Use of the Patient-Reported Outcomes Measurement Information System (PROMIS) for Operative Shoulder Outcomes

Ilona Schwarz,*† MS, John-Rudolph H. Smith,† BS, Darby A. Houck,† BA, Rachel M. Frank,† MD, Jonathan T. Bravman,† MD, and Eric C. McCarty,† MD

Investigation performed at the Division of Sports Medicine and Shoulder Surgery, Department of Orthopedics, University of Colorado School of Medicine, Aurora, Colorado, USA

Background: Few studies have investigated the relationship between the Patient-Reported Outcomes Measurement Information System (PROMIS) and legacy patient-reported outcome (PRO) measurements.

Purpose: To compare patient-reported outcomes from the PROMIS physical function (PF) and upper extremity (UE) platforms against one another and against legacy PRO measurements to assess the potential strengths and weaknesses of the National Institutes of Health PROMIS initiative and expand on the use of PRO measurements in clinical orthopaedic practice.

Study Design: Systematic review; Level of evidence, 4.

Methods: A systematic search of the PubMed, Embase, and Cochrane Library databases was conducted following PRISMA (Preferred Reporting Items for Systematic Meta-Analyses) guidelines. All English-language studies published between 2017 and 2019 using PROMIS to evaluate patients for shoulder surgery were analyzed. PROs were compared based on survey administered and the shoulder condition being investigated. Study quality was evaluated using the Modified Coleman Methodology Score and the Methodological Index for Non-Randomized Studies score.

Results: We included 9 studies (5 studies were level 2; 3 studies were level 3; 1 study was level 4) encompassing a total of 1130 patients (60.2% male; mean age, 52.6 ± 16.5 years; mean BMI, 29.8 ± 2.8 kg/m^2). Of these, 6 studies administered the PROMIS PF, and 6 studies administered the PROMIS UE. The strongest correlation was between PROMIS PF computer adaptive test and the 36-Item Short Form Health Survey Global Health (SF-36 GH) ($r = 0.75$). The highest overall correlation with the PROMIS UE was found with the American Shoulder and Elbow Surgeons (ASES) Shoulder Score ($r = 0.70$). The lowest correlations were found between PROMIS PF and the Marx Shoulder Activity Scale ($r = 0.08$) and the PROMIS UE and the Marx Shoulder Activity Scale ($r = 0.18$).

Conclusion: From available data, the PROMIS PF and PROMIS UE were most closely correlated with outcomes measured by the SF-36 GH. The PROMIS UE alone was most correlated with ASES Shoulder Score. Thus, either PROMIS PF or UE may provide a possible alternative to legacy PRO measurements but with a lower overall number of questions and higher generalizability. Future research should compare the time and question burden of the various PROMIS platforms with a more consistent evaluation of standard PRO measurements.

Keywords: orthopaedic; Patient-Reported Outcomes Measurement Information System; PROMIS; shoulder

Patient-reported outcome (PRO) measurements play an important role in evaluating a patient's perspective on clinical care, clinical research, and health care policy. However, with the development of new PRO instruments, patients may face "survey fatigue" from question burden, and providers may face the challenge of which PRO instrument to administer and to whom, as well as potential ceiling effects, especially as patients age.[4,5,13] To mitigate some of the limitations faced by earlier generation PRO tools, the National Institutes of Health (NIH) has developed a platform PRO measurement that is applicable to the general population and can be administered and scored in a standardized fashion, thereby allowing for the comparison across a wider range of clinical scenarios.[13,18] This initiative, known as the Patient-Reported Outcomes Measurement Information System (PROMIS), attempts to overcome the criticism of many common forms being administered and created without proper statistical validation.[23] There are 2 possible administration platforms: the Short Form (SF), which ranges from 2 to 8 questions, and

The Orthopaedic Journal of Sports Medicine, 8(6), 2325967120924345
DOI: 10.1177/2325967120924345

the computer adaptive test (CAT), which is a computer algorithm–based questionnaire that outputs questions based on previous question input—therefore, the number of questions administered is variable based on the patient.[13]

PROMIS can be subdivided into either the PROMIS physical function (PF) or the PROMIS global health (GH). The PROMIS PF considers social function, pain, fatigue, and emotional distress. The PROMIS GH considers overall physical function, pain, fatigue, emotional distress, and social health.

Shoulder injuries and diseases are common conditions seen by orthopaedists.[22] Often, surgical intervention is required to help a patient regain function and relieve pain in conditions, such as rotator cuff disease, osteoarthritis, and shoulder instability. Given the potential overall increase in quality of life after surgical repair while still considering the potential decrease in quality of life immediately after surgery, PRO instruments can play a role in assessing a patient's perception of the clinical intervention both pre- and postoperatively. Additionally, PRO measurements provide meaningful data on the success of a procedure's outcome based on PF scores, such as pain intensity, pain interference (PI), fatigue, and sleep disturbance—all of which encompass the PF domains in the NIH PROMIS.[3,4,12]

Traditionally, shoulder-related PROs have been assessed via questionnaires, such as the Simple Shoulder Test (SST) and the American Shoulder and Elbow Surgeons (ASES) Shoulder Score, often termed *legacy PROs*.[4,14] However, such PRO instruments may lack the generalizability of PROMIS and can be time-consuming, display ceiling effects, and pose an increased question burden. Considering that the goal of the PROMIS survey is to broaden outcomes reporting among various diagnoses and assess big-picture outcomes in a timely, consistent fashion, its use in patients with orthopaedic shoulder conditions holds significant potential for quantifying patient perspective as well as clinician performance. The purpose of this review is to assess the performance of different PROMIS platforms against one another and compare them to legacy PRO measurements in patients undergoing surgical intervention for shoulder conditions.

METHODS

This systematic review was conducted by 2 independent reviewers (I.S. and J.-R.H.S.) via the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines following the appropriate PRISMA checklist and template. Each reviewer searched and documented results from 3 databases, PubMed, Embase, and the Cochrane Library, using the following search terms with no additional exclusionary criteria: ((“Orthopedic” OR “Orthopaedic” OR “Orthopedics” OR “Orthopaedics”)) AND (“PROMIS” OR “Patient Reported Outcomes Measurement Information System”) AND (“shoulder” OR “shoulders”). All English-language studies published between 2017 and 2019 using PROMIS to evaluate patients for shoulder surgery were analyzed. A total of 123 articles were identified through the database search, and upon the removal of duplicates and abstract screening, 30 studies were determined to be eligible. Eligibility criteria included studies using any PROMIS questionnaire to report outcome measures in patients undergoing surgical intervention for shoulder disease or injury; thus, studies not using PROMIS that initially met the search criteria were eliminated. Exclusion criteria included studies not using the PROMIS score, studies not looking at shoulder outcomes, and studies not looking at patients pre- or postoperatively. After a final screening eliminating studies not reporting on PROMIS outcomes, studies not reporting patient-reported outcomes, or studies with incomplete data, 9 studies were determined to be eligible for the review (Figure 1). Data extraction was performed independently (I.S.) and reviewed by the other reviewer (J.-R.H.S.). Funding and third-party involvement were not required to obtain any of the collected data.

Reporting Outcomes

The outcomes extracted and pooled from the studies included demographic data (age, percentage male, body mass index [BMI]), shoulder condition or main concern, and timing of PROMIS survey administration (pre- or postoperatively). Additional data included scores on all PROMIS domains (PF, CAT, and upper extremity [UE]), ASES Shoulder Score, Marx Shoulder Activity Scale, 36-Item Short Form Health Survey (SF-36) GH or PF, Euro-Qol 5 Dimensions (EQ-5D), Western Ontario Rotator Cuff (WORC) index, Western Ontario Shoulder Instability (WOSI) index, Western Ontario Osteoarthritis Shoulder (WOOS) index, Single Assessment Numeric Evaluation, postoperative visual analog scale for pain, and SST.

Study Method Assessment

Study quality was assessed via the Methodological Index for Non-Randomized Studies (MINORS) score. MINORS is a bias assessment tool based on 8 criteria for noncomparative studies and 12 criteria for comparative studies. Scores range

*Address correspondence to Ilona Schwarz, MS, CU Sports Medicine and Performance Center, 2150 Stadium Drive, Room 222, Boulder, CO 80309, USA (email: ilona.schwarz@cuanschutz.edu).
†University of Colorado School of Medicine, Department of Orthopedics, Division of Sports Medicine and Shoulder Surgery, Aurora, Colorado, USA.
Final revision submitted January 29, 2020; accepted February 13, 2020.
One or more of the authors has declared the following potential conflict of interest or source of funding: R.M.F. has received grant support from Arthrex; educational support from Arthrex, Medwest, and Smith & Nephew; speaking fees from Arthrex; and royalties from Elsevier. J.T.B. has received research support from Biomet and Stryker; consulting fees from DJO, Encore Medical, and Smith & Nephew; and royalties from Shukla Medical. E.C.M. has received research support from Arthrex, Biomet, Breg, Mitek, Ossur, Smith & Nephew, and Stryker; consulting fees from DePuy and Zimmer Biomet; speaking fees from Arthrex; and royalties from Elsevier and Zimmer Biomet. AOSSM checks author disclosures against the Open Payments Database (OPD). AOSSM has not conducted an independent investigation on the OPD and disclaims any liability or responsibility relating thereto.

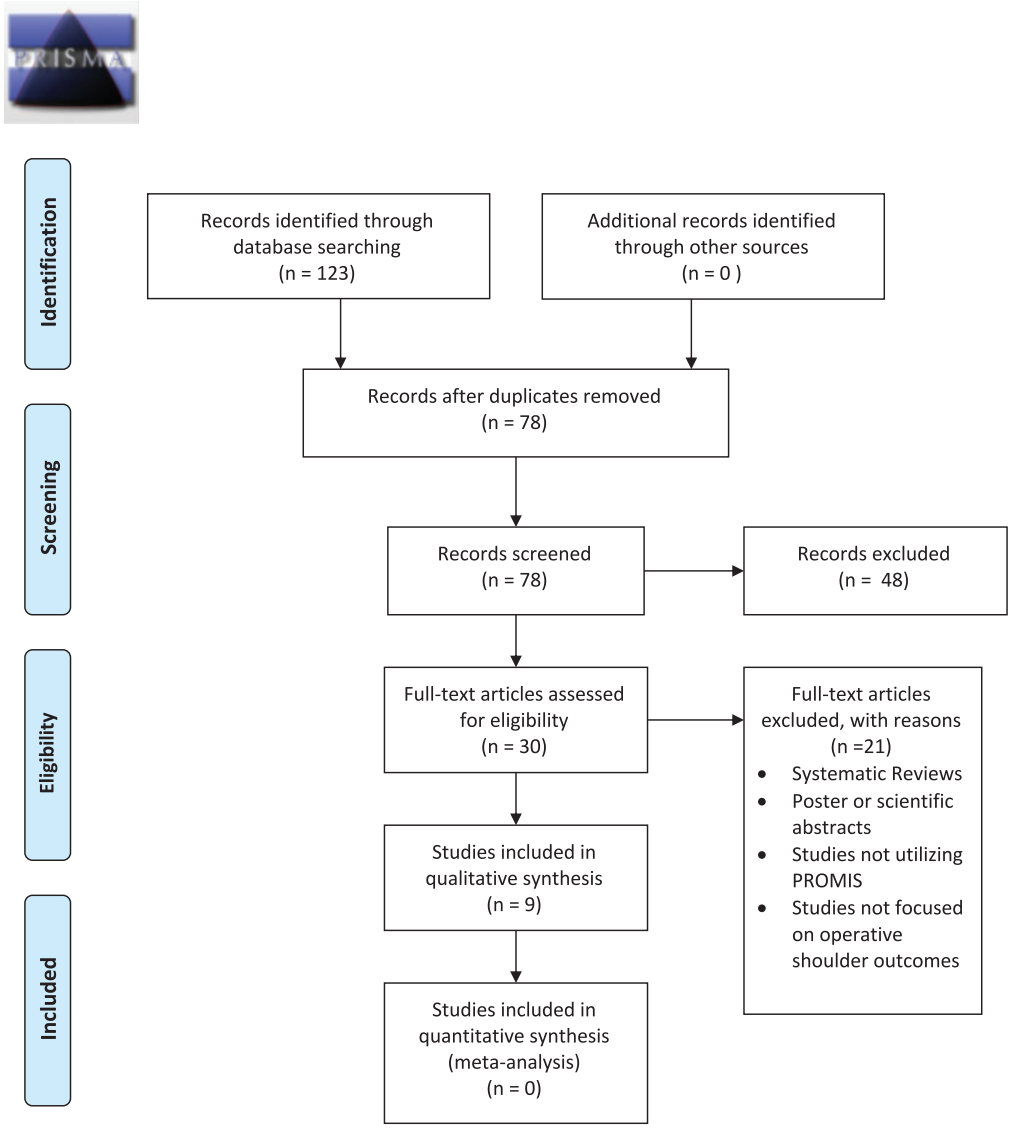

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram in patients undergoing surgical intervention for shoulder conditions using the Patient-Reported Outcomes Measurement Information System (PROMIS) to report outcome measure.

from 0 to 16 for comparative studies and from 0 to 24 for noncomparative studies.[21] For each specified criterion, an independent, numerical score is assigned: 0 (not reported), 1 (reported but inadequately), or 2 (reported adequately).

The quality of study method was evaluated through use of the Modified Coleman Methodology Score (MCMS), which is scored from 0 to 100. Scores of 85-100 are excellent; 70-84, good; 55-69, fair; and < 55, poor.[7] For correlation coefficients, $r \geq 0.7$ was considered a strong or high level of correlation.

Statistical Analysis

Given the limited specifications of the reported outcome measures and the heterogeneity of the included studies, no calculable data or meta-analyses are presented in this review. The intent of this review was to group descriptive statistics, demographics, categorical variables, and outcome measures already presented in each respective study.

RESULTS

Patient Demographics

We included 9 studies (5 studies were level 2,[1,2,6,11,17] 3 studies were level 3,[9,19,20] and 1 study was level 4[15]) with a total of 1130 patients (60.2% male; mean age, 52.6 ± 16.5 years; mean BMI, 29.8 ± 2.8 kg/m^2).[1,2,9,11,19] Of these, 6 studies[1,2,6,9,11,17] administered the PROMIS PF, and 6 studies[1,2,9,11,15,19] administered the PROMIS UE.

A total of 569 patients had a primary concern and/or diagnosis of rotator cuff disease (3 studies)[1,17,19]; 274 patients

TABLE 1
Characteristics of Included Study Participants[a]

Study	No. of Patients Analyzed	Mean Age, y	% Male	Mean BMI, kg/m^2	Patient Population	Survey Timing
Anthony (2017)[1]	82	54	61	31.7 ± 1 7.4	RC disease	Preoperative
Anthony (2017)[2]	70	27 ± 10	74.3	27 ± 5.1	Recurrent SI	Preoperative
Chen (2019)[6]	62	67.6 ± 8.9	54.8	NR	GH OA	Preoperative Postoperative
Dowdle (2017)[9]	51	60.8 ± 12.9	58.5	33.9 ± 6.8	GH OA	Preoperative
Hajewski (2019)[11]	72	22.1	79	26.6	SI	Preoperative Postoperative
Monroe (2019)[15]	145	62.0 ± 9.8	59.3	NR	SSc tear	Postoperative
Nicholson (2019)[17]	323	57.7 ± 13.8	53.9	29.4 ± 6.1	RC disease	Preoperative
Patterson (2018)[19]	164	58 ± 8.3	52	30 ± 6.2	RC disease	Preoperative
Saad (2018)[20]	161	64.5 ± 13.3	52.8	NR	Shoulder arthritis	Preoperative

[a]All values (and SDs when available) are reported based on what was provided in each study. BMI, body mass index; GH OA, glenohumeral osteoarthritis; NR, not reported; RC, rotator cuff; SI, shoulder instability; SSc, subscapularis.

TABLE 2
Included Study Characteristics[a]

Study	Journal	PDR	Country	Study Design	MINORS	MCMS
Anthony (2017)[1]	Arthroscopy	01/2015–09/2015	USA	Prospective; Level 2	11	78
Anthony (2017)[2]	AJSM	NR	USA	Cohort; Level 2	11	75
Chen (2019)[6]	JSES	02/2015–02/2017	USA	Prospective; Level 2	15	90
Dowdle (2017)[9]	OJSM	NR	USA	Cohort; Level 3	11	80
Hajewski (2019)[11]	OJSM	01/2015–11/2018	USA	Cohort; Level 2	16	82
Monroe (2019)[15]	Arthroscopy	01/2010–04/2016	USA	Retrospective case series; Level 4	15	92
Nicholson (2019)[17]	AJSM	01/2015–09/2017	USA	Cohort; Level 2	11	78
Patterson (2018)[19]	JSES	09/2015–12/2016	USA	Cross-sectional; Level 3	11	83
Saad (2018)[20]	JSES	01/2015–10/2017	USA	Prospective cross-sectional; Level 3	11	78

[a]Study design was obtained from the respective publications except for the 2 cross-sectional studies, Patterson et al and Saad et al, whose levels were determined by the reviewers (I.S., J.-R.H.S.) for this review. AJSM, *American Journal of Sports Medicine*; JSES, *Journal of Shoulder and Elbow Surgery*; MCMS, Modified Coleman Methodology Score; MINORS, Methodological Index for Non-Randomized Studies; NR, not reported; OJSM, *Orthopaedic Journal of Sports Medicine*; PDR, procedure date range (MM/YYYY).

had concerns or a diagnosis of shoulder arthritis or glenohumeral osteoarthritis (3 studies)[6,9,20]; 142 patients had concerns of recurrent shoulder instability (2 studies)[2,11]; and 145 patients had a subscapularis (SSc) tear (1 study)[15] (Table 1). The mean age at the time of treatment was 56.6 years for rotator cuff disease, 62.7 years for arthritis, 24.6 years for recurrent shoulder instability, 67.6 years for unspecified condition requiring total shoulder arthroscopy, and 59.3 years for SSc repairs (Table 1).

Study Method Assessment

Table 2 presents the MINORS and MCMS scores of the included studies. The average MINORS score for noncomparative studies was 12.4, indicating overall adequate reporting on the specified criteria. The average MCMS score was 81.8, indicating overall good methodological quality of the included studies.

Country of Origin

All of the included studies[1,2,6,9,11,15,17,19,20] took place in the United States, and their authors specified where they received approval to conduct their study (Table 2).

Conflict of Interest

All studies[1,2,6,9,11,15,17,19,20] included disclosures or disclaimers of potential conflicts of interest.

Surgical Technique

Two studies[9,15] provided details on the operative procedure undertaken. Dowdle et al[9] analyzed only patients undergoing primary total shoulder arthroplasty, so they had no comparative groups. Monroe et al[15] compared PROs among patients undergoing isolated and combined arthroscopic SSc tendon repairs and found that outcomes were similar irrespective of the size of the SSc tear and regardless of whether there were concurrent tears to the supraspinatus or infraspinatus; those investigators also reported that biceps abnormality was common in patients with rotator cuff tears. Each study discussed what intervention was undertaken (Table 3).

Clinical Outcomes

PROMIS. The PROMIS UE and PROMIS PF CAT were correlated well with one another, with an average $r = 0.68$

TABLE 3
Study Populations and Outcomes[a]

Study	Patient Population	Follow-up	PROMIS Outcome
Anthony (2017)[1]	Patients with a preoperative diagnosis of RC disease enrolled at the time of their surgical indication for surgical RCR	None specified	PROMIS UE and PROMIS CAT are valid PRO alternatives that have high correlation with traditional shoulder and UE PRO instruments. PROMIS PF CAT has a decreased question burden.
Anthony (2017)[2]	Patents with a primary diagnosis of shoulder instability scheduled to undergo operative intervention for treatment	None specified	The PROMIS UE and PROMIS PF CAT demonstrated good to excellent correlation with common shoulder and UE PRO instruments (including the SF-36 PF) in patients with diagnosed shoulder instability. However, in patients < 21 y, the PROMIS UE showed a significant ceiling effect.
Chen (2019)[6]	Patients pre- and postoperatively (>3 mo) who underwent primary anatomic TSA	3 mo	Preoperative (within 60 d of surgery) PROMIS PF, depression, and PI scores were strong predictors of postoperative (at 3-mo follow-up) outcomes after shoulder arthroplasty. This study compared the different PROMIS platforms with one another.
Dowdle (2017)[9]	Patients preoperatively with glenohumeral osteoarthritis TSA	None specified	PROMIS UE and PROMIS PF CAT may be valid alternative PRO instruments for patients with operative shoulder osteoarthritis by providing a lower question burden with no ceiling effects. This study also compared across PROMIS platforms.
Hajewski (2019)[11]	Patients undergoing operative interventions for shoulder instability	6 wk 6 mo 2 y	The PROMIS PF CAT demonstrated good to excellent correlation with other previously validated PRO instruments to assess physical function in patients postoperatively after a diagnosis of shoulder instability. The PROMIS UE demonstrated good correlation with other PRO tools but had a significant ceiling effect.
Monroe (2019)[15]	Patients who underwent arthroscopic SSc repair	Average follow-up period, 52.2 ± 19.5 mo	This study compared PROMIS-UE against itself over time to assess PRO metrics, such as compliance and changes in patient reports over time.
Nicholson (2019)[17]	Patients with known RC disease without previous history of RCR	None specified	PF scores of the PROMIS Global-10 showed high correlation with previously validated PRO instruments, suggesting that it is a reliable tool for outcome assessment in patients with RC disease. However, large variability in 95% limits of agreement suggested that the estimated EQ-5D scores from the PROMIS Global-10 cannot replace traditional EQ-5D scores.
Patterson (2018)[19]	Patients undergoing arthroscopic RC repair	None specified	PROMIS UE scores indicated greater impairment and demonstrated a stronger correlation with the legacy shoulder scores than PROMIS PF scores in patients with symptomatic RC tears. PROMIS CATs allowed for more efficient PRO data collection compared with traditional outcome scores.
Saad (2018)[20]	111 Patients with glenohumeral arthritis (69%), 29 patients with arthritis and RC tears (18%), and 21 patients with RC tear arthropathy (13%) before intervention	None specified	PROMIS Global-10 physical scores showed excellent correlation with the EQ-5D and good correlation with the ASES Shoulder Score but showed poor correlation with other gold standard PRO instruments, suggesting that it is an inappropriate instrument for outcome measurement in populations with shoulder arthritis.

[a]See Appendix Table A1 for a summary of the included studies. ASES, American Shoulder and Elbow Surgeons; CAT, computer adaptive test; EQ-5D, Euro-Qol 5 Dimensions; PF, physical function; PI, pain interference; PRO, patient-reported outcome; PROMIS, Patient-Reported Outcomes Measurement Information System; RC, rotator cuff; RCR, rotator cuff repair; SF-36, 36-Item Short Form Health Survey; SSc, subscapularis; TSA, total shoulder arthroscopy; UE, upper extremity.

TABLE 4
Study PROMIS Domains[a]

Study	PROMIS Domains					PROMIS PF CAT No. of Questions	Total Domains
	PF CAT	UE	PI	Dn	Global-10		
Anthony (2017)[1]	X	X				4.3 ± 1.2	2
Anthony (2017)[2]	X	X				4.6 ± 1.8	2
Chen (2019)[6]	X		X	X		NA	3
Dowdle (2017)[9]	X	X				4	2
Hajewski (2019)[11]	X	X				7	2
Monroe (2019)[15]		X				NA	1
Nicholson (2019)[17]					X	NA	1
Patterson (2018)[19]		X	X		X	Up to 121 possible	3
Saad (2018)[20]					X	NA	1

[a]CAT, computer adaptive test; Dn, depression; NA, not applicable; PF, physical function; PI, pain interference; PROMIS, Patient-Reported Outcomes Measurement Information System; UE, upper extremity.

TABLE 5
Association of PROMIS UE (at Earliest Time Point) With Other Outcome Measures[a]

Study	PROMIS Only	PROMIS PF CAT	WORC, WOSI, or WOOS	ASES	Marx Shoulder Activity Scale	SF-36 GH or PF	EQ-5D	Other
Anthony (2017)[1]		$r = 0.7$ $P < .01$	$r = 0.73$ $P < .01$	$r = 0.77$ $P < .01$		$r = 0.66$ $P < .01$	$r = 0.73$ $P < .01$	
Anthony (2017)[2]		$r = 0.63$ $P < .01$	$r = 0.63$ $P < .01$	$r = 0.71$ $P < .01$	$r = 0.06$ $P = .65$	$r = 0.78$ $P < .01$	$r = 0.66$ $P < .01$	
Chen (2019)[6]	PROMIS PF, PI, and Dn							
Dowdle (2017)[9]		$r = 0.81$ $P < .0001$	$r = 0.34$ $P < .01$	$r = 0.55$ $P < .0001$	$r = 0.06$ $P = .62$	$r = 0.53$ $P < .01$	$r = 0.48$ $P < .01$	
Hajewski (2019)[11]		$r = 0.57$ $P < .0001$	$r = 0.60$ $P < .01$	$r = 0.76$ $P < .01$	$r = 0.11$ $P = .37$[b]	$r = 0.70$ $P < .01$	$r = 0.53$ $P < .01$[b]	
Monroe (2019)[15]	PROMIS UE only							
Nicholson (2019)[17]	PROMIS Global-10 only							
Patterson (2018)[19]				$r = 0.59$ $P < .001$				SST $r = 0.62$ $P < .001$
Saad (2018)[20]	PROMIS Global-10 only							

[a]Correlation only at the earliest time point is reported because that was the only time point consistently reported across all of the included studies. All reported *r* and *P* values were obtained from the analyses performed in each respective study. Type of statistical analysis performed is listed in Appendix Table A2. ASES, American Shoulder and Elbows Surgeons Shoulder Score; CAT, computer adaptive test; Dn, depression; EQ-5D, Euro-Qol 5 Dimensions; GH, global health; PF, physical function; PI, pain interference; PROMIS, Patient-Reported Outcomes Measurement Information System; SF-36, 36-Item Short Form Health Survey; SST, Simple Shoulder Test; UE, upper extremity; WOOS, Western Ontario Osteoarthritis Shoulder Index; WORC, Western Ontario Rotator Cuff Index; WOSI, Western Ontario Shoulder Instability Index.
[b]Divergent validity is listed for Hajewski et al comparing PROMIS UE with the Marx Shoulder Activity Score and the EQ-5D.

reported across the 4 studies[1,2,9,11] that used them both. In addition, the PROMIS UE matched moderately ($r < 0.7$) to the ASES Shoulder Score, with an average $r = 0.68$ from the 5 studies[1,2,9,11,19] that used them both, followed by the SF-36 GH, with an average $r = 0.67$ from the 4 studies[1,2,9,11] that compared them (Tables 4 and 5).

The PROMIS PF CAT platform was more closely correlated with the SF-36 GH 0.74 per Table 6 with an average $r = 0.74$ across the 4 studies[1,2,9,11] that compared them, followed by the EQ-5D with an average $r = 0.67$ from the 6 studies[1,2,9,11,17,20] that compared them (Table 6).

Overall, the worst correlations were between the PROMIS UE and the Marx Shoulder Activity Score, with an average $r = 0.08$ among 3 studies[2,9,11] that compared them. The PROMIS PF CAT and Marx Shoulder Activity Scale had an average $r = 0.18$ (Tables 5 and 6).

All studies[1,2,9,11] that compared the PROMIS UE and PROMIS PF CAT, WORC, WOSI, or WOOS, ASES Shoulder Score, SF-36, and the EQ-5D reported correlation coefficients (strength of association between the relative values) with statistical significance ($P < .05$) (Table 5).

All studies that compared the PROMIS PF CAT and ASES Shoulder Score, SF-36 GH, or EQ-5D reported correlation coefficients with statistical significance ($P < .05$) (Table 6).

TABLE 6
Correlation of PROMIS PF CAT (or PROMIS Global-10 With PF) at Earliest Time Point With Other Outcome Measures[a]

Study	PROMIS Only	PROMIS UE	WORC, WOSI, or WOOS	ASES	Marx Shoulder Activity Scale	SF-36 GH or PF	EQ-5D	Other
Anthony (2017)[1]	NR	$r = 0.7$ $P < .01$	$r = 0.61$ $P < .01$	$r = 0.77$ $P < .01$	NR	$r = 0.77$ $P < .01$	$r = 0.65$ $P < .01$	
Anthony (2017)[2]	NR	$r = 0.63$ $P < .01$	$r = 0.49$ $P < .01$	$r = 0.67$ $P < .01$	$r = 0.18$ $P = .14$	$r = 0.72$ $P < .01$	$r = 0.59$ $P < .01$	
Chen (2019)[6]	PROMIS PF, PI, and Dn							
Dowdle (2017)[9]			$r = 0.51$ $P < .01$	$r = 0.62$ $P < .0001$	$r = 0.29$ $P = .02$	$r = 0.81$ $P < .001$	$r = 0.64$ $P < .001$	
Hajewski (2019)[11]		$r = 0.57$ $P < .01$	$r = 0.53$ $P < .01$	$r = 0.67$ $P < .01$	$r = 0.07$ $P = .58$[b]	$r = 0.65$ $P < .0001$	$r = 0.63$ $P < .01$[b]	
Monroe (2019)[15]	PROMIS UE only							
Nicholson (2019)[17]			$r = 0.32$ $P < .001$	$r = 0.62$ $P < .001$			$r = 0.7$ $P < .001$	SANE $r = 0.41$ $P < .005$
Patterson (2018)[19]				$r = 0.43$ $P < .001$				SST $r = 0.51$ $P < .001$
Saad (2018)[20]			$r = 0.09$ $P = .43$	$r = 0.57$ $P < .001$			$r = 0.72$ $P < .001$	SANE $r = 0.23$ $P = .0045$

[a]All reported r and P values were obtained from the analyses performed in each respective study. Type of statistical analysis performed is listed in Appendix Table A2. Correlation only at the earliest time point is reported because that was the only time point consistently reported across all of the included studies. ASES, American Shoulder and Elbows Surgeons Shoulder Score; CAT, computer adaptive test; Dn, depression; EQ-5D, Euro-Qol 5 Dimensions; GH, global health; NR, not reported; PF, physical function; PI, pain interference; PROMIS, Patient-Reported Outcomes Measurement Information System; SANE, Single Assessment Numeric Evaluation; SF-36, 36-Item Short Form Health Survey; SST, Simple Shoulder Test; UE, upper extremity; WOOS, Western Ontario Osteoarthritis Shoulder Index; WORC, Western Ontario Rotator Cuff Index; WOSI, Western Ontario Shoulder Instability Index.
[b]Hajewski et al reported values for comparison with the Marx Shoulder Activity Scale and the EQ-5D as divergent validity.

PROs. Of the 9 studies, 4 studies[1,2,9,11] reported on both the PROMIS PF CAT and the UE domains. Two studies[17,20] used the PROMIS Global-10, 1 study[15] used only the PROMIS UE, and 1 study[6] used the PROMIS PF, PI, and depression domains. The PROMIS Global-10 is a set of 10 questions designed to work for a variety of health conditions and serves to assess multiple aspects of health and functioning including physical and mental health, social health, pain, fatigue, and overall perceived quality of life (Tables 5 and 6).

Patient Satisfaction

Ultimately, a broad goal of PRO measurements is to assess patient satisfaction through tangible domains, such as mental health, physical health, emotional health, and social health.[8] Dynamic measurements, such as PROs, can help clinicians evaluate their performance and make changes to their practice based on direct patient feedback; further, having a consistent platform, with a standardized, established set of outcomes to assess, can make practical review simpler and more likely to elicit relevant changes.

DISCUSSION

Among the studies that reviewed PROMIS against legacy PRO measurements, against its various platforms, and over time, there were a few common points for improvement: increasing accessibility, physician education in order to expand administration rates, and follow-up to track patient progress. With more data at different time points and reports on physical rehabilitation, surgical technique, and consistent BMI logging, the PROMIS domain could surpass many of the nonstandard PRO measurements. The breadth at which the PROMIS PF CAT domain aims to serve could be beneficial outside of orthopaedics and may be applicable for physician, procedure, and even hospital evaluation given how simple the test is to administer.[5,18] The pitfalls of electronic administration need to be considered, but with increasing access to electronic platforms, such as applications on smartphones, tablets, and computers, there is tremendous potential in mass data collection via PRO instruments, such as the PROMIS.[19] Potential adaptations of the PROMIS platform for highly active, athletic populations remains unavailable, but such a form could improve score correlations among PRO measurements by removing ceiling effects.

Additionally, the NIH PROMIS platform was normalized to the average American citizen based on US census data from 2000. All of the studies in this review were published in 2017-2018, although PROMIS first launched in 2004, and this opens potential for discrepancy between the

baseline standardization and the current population 13 years later. The US census data may not adequately reflect the population at large in the United States and may not be generalizable to other populations, meaning that uniformity on a large scale is unlikely.[10] Nonetheless, of the 7 studies[1,2,9,11,17,19,20] that reported on ceiling effects, 3 studies[1,9,20] reported no significant floor or ceiling effects with the PROMIS PF CAT form. Anthony et al,[2] in their study of patients with shoulder instability, reported that with the PROMIS UE, ceiling effects were present in 28.6% of patients aged 18 to 21 years, which is the most significant report among the included studies.

Hajewski et al[11] reported on the ceiling and floor effects of PROMIS UE and PF CAT and found significant ceiling effects in the PROMIS UE at 6 months (68.1% of included PROs) and at 2 years (67.0% of included PROs). Patterson et al[19] reported that the PROMIS UE had ceiling effects in 3% of patients, showcasing that the type of PROMIS administered may be relevant to the population to which it is administered. Because the ceiling data provided were reported only as percentages and were present for only 1 of the conditions considered (instability), insufficient information was available to pool and analyze for an effect in this review.

The results of this systematic review suggested that the PROMIS domains (UE and PF) demonstrated a strong correlation with previous legacy PRO instruments in patients undergoing surgical intervention for shoulder conditions. Using PROMIS to assess patients undergoing surgical intervention for shoulder injury or disease could simplify the administration and analysis of PRO instruments for physicians as well as lower the question burden for patients, as the goal of the NIH PROMIS initiative is to streamline the number of potential PRO measurements used by having a few broader sets of standardized surveys.[16]

Question Burden and Survey Timing

One of the most obvious drawbacks to using surveys to collect data from patients is survey completion related to "survey fatigue," or question burden. Survey developers and administrators need to be sensitive to how much is reasonable to ask of a patient. CATs provide the ability to track how long a patient takes to answer a questionnaire and even how long a patient spends on each question.

Limitations

As with any systematic review, there are limitations to the data provided. In this study, only complete data were analyzed. This means that important variables, such as completion rate, internal consistency, reproducibility, reliability, sensitivity to change, and ceiling effects, do not have pooled data available for comparison. Nonetheless, given that the PROMIS platform is relatively new, this study serves as a starting point for gathering what common data are available.

CONCLUSION

From available data, the PROMIS PF and PROMIS UE were most closely correlated with outcomes measured in SF-36 GH. The PROMIS UE alone was most correlated with ASES Shoulder Score. Thus, either PROMIS PF or UE may provide a possible alternative to legacy PRO measurements but with an overall lower overall question burden and higher generalizability. Future research should compare the time and question burden of the various PROMIS platforms with a more consistent evaluation of standard PRO measurements.

REFERENCES

1. Anthony CA, Glass N, Hancock K, Bollier M, Hettrich CM, Wolf BR. Preoperative performance of the Patient-Reported Outcomes Measurement Information System in patients with rotator cuff pathology. *Arthroscopy*. 2017;33(10):1770-1774.e1771.
2. Anthony CA, Glass NA, Hancock K, Bollier M, Wolf BR, Hettrich CM. Performance of PROMIS instruments in patients with shoulder instability. *Am J Sports Med*. 2017;45(2):449-453.
3. Beach WR. Editorial commentary: Patient Reported Outcomes Measurement Information System (PROMIS) may be our promise for the future. *Arthroscopy*. 2017;33(10):1775-1776.
4. Beckmann JT, Hung M, Bounsanga J, Wylie JD, Granger EK, Tashjian RZ. Psychometric evaluation of the PROMIS physical function computerized adaptive test in comparison to the American Shoulder and Elbow Surgeons score and simple shoulder test in patients with rotator cuff disease. *J Shoulder Elbow Surg*. 2015;24(12):1961-1967.
5. Cella D, Gershon R, Lai JS, Choi S. The future of outcomes measurement: item banking, tailored short-forms, and computerized adaptive assessment. *Qual Life Res*. 2007;16(suppl 1):133-141.
6. Chen RE, Papuga MO, Nicandri GT, Miller RJ, Voloshin I. Preoperative Patient-Reported Outcomes Measurement Information System (PROMIS) scores predict postoperative outcome in total shoulder arthroplasty patients. *J Shoulder Elbow Surg*. 2019;28(3):547-554.
7. Coleman BD, Khan KM, Maffulli N, Cook JL, Wark JD. Studies of surgical outcome after patellar tendinopathy: clinical significance of methodological deficiencies and guidelines for future studies. Victorian Institute of Sport Tendon Study Group. *Scand J Med Sci Sports.* 2000;10(1):2-11.
8. Donaldson G. Patient-reported outcomes and the mandate of measurement. *Qual Life Res*. 2008;17(10):1303-1313.
9. Dowdle SB, Glass N, Anthony CA, Hettrich CM. Use of PROMIS for patients undergoing primary total shoulder arthroplasty. *Orthop J Sports Med*. 2017;5(9):2325967117726044.
10. Fu MC, Chang B, Wong AC, et al. PROMIS physical function underperforms psychometrically relative to American Shoulder and Elbow Surgeons score in patients undergoing anatomic total shoulder arthroplasty. *J Shoulder Elbow Surg*. 2019;28(9):1809-1815.
11. Hajewski CJ, Glass NA, Westermann RW, Bollier M, Wolf BR, Hettrich C. Performance of the PROMIS after operative interventions for shoulder instability. *Orthop J Sports Med*. 2019;7(5):2325967119846920.
12. Kaat AJ, Rothrock NE, Vrahas MS, et al. Longitudinal validation of the PROMIS physical function item bank in upper extremity trauma. *J Orthop Trauma*. 2017;31(10):e321-e326.
13. Makhni EC, Meadows M, Hamamoto JT, Higgins JD, Romeo AA, Verma NN. Patient Reported Outcomes Measurement Information System (PROMIS) in the upper extremity: the future of outcomes reporting? *J Shoulder Elbow Surg*. 2017;26(2):352-357.
14. Minoughan CE, Schumaier AP, Fritch JL, Grawe BM. Correlation of Patient-Reported Outcome Measurement Information System physical function upper extremity computer adaptive testing, with the American Shoulder and Elbow Surgeons shoulder assessment form

and simple shoulder test in patients with shoulder pain. *Arthroscopy*. 2018;34(5):1430-1436.
15. Monroe EJ, Flores SE, Chambers CC, et al. Patient-reported outcomes after isolated and combined arthroscopic subscapularis tendon repairs. *Arthroscopy*. 2019;35(6):1779-1784.
16. Morgan JH, Kallen MA, Okike K, Lee OC, Vrahas MS. PROMIS physical function computer adaptive test compared with other upper extremity outcome measures in the evaluation of proximal humerus fractures in patients older than 60 years. *J Orthop Trauma*. 2015;29(6):257-263.
17. Nicholson AD, Kassam HF, Pan SD, Berman JE, Blaine TA, Kovacevic D. Performance of PROMIS Global-10 compared with legacy instruments for rotator cuff disease. *Am J Sports Med*. 2019;47(1):181-188.
18. National Institutes of Health. *PROMIS*. 2016. http://www.healthmeasures.net/promis. Accessed June 6, 2019.
19. Patterson BM, Orvets ND, Aleem AW, et al. Correlation of Patient-Reported Outcomes Measurement Information System (PROMIS) scores with legacy patient-reported outcome scores in patients undergoing rotator cuff repair. *J Shoulder Elbow Surg*. 2018;27(suppl 6):S17-S23.
20. Saad MA, Kassam HF, Suriani RJ Jr, Pan SD, Blaine TA, Kovacevic D. Performance of PROMIS Global-10 compared with legacy instruments in patients with shoulder arthritis. *J Shoulder Elbow Surg*. 2018;27(12):2249-2256.
21. Slim K, Nini E, Forestier D, Kwiatkowski F, Panis Y, Chipponi J. Methodological Index for Non-Randomized Studies (MINORS): development and validation of a new instrument. *ANZ J Surg*. 2003;73(9):712-716.
22. Vasta S, Papalia R, Albo E, Maffulli N, Denaro V. Top orthopedic sports medicine procedures. *J Orthop Surg Res*. 2018;13(1):190.
23. Wright RW, Baumgarten KM. Shoulder outcomes measures. *J Am Acad Orthop Surg*. 2010;18(7):436-444.

APPENDIX

TABLE A1
Study Summaries[a]

Study	Study Summary	Patients Excluded or Bias
Anthony (2017)[1]	Patients with a preoperative diagnosis of RC disease were asked to fill out multiple verified PRO instruments. The assessment forms were then compared with one another for outcome measure.	No patients excluded from participation; 91 enrolled, and 82 had full data
Anthony (2017)[2]	Patients with a primary diagnosis of shoulder instability were asked to complete a series of validated PRO instruments along with PROMIS surveys to compare PRO instruments.	Patients with incomplete data (4 from the initial 74 scheduled for operative intervention)
Chen (2019)[6]	Patients scheduled for total shoulder arthroplasty were administered PROMIS PF, PI, and depression tests pre- and postoperatively.	Patients excluded if they underwent a reverse total shoulder arthroplasty or if their procedure were a revision shoulder arthroplasty
Dowdle (2017)[9]	Patients with glenohumeral osteoarthritis scheduled to undergo primary total shoulder arthroplasty were asked to complete validated PRO instruments along with PROMIS before surgery.	Those with incomplete PRO data
Hajewski (2019)[11]	Patients scheduled to undergo operative interventions for shoulder instability completed PROMIS along with other verified questionnaires preoperatively and at 6 weeks, 6 months, and 2 years postoperatively.	Those with incomplete PRO data
Monroe (2019)[15]	PROMIS UE was used to compare preoperative variables and outcomes among isolated partial SSc repair, partial SSc with SS and/or IS repair, isolated complete SSc repair, and complete SSc with SS and/or IS repair in 145 shoulders.	Patients who underwent open SSc repair; 1 patient excluded because of previous biceps tenodesis
Nicholson (2019)[17]	Patients with known RC disease were prospectively enrolled before treatment and asked to complete PROMIS Global-10 along with several other PRO instruments for comparison.	Patients with previous RC surgery
Patterson (2018)[19]	Patients undergoing arthroscopic RC repair were asked to complete PRO instruments preoperatively.	Patients excluded if they were undergoing revision RC repair, were younger than 18 years, or incompletely responded to 1 or more surveys such that the survey could not be scored
Saad (2018)[20]	Patients with shoulder arthritis were asked to complete PRO questionnaires before receiving treatment.	None; patients enrolled based on eligibility criteria

[a]IS, infraspinatus; PF, physical function; PI, pain interference; PRO, patient-reported outcome; PROMIS, Patient-Reported Outcomes Measurement Information System; RC, rotator cuff; SS, supraspinatus; SSc, subscapularis; UE, upper extremity.

TABLE A2
Patient-Reported Outcome Instruments Administered in Each Respective Study[a]

Study	Assessment Method (Correlation Coefficient)	PRO Instruments Compared							Total
		PROMIS Only	ASES	Marx Shoulder Activity Scale	SF-36 GH or PF	EQ-5D	WORC, WOSI, or WOOS	Other	
Anthony (2017)[1]	Pearson and/or Spearman		X	X	GH	X	WORC		5
Anthony (2017)[2]	Pearson and/or Spearman		X	X	PF	X	WOSI		5
Chen (2019)[6]	Pearson	X							2
Dowdle (2017)[9]	Pearson and/or Spearman		X	X	PF	X	WOOS		5
Hajewski (2019)[11]	Spearman		X	X	GH and PF		WOSI		4
Monroe (2019)[15]	NA	X						Postoperative VAS	2
Nicholson (2019)[17]	Spearman	X				X	WORC	SANE	4
Patterson (2018)[19]	Pearson		X					SST, PI	3
Saad (2018)[20]	Spearman		X			X	WOOS	SANE	4

[a]ASES, American Shoulder and Elbows Surgeons Shoulder Score; EQ-5D, Euro-Qol 5 Dimensions; GH, global health; NA, not available; PF, physical function; PI, pain interference; PRO, patient-reported outcome; PROMIS, Patient-Reported Outcomes Measurement Information System; SANE, Single Assessment Numeric Evaluation; SF-36, 36-Item Short Form Health Survey; SST, Simple Shoulder Test; WOOS, Western Ontario Osteoarthritis Shoulder Index; WORC, Western Ontario Rotator Cuff Index; WOSI, Western Ontario Shoulder Instability Index; VAS, visual analog scale for pain.