



5-2019

Performance of the PROMIS After Operative Interventions for Shoulder Instability

Christina J. Hajewski
University of Iowa

Natalie A. Glass
University of Iowa

Robert W. Westermann
University of Iowa

Matthew Bollier
University of Iowa

Brian R. Wolf
University of Iowa

See next page for additional authors

Follow this and additional works at: https://uknowledge.uky.edu/orthopaedicsurgery_facpub



Part of the [Orthopedics Commons](#), and the [Sports Sciences Commons](#)

Right click to open a feedback form in a new tab to let us know how this document benefits you.

Repository Citation

Hajewski, Christina J.; Glass, Natalie A.; Westermann, Robert W.; Bollier, Matthew; Wolf, Brian R.; and Hettrich, Carolyn, "Performance of the PROMIS After Operative Interventions for Shoulder Instability" (2019). *Orthopaedic Surgery and Sports Medicine Faculty Publications*. 22.
https://uknowledge.uky.edu/orthopaedicsurgery_facpub/22

This Article is brought to you for free and open access by the Orthopaedic Surgery and Sports Medicine at UKnowledge. It has been accepted for inclusion in Orthopaedic Surgery and Sports Medicine Faculty Publications by an authorized administrator of UKnowledge. For more information, please contact UKnowledge@lsv.uky.edu.

Performance of the PROMIS After Operative Interventions for Shoulder Instability

Abstract

Background: The Patient-Reported Outcomes Measurement Information System (PROMIS) was developed to capture patient-reported outcomes (PROs) in an efficient manner. Few studies have assessed this instrument postoperatively.

Purpose: To compare the PROMIS Physical Function computer adaptive test (PROMIS PF CAT) and Upper Extremity (PROMIS UE) item bank to other previously validated PRO instruments and to evaluate ceiling and floor effects and construct validity responsiveness in patients who underwent operative interventions for shoulder instability.

Study Design: Cohort study (diagnosis); Level of evidence, 2.

Methods: A total of 72 patients who underwent operative interventions for shoulder instability completed the American Shoulder and Elbow Surgeons (ASES) assessment form, Marx shoulder activity scale (Marx), 36-Item Short Form Health Survey physical function (SF-36 PF) and general health (SF-36 GH), Western Ontario Shoulder Instability Index (WOSI), PROMIS PF CAT, and PROMIS UE before surgery and then at 6 weeks and 6 months postoperatively. Correlation coefficients were calculated among these tools. The effect size of change was also calculated for each tool at each time point. A total of 91 patients who had also undergone surgery for shoulder instability completed these PRO instruments 2 years postoperatively. The percentage of patients hitting the ceiling and floor effects of each of the PRO instruments was calculated at all time points.

Results: The PROMIS PF CAT demonstrated excellent-good correlation with the SF-36 PF at all postoperative time points (0.61 at 6 weeks, 0.68 at 6 months, and 0.64 at 2 years; $P < .01$ for all). The PROMIS UE showed excellent correlation with the ASES at 6 weeks postoperatively (0.73, $P < .01$). Both the PROMIS PF CAT and PROMIS UE demonstrated the ability to detect change after surgical interventions with a medium to large effect size. The PROMIS UE demonstrated a ceiling effect at 6 months (68.1%) and 2 years (67.0%) postoperatively. The PROMIS PF CAT demonstrated no ceiling effect at any time point.

Conclusion: The PROMIS PF CAT demonstrated good to excellent correlation with other previously validated PRO instruments that assess physical function in patients with shoulder instability postoperatively. The PROMIS UE demonstrated good correlation with other PRO tools but had a significant ceiling effect and is not recommended for this patient population. Both tools demonstrated an ability to detect change after surgical interventions with a good effect size.

Keywords

outcomes, shoulder, instability, PROMIS

Disciplines

Orthopedics | Sports Sciences

Notes/Citation Information

Published in *Orthopaedic Journal of Sports Medicine*, v. 7, issue 5, p. 1-6.

© The Author(s) 2019

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDeriv 4.0 License (<http://www.creativecommons.org/licenses/by-nc-nd/4.0/>) which permits non-commercial

use, reproduction and distribution of the work as published without adaptation or alteration, without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

Authors

Christina J. Hajewski, Natalie A. Glass, Robert W. Westermann, Matthew Bollier, Brian R. Wolf, and Carolyn Hettrich

Performance of the PROMIS After Operative Interventions for Shoulder Instability

Christina J. Hajewski,^{*†} MD, Natalie A. Glass,[†] PhD, Robert W. Westermann,[†] MD, Matthew Bollier,[†] MD, Brian R. Wolf,[†] MD, MS, and Carolyn Hettrich,[‡] MD, MPH

Investigation performed at the University of Iowa Hospitals and Clinics, Iowa City, Iowa, USA

Background: The Patient-Reported Outcomes Measurement Information System (PROMIS) was developed to capture patient-reported outcomes (PROs) in an efficient manner. Few studies have assessed this instrument postoperatively.

Purpose: To compare the PROMIS Physical Function computer adaptive test (PROMIS PF CAT) and Upper Extremity (PROMIS UE) item bank to other previously validated PRO instruments and to evaluate ceiling and floor effects and construct validity responsiveness in patients who underwent operative interventions for shoulder instability.

Study Design: Cohort study (diagnosis); Level of evidence, 2.

Methods: A total of 72 patients who underwent operative interventions for shoulder instability completed the American Shoulder and Elbow Surgeons (ASES) assessment form, Marx shoulder activity scale (Marx), 36-Item Short Form Health Survey physical function (SF-36 PF) and general health (SF-36 GH), Western Ontario Shoulder Instability Index (WOSI), PROMIS PF CAT, and PROMIS UE before surgery and then at 6 weeks and 6 months postoperatively. Correlation coefficients were calculated among these tools. The effect size of change was also calculated for each tool at each time point. A total of 91 patients who had also undergone surgery for shoulder instability completed these PRO instruments 2 years postoperatively. The percentage of patients hitting the ceiling and floor effects of each of the PRO instruments was calculated at all time points.

Results: The PROMIS PF CAT demonstrated excellent-good correlation with the SF-36 PF at all postoperative time points (0.61 at 6 weeks, 0.68 at 6 months, and 0.64 at 2 years; $P < .01$ for all). The PROMIS UE showed excellent correlation with the ASES at 6 weeks postoperatively (0.73, $P < .01$). Both the PROMIS PF CAT and PROMIS UE demonstrated the ability to detect change after surgical interventions with a medium to large effect size. The PROMIS UE demonstrated a ceiling effect at 6 months (68.1%) and 2 years (67.0%) postoperatively. The PROMIS PF CAT demonstrated no ceiling effect at any time point.

Conclusion: The PROMIS PF CAT demonstrated good to excellent correlation with other previously validated PRO instruments that assess physical function in patients with shoulder instability postoperatively. The PROMIS UE demonstrated good correlation with other PRO tools but had a significant ceiling effect and is not recommended for this patient population. Both tools demonstrated an ability to detect change after surgical interventions with a good effect size.

Keywords: outcomes; shoulder; instability; PROMIS

Patient-reported outcomes (PROs) are becoming increasingly important in our current health care delivery system. Identifying tools to accurately and efficiently capture PROs will help providers better understand patients' perspective of their abnormality and treatment. These tools are also valuable to providers by allowing them to appropriately counsel patients about expected outcomes before and after operative interventions.

Shoulder instability is not uncommon, especially in young and athletic populations. The incidence of anterior instability has been shown to be between 11.2 per 100,000 person-years¹⁸ and 23.9 per 100,000 person-years¹⁹ in the United States. These patients are often

young and male, and collision and overhead athletes have a higher risk.^{7,8}

Previously validated PRO tools used for the evaluation of patients with shoulder instability include the American Shoulder and Elbow Surgeons (ASES) assessment form,¹² the Western Ontario Shoulder Instability Index (WOSI),¹¹ and the Marx shoulder activity scale.⁵ Other PRO instruments that address general health-related quality of life have also been used in the assessment of patients with shoulder instability, such as the 36-Item Short Form Health Survey, which includes subscales for physical function (SF-36 PF), general health (SF-36 GH), vitality, bodily pain, physical role functioning, emotional role functioning, social role functioning, and mental health,¹³ and the Euro-Qol-5 Dimensions (EQ-5D) questionnaire.¹⁰ Each of these instruments includes several domains to evaluate a patient's outcome, including physical function, pain, and

The Orthopaedic Journal of Sports Medicine, 7(5), 2325967119846920

DOI: 10.1177/2325967119846920

© The Author(s) 2019

This open-access article is published and distributed under the Creative Commons Attribution - NonCommercial - No Derivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits the noncommercial use, distribution, and reproduction of the article in any medium, provided the original author and source are credited. You may not alter, transform, or build upon this article without the permission of the Author(s). For article reuse guidelines, please visit SAGE's website at <http://www.sagepub.com/journals-permissions>.

general health. The ideal tool should have a low respondent burden and have the ability to detect change before and after an intervention.

The Patient-Reported Outcomes Measurement Information System (PROMIS) was developed by the National Institutes of Health in an effort to improve the assessment of PROs by creating a tool that is highly reliable and precise while decreasing the burden on both patients and health care providers. Specifically, the PROMIS Physical Function computer adaptive test (PROMIS PF CAT) has been investigated in multiple patient populations with orthopaedic abnormalities.^{3,4,17} The PROMIS Upper Extremity (PROMIS UE) item bank has also been utilized to study PROs in patients with rotator cuff abnormalities^{2,16} and those undergoing shoulder arthroplasty.⁶ Both of these tools have been previously investigated in patients with a diagnosis of shoulder instability before operative interventions. It was shown that there was good to excellent correlation with other legacy upper extremity PRO instruments; however, the PROMIS UE had significant ceiling effects in patients aged ≤ 21 years.³ Fewer studies have looked at these 2 tools to evaluate patients after surgery when activity levels and use of the operative extremity should be higher and pain levels should be improved.

For patients undergoing a surgical intervention, it is of utmost importance to have a PRO instrument that can capture a meaningful change to judge the impact of the surgical intervention. Ideally, a PRO tool should have the ability both to detect change over a time frame (responsiveness) and to correspond to changes in a reference PRO instrument (construct validity). For these reasons, we aimed to investigate the PROMIS PF CAT and PROMIS UE in patients with shoulder instability before and after operative interventions. We hypothesized that (1) these tools can detect a change before and after an intervention and demonstrate a large effect size of that change; (2) the PROMIS UE may show a ceiling effect after shoulder stabilization in young patients; and (3) there would be convergent validity with other measures of physical function, such as the SF-36 PF, ASES, and WOSI, and divergent validity with measures of general health, such as the SF-36 GH and EQ-5D.

METHODS

This study was approved by our institutional review board and was determined to be HIPAA (Health Insurance Portability and Accountability Act) compliant. Patients enrolled in our institution's prospective shoulder instability

database are asked to complete the ASES, WOSI, Marx, SF-36 (PF and GH subscales included), EQ-5D, PROMIS PF CAT, and PROMIS UE preoperatively and at 6 weeks, 6 months, and 2 years postoperatively. Data from patients enrolled in the database from January 21, 2015, to November 28, 2016, were extracted. A total of 96 patients were identified from this data extraction, and 72 were found to have complete preoperative as well as 6-week and 6-month postoperative data available for longitudinal analysis. Only 18 patients from this group had complete longitudinal data from preoperatively to 2 years postoperatively. Therefore, a second group of patients was identified from the database who were eligible for 2-year follow-up from June 9, 2015, to November 26, 2017. A total of 106 patients were identified in this time frame, and 91 patients had complete data available at 2 years postoperatively for cross-sectional analysis. Eighteen of these patients had preoperative data and overlapped with the longitudinal group. Patient demographic data including age, body mass index (BMI), sex, operative side, and procedure type were extracted from patient records. The typical rehabilitation protocol followed the Multicenter Orthopaedic Outcomes Network (MOON) shoulder instability protocols, which included starting physical therapy 2 weeks postoperatively and progressing to active range of motion and resisted isometric exercises by 6 weeks after surgery.¹⁴

Statistical analysis was carried out to determine Spearman correlation coefficients between the PROMIS PF CAT and PROMIS UE with the previously mentioned legacy PRO instruments: ASES, WOSI, SF-36, Marx, and EQ-5D. Correlations were described as excellent (>0.7), excellent-good (0.61-0.7), good (0.4-0.6), or poor (0.2-0.39).⁹ Responsiveness to change was assessed using the effect size (Cohen *d*) and standardized response mean at 6 weeks and 6 months postoperatively. Values were defined as small (0.2), medium (0.5), or large (0.8).¹⁵ The percentage of patients achieving the lowest (floor) and highest (ceiling) possible scores of each PRO tool was calculated, and the presence of a floor or ceiling effect was determined if $>15\%$ of patients were achieving these values. SAS statistical software (version 9.4; SAS Institute) was used for analyses, with $P < .05$ considered statistically significant.

RESULTS

Of the 72 patients with data available for longitudinal analysis, the mean age was 22.1 years (range, 14-44 years), 79%

*Address correspondence to Christina J. Hajewski, MD, University of Iowa Hospitals and Clinics, 200 Hawkins Drive, Iowa City, IA 52242, USA (email: Christina-hajewski@uiowa.edu).

[†]University of Iowa Hospitals and Clinics, Iowa City, Iowa, USA.

[‡]University of Kentucky, Lexington, Kentucky, USA.

One or more of the authors has declared the following potential conflict of interest or source of funding: R.W.W. has received educational support from Arthrex and Smith & Nephew and hospitality payments from Medical Device Business Services. M.B. has received nonconsulting payments from Arthrex. B.R.W. has received research support from OREF, educational support from Arthrex and Smith & Nephew, and consulting payments from Linvatec. C.H. has received research support from Zimmer Biomet, nonconsulting payments from Pacira Pharmaceuticals, and hospitality payments from Arthrex and Tornier. 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.

Ethical approval for this study was obtained from the University of Iowa Institutional Review Board (ID: 201201714).

TABLE 1
Demographic Data and Procedures Performed^a

	Longitudinal Group (n = 72)	Cross-sectional Group (n = 91)
Age, y	22.1 (14-44)	24.8 (17-56)
Body mass index, kg/m ²	26.6 (20-38)	26.4 (20-42)
Male sex, %	79	80
Procedure, n		
Labral repair	40	58
Capsulorrhaphy/capsular plication	10	7
Latarjet	5	5
Mixed	17	21

^aData are reported as mean (range) unless otherwise indicated. Data were available for the longitudinal group preoperatively and at 6 weeks and 6 months postoperatively and for the cross-sectional group at 2 years postoperatively.

were male, and the mean BMI was 26.6 kg/m² (range, 20-38 kg/m²). Of the 91 patients who had complete data at 2 years postoperatively, the mean age was 24.8 years (range, 17-56 years), 80% were male, and the mean BMI was 26.4 kg/m² (range, 20-42 kg/m²). Demographics and procedures performed can be found in Table 1. The median number of questions answered was 7 questions for the PROMIS PF CAT (range, 3-12) and 16 questions for the PROMIS UE. Among the other PRO tools, there were 10 questions for the SF-36 PF, 5 questions for the SF-36 GH, 10 questions for the ASES, 7 questions for the Marx, 5 questions for the EQ-5D, and 21 questions for the WOSI.

PROMIS Changes After Surgery

The PROMIS UE and PROMIS PF CAT demonstrated a statistically significant change between the 3 time points used in longitudinal analysis: preoperatively, 6 weeks postoperatively, and 6 months postoperatively. Scores initially decreased (PROMIS PF CAT: 51.48 to 45.42, $P < .01$; PROMIS UE: 44.02 to 35.14, $P < .01$) at 6 weeks postoperatively and then increased above the preoperative scores at 6 months postoperatively (PROMIS PF CAT: to 57.18, $P < .01$; PROMIS UE: to 52.15, $P < .01$). The effect size (Cohen d) was medium to large at each time point (PROMIS UE: -1.05 ± 0.30 from preoperatively to 6 weeks and 1.09 ± 0.33 from preoperatively to 6 months; PROMIS PF CAT: -0.94 ± 0.33 from preoperatively to 6 weeks and 0.76 ± 0.32 from preoperatively to 6 months). Results are summarized in Table 2.

Floor and Ceiling Effects

Ceiling effects were demonstrated at 6 months (68.1%) and 2 years (67.0%) postoperatively for the PROMIS UE. There was no ceiling effect for the PROMIS PF CAT at any time point. Ceiling effects were also present at 6 months and 2 years postoperatively for the ASES (23.6% and 39.0%, respectively), SF-36 PF (41.7% and 69.0%, respectively), and EQ-5D (48.6% and 39.0%, respectively). A

subgroup analysis of patients aged ≤ 21 years ($n = 38$) demonstrated even larger ceiling effects at 6 months (71.1%) and 2 years (81.0%) postoperatively. The PROMIS UE has been expanded to a CAT to include 46 items (PROMIS UE CAT v 2.0). We had a total of 18 patients who completed this updated version of the PROMIS UE at 2 years postoperatively. Analysis on this group still demonstrated a ceiling effect of 44.0%. These results are summarized in Table 3.

Correlation With Previously Validated Instruments

The PROMIS UE demonstrated the strongest correlation with legacy PRO tools used in shoulder instability, such as the ASES, WOSI, and SF-36 PF. It showed excellent correlation with the ASES at 6 weeks postoperatively (0.73, $P < .01$) and excellent-good correlation with the WOSI (0.62, $P < .01$) and SF-36 PF (0.68, $P < .01$) at 6 weeks postoperatively. It demonstrated good correlation at 6 months and 2 years postoperatively with the ASES (0.54 and 0.58, respectively; $P < .01$), WOSI (0.48 and 0.58, respectively; $P < .01$), SF-36 PF (0.56 and 0.54, respectively; $P < .01$), and EQ-5D (0.46 and 0.55, respectively; $P < .01$). There was poor correlation at all time points with the Marx (6 weeks: 0.06, $P = .61$; 6 months: 0.26, $P = .03$) and SF-36 GH (6 weeks: 0.20, $P = .10$; 6 months: 0.18, $P = .12$).

The PROMIS PF CAT showed excellent-good correlation with the SF-36 PF at all time points (6 weeks: 0.61, $P < .01$; 6 months: 0.68, $P < .01$; 2 years: 0.64, $P < .01$). There was excellent-good correlation at 2 years postoperatively with the ASES (0.64, $P < .01$) and WOSI (0.62, $P < .01$). Results are summarized in Table 4.

DISCUSSION

The present study determined that the PROMIS PF CAT and PROMIS UE are able to adequately detect changes after shoulder stabilization for instability. They both correlated well with shoulder outcome measures that have been previously validated. However, we did find significant ceiling effects for the PROMIS UE. The PROMIS PF CAT did not demonstrate floor or ceiling effects either before or after shoulder stabilization for instability and could be considered for use in this population.

PROs are becoming increasingly important in our health care delivery system and are being incorporated into health care legislation. There is incredible value in identifying PRO tools with a low respondent burden that can detect meaningful change before and after an intervention. The PROMIS was developed by the National Institutes of Health to fulfill these requirements and has been shown to correspond with other previously validated PRO tools for patients with shoulder instability preoperatively; yet, the PROMIS UE was found to have ceiling effects in patients aged ≤ 21 years.³ Few reports in the literature have studied the ability of the PROMIS to detect change before and after surgical interventions, and no studies have investigated this in patients with shoulder instability.

TABLE 2
Effect Size and SRM of PRO Instruments Between Study Time Points^a

	Preoperatively to 6 wk		Preoperatively to 6 mo	
	Effect Size (95% CI)	SRM (95% CI)	Effect Size (95% CI)	SRM (95% CI)
ASES	-0.41 (-0.70 to -0.13)	-0.34 (-0.63 to -0.06)	1.27 (0.93 to 1.60)	1.04 (0.72 to 1.36)
WOSI	0.09 (-0.18 to 0.36)	0.08 (-0.20 to 0.34)	1.83 (1.44 to 2.21)	1.74 (1.37 to 2.12)
Marx	0.25 (0.04 to 0.46)	0.29 (0.08 to 0.50)	0.07 (-0.12 to 0.26)	0.08 (-0.10 to 0.28)
SF-36 PF	-1.20 (-1.52 to -0.89)	-1.14 (-1.45 to -0.83)	0.69 (0.45 to 0.94)	0.73 (0.48 to 0.98)
EQ-5D	-0.12 (-0.38 to 0.14)	-0.11 (-0.38 to -0.15)	0.85 (0.55 to 1.15)	0.73 (0.44 to 1.02)
PROMIS PF CAT	-0.94 (-1.27 to -0.61)	-0.75 (-1.07 to -0.43)	0.76 (0.44 to 1.08)	0.59 (0.29 to 0.91)
PROMIS UE	-1.05 (-1.27 to -0.74)	-0.98 (-1.28 to -0.68)	1.09 (0.77 to 1.41)	0.92 (0.61 to 1.23)

^aValues were defined as small (absolute value of 0.2), medium (absolute value of 0.5), or large (absolute value of 0.8). ASES, American Shoulder and Elbow Surgeons; EQ-5D, EuroQol-5 Dimensions; PRO, patient-reported outcome; PROMIS PF CAT, Patient-Reported Outcomes Measurement Information System Physical Function computer adaptive test; PROMIS UE, Patient-Reported Outcomes Measurement Information System Upper Extremity; SF-36 PF, 36-Item Short Form Health Survey physical function; SRM, standardized response mean; WOSI, Western Ontario Shoulder Instability Index.

TABLE 3
Ceiling and Floor Effects of PRO Instruments^a

	Preoperative		6 wk		6 mo		2 y	
	Ceiling	Floor	Ceiling	Floor	Ceiling	Floor	Ceiling	Floor
ASES	2.8	0.0	0.0	0.0	23.6	0.0	39.0	0.0
WOSI	0.0	0.0	0.0	0.0	2.8	0.0	3.0	0.0
SF-36 PF	13.9	0.0	0.0	0.0	41.7	0.0	69.0	0.0
EQ-5D	13.9	0.0	11.3	0.0	48.6	0.0	39.0	0.0
PROMIS PF CAT	0.0	0.0	0.0	0.0	0.0	0.0	8.0	0.0
PROMIS UE	23.6	0.0	4.2	0.0	68.1	0.0	67.0	0.0
Subgroup analyses								
PROMIS UE CAT v 2.0	—	—	—	—	—	—	44.0	0.0
PROMIS UE ≤ 21 y	29.0	0.0	8.1	0.0	71.1	0.0	81.0	0.0

^aData are reported as percentage of patients. The presence of a ceiling effect was defined as >15% of patients achieving the maximum possible score. ASES, American Shoulder and Elbow Surgeons; PRO, patient-reported outcome; PROMIS PF CAT, Patient-Reported Outcomes Measurement Information System Physical Function computer adaptive test; PROMIS UE, Patient-Reported Outcomes Measurement Information System Upper Extremity; SF-36 PF, 36-Item Short Form Health Survey physical function; WOSI, Western Ontario Shoulder Instability Index.

This study aimed to examine the internal and external responsiveness of 2 PROMIS tools after operative interventions. Internal responsiveness demonstrates the ability of a PRO to change before and after treatment.¹⁵ The PROMIS PF CAT and PROMIS UE both demonstrated a statistically significant change in scores preoperatively compared with 6 weeks and 6 months postoperatively. The PROMIS PF CAT showed a medium to large effect size at 6 weeks and 6 months postoperatively compared with preoperatively, and the PROMIS UE demonstrated a large effect size at both time points. These effect sizes were comparable with or larger than those of the legacy PRO tools evaluated in this study.

External responsiveness can be assessed by examining the correlation between a PRO tool and previously validated PRO instruments that measure a similar health domain. In the current study, we examined the correlation between the PROMIS PF CAT and PROMIS UE and legacy PRO tools used to analyze shoulder function (WOSI, ASES,

and Marx) as well as general health and function (SF-36 PF, SF-36 GH, and EQ-5D). The PROMIS UE generally demonstrated convergent validity with the other measures of shoulder function (ASES, WOSI), while the PROMIS PF CAT showed the most robust correlation with the SF-36 PF at all postoperative time points. Therefore, the PROMIS PF CAT could be administered in place of the SF-36 PF and possibly the ASES and WOSI. There was generally divergent validity between the PROMIS PF CAT and PROMIS UE with other tools that measure other health domains, such as general health and quality of life. This suggests that co-administration of these other tools may need to be continued when investigating these other health domains.

Ceiling and floor effects represent the percentage of patients who achieve the highest or lowest possible scores of a particular PRO instrument. Ideally, a PRO tool would have few patients achieving the extremes to stratify patients appropriately. Our study demonstrated a large ceiling effect with the PROMIS UE, especially at 6 months

TABLE 4
Correlation of PROMIS PF CAT and PROMIS UE With Other PRO Instruments^a

	Preoperative	6 wk	6 mo	2 y
PROMIS UE				
Convergent validity				
ASES	0.76 (<i>P</i> < .01)	0.73 (<i>P</i> < .01)	0.54 (<i>P</i> < .01)	0.58 (<i>P</i> < .01)
WOSI	0.60 (<i>P</i> < .01)	0.62 (<i>P</i> < .01)	0.48 (<i>P</i> < .01)	0.58 (<i>P</i> < .01)
SF-36 PF	0.70 (<i>P</i> < .01)	0.68 (<i>P</i> < .01)	0.56 (<i>P</i> < .01)	0.54 (<i>P</i> < .01)
PROMIS PF CAT	0.57 (<i>P</i> < .01)	0.47 (<i>P</i> < .01)	0.59 (<i>P</i> < .01)	0.54 (<i>P</i> < .01)
Divergent validity				
SF-36 GH	0.32 (<i>P</i> < .10)	0.20 (<i>P</i> = .10)	0.18 (<i>P</i> = .12)	0.38 (<i>P</i> < .01)
EQ-5D	0.63 (<i>P</i> < .01)	0.59 (<i>P</i> < .01)	0.46 (<i>P</i> < .01)	0.55 (<i>P</i> < .01)
Marx	0.11 (<i>P</i> = .37)	0.06 (<i>P</i> = .61)	0.26 (<i>P</i> = .03)	0.30 (<i>P</i> < .01)
PROMIS PF CAT				
Convergent validity				
ASES	0.67 (<i>P</i> < .01)	0.33 (<i>P</i> < .01)	0.52 (<i>P</i> < .01)	0.64 (<i>P</i> < .01)
WOSI	0.53 (<i>P</i> < .01)	0.32 (<i>P</i> < .01)	0.60 (<i>P</i> < .01)	0.62 (<i>P</i> < .01)
SF-36 PF	0.65 (<i>P</i> < .01)	0.61 (<i>P</i> < .01)	0.68 (<i>P</i> < .01)	0.64 (<i>P</i> < .01)
PROMIS UE	0.57 (<i>P</i> < .01)	0.47 (<i>P</i> < .01)	0.59 (<i>P</i> < .01)	0.53 (<i>P</i> < .01)
Divergent validity				
SF-36 GH	0.46 (<i>P</i> < .01)	0.23 (<i>P</i> = .05)	0.33 (<i>P</i> < .01)	0.49 (<i>P</i> < .01)
EQ-5D	0.53 (<i>P</i> < .01)	0.25 (<i>P</i> = .04)	0.43 (<i>P</i> < .01)	0.50 (<i>P</i> < .01)
Marx	0.07 (<i>P</i> = .58)	0.14 (<i>P</i> = .23)	0.21 (<i>P</i> = .07)	0.35 (<i>P</i> < .01)

^aThere was general convergent validity with other instruments focused on physical function and divergent validity with instruments aimed to measure general health. Correlations were defined as excellent (>0.7, black shading), excellent-good (0.61-0.7, dark gray shading), good (0.4-0.6, light gray shading), or poor (0.2-0.39, no shading). ASES, American Shoulder and Elbow Surgeons; EQ-5D, EuroQol-5 Dimensions; PRO, patient-reported outcome; PROMIS PF CAT, Patient-Reported Outcomes Measurement Information System Physical Function computer adaptive test; PROMIS UE, Patient-Reported Outcomes Measurement Information System Upper Extremity; SF-36 GH, 36-Item Short Form Health Survey general health; SF-36 PF, 36-Item Short Form Health Survey physical function; WOSI, Western Ontario Shoulder Instability Index.

and 2 years postoperatively, which is not unexpected if a patient's pain and function improve after surgery. This effect was more dramatic in patients aged ≤ 21 years, which was also detected in preoperative patients.³ Other studies have also suggested that the PROMIS UE seems to focus on lower levels of function, which is problematic when trying to evaluate young and healthy populations.⁹ Such ceiling effects were not present in other studies using the PROMIS UE, including patients with rotator cuff tears^{2,16} or those undergoing shoulder arthroplasty,⁶ who represent older and likely less functional populations. The PROMIS UE has subsequently been expanded to a CAT to include 46 items (PROMIS UE CAT v 2.0). We had a subgroup of patients ($n = 18$) complete this version at 2 years postoperatively, and although this improved the percentage of patients hitting the ceiling effect, 44.0% of this group still achieved the maximum score. This means that the PROMIS UE and PROMIS UE CAT v 2.0 will not be able to detect higher levels of function in these patients and would need to be revised, or other PRO tools such as the Kerlan-Jobe Orthopaedic Clinic overhead athlete shoulder and elbow scale¹ should be utilized in conjunction with this.

There are several limitations to this study. First, we had incomplete longitudinal data among the patients who had outcome scores at 2 years postoperatively, necessitating cross-sectional analysis at this time point. Therefore, although we were able to measure the correlation of the PROMIS with legacy PRO instruments as well as ceiling

and floor effects at 2 years postoperatively, we could not assess responsiveness to change at this time point. Along with this, the longitudinal group was only evaluated to 6 months in the postoperative period, which is near the time to return to full activity. Therefore, conclusions can only be drawn about the correlation of the PROMIS with other PRO instruments at this time point and beyond. Second, all patients were required to complete multiple PRO tools at each visit, and questionnaire fatigue may have had an influence on our results. However, in a study conducted on these patients preoperatively, it was found that the order in which the surveys were filled out had no significant influence on the results.³ In the current study, approximately half the patients filled out PROMIS questions first, and half filled them out last. Third, the population of patients was heterogeneous in terms of the type of shoulder stabilization procedure performed secondary to our methods of data extraction from our prospective database that includes patients based on diagnosis (shoulder instability) rather than procedure. The procedures investigated in the current study were heavily skewed to represent arthroscopic stabilization procedures. Although we were able to perform a small subgroup analysis on patients who had used the expanded PROMIS UE CAT v 2.0, further studies are warranted to assess this updated tool. Finally, this study did not address one of the most important outcomes in shoulder instability: recurrence. An additional study is warranted to assess the correlation of the

PROMIS and other PRO tools in evaluating outcomes in patients with and without recurrent shoulder instability.

CONCLUSION

The PROMIS PF CAT and PROMIS UE correlated well with previously validated tools in patients with shoulder instability. Both tools demonstrated the ability to detect meaningful change before and after operative interventions for shoulder instability. The PROMIS UE demonstrated ceiling effects at all time points, but no ceiling effect was found for the PROMIS PF CAT. The PROMIS PF CAT can be considered in place of the SF-36 PF. The PROMIS UE or PROMIS UE CAT v 2.0 may not have the ability to stratify higher levels of function, especially in patients aged ≤ 21 years, and should not be utilized at this time.

REFERENCES

1. Alberta FG, ElAttrache NS, Bissell S, et al. The development and validation of a functional assessment tool for the upper extremity in the overhead athlete. *Am J Sports Med.* 2010;38(5):903-911.
2. 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.
3. 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.
4. Bernholt D, Wright RW, Matava MJ, Brophy RH, Bogunovic L, Smith MV. Patient Reported Outcomes Measurement Information System scores are responsive to early changes in patient outcomes following arthroscopic partial meniscectomy. *Arthroscopy.* 2018;34(4):1113-1117.
5. Brophy RH, Beauvais RL, Jones EC, Cordasco FA, Marx RG. Measurement of shoulder activity level. *Clin Orthop Relat Res.* 2005;439:101-108.
6. 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.
7. Gibbs DB, Lynch TS, Nuber ED, Nuber GW. Common shoulder injuries in American football athletes. *Curr Sports Med Rep.* 2015;14(5):413-419.
8. Hovelius L, Augustini BG, Fredin H, Johansson O, Norlin R, Thorling J. Primary anterior dislocation of the shoulder in young patients: a ten-year prospective study. *J Bone Joint Surg Am.* 1996;78(11):1677-1684.
9. Hung M, Voss MW, Bounsanga J, Crum AB, Tyser AR. Examination of the PROMIS upper extremity item bank. *J Hand Ther.* 2017;30(4):485-490.
10. Jansson KA, Granath F. Health-related quality of life (EQ-5D) before and after orthopedic surgery. *Acta Orthop.* 2011;82(1):82-89.
11. Kirkley A, Griffin S, McLintock H, Ng L. The development and evaluation of a disease-specific quality of life measurement tool for shoulder instability: the Western Ontario Shoulder Instability Index (WOSI). *Am J Sports Med.* 1998;26(6):764-772.
12. Kocher MS, Horan MP, Briggs KK, Richardson TR, O'Holleran J, Hawkins RJ. Reliability, validity, and responsiveness of the American Shoulder and Elbow Surgeons subjective shoulder scale in patients with shoulder instability, rotator cuff disease, and glenohumeral arthritis. *J Bone Joint Surg Am.* 2005;87(9):2006-2011.
13. 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):247-263.
14. MOON Shoulder Group. MOON shoulder instability: patients undergoing operative treatment of the shoulder capsule or labrum. <http://www.moonshoulder.com/instabilitystudy.xhtml>. Accessed February 23, 2019.
15. Oak SR, Strnad GJ, Bena J, et al. Responsiveness comparison of the EQ-5D, PROMIS global health, and VR-12 questionnaires in knee arthroscopy. *Orthop J Sports Med.* 2016;4(12):2325967116674714.
16. 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(6S):S17-S23.
17. Scott EJ, Westermann R, Glass NA, Hettrich C, Wolf BR, Bollier MJ. Performance of the PROMIS in patients after anterior cruciate ligament reconstruction. *Orthop J Sports Med.* 2018;6(5):2325967118774509.
18. Simonet WT, Melton LJ 3rd, Cofield RH, Ilstrup DM. Incidence of anterior shoulder dislocation in Olmsted County, Minnesota. *Clin Orthop Relat Res.* 1984;186:186-191.
19. Zacchilli MA, Owens BD. Epidemiology of shoulder dislocations presenting to emergency departments in the United States. *J Bone Joint Surg Am.* 2010;92(3):542-549.