self-Efficacy to Manage Chronic Disease Scale (SEMCD)
 - PROMIS Self-Efficacy for Managing Symptoms Short Forms and CAT
 - PROMIS Self-Efficacy for Managing Medications and Treatments Short Forms and CAT
-

Objective 3: PROM Selection Criteria

We identified 15 measure standards documents and related resources which provide PROM evaluation criteria (summarized in Table 3). A content analysis of the evaluation criteria across these 15 documents is presented in online appendix 4. Online appendix 4 also details how the PROM selection criteria described in the 15 measure standards documents relate to the five broad categories of the NQF quality measure evaluation criteria: (1) importance, (2) scientific acceptability, (3) feasibility, (4) usability and use, and (5) related and competing measures. We describe key considerations related to each NQF criterion below as well as additional guidance regarding these criteria provided by key informant groups.

Importance. Primary care PRO domains should be relevant and meaningful to stakeholders. Practice representatives and patients expressed the desire to choose PROMs that meet specific health-related patient needs. Clinical thought leaders, practice representatives, and payer representatives identified the importance of selecting a PROM whose results lead to specific clinical actions, are useful for tracking patients’ health, and identify gaps in care. They recommended that (1) PROM and PRO-PM scores be sensitive to a change in clinical practice and (2) scores distinguish between high and low performing primary care practices.

Scientific Acceptability. Each of the 15 standards documents included scientific acceptability as a PROM selection criterion. Aspects of scientific acceptability included (1) availability of a conceptual and measurement model for the PROM; (2) empirical evidence for the reliability, validity, and responsiveness of PROM scores; and (3) availability of aids to

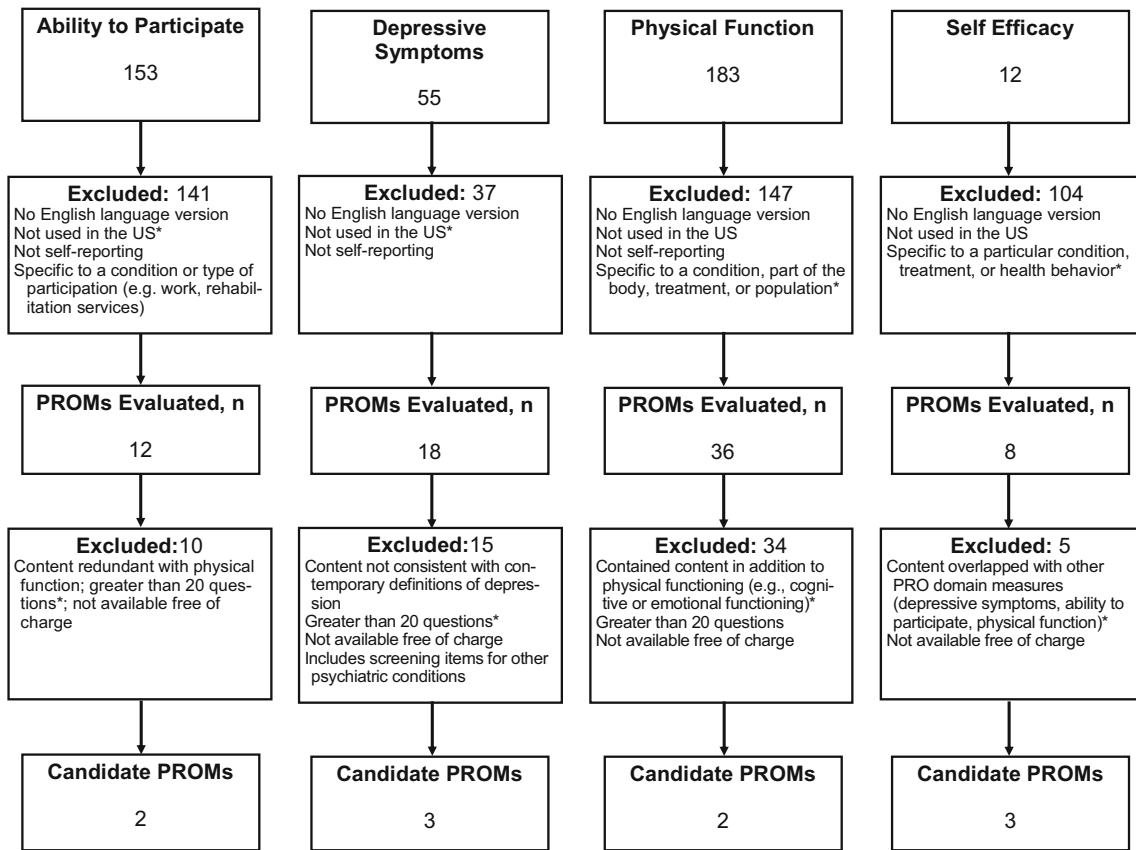


Figure 1 Flow of published evidence, by domain, through the review process. The asterisk indicates the most common reason for exclusion. US, United States of America.

support score interpretation. The availability of interpretive aids improves feasibility and usability as well. For example, clinical thought leaders indicated that a PROM would be useful if it produced results that were easy to understand and had interpretive aids for PROM scores such as graphs.

Feasibility. Feasibility refers to ease of implementing the PROM in primary care practice. For example, PROM length, availability of different formats, including electronic, and whether there is guidance to practitioners about how to use the PROM data will determine feasibility. Key informant groups emphasized the need to keep response time to less than 10 min, offer multiple data collection formats (including computer adaptive tests), and make the instrument accessible to patients with impairments. In addition, to reduce administrative burden, payer representatives recommended that the PROM include methods to reduce missing data and enhance data quality, such as those typically available in electronic formats.

Usability and Use. For this criterion, we focused on the Innovation Center’s goal of developing PRO-PMs to evaluate the quality of care under the CPC+ model. We looked for

evidence that the PROM was currently widely used in primary care or that there already was a PRO-PM based on the PROM.

Related and Competing Measures. For any given PROM, multiple PRO-PMs might be specified. Three of 15 sources referenced the need to consider related and competing PRO-PM measures. We did not evaluate related or competing PRO-PMs at this stage because there currently are so few PROMs for which there is even one PRO-PM.

Objective 4: Selection of the Best PROMs

Table 4 presents the results of our assessment of each of the 10 PROMs identified in the four priority domains, sorted by total score. We based the scores on a subset of the criteria identified in our analysis of the 15 standards documents (see online appendix 4) for which there was the most published evidence. For example, the most common evidence for the scientific acceptability criterion refers to the reliability and validity of the PROMs. The most common types of reliability evidence published are for internal consistency (Cronbach’s alpha) and test-retest reliability. The most common types of validity evidence published are for content and for construct validity. For construct validity, we included the specific instances of

Table 3 Criteria Used To Evaluate PROMs

NQF criteria	Relevance to primary care	Examples of evidence
Importance	The PRO domain targeted by the PROM is relevant	<ul style="list-style-type: none"> • PROM addresses a domain that is relevant to primary care • PROM addresses health experiences that are common across geographic populations (e.g., rural, urban, suburban) • PROM addresses a health experience that is salient to patients and practitioners • PROM produces results that patients understand and value • PROM has meaning for a large portion of the practice’s patient population
Scientific acceptability	Authorities and evidence support the PROM and the PROM is shown to be valid for primary care	<ul style="list-style-type: none"> • PROM scales show high internal consistency reliability (e.g., Cronbach’s alpha) • PROM scores show high test-retest reliability (stability over time) • Content validity of the PROM is documented • PROM scales and scores have high construct validity (e.g., structural, responsiveness, convergent, predictive)
Feasibility	PROM implementation is feasible for primary care practices	<ul style="list-style-type: none"> • PROM scores are easily interpretable • PROM has low respondent/administrative burden (e.g., electronic administration, adaptive administration)
Usability	The PROM is actionable for primary care practices	<ul style="list-style-type: none"> • PROM is easy to obtain • PROM is currently used by primary care practices to guide care • Performance measures based on the PROM have been developed

structural validity (i.e., factor analysis) because that was by far the most common analysis done, and responsiveness because the ability to detect change is important to informing patient care and to evaluating the performance of practices. Other types of construct validity evidence were collapsed into the category “other construct.” For feasibility, we included whether the PROM could be administered electronically and using computer adaptive software because this would shorten the administration time.

Electronic administration also can build in features to improve data quality in real time and immediately populate the PROM data base without the need to enter data manually. The Importance criterion is not included in Table 4 because the domains represented in Table 4 columns were those already determined to be important through the analysis shown in Table 2.

We recommended those PROMs with consistently positive evidence for 70% (8 of 11) of the selection criteria:

Table 4 Scores for Candidate PROMs, by Criterion and Total

Criteria*	Ability to participate		Depressive symptoms			Physical function		Self-efficacy		
	KAP	PROMISAP	PHQ-9	CESD-R	PROMIS-D	PF-10	PROMIS-PF	SEMCD	PROMIS-SE Sx	PROMIS-SE Tx
Reliability										
Cronbach’s alpha coefficient [†]	–	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Test-retest reliability [‡]	–	–	1.0	1.0	1.0	1.0	1.0	–	–	–
Validity										
Content	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Factor analysis [§]	–	1.0	0.5	0.5	1.0	0.5	0.5	1.0	0.5	1.0
Other construct	0.5	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Responsiveness	–	0.5	1.0	–	1.0	1.0	1.0	1.0	–	–
Feasibility										
Low patient burden	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Computer adaptive administration is available	–	1.0	–	–	1.0	–	1.0	–	1.0	1.0
Interpretable	–	1.0	1.0	1.0	1.0	1.0	1.0	–	1.0	1.0
Usability										
Evidence of use in clinical practice	–	–	1.0	–	1.0	1.0	1.0	–	1.0	1.0
Has PRO-PM	–	–	1.0	–	–	–	–	–	–	–
Total score	2.5	7.5	9.5	6.5	10.0	8.5	9.5	6.0	7.5	8.0

*A PROM received 1 point for each criterion for which there was evidence from one or multiple studies that supported the PROM; it received half a point when the evidence for that criterion was mixed and no points when there was no evidence or when the only evidence found was unfavorable to the PROM

[†]Cronbach’s alpha coefficient to estimate internal consistency reliability

[‡]Stability over time

[§]Factor analysis indicates models of the structure underlying the responses to PROM questions—that is, whether the responses all measure one thing. Factor analysis includes exploratory factor analysis, confirmatory factor analysis, Rasch, and item response theory analysis

- PHQ-9^{51, 55, 88–93}
- PROMIS Depression short forms and CAT^{36, 37, 57, 94, 95}
- PROMIS Self-Efficacy for Managing Medications and Treatments Short Forms and CATs^{38, 39}
- PF-10, Short Form-36 (SF-36) Physical Functioning Scale^{40–49}
- PROMIS Physical Function short forms and CAT^{50, 52–54, 56–60}

Detailed descriptions of these PROMs are available in online appendix 5.

DISCUSSION

Although the term “patient-centered” was introduced over 50 years ago, most measures used in primary care are clinician-centered—such as vital signs and laboratory tests based on biomedical and physical science.⁶¹ PROM data can complement conventional clinical measures.¹⁰ Exciting efforts to do this at the healthcare system scale are now taking shape in advanced models of primary care like CPC+. Our systematic process of stakeholder engagement and evidence-based review resulted in our identifying (1) four high-priority, health-related domains for primary care (ability to participate in social roles, depression, physical function, and self-efficacy for managing one’s health); (2) criteria for use in selecting PROMs for those domains; and (3) the five best existing PROMs to consider for measuring three of the four high-priority domains: PHQ-9 or PROMIS Depression (for depression), PF-10 or PROMIS-PF (for physical function), and PROMIS Self-Efficacy for Managing Treatments and Medications (for self-efficacy for managing one’s health). The paucity of evidence supporting PROMs for ability to participate in social roles suggests the need for further research and development of PROMs to address this important domain.

Choosing among these recommended PROMs will require comparing the benefits and drawbacks of each. For example, the content of the two recommended depression measures differs in that the PHQ-9 includes questions that are not about depressed mood but are associated with depressed mood (e.g., questions about sleep, fatigue, appetite, concentration, and behavior). In contrast, the PROMIS Depression items are all specific to sad thoughts and mood. The PHQ’s mixed content represents symptoms characteristic of the depression “syndrome” as specified by the DSM-IV. This diversity of content results in lower internal consistency reliability for the PHQ-9 than that observed for PROMIS depression. Factor analysis shows that the PHQ-9 does not measure a single phenomenon, but that PROMIS depression questions measure depressed mood exclusively. In addition, unlike PROMIS, there is no computerized adaptive test for PHQ-9, although there is the two-stage screening process. However, PHQ-9 has an extensive history of use as a performance measure in clinics, providing information about its feasibility and usefulness. We

provide additional comparative information about measures for each of the other PRO domains in online appendix 5.

Considerable research shows that comorbid depression increases the impact of chronic disease on patient outcomes; thus, supporting the routine assessment of depression in primary care practices.^{62–65} Unlike depressive symptoms, however, physical functioning is not often measured systematically in primary care.⁶⁶ In our study, stakeholder support for measuring physical functioning may have been due to the large percentage of patients with musculoskeletal disorders (e.g., arthritis, back pain) seen by primary care providers. For such patients, routine use of a physical function PROM applicable across conditions could aid in screening, in decision-making and monitoring, and in accounting for variations in clinical performance. Our research showed that both the PF-10 and the PROMIS-PF have strong evidence supporting their use for this purpose. The advantage of the PROMIS-PF is computer adaptive technology which leads to increased precision and, thus, ability to detect improvement or decline over time.^{44, 67}

Self-efficacy for managing one’s health also has not been routinely assessed in primary care perhaps because improving this outcome is a recent and aspirational goal of healthcare.^{68, 69} Yet, a major focus of CMS’ advanced primary care models is to include services such as patient education and coaching to help patients manage their chronic conditions.^{70, 71} The effectiveness of advanced primary care models may be judged, in part, on evidence of patient self-efficacy for managing their own health. Thus, primary care model evaluation research may provide data to describe the measurement properties of self-efficacy PROMs and guide their future use.

One could argue that improving patients’ abilities to participate in valued social roles is the ultimate goal of healthcare; yet this remains a challenging topic for measurement. The ability to participate in social roles is a function of many factors external to healthcare such as opportunity and interest. To the extent that outcomes like depressive symptoms and physical function help determine one’s ability to participate in social roles, the PROMs we recommend for those topics may help to address social role participation as well.

Limitations

The volume of information about PRO domains and PROMs required that we set boundaries on our search for evidence. A different approach to the environmental scan may have surfaced alternative or additional PRO domains and PROMs. Similarly, the volume of evidence that we obtained required that we use a strategy to screen the PRO domains and PROMs. While we described and justified our strategy, the application of different criteria in a different way may have changed the results. For example, the burden associated with each PROM was inferred from the number of items. However, other PROM characteristics also influence burden such as reading level required to comprehend the text, potential sensitivity of the

topic, and the way the PROM is formatted. Additionally, we considered the quality of evidence supporting the PROM (online appendix 5), but did not use a formal method such as GRADE.⁷² Moreover, we used a simple method of scoring PROMs, and criteria were not weighted by importance. For example, evidence for internal consistency reliability for the PROM scores was given the same weight as the existence of a PRO-PM based on the PROM. Our choice of the 70% level of positive support across criteria was arbitrary. If, instead, we had required 75% support, none of the self-efficacy PROMs would qualify, which would mean that just two PRO domains would be measured. On the other hand, the PROMIS Ability to Participate in Social Roles and Activities met 68% of our criteria. If the level of positive support was reduced to 65%, there would be at least one recommended PROM for each of the four PRO domains. Finally, we selected PRO domains according to number of stakeholder groups that recommended them and based on the literature review. Table 2 shows that patients also mentioned cognitive ability, sleep, fatigue, and sexual functioning as important.

Future research with patients should be conducted to evaluate the replicability of our results and supplement them. We focused on identifying consensus PRO domains but did not investigate how the meaning of a PRO domain might differ across stakeholder groups: this is another topic worthy of further research.

Conclusions

We identified two strong candidate measures for each of the core health domains of depression and physical function. Problems in these domains are highly prevalent in primary care patients, especially those with MCCs. Additionally, we identified a strong candidate PROM for a domain important to people living with chronic conditions—self-efficacy for managing their medications and treatments. PCPs can coach and support patients in developing the required skills. Existing evidence suggests that it should be feasible to incorporate these tools into PRO information systems and EHRs to support primary care practice and performance measures. Further studies will be needed to provide evidence that these PROMs produce information that is useful to clinicians, patients, and practices. If this evidence supports the effectiveness of these PROMs and performance measures, they can be used to promote patient centeredness and convey the value and quality of primary care.

Acknowledgments: *The authors would like to acknowledge all participants of patient interview and group discussions, practice representatives focus groups, and the payer representatives survey. We also are grateful to clinical thought leaders, CMMI officers, and the CMS' clinical care team for their invaluable contributions. Finally, we thank Lisa M. Shulman for sharing her, at that time, unpublished materials and colleagues at AIR (Rikki Mangrum, Rachel Shapiro, Coretta Mallery) and Johns Hopkins (Eric Bass, Hadi Kharazzi, Kitty Chan, Najilla Nassery, Christine Weston, Sarah Gensheimer, Cynthia Boyd, Junay Zhu, Zack Berger, Zishan Siddiqui, and Patricia Davidson) who were involved in the work preceding this manuscript preparation.*

Corresponding Author: *San Keller, PhD; American Institutes for Research in the Behavioral and Social Sciences, Chapel Hill, NC, USA (e-mail: sankeller@air.org).*

Compliance with ethical standards:

Conflict of interest: *Dr. Keller reports funding from the Centers for Medicare and Medicaid Services during the conduct of this study. Dr. Snyder reports funding from the American Institutes for Research (AIR) during the conduct of this study. She received funding outside of the submitted work from Genentech and Optum. Dr. Wu reports funding from AIR, during the conduct of the study. In addition he reports grants from PCORI, grants from AHRQ, NIH, Robert Wood Johnson Foundation, GenRe, and Genentech, personal fees from GSK, Gilead, Viiv, and Osmotica, outside the submitted work. All other authors have nothing to disclose*

REFERENCES

1. B S. Primary Care: Concept, Evaluation, and Policy: Oxford University Press; 1992.
2. Institute of Medicine Committee on the Future of Primary C. In: Donaldson MS, Yordy KD, Lohr KN, Vanselow NA, editors. Primary Care: America's Health in a New Era. Washington (DC): National Academies Press (US) Copyright 1996 by the National Academy of Sciences; 1996.
3. **Baron RJ.** The chasm between intention and achievement in primary care. *JAMA.* 2009;301(18):1922–4.
4. **Mechanic D.** How should hamsters run? Some observations about sufficient patient time in primary care. *BMJ (Clinical Research Ed).* 2001;323(7307):266–8.
5. **Feldman MD.** What's in a name?: is it time to retire the term "primary care physician"? *J Gen Intern Med.* 2017;32(9):957–8.
6. **Berenson RA, Hammons T, Gans DN, Zuckerman S, Merrell K, Underwood WS, et al.** A house is not a home: keeping patients at the center of practice redesign. *Health Aff (Project Hope).* 2008;27(5):1219–30.
7. **Barr MS.** The need to test the patient-centered medical home. *JAMA.* 2008;300(7):834–5.
8. Services CfMaM. Center for Medicare and Medicaid Innovation [Available from: <https://innovation.cms.gov/>].
9. **Sinsky CA, Beasley JW, Simmons GE, Baron RJ.** Electronic health records: design, implementation, and policy for higher-value primary care. *Ann Intern Med.* 2014;160(10):727–8.
10. **Wu AW, Cagney KA, St John PD.** Health status assessment. Completing the clinical database. *J Gen Intern Med.* 1997;12(4):254–5.
11. **Wu AW, Kharrazi H, Boulware LE, Snyder CF.** Measure once, cut twice—adding patient-reported outcome measures to the electronic health record for comparative effectiveness research. *J Clin Epidemiol.* 2013;66(8 Suppl):S12–20.
12. **Snyder CF, Aaronson NK, Choucair AK, Elliott TE, Greenhalgh J, Halyard MY, et al.** Implementing patient-reported outcomes assessment in clinical practice: a review of the options and considerations. *Qual Life Res.* 2012;21(8):1305–14.
13. **Users' Guide to Integrating Patient-Reported Outcomes in Electronic Health Records.** Baltimore, MD: Johns Hopkins University; 2017.
14. International Society for Quality of Life Research (prepared by: Aaronson N ET, GreenHalgh J, Haylard M, Hess R, Miller D, Reeve B, Santana M, Snyder C). *User's Guide to Implementing Patient-Reported Outcomes Assessment in Clinical Practice.* 2015.
15. **Basch E, Deal AM, Dueck AC, Scher HI, Kris MG, Hudis C, et al.** Overall survival results of a trial assessing patient-reported outcomes for symptom monitoring during routine cancer treatment. *JAMA.* 2017;318(2):197–8.
16. **Detmar SB, Muller MJ, Schornagel JH, Wever LD, Aaronson NK.** Overall quality-of-life assessments and patient-physician communication: a randomized controlled trial. *JAMA.* 2002;288(23):3027–34.
17. **Velikova G, Booth L, Smith AB, Brown PM, Lynch P, Brown JM, et al.** Measuring quality of life in routine oncology practice improves communication and patient well-being: a randomized controlled trial. *J Clin Oncol.* 2004;22(4):714–24.

18. **Rotenstein LS, Huckman RS, Wagle NW.** Making patients and doctors happier—the potential of patient-reported outcomes. *N Engl J Med.* 2017;377(14):1309–12.
19. **Baumhauer JF.** Patient-reported outcomes—are they living up to their potential? *N Engl J Med.* 2017;377(1):6–9.
20. Forum NQ. Patient Reported Outcomes (PROs) in Performance Measurement. Washington, DC; 2013.
21. **Reuben DB, Tinetti ME.** Goal-oriented patient care—an alternative health outcomes paradigm. *N Engl J Med.* 2012;366(9):777–9.
22. **Boyd CM, Darer J, Boulton C, Fried LP, Boulton L, Wu AW.** Clinical practice guidelines and quality of care for older patients with multiple comorbid diseases: implications for pay for performance. *Jama.* 2005;294(6):716–24.
23. ACP: American College of Physicians Philadelphia, PA: American College of Physicians; 2017 [Available from: <https://www.acponline.org>].
24. AGS: American Geriatrics Society New York, NY: American Geriatrics Society; 2017 [Available from: <https://www.americangeriatrics.org>].
25. Patients' View Institute: Patients' View Institute; 2017 [Available from: <https://govvi.org>].
26. Planetree Derby, CT: Planetree; 2014 [Available from: <http://planetree.org>].
27. PatientsLikeMe: PatientsLikeMe; 2017 [Available from: <https://www.patientslikeme.com>].
28. SGIM: Society of General Internal Medicine Alexandria, VA: Society of General Internal Medicine; [Available from: <http://www.sgim.org>].
29. AMA: American Medical Association: American Medical Association; 2017 [Available from: <https://www.ama-assn.org/>].
30. **Basch E, Spertus J, Dudley RA, Wu A, Chuahan C, Cohen P, et al.** Methods for developing Patient-Reported Outcome-Based Performance Measures (PRO-PMs). *Value Health.* 2015;18(4):493–504.
31. **Albright KS.** Environmental scanning: radar for success. *Inf Manag J.* 2004;38(3):38–45.
32. **Graham P, Evitts T, Thomas-MacLean R.** Environmental scans: how useful are they for primary care research? *Can Fam Physician.* 2008;54(7):1022–3.
33. **Valderas JM, Ferrer M, Mendivil J, Garin O, Rajmil L, Herdman M, et al.** Development of EMPRO: a tool for the standardized assessment of patient-reported outcome measures. *Value Health.* 2008;11(4):700–8.
34. COSMIN methodology for systematic reviews of Patient-Reported Outcome Measures (PROMs): user manual version 1.0 [Available from: https://cosmin.nl/wp-content/uploads/COSMIN-syst-review-for-PROMs-manual_version-1_feb-2018.pdf].
35. **Baker DW.** History of The Joint Commission's Pain Standards: lessons for today's prescription opioid epidemic. *JAMA.* 2017;317(11):1117–8.
36. **Schalet BD, Pilkonis PA, Yu L, Dodds N, Johnston KL, Yount S, et al.** Clinical validity of PROMIS Depression, Anxiety, and Anger across diverse clinical samples. *J Clin Epidemiol.* 2016;73:119–27.
37. **Kim J, Chung H, Askew RL, Park R, Jones SM, Cook KF, et al.** Translating CESD-20 and PHQ-9 Scores to PROMIS Depression. *Assessment.* 2017;24(3):300–7.
38. **Gruber-Baldini AL, Velozo C, Romero S, Shulman LM.** Validation of the PROMIS(R) measures of self-efficacy for managing chronic conditions. *Qual Life Res.* 2017;26(7):1915–24.
39. **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.
40. **Bohannon RW, DePasquale L.** Physical Functioning Scale of the Short-Form (SF) 36: internal consistency and validity with older adults. *J Geriatr Phys Ther.* 2010;33(1):16–8.
41. **Beaton DE, Hogg-Johnson S, Bombardier C.** Evaluating changes in health status: reliability and responsiveness of five generic health status measures in workers with musculoskeletal disorders. *J Clin Epidemiol.* 1997;50(1):79–93.
42. **Ware JE, Jr., Sherbourne CD.** The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care.* 1992;30(6):473–83.
43. **Haley SM, McHorney CA, Ware JE, Jr.** Evaluation of the MOS SF-36 physical functioning scale (PF-10): I. Unidimensionality and reproducibility of the Rasch item scale. *J Clin Epidemiol.* 1994;47(6):671–84.
44. **Rose M, Bjorner JB, Gandek B, Bruce B, Fries JF, Ware JE, Jr.** The PROMIS Physical Function item bank was calibrated to a standardized metric and shown to improve measurement efficiency. *J Clin Epidemiol.* 2014;67(5):516–26.
45. **McHorney CA.** Measuring and monitoring general health status in elderly persons: practical and methodological issues in using the SF-36 Health Survey. *Gerontologist.* 1996;36(5):571–83.
46. **McHorney CA, Haley SM, Ware JE, Jr.** Evaluation of the MOS SF-36 Physical Functioning Scale (PF-10): II. Comparison of relative precision using Likert and Rasch scoring methods. *J Clin Epidemiol.* 1997;50(4):451–61.
47. **McHorney CA, Ware JE, Jr., Lu JF, Sherbourne CD.** The MOS 36-item Short-Form Health Survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. *Med Care.* 1994;32(1):40–66.
48. **McHorney CA, Ware JE, Jr., Raczek AE.** The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. *Med Care.* 1993;31(3):247–63.
49. **Perkins AJ, Stump TE, Monahan PO, McHorney CA.** Assessment of differential item functioning for demographic comparisons in the MOS SF-36 health survey. *Qual Life Res.* 2006;15(3):331–48.
50. **Fries JF, Lingala B, Siemons L, Glas CA, Cella D, Hussain YN, et al.** Extending the floor and the ceiling for assessment of physical function. *Arthritis Rheum.* 2014;66(5):1378–87.
51. **Spitzer RL, Williams JB, Kroenke K, Linzer M, deGruy FV, 3rd, Hahn SR, et al.** Utility of a new procedure for diagnosing mental disorders in primary care. The PRIME-MD 1000 study. *JAMA.* 1994;272(22):1749–56.
52. **Bruce B, Fries JF, Ambrosini D, Lingala B, Gandek B, Rose M, et al.** Better assessment of physical function: item improvement is neglected but essential. *Arthritis Res Ther.* 2009;11(6):R191.
53. **Paz SH, Jones L, Calderon JL, Hays RD.** Readability and Comprehension of the Geriatric Depression Scale and PROMIS(R) Physical Function Items in Older African Americans and Latinos. *Patient.* 2017;10(1):117–31.
54. **Fidai MS, Saltzman BM, Meta F, Lizzio VA, Stephens JP, Bozic KJ, et al.** Patient-Reported Outcomes Measurement Information System and Legacy Patient-Reported Outcome Measures in the Field of Orthopaedics: a systematic review. *Arthroscopy.* 2018;34(2):605–14.
55. **Lowe B, Unutzer J, Callahan CM, Perkins AJ, Kroenke K.** Monitoring depression treatment outcomes with the patient health questionnaire-9. *Med Care.* 2004;42(12):1194–201.
56. PROMIS® Scoring Manuals: Northwestern University: HealthMeasures; 2018 [Available from: <http://www.healthmeasures.net/promis-scoring-manuals>].
57. **Stone AA, Broderick JE, Junghaenel DU, Schneider S, Schwartz JE.** PROMIS fatigue, pain intensity, pain interference, pain behavior, physical function, depression, anxiety, and anger scales demonstrate ecological validity. *J Clin Epidemiol.* 2016;74:194–206.
58. **Rothrock NE, Hays RD, Spritzer K, Yount SE, Riley W, Cella D.** Relative to the general US population, chronic diseases are associated with poorer health-related quality of life as measured by the Patient-Reported Outcomes Measurement Information System (PROMIS). *J Clin Epidemiol.* 2010;63(11):1195–204.
59. **Madsen LP, Evans TA, Snyder KR, Docherty CL.** Patient-Reported Outcomes Measurement Information System Physical Function Item Bank, Version 1.0: physical function assessment for athletic patient populations. *J Athl Train.* 2016;51(9):727–32.
60. **Schalet BD, Hays RD, Jensen SE, Beaumont JL, Fries JF, Cella D.** Validity of PROMIS Physical Function measured in diverse clinical samples. *J Clin Epidemiol.* 2016;73:112–8.
61. **Balint E.** The possibilities of patient-centered medicine. *J R Coll Gen Pract.* 1969;17(82):269–76.
62. **Schmitz N, Wang J, Malla A, Lesage A.** Joint effect of depression and chronic conditions on disability: results from a population-based study. *Psychosom Med.* 2007;69(4):332–8.
63. **Ho C, Feng L, Fam J, Mahendran R, Kua EH, Ng TP.** Coexisting medical comorbidity and depression: multiplicative effects on health outcomes in older adults. *Int Psychogeriatr.* 2014;26(7):1221–9.
64. **Egede LE.** Major depression in individuals with chronic medical disorders: prevalence, correlates and association with health resource utilization, lost productivity and functional disability. *Gen Hosp Psychiatry.* 2007;29(5):409–16.
65. **Guthrie EA, Dickens C, Blakemore A, Watson J, Chew-Graham C, Lovell K, et al.** Depression predicts future emergency hospital admissions in primary care patients with chronic physical illness. *J Psychosom Res.* 2016;82:54–61.
66. **Dy SM, Pfoh ER, Salive ME, Boyd CM.** Health-related quality of life and functional status quality indicators for older persons with multiple chronic conditions. *J Am Geriatr Soc.* 2013;61(12):2120–7.
67. **Hung M, Stuart AR, Higgins TF, Saltzman CL, Kubiak EN.** Computerized adaptive testing using the PROMIS Physical Function item bank reduces test burden with less ceiling effects compared with the short musculoskeletal function assessment in orthopaedic trauma patients. *J Orthop Trauma.* 2014;28(8):439–43.

68. **Fortin M, Chouinard MC, Bouhali T, Dubois MF, Gagnon C, Belanger M.** Evaluating the integration of chronic disease prevention and management services into primary health care. *BMC Health Serv Res.* 2013;13:132.
69. **Battersby M, Von Korff M, Schaefer J, Davis C, Ludman E, Greene SM, et al.** Twelve evidence-based principles for implementing self-management support in primary care. *Jt Comm J Qual Patient Saf.* 2010;36(12):561–70.
70. **Baron RJ.** New pathways for primary care: an update on primary care programs from the innovation center at CMS. *Ann Fam Med.* 2012;10(2):152–5.
71. **Comprehensive Primary Care Plus (CPC+).** Baltimore, MD: U.S. Department of Health & Human Services; Centers for Medicare & Medicaid Services; 2017.
72. **Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al.** GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ (Clinical Research Ed).* 2008;336(7650):924–6.
73. **HealthMeasures: Search & View Measures.** Northwestern University; 2018 [Available from: <http://www.healthmeasures.net/search-view-measures?task=Search.search>].
74. **Rehabilitation Measures Database Chicago, IL:** AbilityLab; 2018 [Available from: <https://www.sralab.org/rehabilitation-measures>].
75. **Services CfMaM.** What Is MACRA? [Available from: <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/MACRA-MIPS-and-APMs/MACRA-MIPS-and-APMs.html>].
76. **Center ER.** Electronic Quality Improvement Resource Center [Available from: <https://ecqi.healthit.gov/>].
77. **Measure Developer Guidebook for Submitting Measures to NQF:** National Quality Forum; 2017 [Available from: http://www.qualityforum.org/Measuring_Performance/Submitting_Standards.aspx].
78. **Blueprint for the CMS measures management system version 13.0.:** Centers for Medicare & Medicaid Services; 2017.
79. **Patient-Reported Outcomes in Performance Measurement:** National Quality Forum; 2017 [Available from: https://www.qualityforum.org/Publications/2012/12/Patient-Reported_Outcomes_in_Performance_Measurement.aspx].
80. **Reeve BB, Wyrwich KW, Wu AW, Velikova G, Terwee CB, Snyder CF, et al.** ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research. *Qual Life Res.* 2013;22(8):1889–905.
81. **PROMIS® Instrument Development and Validation Scientific Standards Version 2.0:** PROMIS®; 2013 [Available from: http://www.healthmeasures.net/images/PROMIS/PROMISStandards_Vers2.0_Final.pdf].
82. **Mokkink LB, Terwee CB, Knol DL, Stratford PW, Alonso J, Patrick DL, et al.** Protocol of the COSMIN study: COnsensus-based Standards for the selection of health Measurement INstruments. *BMC Med Res Methodol.* 2006;6:2.
83. **Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al.** COSMIN checklist manual. Amsterdam, The Netherlands: COSMIN; 2012.
84. **Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al.** The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res.* 2010;19(4):539–49.
85. **Mokkink LB, Terwee CB, Knol DL, Stratford PW, Alonso J, Patrick DL, et al.** The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content. *BMC Med Res Methodol.* 2010;10:22.
86. **Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al.** The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol.* 2010;63(7):737–45.
87. **Butt Z, Reeve B.** Enhancing the Patient's Voice: Standards in the Design and Selection of Patient-Reported Outcomes Measures (PROMs) for Use in Patient-Centered Outcomes Research Washington, DC: Patient Centered Outcomes Research Institute; 2012.
88. **Kroenke K, Spitzer RL, Williams JB, Lowe B.** The Patient Health Questionnaire Somatic, Anxiety, and Depressive Symptom Scales: a systematic review. *Gen Hosp Psychiatry.* 2010;32(4):345–59.
89. **El-Den S, Chen TF, Gan YL, Wong E, O'Reilly CL.** The psychometric properties of depression screening tools in primary healthcare settings: a systematic review. *J Affect Disord.* 2018;225:503–22.
90. **Diagnostic and Statistical Manual of Mental Disorders, 4th Edition.** 4 ed. Arlington, VA: American Psychiatric Association; 2000.
91. **Huang FY, Chung H, Kroenke K, Delucchi KL, Spitzer RL.** Using the Patient Health Questionnaire-9 to measure depression among racially and ethnically diverse primary care patients. *J Gen Intern Med.* 2006;21(6):547–52.
92. **Titov N, Dear BF, McMillan D, Anderson T, Zou J, Sunderland M.** Psychometric comparison of the PHQ-9 and BDI-II for measuring response during treatment of depression. *Cogn Behav Ther.* 2011;40(2):126–36.
93. **NQF Quality Positioning System™ Outcome: PRO-PM, Measure #0209** Washington D.C.: National Quality Forum; [Available from: <https://www.qualityforum.org/QPS/MeasureDetails.aspx?standardID=457&print=0&entityTypeID=1>].
94. **Pilkonis PA, Choi SW, Reise SP, Stover AM, Riley WT, Cella D.** Item banks for measuring emotional distress from the Patient-Reported Outcomes Measurement Information System (PROMIS(R)): depression, anxiety, and anger. *Assessment.* 2011;18(3):263–83.
95. **Pilkonis PA, Yu L, Dodds NE, Johnston KL, Maihofer CC, Lawrence SM.** Validation of the depression item bank from the Patient-Reported Outcomes Measurement Information System (PROMIS) in a three-month observational study. *J Psychiatr Res.* 2014;56:112–9.
96. **Acaster S, Cimms T, Lloyd A.** Development of a Methodological Standards Report: Topic #3: The Design and Selection of Patient-Reported Outcomes Measures (PROMs) for Use in Patient Centered Outcomes Research. San Francisco, CA: Oxford Outcomes; 2012.
97. **Johnson C, Aaronson N, Blazeby JM, Bottomley B, Fayers P, Koller M, et al.** EORTC Quality of Life Group: Guidelines for Developing Questionnaire Modules. European Organization for Research and Treatment of Cancer; 2011.
98. **Guidance for Industry Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims.** Rockville, MD: U.S. Food and Drug Administration; 2009.
99. **A guide to patient reported measures—theory, landscape and uses: mon-mouth partners;** [Available from: <http://www.monmouthpartners.com/assets/pdf/A%20Guide%20to%20Patient%20Reported%20Measures.pdf>].
100. **Coons SJ, Gwaltney CJ, Hays RD, Lundy JJ, Sloan JA, Revicki DA, et al.** Recommendations on evidence needed to support measurement equivalence between electronic and paper-based patient-reported outcome (PRO) measures: ISPOR ePRO Good Research Practices Task Force report. *Value Health.* 2009;12(4):419–29.
101. **Rothman M, Burke L, Erickson P, Leidy NK, Patrick DL, Petrie CD.** Use of existing patient-reported outcome (PRO) instruments and their modification: the ISPOR Good Research Practices for Evaluating and Documenting Content Validity for the Use of Existing Instruments and Their Modification PRO Task Force Report. *Value Health.* 2009;12(8):1075–83.
102. **Wild D, Eremenco S, Mear I, Martin M, Houchin C, Gawlicki M, et al.** Multinational trials—recommendations on the translations required, approaches to using the same language in different countries, and the approaches to support pooling the data: the ISPOR Patient-Reported Outcomes Translation and Linguistic Validation Good Research Practices Task Force report. *Value Health.* 2009;12(4):430–40.
103. **Wild D, Grove A, Martin M, Eremenco S, McElroy S, Verjee-Lorenz A, et al.** Principles of Good Practice for the Translation and Cultural Adaptation Process for Patient-Reported Outcomes (PRO) Measures: report of the ISPOR Task Force for Translation and Cultural Adaptation. *Value Health.* 2005;8(2):94–104.
104. **Regulatory guidance for the use of health-related quality of life (HRQL) measures in the evaluation of medicinal products.** London, UK: European Medicines Agency; 2005. Contract No.: November 17, 2017.
105. **Aaronson N, Alonso J, Burnam A, Lohr KN, Patrick DL, Perrin E, et al.** Assessing health status and quality-of-life instruments: attributes and review criteria. *Qual Life Res.* 2002;11(3):193–205.
106. **Hahn EA, DeWalt DA, Bode RK, Garcia SF, DeVellis RF, Correia H, et al.** New English and Spanish social health measures will facilitate evaluating health determinants. *Health Psychol.* 2014;33(5):490–9.
107. **Hermesen LA, Terwee CB, Leone SS, van der Zwaard B, Smalbrugge M, Dekker J, et al.** Social participation in older adults with joint pain and comorbidity: testing the measurement properties of the Dutch Keele Assessment of Participation. *BMJ Open.* 2013;3(8):e003181.
108. **van der Meij E, Anema JR, Huirne JAF, Terwee CB.** Using PROMIS for measuring recovery after abdominal surgery: a pilot study. *BMC Health Serv Res.* 2018;18(1):128.
109. **Wilkie R, Peat G, Thomas E, Hooper H, Croft PR.** The Keele Assessment of Participation: a new instrument to measure participation

- restriction in population studies. Combined qualitative and quantitative examination of its psychometric properties. *Qual Life Res.* 2005;14(8):1889–99.
110. **Eaton WW, Muntaner C, Smith C, Tien A, Ybarra M.** Center for Epidemiologic Studies Depression Scale: review and revision (CESD and CESD-R). In: Maruish ME, editor. *The Use of Psychological Testing for Treatment Planning and Outcomes Assessment.* 3rd ed. Mahwah, NJ: Lawrence Erlbaum; 2004. p. 363–77.
111. **Bartlett SJ, Orbai AM, Duncan T, DeLeon E, Ruffing V, Clegg-Smith K, et al.** Reliability and validity of selected PROMIS measures in people with rheumatoid arthritis. *PLoS One.* 2015;10(9):e0138543.
112. **Kroenke K, Yu Z, Wu J, Kean J, Monahan PO.** Operating characteristics of PROMIS four-item depression and anxiety scales in primary care patients with chronic pain. *Pain Med.* 2014;15(11):1892–901.
113. **Van Dam NT, Earleywine M.** Validation of the Center for Epidemiologic Studies Depression Scale—Revised (CESD-R): pragmatic depression assessment in the general population. *Psychiatry Res.* 2011;186(1):128–32.
114. **Williams JR, Hirsch ES, Anderson K, Bush AL, Goldstein SR, Grill S, et al.** A comparison of nine scales to detect depression in Parkinson disease: which scale to use? *Neurology.* 2012;78(13):998–1006.
115. **Cook KF, Jensen SE, Schalet BD, Beaumont JL, Amtmann D, Czajkowski S, et al.** PROMIS measures of pain, fatigue, negative affect, physical function, and social function demonstrated clinical validity across a range of chronic conditions. *J Clin Epidemiol.* 2016;73:89–102.
116. **Haskell A, Kim T.** Implementation of Patient-Reported Outcomes Measurement Information System data collection in a private orthopedic surgery practice. *Foot Ankle Int.* 2018;1071100717753967.
117. **Lee AC, Driban JB, Price LL, Harvey WF, Rodday AM, Wang C.** Responsiveness and minimally important differences for 4 Patient-Reported Outcomes Measurement Information System Short Forms: Physical Function, Pain Interference, Depression, and Anxiety in Knee Osteoarthritis. *J Pain.* 2017;18(9):1096–110.
118. **Murphy M, Hollinghurst S, Salisbury C.** Identification, description and appraisal of generic PROMs for primary care: a systematic review. *BMC Fam Pract.* 2018;19(1):41.
119. **Garratt AM, Ruta DA, Abdalla MI, Russell IT.** SF 36 health survey questionnaire: II. Responsiveness to changes in health status in four common clinical conditions. *Qual Health Care.* 1994;3(4):186–92.
120. **Bjorner JB, Rose M, Gandek B, Stone AA, Junghaenel DU, Ware JE, Jr.** Method of administration of PROMIS scales did not significantly impact score level, reliability, or validity. *J Clin Epidemiol.* 2014;67(1):108–13.
121. **Shulman L.** Prepublication results provided by the developer, Dr. Lisa Shulman, in an email dated February 16, 2018. 2018.
122. **Ritter PL, Lorig K.** The English and Spanish Self-Efficacy to Manage Chronic Disease Scale measures were validated using multiple studies. *J Clin Epidemiol.* 2014;67(11):1265–73.
123. **Amtmann D, Bamer AM, Cook KF, Askew RL, Noonan VK, Brockway JA.** University of Washington self-efficacy scale: a new self-efficacy scale for people with disabilities. *Arch Phys Med Rehabil.* 2012;93(10):1757–65.
124. **Bandura A.** *Self-efficacy: the exercise of control.* New York, NY: WH Freeman and Company; 1997.
125. **Riehm KE, Kwakkenbos L, Carrier ME, Bartlett SJ, Malcarne VL, Mouthon L, et al.** Validation of the Self-Efficacy for Managing Chronic Disease Scale: a scleroderma patient-centered intervention network cohort study. *Arthritis Care Res.* 2016;68(8):1195–200.
126. **Lorig KR, Sobel DS, Stewart AL, Brown BW, Jr., Bandura A, Ritter P, et al.** Evidence suggesting that a chronic disease self-management program can improve health status while reducing hospitalization: a randomized trial. *Med Care.* 1999;37(1):5–14.
127. **Shulman L.** Personal communication. 2018.

Publisher's Note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.