4-2019

Do Patient Reported Outcome Measurement Information System (PROMIS) Scales Demonstrate Responsiveness as Well as Disease-Specific Scales in Patients Undergoing Knee Arthroscopy?

Raymond J. Kenney

Jeff Houck

Brian D. Giordano

Judith F. Baumhauer

Meghan Herbert

See next page for additional authors

Authors

Raymond J. Kenney, Jeff Houck, Brian D. Giordano, Judith F. Baumhauer, Meghan Herbert, and Michael D. Maloney

Do Patient Reported Outcome Measurement Information System (PROMIS) Scales Demonstrate Responsiveness as Well as Disease-Specific Scales in Patients Undergoing Knee Arthroscopy?

Raymond J. Kenney,*† MD, Jeff Houck,‡ PT, PhD, Brian D. Giordano,† MD, Judith F. Baumhauer,† MD, MPH, Meghan Herbert,† PA, and Michael D. Maloney,† MD
Investigation performed at University of Rochester, Rochester, New York, USA

Background: The Patient Reported Outcomes Information System (PROMIS) is an efficient metric able to detect changes in global health.

Purpose: To assess the responsiveness, convergent validity, and clinically important difference (CID) of PROMIS compared with disease-specific scales after knee arthroscopy.

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

Methods: A prospective institutional review board–approved study collected PROMIS Physical Function (PF), PROMIS Pain Interference (PI), International Knee Documentation Committee (IKDC), and Knee injury and Osteoarthritis Outcome Score (KOOS) results in patients undergoing knee arthroscopy. The change from preoperative to longest follow-up was used in analyses performed to determine responsiveness, convergent validity, and minimal and moderate CID using the IKDC scale as the anchor.

Results: Of the 100 patients enrolled, 76 were included. Values of the effect size index (ESI) ranged from near 0 to 1.69 across time points and were comparable across scales. Correlations of the change in KOOS and PROMIS with IKDC ranged from *r* values of 0.61 to 0.79. The minimal CID for KOOS varied from 12.5 to 17.5. PROMIS PF and PI minimal CID were 3.3 and −3.2. KOOS moderate CID varied from 14.3 to 18.8. PROMIS PF and PI moderate CID were 5.0 and −5.8.

Conclusion: The PROMIS PF and PI showed similar responsiveness and CID compared with disease-specific scales in patients after knee arthroscopy. PROMIS PI, PROMIS PF, and KOOS correlations with IKDC demonstrate that these scales are measuring a similar construct. The ESIs of PROMIS PF and PI were similar to those of KOOS and IKDC, suggesting similar responsiveness at 6 months or longer (ESI >1.0). Minimum and moderate CID values calculated for PROMIS PF and PI using IKDC as an anchor were sufficiently low to suggest clinical usefulness.

Clinical Relevance: PROMIS PF and PI can be accurately used to determine improvement or lack thereof with clinically important changes after knee arthroscopy.

Keywords: knee; general; epidemiology; statistics; medical aspects of sports; PROMIS

Knee arthroscopy is the most commonly performed ambulatory surgical procedure by orthopaedic surgeons, reaching nearly 1 million procedures annually in the US in 2006.[21,25] Knee arthroscopy techniques have evolved over the years and are used to treat a variety of knee conditions including meniscal tear, chondral defects, ligamentous injury, loose body, and synovial hyperplasia.[25] Current editorials call for orthopaedic providers to assess varied orthopaedic surgery procedures using patient-reported outcomes (PROs).[5,29] Regulations such as the Patient Protection and Affordable Care Act in 2010 expanded federal oversight of tracking and reporting quality measures including PROs.[7] Likewise, current Medicare policy has linked up to 9% of hospital payment to performance.[7] Given these current research, regulatory, and reimbursement incentives, PROs are now an important component for tracking the success of orthopaedic procedures, including arthroscopic knee surgery.

Generic global health PRO scales have distinct advantages over disease-specific scales but are not fully validated. A multitude of disease-specific PROs are available,

including the International Knee Documentation Committee (IKDC), Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC), Western Ontario Meniscal Evaluation Tool, and Knee injury and Osteoarthritis Outcome Score (KOOS).[2,11,12,18,19,30,31] Disease-specific scales are designed to measure pain and functional deficits for a particular joint (eg, IKDC for the knee) or type of problem (eg, WOMAC for osteoarthritis). In contrast, global health PRO scales such as the Patient Reported Outcomes Information System (PROMIS) are not limited to a particular joint or type of disease.[8,20] A concern is that because the PROMIS scales are not disease specific, the scales may not demonstrate responsiveness or detect change well. Another concern is that other, non–knee related problems that also influence pain and function might mask benefits that patients experience after procedures like knee arthroscopy.

Documenting responsiveness and detectable change may significantly affect the clinical application and address reasonable concerns regarding global health scales. Responsiveness is defined as the ability of a scale to detect change over time. Responsiveness is assessed across scales using the effect size index (ESI), where the change in a particular PRO scale is divided by the SD of the change.[17] The ESI essentially normalizes the PRO change scores for comparison across scales.[17] When PRO scales are applied in patient care, it is also beneficial to know what level of change is needed such that the patient will note a perceived benefit.[24] Minimal and moderate clinically important differences (CIDs) inform a clinician of what change score is associated with a minimal and moderate level of perceived improvement. The CID for minimal improvement on the IKDC is 11.5 and for moderate improvement is 20.5.[19] The IKDC is commonly used and well validated.[9,11,22,23] However, minimal and moderate CID for the KOOS is not well established in patients after knee arthroscopy. The PROMIS scales are new, with no studies of responsiveness and minimal/moderate CID addressing knee arthroscopy patients directly.[10,13] Also, to date most studies of PROMIS CID used distribution method approaches rather than preferred anchor method approaches.[4,13]

The appeal of generic health measures such as PROMIS is great, as these scales capture overall health rather than disease-specific effects, are quick to administer, and allow comparison across disease conditions with limited patient burden, lending themselves to the addition of other health domains that affect health care, such as mental or social health. The ability to measure generic health domains informs clinicians of how procedures like knee arthroscopy are influencing overall health rather than only focal disease-specific effects.[16] Further, because PROMIS scales use computer adaptive test (CAT) algorithms, they take less time than most disease-specific scales.[14,15,27] For example, previously published data have determined that the CAT for the PROMIS Physical Function (PF) scale is administered in 6 to 12 questions per patient and the IKDC consists of 18 static questions with an average completion time of 85 seconds and 195 seconds, respectively.[27]

The purpose of this study was to determine whether the PROMIS scales are responsive and able to detect meaningful clinical change similarly to disease-specific scales (IKDC and KOOS) across the continuum of care from 2 weeks to 12 months in patients who underwent knee arthroscopy. The first hypothesis was that the selected PROMIS scales would show similar responsiveness as measured by the ESI compared with the KOOS and IKDC scales from preoperative to longest follow-up. The second hypothesis was that the change in scores (preoperative to longest follow-up) for the PROMIS scales and KOOS subscales would show similar convergence (correlation) with the change in IKDC scales. The third hypothesis was that the minimal and moderate CID for the PROMIS and KOOS scales would show similar accuracy in determining the minimal and moderate CID on the IKDC scale.

METHODS

Patients were prospectively enrolled into an institutional review board–approved study (the University of Rochester approved the study protocol) to determine the responsiveness of PROMIS, KOOS, and IKDC scales for patients undergoing knee arthroscopy for medial meniscal tear, lateral meniscal tear, chondromalacia, loose body, and/or synovial hyperplasia. Included patients were over 18 years old and underwent primary knee arthroscopy with partial medial meniscectomy, partial lateral meniscectomy, chondroplasty, loose body removal, and/or synovectomy. Patients were excluded who had revision surgery or ligamentous injury or were non–English speaking. We enrolled 100 patients, accounting for a possible dropout of 30 patients, based on a Pearson correlation power analysis showing adequate power of 70 with a low-moderate correlation r value of 0.33, $\alpha = .05$, and $\beta = .20$. PROMIS

*Address correspondence to Raymond J. Kenney, MD, University of Rochester Medical Center, Department of Orthopaedics and Rehabilitation, 601 Elmwood Ave, Box 665, Rochester, NY 14642, USA (email: raymond_kenney@urmc.rochester.edu).

†Department of Orthopaedics and Rehabilitation, University of Rochester, Rochester, New York, USA.

‡Department of Physical Therapy, George Fox University, Newberg, Oregon, USA.

One or more of the authors has declared the following potential conflict of interest or source of funding: B.D.G. has received research support, consulting fees, speaking fees, and royalties from Arthrex; consulting fees from Carticept Medical; compensation for services other than consulting from Smith & Nephew; and hospitality payments from Exactech Inc. J.F.B. has received research support from Cartiva, Ferring Pharmaceuticals, and Wright Medical Technology; has received consulting fees from Cartiva, DJ Orthopaedics, Ferring Pharmaceuticals, Nextremity Solutions, Stryker, Trimed Inc, EMPI, Medtronic Sofamor Danek, Biomimetic Therapeutics, and Zimmer; has received hospitality payments from Encore Medical; and is a board or committee member for the American Orthopaedic Foot & Ankle Society and the PROMIS Health Organization. M.D.M. has received consulting and speaking fees from Arthrex, has stock or stock options in Telephus, and has received compensation for services other than consulting from Prodigy Surgical Distribution. 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.

TABLE 1
Demographics and Follow-up[a]

Variable	Finding
Age, y	48.9 ± 11.1 (20-74)
Male sex, n (%)	40 (52.6)
Height, cm	171.5 ± 9.6 (152.0-196.0)
Weight, kg	89.7 ± 20.6 (49.4-143.8)
Body mass index, kg/m^2	30.5 ± 6.6 (19.3-48.0)
Injured knee (right), n (%)	39 (51.3)
Follow-up, n (%)	
2 weeks	11 (14.5)
3 months	15 (19.7)
6 months	26 (34.2)
12 months	24 (31.6)
International Knee Documentation Committee scale, n (%)	
Minimal CID (>11.5)	51 (67.1)
Moderate CID (>20.5)	37 (48.7)

[a]Values are expressed as average ± SD (range) unless otherwise noted. CID, clinically important difference.

(PF) and Pain Interference (PI) scales, IKDC, and KOOS were collected at preoperative and postoperative visits through use of iPads and were stored using the REDcap system.[28] The time to collect PROMIS data has been reported in previous studies to be 2.4 minutes, and the data are instantaneously viewed in the electronic record for patient engagement and shared decision making on treatment plans.[14,15,27,28] All PROMIS CAT scores are presented as t scores with a score of 50 linked to the mean of the 2010 US Census data. Every 10 points represent 1 SD. Higher scores on the PROMIS PF indicate better function. In contrast, lower scores on the PROMIS PI indicate less pain.[6]

Statistical Analysis

The change from preoperative to the longest follow-up point on average was evaluated across all participants using analysis of variance (ANOVA) and t tests to examine the sample and follow-up period. The longest follow-up period varied widely, from 2 weeks to 12 months (Table 1). This ensured a wide range of change scores across the sample. To describe the preoperative to longest follow-up characteristics of the sample on the PRO scales, separate 2-way ANOVA models were used for KOOS and PROMIS, and a paired t test was used for the IKDC scale. The 2 factors were time (preoperative, longest follow-up) and each outcome's subscales or domains. The 5 subscales of the KOOS were Symptoms, Pain, Activities of Daily Living (ADL), Sports, and Quality of Life. The PROMIS had 2 domains: PF and PI. A main effect for KOOS was consistent with improvement across subscales. However, this was also followed by pairwise comparisons to verify that each subscale documented improvement over the time interval. An interaction effect was consistent with improvement across scales for PROMIS because the scales show improvement differently (improvement for PF is a higher t score and improvement for PI is a lower t score). This was followed by pairwise comparisons for each domain.

The ESI was used for all changes in scores to determine responsiveness.[17] The ESI is a common index of change to judge responsiveness of scales, where higher values indicate better responsiveness.[17] Values of 0.2 are considered low effects, 0.5 medium effects, and above 0.8 large effects.[17] Scales that show higher ESIs are able to detect change better than scales with lower ESI values. The change evaluated was the preoperative to longest follow-up (all participants) and each subgroup of time intervals (2 weeks, 3 months, 6 months, and 12 months). Although group membership was small for some time intervals (n = 11-23 for subgroups), the subgroup analysis provides preliminary data for exploring whether responsiveness may vary for some time intervals.

Pearson correlation coefficients were used to determine the convergent validity of the different scales to detect IKDC change. Pearson correlation coefficients (r values) were calculated for the change from preoperative to longest follow-up for the KOOS and PROMIS with the IKDC total scores. Because the KOOS and PROMIS PF and PI scales measure similar constructs to the IKDC scale, convergence or higher correlations were expected.

To evaluate what values might be used to judge minimal and moderate CID, the IKDC scale was used as the anchor. The IKDC scale is validated, and both minimal and moderate CID have been established.[18,19] Preoperative to postoperative change in IKDC scores of 11.5 was used to indicate a minimal CID and 20.5 was used to indicate a moderate CID.[18] Only improvement was evaluated in this study. After coding (0,1) minimal and moderate CID, receiver operating characteristic (ROC) curve analysis was used to determine the accuracy of each scale in documenting a minimal or moderate CID. The area under the curve (AUC) is a global measure of accuracy that can be used to compare scales.[4,13,17] An AUC above 0.7 is considered clinically useful and above 0.9 highly accurate.[26] The ROC curve analysis was sensitive to prevalence of minimal and moderate change as defined by the IKDC scale. Therefore, the prevalence of CID (moderate) was used to estimate the adequacy of the sample size. With a minimum sample of 74 participants and an AUC of 0.7, the estimated 95% CI for the AUC at 10% increments for the prevalence of CID was calculated. A prevalence of 20% or higher resulted in the 95% CI for the AUC excluding 0.5.

The threshold for change was chosen based on the closest point (ie, shortest distance) to no errors on the ROC curve (sensitivity = 1, specificity = 0) to determine what change was ideal. Although different approaches can be used, preliminary analysis showed that for most scales there was a clear shortest distance.[13,19] Where there was no clear shortest distance, the threshold lying in the middle of thresholds that resulted in similar shortest distance was used. The thresholds for each KOOS subscale and PROMIS scale are reported along with their sensitivity and specificity.

TABLE 2
Patient-Reported Outcomes at Preoperative Point and Longest Follow-up, and Changes Between These Values[a]

	Preoperative			Longest Follow-up			Change		
	n	Mean ± SD	Range	n	Mean ± SD	Range	Mean ± SD	Range	*P* Value
KOOS									
Symptoms	76	54.5 ± 20.3	0-89.3	75	72.8 ± 19.8	10.7-100	18.4 ± 19.4	–35.7 to 57.1	<.01
Pain	76	54.0 ± 16.6	0-91.7	74	75.5 ± 17.9	25.0-100	21.5 ± 18.1	–22.0 to 69.4	<.01
ADL	74	63.1 ± 17.9	1.5-100	68	80.2 ± 17.8	30.9-100	17.5 ± 16.7	–30.9 to 55.9	<.01
Sports	74	31.8 ± 21.7	0-80	74	56.1 ± 31.6	0-100	24.0 ± 30.0	–50.0 to 100	<.01
QOL	76	31.2 ± 15.8	0-75	75	53.0 ± 21.9	0-100	22.0 ± 21.3	–37.5 to 68.8	<.01
PROMIS									
PF	76	40.1 ± 6.2	27.2-63.5	75	46.4 ± 8.7	28.8-70.3	6.3 ± 7.3	–7.8 to 31.3	<.01
PI	76	59.7 ± 6.1	38.7-74.1	73	53.6 ± 8.0	38.7-74.1	–6.2 ± 6.7	–25.5 to 11.2	<.01
IKDC									
Total	76	41.1 ± 15.8	6.9-86.2	76	60.8 ± 21.2	16.1-100	20.2 ± 18.4	–15.0 to 66.7	<.01

[a]ADL, Activities of Daily Living; IKDC, International Knee Documentation Committee; KOOS, Knee injury and Osteoarthritis Outcome Score; PF, Physical Function; PI, Pain Interference; PROMIS, Patient Reported Outcomes Information System; QOL, Quality of Life.

RESULTS

Participants

Participants varied on several demographic and clinical variables (Table 1). After exclusion of participants who lacked demographic data, preoperative data, and follow-up data points, a total of 76 records were available for analysis. All 76 patients had completed preoperative data and at least one set of complete postoperative data. Participants ranged in follow-up time from 2 weeks (13.6%) to 12 months (29.6%). Average ± SD age was 48.9 ± 11.1 years (range, 20-74 years). Body mass index also indicated variability with an average ± SD of 30.5 ± 6.6 kg/m^2 (range, 19.3-48 kg/m^2). The proportion of participants experiencing a minimal CID improvement on the IKDC scale was 67.1%. The proportion experiencing a moderate CID improvement on the IKDC scale was slightly lower at 48.7%. With the exception of the KOOS ADL scale, which has 9 missing values, all other scales were missing fewer than 4 values (see values listed next to each scale in tables).

Change From Preoperative to Longest Follow-up

Despite the varied follow-up times, the average values preoperatively were significantly improved on all PROs at longest follow-up (Table 2). The ANOVA model main effect for KOOS ($P < .01$) and interaction effect for PROMIS ($P < .01$) scales were both significant. The KOOS scales all statistically improved ($P < .05$), with improvement ranging from 17.5 ± 16.7 on the ADL subscale to 24.0 ± 30.0 on the Sports subscale (Table 2). The PROMIS scales also statistically improved for each scale, where the PROMIS PF improved 6.3 ± 7.3 and PI improved −6.2 ± 6.7 (Table 2). The paired *t* test comparing preoperative IKDC with longest follow-up on the IKDC scale showed significant improvement of 20.2 ± 18.4 points ($P < .01$).

Responsiveness as Indicated by the Effect Size Index

Values of the ESI ranged from near 0 to 1.69 across the various time points (Table 3). At the 2-week time point, ESI values above 0.5 included IKDC (0.54), KOOS Symptoms (0.51) and KOOS Pain (0.68). All other scales were between 0.2 and 0.5 except PROMIS PF (0.04) and KOOS Sports (0.04). At 3 months, all scales showed an ESI above 0.5. Three of the ESI values for the KOOS subscales were above 0.8: Pain (1.05), ADL (0.84), and Sports (0.84). At 6 and 12 months, all scales were higher than 1.13. For the interval from preoperative to longest follow-up, all scales were higher than 0.8.

KOOS and PROMIS Correlation With IKDC

The KOOS and PROMIS scales all showed significant correlations with the IKDC scale. The correlations ranged from 0.61 to 0.79. The PROMIS domain scales PF and PI, which were 0.76 and −0.67, were within the range of KOOS subscales (*r* values ranged from 0.61 to 0.79) (Table 4).

Minimal Clinically Important Difference

The AUCs for all PROs were significant for predicting patients who experienced a minimal CID (Table 5). The KOOS subscales showed AUC between 0.78 and 0.92 (Figure 1). The minimal CID for the KOOS subscales varied from 12.5 to 17.5 on a 100-point scale. The sensitivity and specificity varied from 63.2% to 91.3%. The PROMIS PF and PI scales showed AUC values of 0.88 and 0.85, respectively (Figure 2). The minimal CIDs for the PROMIS PF and PI scales were 3.3 and −3.2, respectively. The sensitivity and specificity varied from 75.0% to 86.0% (Table 5).

Moderate Clinically Important Difference

The AUCs for all PROs were significant for predicting patients who experienced a moderate CID (Table 6). The KOOS subscales showed AUCs between 0.79 and 0.86

TABLE 3
Effect Size Index Values for the Entire Group and at Each Time Point[a]

	2 Weeks	3 Months	6 Months	12 Months	All
KOOS					
Symptoms	0.51 (n = 11)	0.49 (n = 14)	1.13 (n = 26)	1.56 (n = 23)	0.95 (n = 74)
Pain	0.68 (n = 10)	1.05 (n = 15)	1.46 (n = 25)	1.36 (n = 23)	1.19 (n = 73)
ADL	0.47 (n = 10)	0.84 (n = 14)	1.26 (n = 23)	1.58 (n = 20)	1.05 (n = 67)
Sports	0.04 (n = 11)	0.84 (n = 14)	1.15 (n = 25)	1.15 (n = 22)	0.80 (n = 72)
QOL	0.75 (n = 11)	0.48 (n = 15)	1.30 (n = 25)	1.44 (n = 24)	1.03 (n = 75)
PROMIS					
PF	0.04 (n = 11)	0.54 (n = 15)	1.14 (n = 25)	1.33 (n = 24)	0.86 (n = 75)
PI	0.44 (n = 11)	0.54 (n = 14)	1.25 (n = 25)	1.27 (n = 23)	0.93 (n = 73)
IKDC					
Total	0.54 (n = 11)	0.64 (n = 15)	1.33 (n = 26)	1.69 (n = 24)	1.10 (n = 76)

[a]ADL, Activities of Daily Living; IKDC, International Knee Documentation Committee; KOOS, Knee injury and Osteoarthritis Outcome Score; PF, Physical Function; PI, Pain Interference; PROMIS, Patient Reported Outcomes Information System; QOL, Quality of Life.

TABLE 4
Correlations of the Change From Preoperative to Longest Follow-up With the Change of the IKDC Scale[a]

	n	*r* Value	*P* Value
KOOS			
Symptoms	74	0.61	<.01
Pain	73	0.74	<.01
ADL	67	0.70	<.01
Sports	72	0.79	<.01
QOL	75	0.72	<.01
PROMIS			
PF	75	0.76	<.01
PI	73	–0.67	<.01

[a]ADL, Activities of Daily Living; IKDC, International Knee Documentation Committee; KOOS, Knee injury and Osteoarthritis Outcome Score; PF, Physical Function; PI, Pain Interference; PROMIS, Patient Reported Outcomes Information System; QOL, Quality of Life.

(Figures 1 and 2). The moderate CID for the KOOS subscales, determined by taking the shortest distance to a perfect score on the ROC, varied from 14.3 to 18.8 on a 100-point scale. The sensitivity and specificity varied from 61.8% to 88.8%. The PROMIS PF and PI scales showed AUC values of 0.89 for both scales (Figures 3 and 4). The moderate CIDs for the PROMIS PF and PI scales were 5.0 and −5.8, respectively. The sensitivity and specificity varied from 71.8% to 88.9%.

DISCUSSION

The new findings of this analysis are that PROMIS PF and PI scales show similar responsiveness to selected disease-specific scales in knee arthroscopy patients. Except for the 2-week time point, all scales demonstrated ESI above 0.5, suggesting at least medium ability or above to detect change. The correlations between the PROMIS and KOOS scales with the IKDC scales demonstrated convergence. This finding supports the construct of the PROMIS PF and PI scales as the comparable disease-specific scales (ie, KOOS and IKDC). The minimal and moderate CID across the PROMIS and KOOS scales suggest that both scales are likely useful to detect change. Aside from the 2-week time point, the data support the ability of PROMIS PF and PI to detect change in patients after knee arthroscopy.

Rather than use a standardized follow-up point, this analysis used a wide range of follow-up points (2 weeks to a minimum of 12 months). This sampling strategy ensures a wide range of change scores across the continuum of recovery (see Table 2). This sampling strategy resulted in 67.1% and 48.7% of participants meeting minimal and moderate CID criteria on the IKDC scale, respectively. The advantage of this approach is that a wide range of change scores are represented in the sample. Therefore, the ability of the scales to detect change across time points is evaluated. The disadvantage is that average improvements hold little meaning relative to overall outcome. However, all PRO scales showed significant aggregate improvement (see Table 2). The current sample is representative of a wide range of clinical responses across the continuum of care that led to a higher proportion of patients meeting criteria for "improved" and "not improved" in this analysis.

The correlation of the PROMIS scale change scores demonstrates convergent validity with the IKDC scale. The hypothesis that the PROMIS PF and PI scales share similar constructs with the IKDC scales was supported. The convergences of PROMIS PF and PI were 0.76 and −0.67, respectively. The negative correlation between PROMIS PI and IKDC was consistent with the fact that higher scores on the PROMIS PI scale indicate worse pain interference while lower scores on IKDC indicate worse function. The strength of the PROMIS PF and PI correlations were similar to the KOOS subscales, which varied from 0.61 to 0.79. This suggests that despite including only global health items, the PROMIS PF and PI scales showed similar convergence as a disease-specific scale. These correlation data suggest that the PROMIS PF and PI capture similar constructs associated with patient functioning and pain as the IKDC scale.

TABLE 5
Receiver Operating Characteristic Curve Analysis Based on 11.5 Change in IKDC Scale at Longest Follow-up (2 weeks to 12 months)[a]

Scale	Threshold	AUC	95% CI	*P* Value	Sensitivity	Specificity
PROMIS						
PF	3.3	0.88 (0.04)	0.81-0.96	<.001	86.0	76.0
PI	–3.2	0.85 (0.04)	0.77-0.94	<.001	83.7	75.0
KOOS						
Symptoms	16.0	0.78 (0.06)	0.66-0.90	<.001	72.2	73.9
Pain	12.5	0.84 (0.05)	0.73-0.94	<.001	81.1	63.2
ADL	12.6	0.86 (0.05)	0.76-0.97	<.001	86.1	78.3
Sports	17.5	0.88 (0.04)	0.80-0.97	<.001	80.6	82.6
QOL	15.6	0.92 (0.04)	0.84-1.00	<.001	83.3	91.3

[a]ADL, Activities of Daily Living; AUC, area under the curve; IKDC, International Knee Documentation Committee; KOOS, Knee injury and Osteoarthritis Outcome Score; PF, Physical Function; PI, Pain Interference; PROMIS, Patient Reported Outcomes Information System; QOL, Quality of Life. Values in parentheses are standard error.

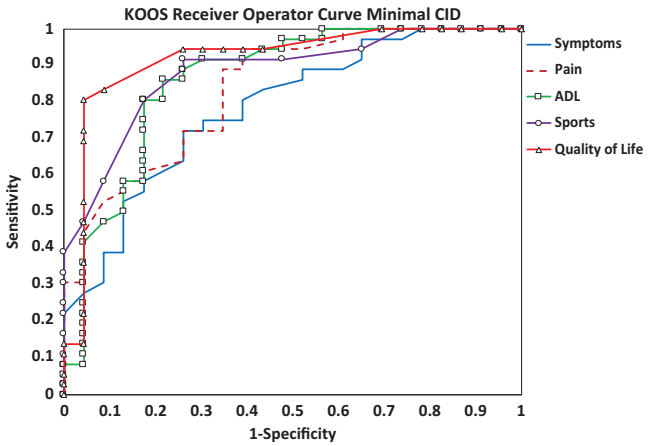

Figure 1. Receiver operating characteristic curve of minimal clinically important differences (CIDs) for Knee injury and Osteoarthritis Outcome Score (KOOS) scales. ADL, Activities of Daily Living.

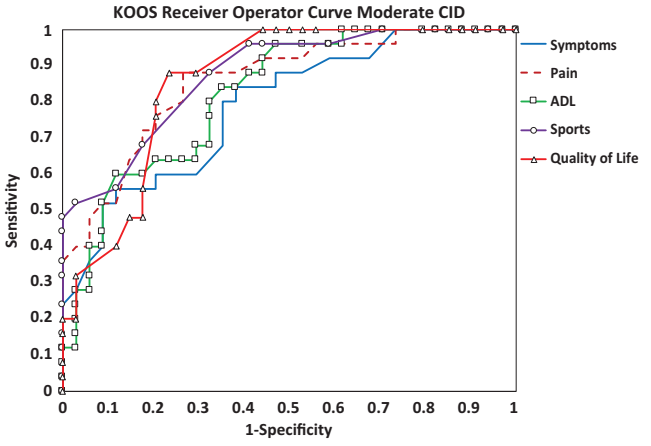

Figure 2. Receiver operating characteristic curve of moderate clinically important differences (CID) for Knee injury and Osteoarthritis Outcome Score (KOOS) scales. ADL, Activities of Daily Living.

This study improves on previous studies by reporting CID changes using the IKDC as an anchor or reference for change (Tables 5 and 6). Previous studies have used the distribution method, which determines the minimal CID as one-half SD of the change for the PROMIS PF and PI scales.[4,13] Their values for minimal CID were as high as 5.8.[4,13] This contrasts with an anchor-based method applied to patients with low back pain. PROMIS PI values of 3.5 to 5.5 were considered meaningful for these participants corresponding to a minimal CID.[1] The data from the present study suggest that a minimal CID is 3.2 to 3.3 and a moderate CID is 5.0 to 5.8 for PROMIS PF and PI, which are similar to other anchor-based estimates[1] and lower than distribution-based estimates.[4,13] It remains unclear whether there are distinct differences in minimal and moderate CID across diagnoses. The AUC values for the PROMIS PF and PI scales were above 0.85 for both minimal (Table 5) and moderate (Table 6) CID. This was similar to the KOOS subscales, whose AUC values ranged from 0.78 to 0.92 for minimal CID and 0.79 to 0.87 for moderate CID. This suggests that both the PROMIS PF and PI scales had similar accuracy to a disease-specific scale in identifying change on the IKDC. The sensitivity and specificity of the PROMIS PF and PI scales were also similar for the identified CID values compared with the KOOS subscales. Overall, the PROMIS PI and PF scales' minimal and moderate CID values appear to perform similarly to those of a comparative disease-specific scale (ie, KOOS).

The clinical implication of this study is that PROMIS PF and PI are likely equally as effective at tracking patient recovery after knee arthroscopy as disease-specific scales. An advantage to using PROMIS PF and PI clinically is that improvements due to arthroscopic knee surgery are associated with global health rather than specific to the knee. The only caution seen in these data is the low ESI at 2 weeks. PROMIS may also be used to satisfy federal regulation and performance metrics used in evolving reimbursement models. Advantages to the PROMIS scales are

TABLE 6
Receiver Operating Characteristic Curve Analysis Based on 20.5 Change in IKDC Scale at Longest Follow-up (2 weeks to 12 months)[a]

Scale	Threshold	AUC	95% CI	*P* Value	Sensitivity	Specificity
PROMIS						
PF	5.0	0.89 (0.04)	0.81-0.96	<.001	88.9	71.8
PI	–5.8	0.89 (0.04)	0.82-0.96	<.001	85.7	78.9
KOOS						
Symptoms	14.3	0.79 (0.06)	0.67-0.90	<.001	84.0	61.8
Pain	18.1	0.86 (0.05)	0.76-0.95	<.001	88.0	74.5
ADL	15.4	0.82 (0.05)	0.72-0.93	<.001	80.0	67.6
Sports	17.5	0.87 (0.04)	0.79-0.96	<.001	88.0	67.6
QOL	18.8	0.86 (0.04)	0.76-0.95	<.001	88.0	76.5

[a]ADL, Activities of Daily Living; AUC, area under the curve; IKDC, International Knee Documentation Committee; KOOS, Knee injury and Osteoarthritis Outcome Score; PF, Physical Function; PI, Pain Interference; PROMIS, Patient Reported Outcomes Information System; QOL, Quality of Life. Values in parentheses are standard error.

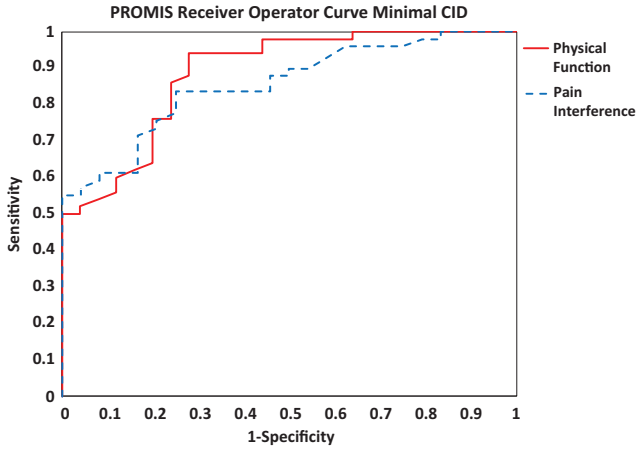

Figure 3. Receiver operating characteristic curve of minimal clinically important differences (CIDs) for the Patient Reported Outcomes Information System (PROMIS).

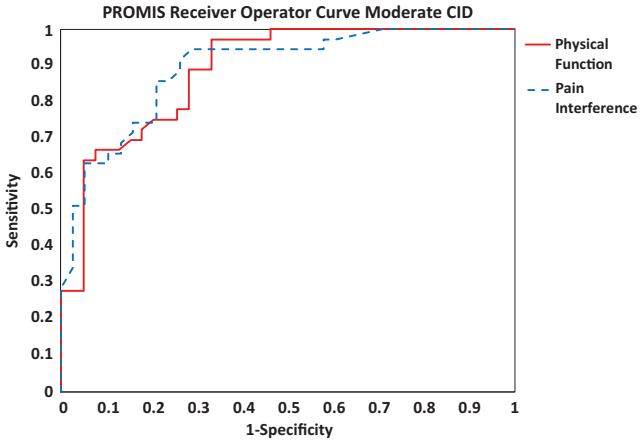

Figure 4. Receiver operating characteristic curve of moderate clinically important differences (CIDs) for the Patient Reported Outcomes Information System (PROMIS).

that they are agnostic to diagnosis and therefore more easily applied than disease-specific scales across a wide spectrum of patients. In addition, use of the CAT approach decreases data collection time and the results are displayed in the electronic medical record, making it a useful tool in high-flow practice models.[28]

Limitations

The limitations of this study are also important to consider. The advantage of a sample with varied follow-up was a wide variance in the change scores. An alternative sampling approach would be prospective data of a larger sample at standardized time points. This directly affects the ESI data shown at different time points. The 2-week time point is undersampled but demonstrates how prospective data at multiple time points would be useful to understand the responsiveness of scales across recovery. Applying the current minimal and moderate CID values to a separate sample would also validate the generalizability of the CID values.[3,4] Other approaches to determining the CID value include the interquartile range and using other Likert scale approaches.[1] The results of this study are dependent on the validity of the IKDC scale as an anchor.[9,11,22,23]

CONCLUSION

The global health PROMIS PF and PI scales showed similar ability to detect change to that of disease-specific measures for patients recovering from knee arthroscopy. Specifically, the responsiveness of the PROMIS PF and PI scales was similar to that of the KOOS and IKDC. The PROMIS PF and PI showed similar convergence with IKDC to the KOOS, suggesting that these scales measure similar constructs associated with physical ability and pain. Finally, minimal and moderate CID values were calculated for PROMIS PF and PI and KOOS; such values will assist with clinical decisions regarding whether

patients are significantly improved as they recover from knee arthroscopy and can guide patients in clinical decision making with recovery.

ACKNOWLEDGMENT

The authors thank Kaili Widrick and Emily Palmeri for research coordination and patient enrollment.

REFERENCES

1. Amtmann D, Kim J, Chung H, Askew RL, Park R, Cook KF. Minimally important differences for Patient Reported Outcomes Measurement Information System pain interference for individuals with back pain. *J Pain Res*. 2016;9:251-255.
2. Anderson AF, Irrgang JJ, Kocher MS, Mann BJ, Harrast JJ; International Knee Documentation Committee. The International Knee Documentation Committee Subjective Knee Evaluation Form: normative data. *Am J Sports Med*. 2006;34(1):128-135.
3. Anderson MR, Baumhauer JF, DiGiovanni BF, et al. Determining success or failure after foot and ankle surgery using Patient Acceptable Symptom State (PASS) and Patient Reported Outcome Information System (PROMIS). *Foot Ankle Int*. 2018;39(8):894-902.
4. Anderson MR, Houck JR, Saltzman CL, et al. Validation and generalizability of preoperative PROMIS scores to predict postoperative success in foot and ankle patients. *Foot Ankle Int*. 2018;39(7):763-770.
5. Baumhauer JF. Patient-reported outcomes—are they living up to their potential? *N Engl J Med*. 2017;377(1):6-9.
6. Brodke DJ, Saltzman CL, Brodke DS. PROMIS for orthopaedic outcomes measurement. *J Am Acad Orthop Surg*. 2016;24(11):744-749.
7. Bumpass DB, Samora JB, Butler CA, Jevsevar DS, Moffatt-Bruce SD, Bozic KJ. Orthopaedic quality reporting: a comprehensive review of the current landscape and a roadmap for progress. *JBJS Rev*. 2014;2(8):e5.
8. Cella D, Yount S, Rothrock N, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS): progress of an NIH Roadmap cooperative group during its first two years. *Med Care*. 2007;45(5)(suppl 1):S3-S11.
9. Chatain F, Adeleine P, Chambat P, Neyret P. A comparative study of medial versus lateral arthroscopic partial meniscectomy on stable knees: 10-year minimum follow-up. *Arthroscopy*. 2003;19(8):842-849.
10. Cook KF, Jensen SE, Schalet BD, et al. PROMIS measures of pain, fatigue, negative affect, physical function, and social function demonstrated clinical validity across a range of chronic conditions. *J Clin Epidemiol*. 2016;73:89-102.
11. Crawford K, Briggs KK, Rodkey WG, Steadman JR. Reliability, validity, and responsiveness of the IKDC score for meniscus injuries of the knee. *Arthroscopy*. 2007;23(8):839-844.
12. Greco NJ, Anderson AF, Mann BJ, et al. Responsiveness of the International Knee Documentation Committee Subjective Knee Form in comparison to the Western Ontario and McMaster Universities Osteoarthritis Index, modified Cincinnati Knee Rating System, and Short Form 36 in patients with focal articular cartilage defects. *Am J Sports Med*. 2010;38(5):891-902.
13. Ho B, Houck JR, Flemister AS, et al. Preoperative PROMIS scores predict postoperative success in foot and ankle patients. *Foot Ankle Int*. 2016;37(9):911-918.
14. Hung M, Baumhauer JF, Brodsky JW, et al. Psychometric comparison of the PROMIS physical function CAT with the FAAM and FFI for measuring patient-reported outcomes. *Foot Ankle Int*. 2014;35(6):592-599.
15. Hung M, Baumhauer JF, Latt LD, et al. Validation of PROMIS (R) Physical Function computerized adaptive tests for orthopaedic foot and ankle outcome research. *Clin Orthop Relat Res*. 2013;471(11):3466-3474.
16. 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;35(1):1-7.
17. Husted JA, Cook RJ, Farewell VT, Gladman DD. Methods for assessing responsiveness: a critical review and recommendations. *J Clin Epidemiol*. 2000;53(5):459-468.
18. Irrgang JJ, Anderson AF, Boland AL, et al. Development and validation of the International Knee Documentation Committee Subjective Knee Form. *Am J Sports Med*. 2001;29(5):600-613.
19. Irrgang JJ, Anderson AF, Boland AL, et al. Responsiveness of the International Knee Documentation Committee Subjective Knee Form. *Am J Sports Med*. 2006;34(10):1567-1573.
20. Jones RS, Stukenborg GJ. Patient-Reported Outcomes Measurement Information System (PROMIS) use in surgical care: a scoping study. *J Am Coll Surg*. 2017;224(3):245-254.e241.
21. Kim S, Bosque J, Meehan JP, Jamali A, Marder R. Increase in outpatient knee arthroscopy in the United States: a comparison of National Surveys of Ambulatory Surgery, 1996 and 2006. *J Bone Joint Surg Am*. 2011;93(11):994-1000.
22. Kon E, Filardo G, Berruto M, et al. Articular cartilage treatment in high-level male soccer players: a prospective comparative study of arthroscopic second-generation autologous chondrocyte implantation versus microfracture. *Am J Sports Med*. 2011;39(12):2549-2557.
23. Koyonos L, Yanke AB, McNickle AG, et al. A randomized, prospective, double-blind study to investigate the effectiveness of adding DepoMedrol to a local anesthetic injection in postmeniscectomy patients with osteoarthritis of the knee. *Am J Sports Med*. 2009;37(6):1077-1082.
24. Kvien TK, Heiberg T, Hagen KB. Minimal clinically important improvement/difference (MCII/MCID) and patient acceptable symptom state (PASS): what do these concepts mean? *Ann Rheum Dis*. 2007;66(suppl 3):iii40-iii41.
25. Lubowitz JH, Appleby D. Cost-effectiveness analysis of the most common orthopaedic surgery procedures: knee arthroscopy and knee anterior cruciate ligament reconstruction. *Arthroscopy*. 2011;27(10):1317-1322.
26. Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. *J Thorac Oncol*. 2010;5(9):1315-1316.
27. Papuga MO, Beck CA, Kates SL, Schwarz EM, Maloney MD. Validation of GAITRite and PROMIS as high-throughput physical function outcome measures following ACL reconstruction. *J Orthop Res*. 2014;32(6):793-801.
28. Papuga MO, Dasilva C, McIntyre A, Mitten D, Kates S, Baumhauer JF. Large-scale clinical implementation of PROMIS computer adaptive testing with direct incorporation into the electronic medical record. *Health Systems*. 2017;7(1):1-12.
29. Porter ME, Larsson S, Lee TH. Standardizing patient outcomes measurement. *N Engl J Med*. 2016;374(6):504-506.
30. Roos EM, Lohmander LS. The Knee injury and Osteoarthritis Outcome Score (KOOS): from joint injury to osteoarthritis. *Health Qual Life Outcomes*. 2003;1:64.
31. Wright RW. Knee injury outcomes measures. *J Am Acad Orthop Surg*. 2009;17(1):31-39.