JB&JS OPEN ACCESS

Floor and Ceiling Effects, Time to Completion, and Question Burden of PROMIS CAT Domains Among Shoulder and Knee Patients Undergoing Nonoperative and Operative Treatment

Caleb M. Gulledge, BS, D. Grace Smith, BS, Alexander Ziedas, BS, Stephanie J. Muh, MD, Vasilios Moutzouros, MD, and Eric C. Makhni, MD, MBA

Investigation performed at the Department of Orthopaedic Surgery, Henry Ford Health System, Detroit, Michigan

Background: The Patient-Reported Outcomes Measurement Information System (PROMIS) computer adaptive tests (CATs) have emerged as an efficient technique for measuring patient-reported outcomes among orthopaedic patients. The purpose of this study was to investigate the floor and ceiling (F/C) effects, time to completion (TTC), and question burden of PROMIS CATs administered to patients presenting to a shoulder and sports medicine orthopaedic clinic.

Methods: Patients prospectively completed PROMIS CATs including the physical function (PROMIS-PF) or upper-extremity function (PROMIS-UE), pain interference (PROMIS-PI), and depression (PROMIS-D) domains at their initial encounter and were retrospectively included in this study. Adult patients indicating a single problem involving either the shoulder or knee were included. Patients were also grouped as either preoperative or nonoperative. F/C effects were defined as the proportion of respondents scoring the highest (ceiling) or lowest (floor) possible score across a given domain.

Results: Included were 2,952 patients (average age, 51.0 ± 16.9 years). The PROMIS-UE, PROMIS-PF, and PROMIS-PI demonstrated negligible F/C effects across all shoulder and knee patients (<2%). The PROMIS-D displayed moderate to significant floor effects (13.9% to 18.9%) and a 0% ceiling effect in all main patient groups. The mean TTC and mean question burden of the PROMIS-UE, PROMIS-PF, and PROMIS-PI ranged from 45.3 to 54.4 seconds and 4.1 to 4.9 questions for all patient groups, while the PROMIS-D exhibited a TTC ranging from 20.9 to 38.6 seconds for all groups and a question burden that ranged from 6.2 to 6.7 questions.

Conclusions: The PROMIS-PF, PROMIS-UE, and PROMIS-PI demonstrated favorable F/C effects, TTC, and question burden among both nonoperative and preoperative patients. These findings justify consideration of the PROMIS-PF, PROMIS-UE, and PROMIS-PI for clinical and research applications involving shoulder and knee sports medicine patients. Additionally, we found moderate to significant floor effects for the PROMIS-D in all patient groups, which may be multifactorial in nature and may not be unexpected in patients with an isolated joint concern.

Clinical Relevance: This study highlights the psychometric properties of PROMIS CAT forms for knee and shoulder patients. Understanding these basic properties is important in considering the adoption of PROMIS CAT forms for patients with musculoskeletal conditions.

The U.S. National Institutes of Health (NIH) Patient-Reported Outcomes Measurement Information System (PROMIS) has emerged as a dynamic, efficient technique for measuring patient-reported outcomes among orthopaedic patients across a number of health and functional domains. In particular, computer adaptive test (CAT) versions of these

Disclosure: The authors indicated that no external funding was received for any aspect of this work. On the **Disclosure of Potential Conflicts of Interest** forms, *which are provided with the online version of the article*, one or more of the authors checked "yes" to indicate that the author had a relevant financial relationship in the biomedical arena outside the submitted work (http://links.lww.com/JBJSOA/A127).

JBJS Open Access • 2019:e0015. openaccess.jbjs.org

questionnaires have demonstrated favorable psychometric properties, with improved efficiency (decreased question burden and shorter completion times), increased convergent validity (high correlation with previously established patient-reported outcome measures [PROMs]), and increased reliability (consistency of the questionnaire in ranking patients according to their responses) when compared with traditional "legacy" PROMs[1-10]. As 1 example, Hancock et al.[2] found that the PROMIS Physical Function CAT had high convergent validity (Pearson correlation of 0.82) with the Short-Form Health Survey (SF-36) Physical Function score among patients with anterior cruciate ligament (ACL) tears but with decreased question burden (4 versus 36, respectively) and improved floor (0% versus 1%, respectively) and ceiling (0% versus 5%, respectively) effects. However, these assessments are still relatively new, and they require rigorous validation in order to justify widespread adoption in clinical and research efforts.

One important component of this validation is the identification of floor and ceiling (F/C) effects of PROMIS CAT assessments, which indicate the ability of a questionnaire to distinguish between respondents at the extreme ends of the scale[11]. F/C effects are defined as the proportion of respondents scoring the highest (ceiling) or lowest (floor) possible score across any given domain, measuring the sensitivity and coverage of a questionnaire at each end of the scale[11]. For example, if a large proportion of patients receive the lowest possible score on a questionnaire, then that suggests that all of those patients have the same level of health, which in turn indicates the inability of that instrument to differentiate among those at the low end of the spectrum. High F/C effects also may suggest limited instrument range, measurement inaccuracy, and response bias[12], all of which indicate inadequate questionnaire performance. Significant F/C effects have historically been set at 15%[11]; however, others have stated that <10% or even 5% is an acceptable benchmark[8,13]. Recently published studies have demonstrated variable findings with regard to the F/C effects of PROMIS CAT domains[5,6,8,14-17]. However, these studies have typically focused on condition-based patient cohorts as opposed to large groups representing heterogenous patient populations.

The primary purpose of the current study, therefore, was to determine the F/C effects of PROMIS CAT domains (physical function, upper-extremity physical function, pain interference, and depression) among patients presenting to a shoulder and sports medicine ambulatory orthopaedic clinic. Our secondary purpose was to determine whether the F/C effects as well as the time to completion (TTC) and question burden (number of questions) per PROMIS assessment differed between patients indicated for surgery and those indicated for nonoperative treatment of the shoulder and knee, and among demographic groupings of age, sex, and race. We hypothesized that the PROMIS CATs would demonstrate generally favorable F/C effects and measures of efficiency in all patient groups and subgroups, thus justifying consideration for widespread use in orthopaedic clinical and research applications.

Materials and Methods

Study Design and Participants

This study was approved by our institutional review board. All patients presenting to the ambulatory clinic setting between May 1, 2017, and February 27, 2019, and treated by 1 of 3 fellowship-trained orthopaedic surgeons (2 shoulder and sports medicine surgeons and 1 shoulder and elbow surgeon) were prospectively enrolled in a PROMIS registry. Our inclusion criteria were as follows: completion of an intake form, completion of at least 1 PROMIS CAT domain (within 1 year before surgery for preoperative patients), indication of a single problem involving either the shoulder or knee, and the ability to communicate in written and spoken English. Patients <18 years of age (303 patients) were excluded. More than 99.5% of the included patients completed all PROMIS CAT domains.

TABLE I Patient Characteristics

	Shoulder		Knee		
	Preop.	Nonop.	Preop.	Nonop.	P Value†
No. of patients	520	1,088	404	940	
Age* *(yr)*	54.9 ± 15.9	56.0 ± 16.0	41.3 ± 15.7	47.1 ± 16.2	**<0.001**
Female sex *(%)*	42.9	50.5	40.8	49.1	**0.003**
Caucasian race *(%)*	67.5	61.0	64.4	60.5	0.070
T-score*					
PROMIS-UE	30.6 ± 7.2	33.4 ± 8.5	—	—	**<0.001**
PROMIS-PF	—	—	38.9 ± 7.6	41.6 ± 7.3	**<0.001**
PROMIS-PI	62.9 ± 5.9	60.8 ± 7.0	62.1 ± 6.8	60.4 ± 7.1	**<0.001**
PROMIS-D	49.0 ± 9.1	48.2 ± 9.8	48.5 ± 9.8	48.0 ± 9.4	0.337

*The values are given as the mean and standard deviation. †P values from comparison of patient characteristics among the 4 groups for the given row. Sex and race were compared using chi-square tests, and age was compared using 1-way ANOVA. Bold indicates a significant value.

Patient data, including age, sex, race, and anatomic location of the problem, were also collected at the office visit using a standardized intake form. PROMIS forms and patient data from each patient's initial encounter were then retrospectively obtained from the registry for this study. Patients were grouped by indication (or nonindication) for surgery such that patients with a Current Procedural Terminology (CPT) surgical code were identified and grouped as "preoperative" and those without a CPT surgical code were identified and grouped as "nonoperative." Patients were further classified according to the involved anatomic location (knee or shoulder), such that 4 groups were established: preoperative knee, preoperative shoulder, nonoperative knee, and nonoperative shoulder. Additionally, age groups were classified as follows: 18 to 39, 40 to 59, and ≥60 years.

Instruments

Each patient completed PROMIS forms for physical function, pain interference (the impact of pain on patient quality of life), and depression. Patients with a knee problem completed the PROMIS Physical Function-CAT version 2.0 (PROMIS-PF), while those with a shoulder problem completed the PROMIS Upper Extremity Physical Function-CAT version 2.0 (PROMIS-UE). Additionally, all patients completed the PROMIS Pain Interference-CAT version 1.1 (PROMIS-PI), as well as the PROMIS Depression-CAT version 1.0 (PROMIS-D). The PROMIS-PF pulls questions from a 165-item bank and covers both upper and lower-extremity function, with a score range of 14.7 to 75.6[18]. The PROMIS-UE pulls questions from a 46-item bank covering only upper-extremity function, with a score range of 14.7 to 61.0[18]. The PROMIS-PI draws from a 40-item bank, with a score range of 38.7 to 83.8, while the PROMIS-D draws from a 28-item bank of questions that measure depressive symptoms over the preceding 7 days, with a score range of 34.2 to 84.4[18].

All PROMIS instruments have a standardized scoring system, with a reference-population mean T-score of 50 and a standard deviation of 10 and with higher scores indicating more of the health domain in question. For example, a high score on the physical function forms denote a high functional ability, while a high score on the depression and pain forms indicate high levels of depressive symptoms and pain[7]. In the CAT format, question prompts are made in response to answers on the prior item, thus creating a dynamic and efficient scoring algorithm. All demographic and PROMIS information was collected on tablet computers using Research Electronic

TABLE II Psychometric Properties of PROMIS CAT Domains by Anatomic Location and Surgical Status

	Shoulder		Knee		
PROMIS CAT Domain and Variable	Preop.	Nonop.	Preop.	Nonop.	P Value*
Upper extremity					
Floor effects *(%)*	1.3	0.6	—	—	0.218
Ceiling effects *(%)*	0.4	1.4	—	—	0.123
Time to completion (TTC)† *(sec)*	50.0 ± 33.7	54.4 ± 43.7	—	—	0.083
No. of questions†‡	4.4 ± 1.4	4.9 ± 2.0	—	—	**<0.001**
Physical function					
Floor effects *(%)*	—	—	0.0	0.0	—
Ceiling effects *(%)*	—	—	0.0	0.2	—
Time to completion (TTC)† *(sec)*	—	—	45.3 ± 40.2	48.1 ± 41.4	0.340
No. of questions†‡	—	—	4.1 ± 0.7	4.1 ± 0.8	0.798
Pain interference					
Floor effects *(%)*	0.2	1.4	0.7	1.9	0.060
Ceiling effects *(%)*	0.2	0.1	0.0	0.0	0.625
Time to completion (TTC)† *(sec)*	50.0 ± 37.1	54.3 ± 43.1	45.4 ± 38.8	48.4 ± 41.4	**<0.001**
No. of questions†‡	4.1 ± 0.5	4.2 ± 1.2	4.2 ± 0.9	4.2 ± 1.2	**0.025**
Depression					
Floor effects *(%)*	13.9	**18.9**	**17.1**	**18.4**	0.145
Ceiling effects *(%)*	0.0	0.0	0.0	0.0	—
Time to completion (TTC)† *(sec)*	20.9 ± 40.7	27.0 ± 73.5	32.2 ± 94.2	38.6 ± 109.7	**0.003**
No. of questions†‡	6.2 ± 3.3	6.7 ± 3.5	6.4 ± 3.4	6.5 ± 3.5	0.123

*P values from comparison of psychometric properties among groups for the given row. For the upper-extremity and physical function domains, T-scores, TTC, and number of questions were compared using independent samples t tests, and floor and ceiling effects were compared using chi-square tests. For the pain interference and depression domains, T-scores, TTC, and number of questions were compared using 1-way ANOVA, while the floor and ceiling effects were compared with chi-square tests. Bold indicates a significant value. †The values are given as the mean and standard deviation. ‡Question burden.

TABLE III Psychometric Properties of PROMIS CAT Domains by Age, Sex, and Race

PROMIS CAT Domain and Variable	Age in Yr				Sex			Race		
	18-39	40-59	≥60	P Value*	Female	Male	P Value*	Caucasian	Not Caucasian	P Value*
Upper extremity										
Floor effects *(%)*	0.0	0.8	1.1	0.333	1.2	0.5	0.205	0.9	0.7	0.716
Ceiling effects *(%)*	3.1	1.0	0.5	**0.009**	0.5	1.7	0.055	1.4	0.7	0.279
Time to completion (TTC)† *(sec)*	41.2 ± 31.6	51.5 ± 40.6	59.2 ± 43.3	**<0.001**	53.9 ± 44.5	52.8 ± 37.6	0.659	52.2 ± 41.2	55.4 ± 40.7	0.205
No. of questions†‡	5.4 ± 2.6	4.7 ± 1.8	4.6 ± 1.7	**<0.001**	4.4 ± 1.4	5.1 ± 2.2	**<0.001**	4.8 ± 1.9	4.7 ± 1.9	0.337
Physical function										
Floor effects *(%)*	0.0	0.0	0.0	—	0.0	0.0	—	0.0	0.0	—
Ceiling effects *(%)*	0.6	0.0	0.0	0.139	0.2	0.3	0.716	0.1	0.4	0.392
Time to completion (TTC)† *(sec)*	46.0 ± 42.2	47.3 ± 44.6	49.1 ± 29.6	0.659	46.1 ± 38.0	47.6 ± 40.7	0.569	44.4 ± 31.7	51.6 ± 51.9	**0.014**
No. of questions†‡	4.3 ± 1.1	4.1 ± 0.5	4.1 ± 0.4	**<0.001**	4.1 ± 0.7	4.2 ± 0.8	0.370	4.2 ± 0.8	4.2 ± 0.8	0.948
Pain interference										
Floor effects *(%)*	2.4	0.9	0.9	**0.015**	0.9	1.5	0.205	1.6	0.7	0.083
Ceiling effects *(%)*	0.0	0.1	0.1	0.765	0.1	0.0	0.205	0.1	0.1	0.769
Time to completion (TTC)† *(sec)*	44.9 ± 43.4	47.4 ± 33.6	58.0 ± 45.8	**<0.001**	49.6 ± 41.4	51.2 ± 40.8	0.391	49.5 ± 39.9	52.1 ± 42.9	0.163
No. of questions†‡	4.3 ± 1.4	4.2 ± 0.9	4.2 ± 0.9	**0.009**	4.2 ± 0.9	4.2 ± 1.2	0.229	4.2 ± 1.1	4.2 ± 1.0	0.621
Depression										
Floor effects *(%)*	**21.0**	**17.4**	**15.4**	**0.020**	**15.6**	**19.5**	**0.015**	**17.6**	**17.6**	0.989
Ceiling effects *(%)*	0.0	0.0	0.0	—	0.0	0.0	—	0.0	0.0	—
Time to completion (TTC)† *(sec)*	39.2 ± 110.5	28.1 ± 76.4	26.3 ± 73.8	**0.014**	26.5 ± 64.9	33.9 ± 101.0	**0.044**	30.5 ± 85.0	30.1 ± 87.0	0.931
No. of questions†‡	6.8 ± 3.6	6.4 ± 3.4	6.4 ± 3.4	0.085	6.3 ± 3.3	6.7 ± 3.5	**0.003**	6.5 ± 3.4	6.5 ± 3.5	0.864

*P values from comparison of psychometric properties among groups for the given row. When comparing age groups, the T-scores, TTC, and number of questions were compared using 1-way ANOVA, and floor and ceiling effects were compared using chi-square tests. When comparing sex and race groups, the T-scores, TTC, and number of questions were compared with independent samples t tests, and floor and ceiling effects were compared with chi-square tests. Bold indicates a significant value. †The values are given as the mean and standard deviation. ‡Question burden.

Data Capture (REDCap), a HIPAA (Health Insurance Portability and Accountability Act)-compliant data collection application maintained by Vanderbilt University[19].

Statistical Analyses

Patient demographics and psychometric variables are presented with descriptive statistics. Sex and race are presented as the percentage of females and the percentage of Caucasians, respectively. F/C effects were defined as the proportion of respondents scoring the highest (ceiling) or lowest (floor) possible score across any given domain. F/C effects were classified as significant if ≥15%, moderate if 10% to <15%, minor if 5% to <10%, and negligible if <5%. The evaluation of continuous variables (mean age, T-scores, TTC, and number of questions) was performed using independent samples t tests for 2-group comparisons and analysis of variance (ANOVA) for comparisons of ≥3 groups. Comparisons between categorical variables (F/C effects, sex, and race) were analyzed using chi-square tests. Because of the numerous statistical tests in this study, the Benjamini-Hochberg method of adjusting p values to control for the false discovery rate (set at 0.05) was used for all tests. Significance was set as alpha = 0.05. SPSS Statistics for Windows (version 25.0; IBM) was used for all data analysis.

Results

Included were data from 2,952 patients (average age [and standard deviation] of 51.0 ± 16.9 years); 924 were indicated for surgery and 2,028 received nonoperative treatment. Generally, shoulder patients were older than knee patients, and preoperative knee patients were younger than nonoperative knee patients. Additionally, among both shoulder and knee patients, the nonoperative patients were more likely to be female. Also, preoperative patients had more pain interference and less function than nonoperative patients, both among shoulder and knee patients. Demographic data and mean T-scores are shown in Table I.

Table II shows a comparison of the F/C effects, TTC, and number of questions answered for each of the 4 PROMIS CAT domains, by anatomic location and surgical status. With the exception of moderate to significant floor effects for the depression domain, the F/C effects were negligible for all domains and did not differ by anatomic location or surgical status.

The psychometric properties for each PROMIS CAT domain were also compared by age group, sex, and race (Table III). Generally, younger patients completed the PROMIS-UE and PROMIS-PI in less time than older patients, but they took longer to complete the PROMIS-D. Finally, while

the floor or ceiling effect differed by age group for the PROMIS-UE, PROMIS-PI, and PROMIS-D, and by sex for the PROMIS-D, all F/C effects were negligible, with the exception of significant floor effects for the PROMIS-D, in all demographic subgroups.

Discussion

Our results indicated that the PROMIS-PF, PROMIS-UE, and PROMIS-PI demonstrated favorable F/C effects for both nonoperative and preoperative shoulder and knee patients. Conversely, moderate to significant floor effects were found for the PROMIS-D for all 4 main patient groups. Additionally, the TTC and question burden were favorable for all PROMIS domains and all patient groups.

In our study, we found that the PROMIS-PF demonstrated negligible F/C effects among nonoperative and preoperative knee patients (<1%). Additionally, both patient groups and all subgroups were able to complete the PROMIS-PF in <1 minute, answering an average of only 4 questions. Our results agree with those of other studies that found <1% F/C effects when used for several other lower-extremity patient populations, including those presenting for knee conditions[20], ACL reconstruction[2,21], hip conditions[22], meniscal repair or meniscectomy[17], and foot and ankle conditions[23-27]. In comparison, the International Knee Documentation Committee (IKDC) Subjective Knee Form, a validated and frequently utilized PROM for knee patients[28,29], has shown high correlation with the PROMIS-PF, with comparably low floor (0% to 1%) and ceiling (0% to 6%) effects[20,30]. However, the PROMIS-PF is much quicker to administer[20], with reduced test burden[4]. Furthermore, the PROMIS-PF has been successfully utilized and validated for many other anatomic locations, including the hand[12], shoulder[5,31], spine[14,32], neck[33,34], and back[9], making it efficient and practical for routine PROMs administration in the ambulatory setting. These favorable psychometric and administrative properties justify consideration of the PROMIS-PF for widespread adoption as a physical function measure for orthopaedic patients.

With regard to upper-extremity function, we found that the PROMIS-UE demonstrated negligible F/C effects among both preoperative and nonoperative shoulder patients (<2%). As with the PROMIS-PF, both patient groups and all subgroups were able to complete the PROMIS-UE in <1 minute, answering an average of only 4 to 5 questions. Several prior studies also demonstrated negligible F/C effects for shoulder patients[7,31,35,36]. However, 2 recent studies reported minor ceiling effects when used in non-shoulder upper-extremity and hand clinics (6.9%[10] and 7.6%[12]). Prior studies have consistently reported a ceiling effect (7.2% to 28.2%) for earlier versions of the PROMIS-UE[15,16,37-39]. Our findings demonstrated that the newer PROMIS-UE (version 2.0), which was recently modified for greater sensitivity toward the upper end of the scale, has improved ceiling effects for shoulder patients. Yet, the 2 aforementioned studies[10,12] suggest that minor ceiling effects may still be observed when used for non-shoulder upper-extremity patients. However, these minor ceiling effects are within reasonable bounds (well below 15%) and should not preclude consideration of the PROMIS-UE version 2.0 for clinical and research purposes.

When considering measures of pain, there were negligible F/C effects for the PROMIS-PI in all patient groups. Several studies have also noted negligible ceiling effects for upper-extremity[16,40], lower-extremity[22,23], spine[14,41,42], neck[33], and trauma patients[43]. In fact, only 1 study noted a ceiling effect of >3% (4.7%)[5]. However, there is variation among reported PROMIS-PI floor effects that appears to depend on patient population. For example, 2 longitudinal studies found negligible floor effects prior to surgery but moderate to significant floor effects postoperatively[23,32]. Additionally, no floor effects were exhibited in 1 study of trauma patients[43], while a moderate floor effect was exhibited in another study of hand and upper-extremity patients[40]. Our findings suggest that the PROMIS-PI has reliable coverage for preoperative and nonoperative shoulder and knee patients. Because of the variability within the literature based on patient population, further research is needed to better determine the floor effects of the PROMIS-PI for specific populations.

With regard to depression, the PROMIS-D exhibited moderate to significant floor effects (13.9% to 18.9%) and a 0% ceiling effect for all main patient groups. Significant floor effects have also been reported when used in an academic orthopaedic center[44] and a tertiary hip center[22] as well as when used for patients with glenohumeral osteoarthritis[31]. Guattery et al.[44] suggested that "hasty" completion of the PROMIS-D may contribute to the floor effect. Our data support this finding, with the PROMIS-D requiring more questions but less TTC than the pain and function domains for all patient groups. Additionally, given the high prevalence of depressive symptoms in the United States[45], it is likely that patients are not forthcoming about experiencing depressive symptoms and may be prone to minimizing these symptoms. Moreover, depression is rarely the primary concern for a patient presenting with a shoulder or knee problem, and it could be expected that this domain would exhibit different test characteristics than other health domains for these populations. Furthermore, PROMIS depression measures have demonstrated high correlations with previously established mental health scores, including the Scoliosis Research Society (SRS)-30 mental health score, the Patient Health Questionnaire (PHQ) 9-item scale, the PHQ 8-item scale, the SF-12 mental component summary, and the SF-36 mental health scale[42,46-48], among orthopaedic spine patients and veterans after orthopaedic surgery. However, to our knowledge, the PROMIS-D has yet to be compared with other depression measures for shoulder and knee orthopaedic patients. The PROMIS-D is important for orthopaedic providers because it can help provide a holistic view of a patient and may also inform clinical decisions. In fact, recent evidence indicates that preoperative PROMIS-D scores indicating depression are associated with worse function and pain outcomes after lumbar spine decompression[49]. However, the clinical utility of the PROMIS-D may be limited in patients with isolated shoulder and knee problems until additional studies validate the PROMIS-D in these populations and determine how to minimize the floor effect.

This study had pertinent limitations. Only patients reporting isolated shoulder and knee problems presenting to a shoulder and sports medicine clinic were included in this study; therefore, we cannot draw conclusions regarding the F/C effects for other patient populations. Additionally, this study included PROMIS forms administered in English, thus limiting generalizability to patients who do not speak or communicate in English. Also, nonoperative patients were not followed to other health systems, and thus, it is possible that patients included in our nonoperative groups elected to undergo surgery in a different health system. Lastly, many diagnoses may fall within a single anatomic location, and thus, the F/C effects cannot be directly related to a specific diagnosis.

Conclusions

The PROMIS-PF, PROMIS-UE, and PROMIS-PI demonstrated favorable F/C effects, TTC, and question burden among both nonoperative and preoperative patients. These findings justify consideration of the PROMIS-PF, PROMIS-UE, and PROMIS-PI for clinical and research applications involving shoulder and knee sports medicine patients. Additionally, we found moderate to significant floor effects for the PROMIS-D in all patient groups, which may be multifactorial in nature and may not be unexpected for patients with an isolated joint concern. ■

NOTE: The authors acknowledge Vincent A. Lizzio, MD, and Nabil Mehta, MD, for contributing to the initial design of this study.

Caleb M. Gulledge, BS[1]
D. Grace Smith, BS[1]
Alexander Ziedas, BS[1]
Stephanie J. Muh, MD[1]
Vasilios Moutzouros, MD[1]
Eric C. Makhni, MD, MBA[1]

[1]Department of Orthopaedic Surgery, Henry Ford Health System, Detroit, Michigan

Email address for C.M. Gulledge: cgulledg@med.wayne.edu
Email address for E.C. Makhni: ericmakhnimd@gmail.com

ORCID iD for C.M. Gulledge: 0000-0002-6883-9107
ORCID iD for D.G. Smith: 0000-0002-0732-1953
ORCID iD for A. Ziedas: 0000-0002-6009-3195
ORCID iD for S.J. Muh: 0000-0001-6617-4116
ORCID iD for V. Moutzouros: 0000-0002-9353-6239
ORCID iD for E.C. Makhni: 0000-0003-4408-1175

References

1. Kadri O, Jildeh TR, Meldau JE, Blanchett J, Borowsky P, Muh S, Moutzouros V, Makhni EC. How long does it take for patients to complete PROMIS scores?: an assessment of PROMIS CAT questionnaires administered at an ambulatory sports medicine clinic. Orthop J Sports Med. 2018 Aug 14;6(8):2325967118791180.

2. Hancock KJ, Glass N, Anthony CA, Wolf BR, Hettrich CM, Albright J, Bollier M, Amendola A. PROMIS: a valid and efficient outcomes instrument for patients with ACL tears. Knee Surg Sports Traumatol Arthrosc. 2019 Jan;27(1):100-4. Epub 2018 Jul 4.

3. 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 May;67(5):516-26.

4. 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 Aug;28(8):439-43.

5. Strong B, Maloney M, Baumhauer J, Schaffer J, Houck JR, Hung M, Bounsanga J, Voss MW, Gu Y, Voloshin I. Psychometric evaluation of the Patient-Reported Outcomes Measurement Information System (PROMIS) physical function and pain interference computer adaptive test for subacromial impingement syndrome. J Shoulder Elbow Surg. 2019 Feb;28(2):324-9. Epub 2018 Oct 18.

6. 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 Dec;24(12):1961-7. Epub 2015 Aug 28.

7. Patterson BM, Orvets ND, Aleem AW, Keener JD, Calfee RP, Nixon DC, Chamberlain AM. 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 Jun;27(6S):S17-23.

8. Hung M, Voss MW, Bounsanga J, Gu Y, Granger EK, Tashjian RZ. Psychometrics of the Patient-Reported Outcomes Measurement Information System physical function instrument administered by computerized adaptive testing and the Disabilities of Arm, Shoulder and Hand in the orthopedic elbow patient population. J Shoulder Elbow Surg. 2018 Mar;27(3):515-22. Epub 2018 Jan 9.

9. Brodke DS, Goz V, Voss MW, Lawrence BD, Spiker WR, Hung M. PROMIS PF CAT outperforms the ODI and SF-36 physical function domain in spine patients. Spine (Phila Pa 1976). 2017 Jun 15;42(12):921-9.

10. Tyser AR, Hung M, Bounsanga J, Voss MW, Kazmers NH. Evaluation of version 2.0 of the PROMIS upper extremity computer adaptive test in nonshoulder upper extremity patients. J Hand Surg Am. 2019 Apr;44(4):267-73. Epub 2019 Feb 26.

11. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, Bouter LM, de Vet HC. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007 Jan;60(1):34-42. Epub 2006 Aug 24.

12. Bernstein DN, Houck JR, Hammert WC. A comparison of PROMIS UE versus PF: correlation to PROMIS PI and depression, ceiling and floor effects, and time to completion. J Hand Surg Am. 2019 Feb 4:S0363-5023(18)30751-2. Epub 2019 Feb 4.

13. Hung M, Hon SD, Franklin JD, Kendall RW, Lawrence BD, Neese A, Cheng C, Brodke DS. Psychometric properties of the PROMIS physical function item bank in patients with spinal disorders. Spine (Phila Pa 1976). 2014 Jan 15;39(2):158-63.

14. Boody BS, Bhatt S, Mazmudar AS, Hsu WK, Rothrock NE, Patel AA. Validation of Patient-Reported Outcomes Measurement Information System (PROMIS) computerized adaptive tests in cervical spine surgery. J Neurosurg Spine. 2018 Mar;28(3):268-79. Epub 2018 Jan 5.

15. Döring AC, Nota SP, Hageman MG, Ring DC. Measurement of upper extremity disability using the Patient-Reported Outcomes Measurement Information System. J Hand Surg Am. 2014 Jun;39(6):1160-5. Epub 2014 May 3.

16. Gausden EB, Levack AE, Sin DN, Nwachukwu BU, Fabricant PD, Nellestein AM, Wellman DS, Lorich DG. Validating the Patient Reported Outcomes Measurement Information System (PROMIS) computerized adaptive tests for upper extremity fracture care. J Shoulder Elbow Surg. 2018 Jul;27(7):1191-7. Epub 2018 Mar 19.

17. Hancock KJ, Glass N, Anthony CA, Hettrich CM, Albright J, Amendola A, Wolf BR, Bollier M. Performance of PROMIS for healthy patients undergoing meniscal surgery. J Bone Joint Surg Am. 2017 Jun 7;99(11):954-8.

18. HealthMeasures. PROMIS. 2018. Accessed 2019 Sep 11. http://www.healthmeasures.net/explore-measurement-systems/promis

19. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research Electronic Data Capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform. 2009 Apr;42(2):377-81. Epub 2008 Sep 30.

20. Robins RJ, Anderson MB, Zhang Y, Presson AP, Burks RT, Greis PE. Convergent validity of the Patient-Reported Outcomes Measurement Information System's physical function computerized adaptive test for the knee and shoulder injury sports medicine patient population. Arthroscopy. 2017 Mar;33(3):608-16. Epub 2016 Dec 10.

21. 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 May 25;6(5):2325967118774509.

22. Kollmorgen RC, Hutyra CA, Green C, Lewis B, Olson SA, Mather RC 3rd. Relationship between PROMIS computer adaptive tests and legacy hip measures among patients presenting to a tertiary care hip preservation center. Am J Sports Med. 2019 Mar;47(4):876-84. Epub 2019 Feb 12.

23. Bernstein DN, Kelly M, Houck JR, Ketz JP, Flemister AS, DiGiovanni BF, Baumhauer JF, Oh I. PROMIS pain interference is superior vs numeric pain rating scale for pain assessment in foot and ankle patients. Foot Ankle Int. 2019 Feb; 40(2):139-44. Epub 2018 Oct 4.

24. Gausden EB, Levack A, Nwachukwu BU, Sin D, Wellman DS, Lorich DG. Computerized adaptive testing for patient reported outcomes in ankle fracture surgery. Foot Ankle Int. 2018 Oct;39(10):1192-8. Epub 2018 Jul 4.

25. Koltsov JCB, Greenfield ST, Soukup D, Do HT, Ellis SJ. Validation of Patient-Reported Outcomes Measurement Information System computerized adaptive tests against the Foot and Ankle Outcome Score for 6 common foot and ankle pathologies. Foot Ankle Int. 2017 Aug;38(8):870-8. Epub 2017 Jun 10.

26. Hung M, Baumhauer JF, Latt LD, Saltzman CL, SooHoo NF, Hunt KJ; National Orthopaedic Foot & Ankle Outcomes Research Network. Validation of PROMIS ® physical function computerized adaptive tests for orthopaedic foot and ankle outcome research. Clin Orthop Relat Res. 2013 Nov;471(11):3466-74.

27. Hung M, Franklin JD, Hon SD, Cheng C, Conrad J, Saltzman CL. Time for a paradigm shift with computerized adaptive testing of general physical function outcomes measurements. Foot Ankle Int. 2014 Jan;35(1):1-7. Epub 2013 Oct 7.

28. Makhni EC, Padaki AS, Petridis PD, Steinhaus ME, Ahmad CS, Cole BJ, Bach BR Jr. High variability in outcome reporting patterns in high-impact ACL literature. J Bone Joint Surg Am. 2015 Sep 16;97(18):1529-42.

29. Makhni EC, Meyer MA, Saltzman BM, Cole BJ. Comprehensiveness of outcome reporting in studies of articular cartilage defects of the knee. Arthroscopy. 2016 Oct; 32(10):2133-9. Epub 2016 Jun 17.

30. Jacobs CA, Peabody MR, Lattermann C, Vega JF, Huston LJ, Spindler KP, Amendola A, Andrish JT, Brophy RH, Dunn WR, Flanigan DC, Jones MH, Kaeding CC, Marx RG, Matava MJ, McCarty EC, Parker RD, Reinke EK, Wolcott ML, Wolf BR, Wright RW, Vidal AF; MOON Knee Group. Development of the $KOOS_{global}$ platform to measure patient-reported outcomes after anterior cruciate ligament reconstruction. Am J Sports Med. 2018 Oct;46(12):2915-21. Epub 2018 Aug 3.

31. Kohan EM, Hill JR, Schwabe M, Aleem AW, Keener JD, Chamberlain AM. The influence of mental health on Patient-Reported Outcomes Measurement Information System (PROMIS) and traditional outcome instruments in patients with symptomatic glenohumeral arthritis. J Shoulder Elbow Surg. 2019 Feb;28(2):e40-8. Epub 2018 Dec 12.

32. Patel AA, Dodwad SM, Boody BS, Bhatt S, Savage JW, Hsu WK, Rothrock NE. Validation of Patient Reported Outcomes Measurement Information System (PROMIS) computer adaptive tests (CATs) in the surgical treatment of lumbar spinal stenosis. Spine (Phila Pa 1976). 2018 Nov 1;43(21):1521-8.

33. Moses MJ, Tishelman JC, Stekas N, Jevotovsky DS, Vasquez-Montes D, Karia R, Errico T, Buckland A, Protopsaltis TS. Comparison of Patient Reported Outcome Measurement Information System with Neck Disability Index and Visual Analog Scale in patients with neck pain. Spine (Phila Pa 1976). 2019 Feb 1;44(3):E162-7.

34. Tishelman JC, Vasquez-Montes D, Jevotovsky DS, Stekas N, Moses MJ, Karia RJ, Errico T, Buckland AJ, Protopsaltis TS. Patient-Reported Outcomes Measurement Information System instruments: outperforming traditional quality of life measures in patients with back and neck pain. J Neurosurg Spine. 2019 Feb 1;1:1-6. Epub 2019 Feb 1.

35. Minoughan CE, Schumaier AP, Fritch JL, Grawe BM. Correlation of PROMIS physical function upper extremity computer adaptive test with American Shoulder and Elbow Surgeons shoulder assessment form and Simple Shoulder Test in patients with shoulder arthritis. J Shoulder Elbow Surg. 2018 Apr;27(4):585-91. Epub 2017 Dec 20.

36. 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 May;34(5):1430-6. Epub 2018 Feb 13.

37. Beckmann JT, Hung M, Voss MW, Crum AB, Bounsanga J, Tyser AR. Evaluation of the Patient-Reported Outcomes Measurement Information System upper extremity computer adaptive test. J Hand Surg Am. 2016 Jul;41(7):739-744.e4. Epub 2016 Jun 3.

38. Beleckas CM, Padovano A, Guattery J, Chamberlain AM, Keener JD, Calfee RP. Performance of Patient-Reported Outcomes Measurement Information System (PROMIS) upper extremity (UE) versus physical function (PF) computer adaptive tests (CATs) in upper extremity clinics. J Hand Surg Am. 2017 Nov;42(11):867-74. Epub 2017 Jul 12.

39. Kaat AJ, Rothrock NE, Vrahas MS, O'Toole RV, Buono SK, Zerhusen T Jr, Gershon RC. Longitudinal validation of the PROMIS physical function item bank in upper extremity trauma. J Orthop Trauma. 2017 Oct;31(10):e321-6.

40. Kortlever JT, Janssen SJ, van Berckel MM, Ring D, Vranceanu AM. What is the most useful questionnaire for measurement of coping strategies in response to nociception? Clin Orthop Relat Res. 2015 Nov;473(11):3511-8.

41. Bernstein DN, Bakhsh W, Papuga MO, Menga EN, Rubery PT, Mesfin A. An evaluation of PROMIS in patients with primary or metastatic spine tumors. Spine (Phila Pa 1976). 2019 May 15;44(10):747-52.

42. Bernstein DN, Papuga MO, Sanders JO, Rubery PT, Menga EN, Mesfin A. Evaluating the correlation and performance of PROMIS to SRS questionnaires in adult and pediatric spinal deformity patients. Spine Deform. 2019 Jan;7(1):118-24.

43. Alvarez-Nebreda ML, Heng M, Rosner B, McTague M, Javedan H, Harris MB, Weaver MJ. Reliability of proxy-reported Patient-reported Outcomes Measurement Information System physical function and pain interference responses for elderly patients with musculoskeletal injury. J Am Acad Orthop Surg. 2019 Feb 15;27(4): e156-65.

44. Guattery JM, Dardas AZ, Kelly M, Chamberlain A, McAndrew C, Calfee RP. Floor effect of PROMIS depression CAT associated with hasty completion in orthopaedic surgery patients. Clin Orthop Relat Res. 2018 Apr;476(4):696-703.

45. Shim RS, Baltrus P, Ye J, Rust G. Prevalence, treatment, and control of depressive symptoms in the United States: results from the National Health and Nutrition Examination Survey (NHANES), 2005-2008. J Am Board Fam Med. 2011 Jan-Feb;24(1):33-8.

46. Hadlandsmyth K, Dindo LN, St Marie BJ, Wajid R, Embree JL, Noiseux NO, Tripp-Reimer T, Zimmerman MB, Rakel BA. Patient-Reported Outcomes Measurement Information System (PROMIS) instruments: reliability and validity in veterans following orthopedic surgery. Eval Health Prof. 2019 Jun 13:163278719856406. Epub 2019 Jun 13.

47. Kroenke K, Baye F, Lourens SG. Comparative validity and responsiveness of PHQ-ADS and other composite anxiety-depression measures. J Affect Disord. 2019 Mar 1;246:437-43. Epub 2018 Dec 25.

48. Purvis TE, Neuman BJ, Riley LH, Skolasky RL. Comparison of PROMIS Anxiety and Depression, PHQ-8, and GAD-7 to screen for anxiety and depression among patients presenting for spine surgery. J Neurosurg Spine. 2019 Jan 18:1-8. Epub 2019 Jan 18.

49. Merrill RK, Zebala LP, Peters C, Qureshi SA, McAnany SJ. Impact of depression on patient-reported outcome measures after lumbar spine decompression. Spine (Phila Pa 1976). 2018 Mar 15;43(6):434-9.