

Preoperative Evaluation of the Lower Extremity-Specific *PROMIS Mobility* Bank in Patients with ACL Tears



Molly A. Day, M.D., A.T.C., Kyle J. Hancock, M.D., Vinicius C. Antao, M.D., M.Sc., Ph.D., Joseph D. Lamplot, M.D., Russell F. Warren, M.D., Benedict U. Nwachukwu, M.D., M.B.A., and Andrew D. Pearle, M.D.

Purpose: The purpose of this study was to preoperatively assess the Patient-Reported Outcomes Measurement Information System (PROMIS) Item Bank v2.0–Mobility in patients with anterior cruciate ligament (ACL) tears to (1) determine construct validity by correlating with legacy patient-reported outcomes (PROs), (2) evaluate test burden, (3) determine the presence of floor or ceiling effects, and (4) revisit the conventional threshold for inclusiveness (floor/ceiling effects) in the modern era of computer adaptive testing (CAT)–based PROs. **Methods:** Patients at a large academic musculoskeletal specialty center diagnosed with ACL tears indicated for surgery were administered the following outcomes measures before surgery: PROMIS Mobility CAT, PROMIS Pain Interference CAT (PROMIS PI), International Knee Documentation Committee (IKDC), the Marx Knee Activity Rating Scale (Marx), and Single Assessment Numeric Evaluation (SANE). Construct validity was evaluated using Spearman correlation coefficients. Correlation strengths were defined as high (≥ 0.7), high-moderate (0.61–0.69), moderate (0.4–0.6), moderate-weak (0.31–0.39) and weak (≤ 0.3). Number of questions to completion were recorded as a marker of test burden. The percentage of patients scoring at the extreme high (ceiling) or low (floor) of each measure was recorded to measure inclusivity. **Results:** A total of 1126 patients were evaluated. The mean number of questions answered (\pm standard deviation) was 4.7 ± 2.1 for PROMIS Mobility and 4.5 ± 1.9 for PROMIS PI. PROMIS Mobility demonstrated a high correlation with IKDC, ($r = 0.81, P < .001$), a high-moderate correlation with PROMIS PI ($r = -0.63$), and a moderate correlation with SANE ($r = 0.46, P < .01$). Neither PROMIS Mobility nor PROMIS PI met conventional criteria for floor or ceiling effects ($>15\%$). **Conclusions:** The PROMIS Mobility measure maintains construct validity, because its scores correlate strongly with other PROs measuring physical function with high efficiency among preoperative patients with ACL injuries. Although ceiling effects of PROMIS Mobility CAT were below the conventional significance threshold of 15% at the preoperative timepoint in this population, this study provides critical feedback for redesigning the Mobility bank. **Level of Evidence:** Level III (Diagnostic study).

Although the fundamental importance of well-designed outcome measures has become more clearly defined over the past several decades, there are still many measures in use today that are suboptimal.^{1–4} Recently, there has been a substantial number of

orthopaedic studies defining and validating the characteristics of various Patient-Reported Outcomes Measurement Information System (PROMIS) measures, with particular attention to the physical function (PROMIS PF) and pain interference (PROMIS PI)

From the Sports Medicine Institute, Department of Orthopaedic Surgery, Hospital for Special Surgery (M.A.D., R.F.W., B.U.N., A.D.P.), New York, New York, Desert Orthopaedic Center (K.J.H.), Las Vegas, Nevada, Center for the Advancement of Value in Musculoskeletal Care, Hospital for Special Surgery (V.C.A.), New York, New York, and Emory Orthopaedics & Spine Center—Johns Creek (J.D.L.), Johns Creek, Georgia, U.S.A.

The authors report the following potential conflicts of interest or sources of funding: B.U.N. reports royalties from Remote Health; A.D.P. reports consultancy fees from Exactech, Inc, Smith & Nephew, Stryker, and Zimmer, payment for manuscript preparation from Stryker, stock/stock options from Engage. Full ICMJE author disclosure forms are available for this article online, as [supplementary material](#).

Received October 26, 2020; accepted March 14, 2021.

Address correspondence to Molly Ann Day, M.D., Sports Medicine Institute, Hospital for Special Surgery, 535 E 70th St., New York, NY 10021. E-mail: molly.a.day@gmail.com

© 2021 THE AUTHORS. Published by Elsevier Inc. on behalf of the Arthroscopy Association of North America. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>). 2666-061X/201726

<https://doi.org/10.1016/j.asmr.2021.03.007>

measures, because they are highly relevant to orthopaedic patients.⁴⁻¹³ A key advantage of many PROMIS measures is that they can be administered using computer adaptive tests (CATs), which selectively draw questions from a large bank according to an internal algorithm. This design results in improved efficiency of administration compared to legacy fixed-length measures, a feature that has repeatedly been demonstrated.^{4,7,8,12-14} They have also been shown to have high content validity and responsiveness to change,⁸ good reliability,⁷ and overall construct validity.^{12,13}

As the number of PROMIS measures have expanded over time, researchers and clinicians have suggested a need for anatomic region-specific measures of physical function. The first of these was the PROMIS Upper Extremity bank, which has been studied and refined over time to its current version. This measure is calibrated independently of the PROMIS PF measure and is more applicable to patients with upper extremity problems.¹⁵ With an analogous goal of creating a lower-extremity specific physical function measure, the PROMIS Mobility bank was recently introduced. This measure relies on a subset of questions from the PROMIS PF bank that specifically relates to lower extremity function and excludes those relating more to core and upper extremity function.^{16,17} To date, the PROMIS Mobility bank has not been formally evaluated in patients with ACL tears, and the latest version (v2.0) has never been assessed in any sports medicine population. The early version 1.0 of the Mobility bank has been demonstrated to correlate well with common fixed-length PROs and without floor or ceiling effects meeting the significance threshold in patients with lower extremity fractures or multiligamentous knee injuries.^{16,17} These studies, however, were performed in cohorts of patients who are typically more disabled after surgery than patients undergoing ACL reconstruction who may ultimately achieve higher function and therefore may be at a higher risk of floor or ceiling effects. Although the conventional definition for floor or ceiling effects is $\geq 15\%$ of patients obtaining the minimum or maximum possible score,^{12,13} respectively, this definition may not be appropriate in the modern era of CATs.

The purpose of this study was to preoperatively assesses the Patient-Reported Outcomes Measurement Information System (PROMIS) Item Bank v2.0–Mobility in patients with anterior cruciate ligament tears in order to: (1) determine construct validity by correlating with legacy patient reported outcomes (PROs), (2) evaluate test burden, (3) determine the presence of floor or ceiling effects, and (4) revisit the conventional threshold for inclusiveness (floor/ceiling effects) in the modern era of computer adaptive testing-based PROs. We hypothesized that the PROMIS Mobility CAT would maintain construct validity and

correlate well with other measures with similar domains. Furthermore, we hypothesized that PROMIS Mobility CAT would demonstrate low test burden compared to legacy measures with no floor or ceiling effects above the traditional threshold of significance.

Materials and Methods

All patients at a single musculoskeletal specialty academic institution who were diagnosed with ACL tears and indicated for operative management from January 1, 2019, to December 31, 2019, were enrolled prospectively in the institutional ACL registry. Patients with concomitant meniscal or chondral injuries, and recurrent ACL tear after ACL reconstruction were included; those with simultaneous injuries to the posterior cruciate ligament or collateral ligaments requiring operative treatment were excluded. Baseline outcomes scores were collected from each patient using an electronic registry platform (OBERD, Columbia, MO) during one data collection session within the month preceding surgery for the following: PROMIS Item Bank v2.0 – Mobility (PROMIS Mobility), PROMIS Item Bank v1.1 – Pain Interference (PROMIS PI), International Knee Documentation Committee (IKDC), The Marx Knee Activity Rating Scale (Marx), and the Single Assessment Numeric Evaluation (SANE). Patients who completed all 5 PROs were included for final analysis in the study. Patients who did not complete all 5 surveys were excluded. When necessary, chart checking was performed by a research assistant or orthopaedic fellow (M.D., K.H.). This study was approved by the institutional Registry Steering Committee and deemed exempt from full Institutional Review Board review.

Statistical Analysis

Patient demographic data was recorded. A test for normality was performed using the Shapiro-Wilk method. Spearman correlation coefficients were used to evaluate associations between the PROMIS Mobility measure and all other PROs. Statistical significance level was set at $P < .05$. Correlation strengths were defined as high (≥ 0.7), high-moderate (0.61-0.69), moderate (0.4-0.6), moderate-weak (0.31-0.39), and weak (≤ 0.3).¹⁸ The correlations between PROMIS Mobility and the other PROs were compared to evaluate convergent validity for measures that also measure physical function and divergent validity for measures that are generic or measure some other health domain. An a priori power analysis was performed with an alpha of 0.05 and discriminatory power of 80% suggesting that a sample size of 36 was required to detect a correlation difference of 0.6 from 0.2. Inclusiveness was evaluated by determining the proportion of patients who obtained either the highest (ceiling) or lowest (floor) possible score on a specific instrument. A floor or

ceiling effect was defined as $\geq 15\%$ of patients obtaining the highest or lowest, possible score, respectively, for a specific instrument. PROMIS measures were administered as CATs according to standard rules for PROMIS, and patients completed the instrument when the standard predetermined level of significance was reached.⁹ For each patient, the number of items until completion was recorded as a measure of test burden.

Results

During the study period, 1423 patients were diagnosed with ACL tear, evaluated by 26 orthopaedic surgeons. There was a total of 1126 patients included in final analysis (49.8% female; 50.2% male), ages 12 to 76 years with a mean age (\pm standard deviation [SD]) of 30.6 ± 12.6 years (Table 1); 297 patients were excluded for not having completed all 5 preoperative PRO surveys. Mean results for each instrument are provided in Table 2. The mean number of questions answered (\pm SD) was 4.7 ± 2.1 for PROMIS Mobility and 4.5 ± 1.9 for PROMIS PI (range 4-12 for each).

PROMIS Mobility demonstrated a high correlation with IKDC, ($r = 0.81$, $P < .001$), a high-moderate correlation with PROMIS PI ($r = -0.63$), and a moderate correlation with SANE ($r = 0.46$, $P < .01$) (Table 3). PROMIS PI demonstrated a high correlation ($r = -0.75$, $P < .001$) with IKDC and a moderate correlation with SANE ($r = -0.4$, $P < .01$). Marx demonstrated weak correlations with each PROMIS instrument.

Neither PROMIS Mobility nor PROMIS PI had $\geq 15\%$ floor or ceiling effects (Table 4). Marx was the only measure reaching criteria for a significant ceiling effect, with 29.4% of patients achieving the highest possible score.

Discussion

This study demonstrated PROMIS Mobility CAT to have excellent correlation with existing legacy measures, no floor or ceiling effects above the traditional threshold of significance, and minimal patient burden, with patients only required to answer an average of 4 to

Table 2. PRO Measure Results

PRO Measure	Mean (Standard Deviation)
PROMIS Mobility	41.7 (7.1%)
PROMIS PI	57.2 (7.2%)
IKDC	47.9 (16.7%)
SANE	40.9 (26.5%)
MARX	11.1 (4.9%)

IKDC, International Knee Documentation Committee Score; MARX, Marx Knee Activity Rating Scale; PRO, patient-reported outcome; PROMIS Mobility, PROMIS Mobility Computer Adaptive Test; PROMIS PI, PROMIS Pain Interference Computer Adaptive Test; SANE, Single Assessment Numeric Evaluation.

5 questions. The overarching goal of developing and refining a more anatomic-specific measure is to improve the clinical relevance when investigating problems that specifically affect lower extremity function. The version 2.0 PROMIS Mobility instrument appears to assess physical function similarly to other legacy measures with less patient burden.

The use of PROs has become increasingly prevalent in clinical research, with an increased focus on constructing quality instruments. Yet, there are still many measures in use today that are suboptimal. This concern was highlighted in a recent study comparing legacy fixed-length PRO scores such as the Short Form-36 or the Knee Injury and Osteoarthritis Outcomes Score, to the CAT version of the PROMIS PF (PROMIS PF CAT).⁴ As early as 6 months after ACL reconstruction, one third to one half of patients obtained the highest possible score on several legacy measures, suggesting a ceiling effect. This effectively means that, among patients scoring at or above the fiftieth percentile, these legacy PROs are unable to discriminate between patients' outcomes after surgery. The ceiling effects observed hinder our ability to assess successful clinical outcomes, because the efficacy of treatment may be beyond the measure of these legacy PROs. Furthermore, these ceiling effects may also preclude adequate assessment of new techniques or devices, which strive to improve outcomes within this upper half of patients. These limitations ultimately led to the National Institutes of Health-funded development of PROMIS with the goal of creating more optimal PROs.

Among this cohort of patients with symptomatic ACL tears who subsequently underwent ACL surgery, PROMIS Mobility scores were highly correlated with IKDC scores, a commonly used fixed-length, anatomic-specific outcomes instrument that largely measures knee function.¹⁹ This suggests convergent validity and is consistent with previously published studies of patients with lower extremity fractures and multi-ligamentous knee injuries.^{16,17} PROMIS Mobility correlated to a lesser extent with SANE and PROMIS PI scores, an expected finding that suggests divergent

Table 1. Demographics

Variable	n	%	Mean	SD
Patients	1126	100.0	—	—
Gender				
Female	561	49.8	—	—
Male	565	50.2	—	—
Laterality				
Right	551	49.1	—	—
Left	571	50.9	—	—
Age	—	—	30.6	12.6
BMI	—	—	25.0	4.1

*N, number; SD, standard deviation; BMI, body mass index.

Table 3. Correlation Between PROMIS Measures and Other Patient-Reported Outcome Measures

PRO Measure	PROMIS Mobility			PROMIS PI		
	<i>r</i> Value	<i>P</i> Value	Correlation Strength	<i>r</i> Value	<i>P</i> Value	Correlation Strength
Measuring physical function						
IKDC	0.81	<.01	High	-0.75	<.01	High
PROMIS mobility	1	—	—	-0.63	<.01	High-Moderate
Not measuring physical function						
PROMIS PI	-0.63	<.01	High-Moderate	1	—	—
SANE	0.46	<.01	Moderate	-0.40	<.01	Moderate
MARX	-0.01	.64	Weak	0.04	.18	Weak

IKDC, International Knee Documentation Committee Score; MARX, Marx Knee Activity Rating Scale; PROMIS Mobility, PROMIS Mobility Computer Adaptive Test; PROMIS PI, PROMIS Pain Interference Computer Adaptive Test; PRO, Patient reported outcome; SANE, Single Assessment Numeric Evaluation.

validity. Additionally, the test burden of the PROMIS Mobility instrument was low, with patients answering a mean of 4.7 items, comparatively fewer than the 19 items (10 questions with 1 question having 9 and 1 having 2 subsections, respectively) required to complete the IKDC. These findings indicate that the PROMIS Mobility CAT likely provides similar information as the IKDC but with improved efficiency. Our findings also suggest that PROMIS Mobility is by traditional definitions inclusive in preoperative patients with ACL tears, because there were no floor or ceiling effects meeting the $\geq 15\%$ conventional threshold of significance, consistent with other recently published studies.^{16,17}

In this study, we used an injured preoperative cohort of patients who are expected to return to a higher level of functioning after surgery. As such, the potential exists for ceiling effects at postoperative follow-up time points. Although ceiling effects greater than 15% have not been observed in other studies of PROMIS Mobility, these studies were performed in older, more overweight, and more impaired patient populations^{16,17} who may not attain the same postoperative level of function as those undergoing ACL reconstruction. Importantly, the authors believe that the conventionally used 15% threshold for ceiling effects limits the ability to discern differences among patients reaching the maximum score for a given measure, and that in the era of CATs, this threshold may be unacceptably generous. In a study

investigating the 1.0 version of PROMIS Mobility in patients with lower extremity fractures, Rothrock et al.¹⁷ found that 10% of patients attained the highest score possible on the PROMIS Mobility instrument at final follow-up. Although this does not meet the conventional threshold of 15%, it may still represent an unacceptably high proportion of patients. Patients undergoing ACL reconstruction frequently seek to return to high-level cutting and pivoting activities, and thus having a PRO measure that measures the highest performers discriminately is essential. Efficient PROs with low ceiling effects are also vital to the evaluation of new techniques and devices aimed at improving already highly successful operations such as ACL reconstruction. If one cannot measure those with the best outcomes, then improving on already high success rates is unlikely to be accomplished effectively.

Limitations

There are several limitations to this study. First, the generalizability of the results to other lower extremity conditions and patient populations may be limited because of the specific cohort of patients chosen and that patients evaluated were from a single musculoskeletal academic institution. Second, the number of questions to completion was used as a metric of test burden rather than time to completion, which may be a more relevant metric and was not recorded. Finally, this study only assessed preoperative scores, and therefore were unable to assess the performance of PROMIS Mobility CAT after ACL reconstruction.

Table 4. PRO Measure Floor and Ceiling Effects

PRO Measure	Floor Effects	Ceiling Effects
PROMIS Mobility	4 (0.4%)	41 (3.6%)
PROMIS PI	54 (4.9%)	0 (0.0%)
IKDC	0 (0.0%)	0 (0.0%)
SANE	121 (10.8%)	12 (1.1%)
MARX	68 (6.0%)	331 (29.4%)

IKDC, International Knee Documentation Committee Score; MARX, Marx Knee Activity Rating Scale; PRO, Patient reported outcome; PROMIS Mobility, PROMIS Mobility Computer Adaptive Test; PROMIS PI, PROMIS Pain Interference Computer Adaptive Test; SANE, Single Assessment Numeric Evaluation.

Conclusions

The PROMIS Mobility measure maintains construct validity, as its scores correlate strongly with other PROs measuring physical function with a high efficiency among preoperative patients with ACL injuries. Although ceiling effects of the PROMIS Mobility CAT were below the conventional significance threshold of 15% at the preoperative timepoint in this population,

this study provides critical feedback for redesigning the Mobility bank.

References

1. Kane LT, Namdari S, Plummer OR, Beredjikian P, Vaccaro A, Abboud JA. Use of computerized adaptive testing to develop more concise patient-reported outcome measures. *JBJS Open Access* 2020;5:e0052.
2. Kitaoka HB, Meeker JE, Phisitkul P, Adams SB Jr, Kaplan JR, Wagner E. AOFAS position statement regarding patient-reported outcome measures. *Foot Ankle Int* 2018;39:1389-1393.
3. Svantesson E, Hamrin Senorski E, Webster KE, et al. Clinical outcomes after anterior cruciate ligament injury: panther symposium ACL injury clinical outcomes consensus group. *Knee Surg Sports Traumatol Arthrosc* 2020;28:2415-2434.
4. 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;6:2325967118774509.
5. Cheung EC, Moore LK, Flores SE, Lansdown DA, Feeley BT, Zhang AL. Correlation of PROMIS with orthopaedic patient-reported outcome measures. *JBJS Rev* 2019;7:e9.
6. Fisk F, Franovic S, Tramer J, et al. PROMIS CAT forms demonstrate responsiveness in patients following arthroscopic rotator cuff repair across numerous health domains. *J Shoulder Elbow Surg* 2019;28:2427-2432.
7. Cremers T, Zoulfi Khatiri M, van Maren K, Ring D, Teunis T, Fatehi A. Moderators and mediators of activity intolerance related to pain. *J Bone Joint Surg Am* 2021;103:205-212.
8. Rojas EO, Glass N, Owens J, et al. Performance of the PROMIS in patients undergoing 3 common elbow procedures. *Orthop J Sports Med* 2019;7:2325967119852595.
9. Kortlever JTP, Leyton-Mange A, Keulen MHF, et al. PROMIS physical function correlates with KOOS, JR in patients with knee pain. *J Knee Surg* 2020;33:903-911.
10. Anthony CA, Glass NA, Hancock KJ, Bollier M, Hettrich CM, Wolf BR. Preoperative performance of the patient-reported outcomes measurement information system in patients with rotator cuff pathology. *Arthroscopy* 2017;33:1770-1774.e1.
11. Anthony CA, Glass NA, Hancock KJ, Bollier M, Wolf BR, Hettrich CM. Performance of PROMIS instruments in patients with shoulder instability. *Am J Sports Med* 2017;45:449-453.
12. Hancock KJ, Glass NA, Anthony CA, et al. Performance of PROMIS for healthy patients undergoing meniscal surgery. *J Bone Joint Surg Am* 2017;99:954-958.
13. Hancock KJ, Glass N, Anthony CA, Wolf BR, et al. PROMIS: a valid and efficient outcomes instrument for patients with ACL tears. *Knee Surg Sports Traumatol Arthrosc* 2019;27:100-104.
14. 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;27:515-522.
15. Kaat AJ, Buckenmaier CT 3rd, Cook KF, et al. The expansion and validation of a new upper extremity item bank for the Patient-Reported Outcomes Measurement Information System (PROMIS). *J Patient Rep Outcomes* 2019;3:69.
16. Trasolini NA, Korber S, Gipsman A, San AE, Weber AE, Hatch GF 3rd. Performance of PROMIS computer adaptive testing as compared with established instruments for multiple-ligament knee injuries. *Orthop J Sports Med* 2019;7:2325967119867419.
17. Rothrock NE, Kaat AJ, Vrahas MS, et al. Validation of PROMIS physical function instruments in patients with an orthopaedic trauma to a lower extremity. *J Orthop Trauma* 2019;33:377-383.
18. Schober P, Boer C, Schwarte LA. Correlation coefficients: appropriate use and interpretation. *Anesth Analg* 2018;126:1763-1768.
19. MOON Knee Group. Ten-year outcomes and risk factors after anterior cruciate ligament reconstruction: A MOON longitudinal prospective cohort study. *Am J Sports Med* 2018;46:815-825.