Defining Clinically Significant Outcome Thresholds for the Patient-Reported Outcomes Measurement Information System (PROMIS) at 2 Years After Gluteus Medius and/or Minimus Repair

Ron Gilat,*† MD, Michael J. Vogel,*‡ BS, Omair Kazi,* MS, and Shane J. Nho,* MD, MS
Investigation performed at Rush University Medical Center, Chicago, Illinois, USA

Background: Clinically significant outcome (CSO) thresholds are invaluable to the interpretation of patient-reported outcomes (PROs). The Patient-Reported Outcomes Measurement Information System (PROMIS) is gaining popularity among the orthopaedic community; however, CSO thresholds for PROMIS are yet to be defined for outcomes after gluteus medius and/or minimus (GM) repair.

Purpose: To (1) define CSO thresholds for PROMIS–Pain Interference (PROMIS-PI) and PROMIS–Physical Function (PROMIS-PF) after GM repair, (2) correlate these PROMIS scores with legacy hip-specific PROs, and (3) quantify their floor and ceiling effects.

Study Design: Case series; Level of evidence, 4.

Methods: Consecutive patients who underwent primary GM repair between September 2017 and June 2021 with completed PROMIS at minimum 2-year follow-up were evaluated. The minimal clinically important difference, Patient Acceptable Symptom State, and substantial clinical benefit thresholds were defined for the PROMIS-PI and PROMIS-PF as well as for legacy PROs: Hip Outcome Score–Activities of Daily Living (HOS-ADL) and Hip Outcome Score–Sports Subscale (HOS-SS); modified Harris Hip Score (mHHS); 12-item International Hip Outcome Tool (iHOT-12); and the visual analog scale (VAS) for pain and satisfaction. Pearson correlations were performed between PROMIS scores and legacy PROs. Rates of floor and ceiling effects were quantified.

Results: Overall, 107 patients (81.7% follow-up compliance; mean age, 59.8 ± 8.8 years; 92.5% female; mean body mass index, 28.6 ± 6.3 kg/m^2) were included in the analysis. GM tears were partial thickness in 56.1% of cases and treated endoscopically in 64.5% of cases. The minimal clinically important difference, Patient Acceptable Symptom State, and substantial clinical benefit thresholds, respectively, were as follows: PROMIS-PI (–4.6, 56.0, 52.6), PROMIS-PF (3.5, 42.7, 43.7), HOS-ADL (10.7, 68.2, 78.6), HOS-SS (16.5, 58.6, 60.6), mHHS (9.0, 64.3, 71.5), iHOT-12 (14.2, 63.6, 69.4), VAS pain (–16.1, 34.9, 28.1), and VAS satisfaction (not applicable, 70.9, 93.6). Moderate to strong correlations were observed between the PROMIS and legacy PROs. The PROMIS-PI showed a significant postoperative floor effect of 18.7%.

Conclusion: Study findings indicated that the PROMIS is effective for use in GM repair patients, given the moderate-to-strong correlations between the PROMIS and legacy hip-specific PROs, the mostly limited floor and ceiling effects, and large effect sizes. Use of PROMIS instead of legacy PROs may aid in limiting survey burden.

Keywords: PROMIS; gluteal repair; patient-reported outcomes

The gluteus medius and minimus muscles act as the hip abductors and are instrumental in stabilizing the pelvis during ambulation and ensuring correct hip orientation.[38] Gluteus medius and/or minimus (GM) tears are a relatively common phenomenon that may manifest as lateral hip pain, hip abduction weakness, Trendelenburg gait, and significant dysfunction.[39] GM tears are often seen in middle-aged women and have been regarded as the

The Orthopaedic Journal of Sports Medicine, 12(11), 23259671241281746
DOI: 10.1177/23259671241281746

"rotator cuff of the hip" due to several similarities in symptoms as well as treatment approaches.[39] GM tears can be treated conservatively with anti-inflammatory medications, physical therapy, and corticosteroid injections in the chronic setting, but in some cases in which conservative therapy had failed or in the more acute setting, surgical repair is indicated.[29]

Surgical repair of GM tears has produced statistically significant improvement in patient-reported outcomes (PROs) after the procedure.[7,11,33] Recent advancement in our understanding of clinically significant outcomes (CSOs) has shifted the focus of orthopaedic literature to establish CSO thresholds such as the minimal clinically important difference (MCID), Patient Acceptable Symptom State (PASS), and substantial clinical benefit (SCB). Therefore, subsequent studies have focused on establishing CSO thresholds for legacy hip-specific PROs such as the Hip Outcome Score–Activities of Daily Living (HOS-ADL) and Hip Outcome Score–Sports Subscale (HOS-SS), the modified Harris Hip Score (mHHS), and the 12-item International Hip Outcome Tool (iHOT-12).[29,32,36]

In 2004, the US National Institutes of Health developed the Patient-Reported Outcomes Measurement Information System (PROMIS).[6] PROMIS utilizes integrated item-response theory with computerized adaptive testing (CAT) to allow for a single, generalizable, and validated PRO. PROMIS is also a standardized tool designed to minimize floor or ceiling effects, with fewer individual questions when compared with legacy PROs to minimize questionnaire fatigue.[5] A PROMIS score of 50 is intended to represent the mean score of a "typical population."[13] A higher score on PROMIS–Pain Interference (PROMIS-PI) denotes a higher level of pain burden, while a higher score on PROMIS–Physical Function (PROMIS-PF) denotes a higher level of functioning. The PROMIS questionnaires' correlation with legacy PROs and CSO thresholds of PROMIS have been established for several pathologies, including hip arthroscopy for femoroacetabular impingement syndrome (FAIS).[4,8] However, to our knowledge, no study has reported on PROMIS questionnaires' correlation with legacy questionnaires and minimum 2-year CSO thresholds of PROMIS after GM repair.

The purpose of this study was to (1)define MCID, PASS, and SCB thresholds for the PROMIS-PI and PROMIS-PF after primary GM repair at minimum 2-year follow-up; (2)correlate PROMIS scores with those of the legacy hip-specific PROs; and (3) quantify the floor and ceiling effects of the included PROs. We hypothesized success in defining CSO thresholds for the PROMIS-PI and PROMIS-PF, with similar CSO achievement between the PROMIS and legacy PROs. We also hypothesized there would be moderate to strong correlations between the PROMIS and legacy PROs, with PROMIS showing lower rates of floor and ceiling effects.

METHODS

Patient Selection

Institutional review board approval was obtained before initiating the study. Consecutive patients who underwent primary GM repair between September 2017 and June 2021 by the senior author (S.J.N.) were retrospectively identified in the senior author's prospectively maintained clinical repository. Patients eligible for inclusion were those who underwent primary GM repair with or without concomitant intra-articular procedures with completion of PROMIS at minimum 2-year follow-up. Excluded patients were those who underwent revision GM repair, gluteus maximus transfer, or allograft augmented repair as well as those with a history of prior ipsilateral hip arthroscopy, hip resurfacing, or total hip arthroplasty.

Demographic Characteristics

The clinical charts of the included patients were evaluated for demographic information, including age at the time of surgery, sex, body mass index (BMI), endorsement of regular physical activity, endorsement of preoperative lower back pain, history of a psychiatric diagnosis, duration of preoperative hip pain, current or former tobacco use, and workers' compensation insurance status.

Procedures Performed

All surgical procedures were performed by the senior author (S.J.N.). His general algorithm was to use open repair for cases of full-thickness GM tears and cases of partial-thickness GM tears involving the anterior and superoposterior facets of the greater trochanter, as the open approach improved visualization and access to these tears. Otherwise, for patients with partial-thickness GM tears not involving the anterior and superoposterior facets and with magnetic resonance imaging evidence of a labral tear, he preferred an endoscopic approach for the added benefit of first accessing the central compartment of the

‡Address correspondence to Michael J. Vogel, BS, Rush University Medical Center, 1611 W Harrison Street, Chicago, IL 60612, USA (email: nho.research@rushortho.com).

*Section of Young Adult Hip Surgery, Division of Sports Medicine, Department of Orthopaedic Surgery, Rush Medical College of Rush University, Rush University Medical Center, Chicago, Illinois, USA.

†Department of Orthopaedic Surgery, Shamir Medical Center and Tel Aviv University, Tel Aviv, Israel.

Final revision submitted March 25, 2024; accepted April 5, 2024.

One or more of the authors has declared the following potential conflict of interest or source of funding: S.J.N. has received research support from Mitek and Stryker; consulting fees from SI-Bone and Stryker; and royalties from Ossur, Springer, and Stryker. AOSSM checks author disclosures against the Open Payments Database (OPD). AOSSM has not conducted an independent investigation on the OPD and disclaims any liability or responsibility relating thereto.

Ethical approval for this study was obtained from Rush University Medical Center (ref No. 23092101-IRB01).

hip to perform labral debridement or repair as indicated by the morphology of the labrum. Labral debridement was performed for degenerative tears and uniplanar tears involving <50% of the labrum, while labral repair was performed for complex tears with healthy labral tissue and disruption of the chondrolabral junction.[20,24]

For the study patients, the operative extremity was confirmed and marked in the preoperative area. All patients received general endotracheal anesthesia. Open repairs were performed through a direct lateral approach. GM tears were identified after blunt and sharp dissection. Trochanteric bursectomy was performed. The greater trochanter was decorticated with a burr before single- or double-row suture anchor repair.[31,36] In endoscopic repairs, the hip joint was first visualized and pathology addressed accordingly, including labral debridement or repair based on the labral morphology described above. When present, chondral defects were debrided to a stable edge. Attention was then turned to the GM tear. The GM tear was identified and trochanteric bursectomy performed. Decortication of the greater trochanter was performed before single- or double-row repair.[22,32] Standard postoperative rehabilitation protocols were used for all included patients.[22,29] Briefly, phase 1 included 6 weeks of limited weightbearing with assisted ambulation, a hip abduction orthosis limiting hip flexion to 90° during the day, and an abduction pillow at night. Phase 2 included weaning from assisted ambulation and progression to a noncompensatory gait. After gait stabilization, phase 3 included further hip strengthening to support return to activities of daily living and patient-specific functioning.[22,32]

Patient-Reported Outcomes

PROs were collected preoperatively and at minimum 2-year postoperative follow-up after surgery and included the PROMIS-PI and PROMIS-PF as well as the HOS-ADL and HOS-SS, the mHHS, the iHOT-12, and the visual analog scale (VAS) for pain. A higher score on PROMIS-PI and VAS pain denoted a higher level of pain, while a higher score on PROMIS-PF, HOS-ADL, HOS-SS, mHHS, and iHOT-12 denoted a higher level of functioning. In addition, the VAS for satisfaction was collected postoperatively.

CSO Thresholds

Cohort-specific MCID, PASS, and SCB were defined for PROMIS-PI, PROMIS-PF, HOS-ADL, HOS-SS, mHHS, iHOT-12, VAS pain, and VAS satisfaction accordingly to previously described methods.[5,18,25,27,30,37]

The MCID was calculated through the distribution method as previously described, whereas half the standard deviation of the difference between the preoperative and postoperative scores defined the MCID.[5,18,25,30]

The PASS was determined through the anchor question method, where at the 2-year follow-up, patients were asked the dichotomous question, "Taking into consideration your current state of pain and function, do you consider you current state satisfactory?"[5,26,29] Receiver operating characteristic curve analysis was then performed to determine the PRO score threshold with the maximal sensitivity and specificity for differentiating patients that were and were not in a satisfactory state.[15,23] An area under the curve (AUC) of >0.70 was deemed clinically significant.[23] The Youden J value was employed to define the optimal PASS threshold as previously described.[17,23]

SCB was also defined through the anchor question method. At the 2-year follow-up, patients were asked, "Since your surgery, has there been any change in your daily functioning and nonsporting activities related to your treated hip?"[5] Patients responded with 1 of 15 options: (1) *a very great deal worse*, (2) *a great deal worse*, (3) *a good deal worse*, (4) *moderately worse*, (5) *a little worse*, (6) *somewhat worse*, (7) *almost the same/hardly any worse*, (8) *no change*, (9) *almost the same/hardly any better*, (10) *somewhat better*, (11) *a little better*, (12) *moderately better*, (13) *a good deal better*, (14) *a great deal better*, or (15) *a very great deal better*.[5] Receiver operating characteristic curve analysis was then used to compare PRO scores between those patients selecting options 12 to 15 and those selecting options 7 to 9, as previously described.[5] An AUC of >0.70 was deemed clinically significant.[23] The Youden J value was used to determine the optimal threshold for the SCB.[17,23]

Achievement of MCID, PASS, and SCB for the included PROs was quantified with their respective thresholds. Achievement of MCID, PASS, or SCB for any legacy PRO was recorded as well as achievement of MCID, PASS, or SCB for either PROMIS score and achievement of MCID, PASS, or SCB for any PRO.

Pearson Correlations

Pearson correlations were performed to compare minimum 2-year postoperative scores between the included PROs.[10] The significance of the correlations was recorded. The strength of the correlations was interpreted as follows: none (r = 0.00–0.29), weak (r = 0.30–0.49), moderate (r = 0.50–0.69), and strong (r = 0.70–1.00).[27]

Floor and Ceiling Effects

The prevalence of preoperative and postoperative floor and ceiling effects was evaluated for the included PROs according to previously described methods.[2,14,34] For PROMIS-PI and PROMIS-PF, any percentage of patients ≥15% of the study population scoring within either the bottom 5% or the top 5% of the study sample was deemed a significant floor and ceiling effect, respectively.[2,34] For the legacy PROs, any percentage of patients ≥15% of the study population scoring within either the bottom 5% or top 5% of possible scores was deemed a significant floor and ceiling effect, respectively.[2,34]

Effect Sizes

The effect size (Cohen d) of the included PROs was evaluated to understand the magnitude of preoperative to 2-year

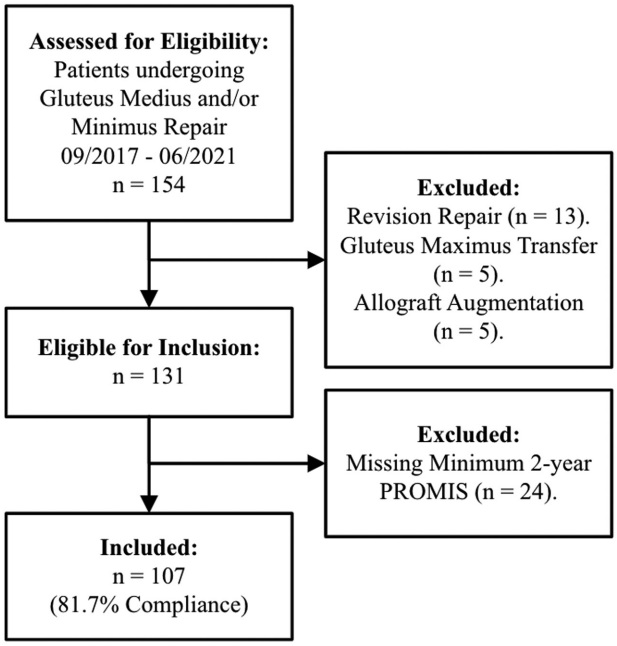

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) diagram showing the inclusion and exclusion of study participants. PROMIS, Patient-Reported Outcome Measurement Information System.

postoperative score changes in relation to the standard deviation for that PRO. The effect size was calculated by dividing the difference in mean scores by the standard deviation of the respective preoperative score.[35] Effect size was interpreted as small (0.20–0.49), moderate (0.50–0.79), or large (≥0.80).[9,12]

Statistical Analysis

Continuous variables were reported as means with standard deviations, and nominal variables were reported as counts with frequencies. Paired *t* tests were performed to evaluate for statistically significant improvements in PRO scores from preoperative to 2-year follow-up. A subgroup analysis of 2-year PRO scores was performed between the patients treated with endoscopic versus open repair. An a priori power analysis for the subgroup analysis revealed that 29 patients would be required in each group to achieve 80% power.[16] All statistical analysis was performed in R 4.1.0 (R Core Team).

RESULTS

A total of 154 GM repairs were performed between September 2017 and June 2021. Patient exclusions were as follows: 13 for revision GM repair, 5 for gluteus maximus transfer, and 5 for allograft augmentation. Of the 131 patients eligible for inclusion, 24 were missing minimum 2-year PROMIS scores and were excluded as such, leaving 107 patients meeting inclusion and exclusion criteria (81.7% compliance rate) (Figure 1).

TABLE 1
Patient Demographic Characteristics (N = 107)[a]

Characteristic	Value
Age, y	59.8 ± 8.8
Sex, female	99 (92.5)
BMI, kg/m^2	28.6 ± 6.3
Follow-up duration, y	3.1 ± 1.4
Regular physical activity	50 (46.7)
Preoperative back pain	49 (45.8)
Psychiatric history	24 (22.4)
Preoperative pain duration ≥2 y	47 (43.9)
Tobacco use, current or former	26 (24.3)
Workers' compensation	4 (3.7)

[a]Data are presented as mean ± SD or n (%). BMI, body mass index.

TABLE 2
Gluteal Tear Characteristics and Procedures Performed[a] (N = 107)

Characteristic	n (%)
Tear thickness	
Full	47 (43.9)
Partial	60 (56.1)
Tendons involved	
Gluteus medius	13 (12.1)
Gluteus minimus	4 (3.7)
Both	90 (84.1)
Repair approach	
Open	38 (35.5)
Endoscopic	69 (64.5)
Labral treatment	
None	48 (44.9)
Debridement	41 (38.3)
Repair	18 (16.8)

The 107 included patients were 92.5% female and were a mean ± SD age 59.8 ± 8.8 years and BMI 28.6 ± 6.3 kg/m^2 (Table 1). The mean follow-up duration was 3.1 ± 1.4 years. Less than half the patients endorsed regular physical activity (46.7%), preoperative backpain (45.8%), psychiatric history (22.4%), preoperative hip pain duration ≥2 years (43.9%), current or former tobacco use (24.3%), and workers' compensation status (3.7%).

Intraoperatively, tears involving both the gluteus medius and gluteus minimus tendons occurred in 84.1% of patients. Tears were partial thickness in 56.1% of cases. Most patients underwent endoscopic repair (64.5%), and most patients underwent no labral treatment (44.9%), while some underwent labral debridement (38.3%) and fewer underwent labral repair (16.8%) (Table 2).

Preoperative and 2-year postoperative PRO scores are presented in Figure 2. Preoperative scores were as follows: PROMIS-PI (63.3 ± 5.3), PROMIS-PF (38.3 ± 5.6), HOS-ADL (51.5 ± 15.7), HOS-SS (28.0 ± 19.2), mHHS (51.5 ±

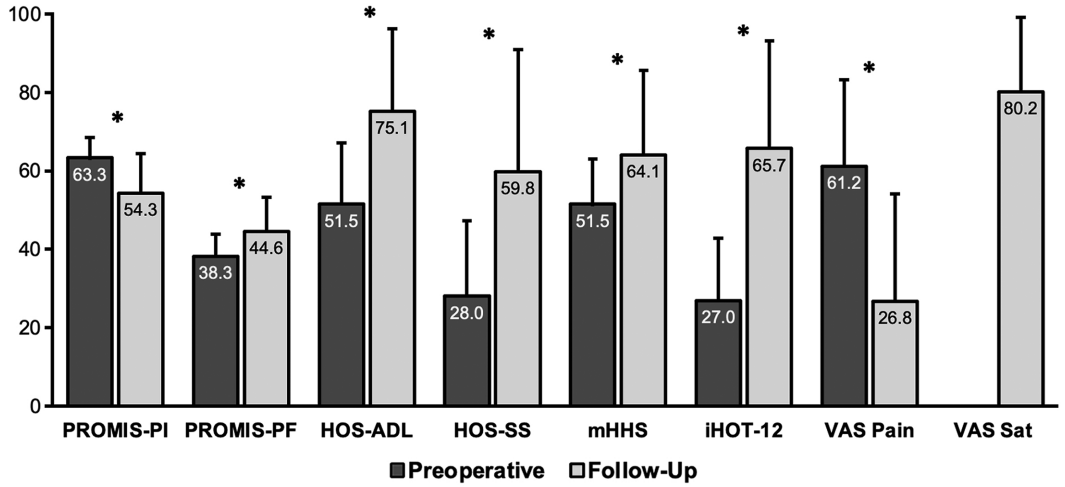

Figure 2. Preoperative and minimum 2-year postoperative follow-up patient-reported outcome scores. Bars indicate means, and error bars indicate SDs. *Statistically significant difference between preoperative and 2-year follow-up scores ($P < .001$). HOS-ADL, Hip Outcome Score–Activities of Daily Living; HOS-SS, Hip Outcome Score–Sports Subscale; iHOT-12, 12-item International Hip Outcome Tool; mHHS, modified Harris Hip Score; PROMIS-PF, Patient-Reported Outcome Measurement Information System–Physical Function; PROMIS-PI, Patient-Reported Outcome Measurement Information System–Pain Interference; Sat, satisfaction; VAS, visual analog scale.

11.5), iHOT-12 (27.0 ± 15.9), and VAS pain (61.2 ± 22.1). Postoperative scores were as follows: PROMIS-PI (54.3 ± 10.1), PROMIS-PF (44.6 ± 8.8), HOS-ADL (75.1 ± 21.2), HOS-SS (59.8 ± 31.1), mHHS (64.1 ± 21.5), iHOT-12 (65.7 ± 27.5), VAS pain (26.8 ± 27.3), and VAS satisfaction (80.2 ± 19.0).

The subgroup analysis comparing 2-year PRO scores of the 69 patients who underwent endoscopic repair and the 38 patients who underwent open repair demonstrated no significant group differences ($P \geq .167$ for all). Comparisons of specific scores are as follows (endoscopic vs open repair): HOS-ADL (73.4 ± 23.3 vs 78.8 ± 15.6; $P = .213$), HOS-SS (59.7 ± 33.6 vs 60.1 ± 25.8; $P = .956$), mHHS (62.6 ± 22.6 vs 69.1 ± 17.2; $P = .345$), iHOT-12 (63.7 ± 29.1 vs 69.2 ± 17.2; $P = .336$), VAS pain (28.6 ± 27.6 vs 23.6 ± 25.2; $P = .393$), VAS satisfaction (76.9 ± 30.1 vs 85.1 ± 17.2; $P = .167$), PROMIS-PI (54.5 ± 10.8 vs 53.8 ± 9.0; $P = .737$), and PROMIS-PF (44.8 ± 9.9 vs 44.2 ± 6.4; $P = .748$).

MCID, PASS, and SCB threshold scores are listed in Table 3. Clinically significant AUCs (>0.70) were observed for all PASS and SCB thresholds.

The achievement rates for the MCID, PASS, and SCB are shown in Table 4. Patients achieved MCID for any legacy hip-specific PRO in 98% of cases compared with 86% for any PROMIS. Patients achieved PASS for any legacy hip-specific PRO in 87% of cases compared with 63% for any PROMIS. Patients achieved SCB for any legacy hip-specific PRO in 84% of cases compared with 53% for any PROMIS.

Moderate to strong correlations were observed between the PROMIS and legacy hip-specific PROs (Table 5). The PROMIS-PI showed a strong correlation with VAS pain ($r = 0.702$) and strong inverse correlations with the iHOT-12 ($r = -0.733$) and PROMIS-PF ($r = -0.736$). The PROMIS-PF showed a strong correlation with the HOS-SS ($r = 0.733$).

A significant preoperative floor effect was observed for the HOS-SS (22.9%) (Table 6). No significant preoperative ceiling effects were observed for any PRO. At 2-year follow-up, significant floor effects were observed for the PROMIS-PI (18.7%) and VAS pain (27.9%), while significant ceiling effects were observed for the HOS-ADL (21.4%), HOS-SS (17.9%), and VAS satisfaction (45.2%).

Large effect sizes (>0.80) were seen on both included PROMIS CATs and all legacy hip-specific PROs (Table 7). The iHOT-12 showed the greatest effect size of 2.4, while the PROMIS-PF and mHHS showed the smallest effect size of 1.1. Nevertheless, all measured PROs showed large effect sizes appropriate for distinguishing preoperative to postoperative change.

DISCUSSION

The primary findings of this study were as follows: (1) The MCID, PASS, and SCB thresholds were defined for PROMIS-PI and PROMIS-PF at minimum 2-year follow-up after primary GM repair; (2) moderate to strong correlations were observed between legacy hip-specific PROs and the included PROMIS CATs; and (3) an overall low prevalence of floor and ceiling effects was observed for the PROMIS CATs, except for PROMIS-PI showing a significant postoperative floor effect.

Previous studies have established CSO thresholds and assessed their achievement rate after GM repair. Uppstrom et al[36] retrospectively evaluated 47 patients who underwent open GM repair with minimum 10-month

TABLE 3
Minimum 2-Year Postoperative Clinically Significant Outcome Threshold Scores[a]

	MCID	PASS		SCB	
Outcome Measure	Threshold	Threshold	AUC[b]	Threshold	AUC[b]
HOS-ADL	10.7	68.2	0.950	78.6	0.856
HOS-SS	16.5	58.6	0.956	60.6	0.763
mHHS	9.0	64.3	0.901	71.5	0.741
iHOT-12	14.2	63.6	0.942	69.4	0.844
VAS pain	–16.1	34.9	0.920	28.1	0.721
VAS satisfaction	NA	70.9	0.951	93.6	0.826
PROMIS-PI	–4.6	56.0	0.921	52.6	0.836
PROMIS-PF	3.5	42.7	0.881	43.7	0.750

[a]AUC, area under the curve; HOS-ADL, Hip Outcome Score–Activities of Daily Living; HOS-SS, Hip Outcome Score–Sports Subscale; iHOT-12, 12-item International Hip Outcome Tool; MCID, minimal clinically important difference; mHHS, modified Harris Hip Score; NA, not applicable; PASS, Patient Acceptable Symptom State; PROMIS-PF, Patient-Reported Outcome Measurement Information System–Physical Function; PROMIS-PI, Patient-Reported Outcome Measurement Information System–Pain Interference; SCB, substantial clinical benefit; VAS, visual analog scale.
[b]All AUC values were considered clinically significant (>0.70).[23]

TABLE 4
Achievement of Clinically Significant Outcomes[a]

Outcome Measure	Achieved MCID	Achieved PASS	Achieved SCB
PROMIS-PI	72	55	39
PROMIS-PF	74	53	47
HOS-ADL	76	66	52
HOS-SS	62	55	28
mHHS	57	50	43
iHOT-12	82	59	52
VAS pain	75	69	65
VAS satisfaction	NA	68	45
Any legacy PRO	98	87	84
Any PROMIS	86	63	53
Any PRO	100	88	86

[a]Data are presented as percentage of patients. HOS-ADL, Hip Outcome Score–Activities of Daily Living; HOS-SS, Hip Outcome Score–Sports Subscale; iHOT-12, 12-item International Hip Outcome Tool; MCID, minimal clinically important difference; mHHS, modified Harris Hip Score; NA, not applicable; PASS, Patient Acceptable Symptom State; PRO, patient-reported outcome; PROMIS-PF, Patient-Reported Outcome Measurement Information System–Physical Function; PROMIS-PI, Patient-Reported Outcome Measurement Information System–Pain Interference; SCB, substantial clinical benefit; VAS, visual analog scale.

follow-up. They found the MCIDs of mHHS and 33-item iHOT to be 9.9 and 14.3, respectively, and reported a >80% achievement rate for both outcome measures. Okoroha et al[29] evaluated 60 patients after isolated endoscopic GM repair with minimum 2-year follow-up. They found the MCID threshold scores of HOS-ADL, HOS-SS, and mHHS to be 15.02, 14.53, and 14.13, respectively, and reported PASS threshold scores for HOS-ADL, HOS-SS, and mHHS to be 81.32, 67.71, and 77.5, respectively. They found that 76.7% of patients achieved any MCID or PASS, and 77.8% and 69.0% of patients achieved ≥1 MCID or PASS, respectively. Further, 48.3% of patients achieved both MCID and PASS.[29] Allahabadi et al,[1] in a case series of 174 patients, including 41 treated through an open approach and 133 treated endoscopically, defined the following MCID thresholds: HOS-ADL (10.93), HOS-SS (15.28), mHHS (10.52), iHOT-12 (13.34), and VAS pain (15.49).[1] In the present study, we found MCID thresholds of HOS-ADL (10.7), HOS-SS (16.5), mHHS (9.0), iHOT-12 (14.2), VAS pain (–16.1), PROMIS-PI (–4.6), and PROMIS-PF (3.5). Further, we found PASS thresholds of HOS-ADL (68.2), HOS-SS (58.6), mHHS (64.3), iHOT-12 (63.6), VAS pain (34.9), VAS satisfaction (70.9), PROMIS-PI (56.0), and PROMIS-PF (42.7).

In large part, our MCID thresholds were comparable with those of Allahabadi et al,[1] Uppstrom et al,[36] and Okoroha et al.[29] We found lower HOS-ADL and mHHS MCID thresholds compared with Okoroha et al,[29] likely secondary to variations in the calculation of MCID. Whereas the present study denoted MCID as half the standard deviation of the difference between preoperative and postoperative scores, as previously described,[5] Okoroha et al[29] identified MCID as half the standard deviation of the outcome score. Furthermore, we found lower PASS thresholds to those of Okoroha et al[29] and Rice et al,[32] with comparable CSO achievements. While the methodology for PASS calculation was similar between the present study and those of Okoroha et al[29] and Rice et al,[32] we included both endoscopic and open repair patients, while the other 2 studies were restricted to only endoscopic repair.[1,36] Despite the subgroup comparison in the present study demonstrating no significant difference in 2-year postoperative PRO scores between the open and endoscopic groups, patients indicated for their respective procedures may have shown differing interpretations of acceptable postoperative symptom state.

TABLE 5
Pearson Correlation Coefficients for 2-Year Minimum Follow-up Scores[a]

	PROMIS-PI	PROMIS-PF	VAS Satisfaction	VAS Pain	iHOT-12	mHHS	HOS-SS
HOS-ADL	–0.656	0.676	0.666	–0.675	0.768	0.653	0.839
HOS-SS	–0.663	0.733	0.696	–0.658	0.751	0.554	—
mHHS	–0.562	0.539	0.538	–0.644	0.606	—	—
iHOT-12	–0.733	0.673	0.631	–0.781	—	—	—
VAS pain	0.702	–0.628	–0.611	—	—	—	—
VAS satisfaction	–0.487	0.538	—	—	—	—	—
PROMIS-PF	–0.736	—	—	—	—	—	—

[a]Data are presented as Pearson *r*s. All correlations were significant at $P < .001$. HOS-ADL, Hip Outcome Score–Activities of Daily Living; HOS-SS, Hip Outcome Score–Sports Subscale; iHOT-12, 12-item International Hip Outcome Tool; mHHS, modified Harris Hip Score; PRO, patient-reported outcome; PROMIS-PF, Patient-Reported Outcome Measurement Information System–Physical Function; PROMIS-PI, Patient-Reported Outcome Measurement Information System–Pain Interference; VAS, visual analog scale.

TABLE 6
Prevalence of Floor and Ceiling Effects for Each PRO[a]

	Preoperative		2-Year Follow-up	
Outcome Measure	Floor	Ceiling	Floor	Ceiling
PROMIS-PI	3.5	1.8	18.7[b]	7.7
PROMIS-PF	1.5	1.5	1.1	1.1
HOS-ADL	0.0	0.0	0.0	21.4[b]
HOS-SS	22.9[b]	0.0	6.4	17.9[b]
mHHS	0.0	0.0	0.0	4.5
iHOT-12	9.0	0.0	1.1	8.6
VAS pain	0.0	6.1	27.9[b]	2.3
VAS satisfaction	NA	NA	0.0	45.2[b]

[a]Data are presented as percentage of patients. HOS-ADL, Hip Outcome Score–Activities of Daily Living; HOS-SS, Hip Outcome Score–Sports Subscale; iHOT-12, 12-item International Hip Outcome Tool; mHHS, modified Harris Hip Score; NA, not applicable; PRO, patient-reported outcome; PROMIS-PF, Patient-Reported Outcome Measurement Information System–Physical Function; PROMIS-PI, Patient-Reported Outcome Measurement Information System–Pain Interference; VAS, visual analog scale.
[b]Significant floor and ceiling effects (≥15% of the included patients).

TABLE 7
Mean Preoperative to Postoperative Differences in PRO Scores With Effect Sizes[a]

Outcome Measure	Δ preoperative to 2-year score	Cohen *d*[b]
PROMIS-PI	–9.0	1.7
PROMIS-PF	6.3	1.1
HOS-ADL	23.6	1.5
HOS-SS	31.8	1.7
mHHS	12.6	1.1
iHOT-12	38.7	2.4
VAS pain	–34.4	1.6

[a]HOS-ADL, Hip Outcome Score–Activities of Daily Living; HOS-SS, Hip Outcome Score–Sports Subscale; iHOT-12, 12-item International Hip Outcome Tool; mHHS, modified Harris Hip Score; PRO, patient-reported outcome; PROMIS-PF, Patient-Reported Outcome Measurement Information System–Physical Function; PROMIS-PI, Patient-Reported Outcome Measurement Information System–Pain Interference; VAS, visual analog scale.
[b]All effect sizes were considered large (≥0.80).[9,12]

The present study further expands on the literature through the addition of 2-year postoperative SCB thresholds after gluteal repairs. The SCB thresholds were defined as follows: HOS-ADL (78.6), HOS-SS (60.6), mHHS (71.5), iHOT-12 (69.4), VAS pain (28.1), VAS satisfaction (93.6), PROMIS-PI (52.6), and PROMIS-PF (43.7). With regard to a PROMIS-PI and PROMIS-PF score of 50 intended to represent the mean score of an average population,[13] it is interesting to note that, while our case series demonstrated a mean greater pain interference (>50 for 2-year PROMIS-PI; 54.3 ± 10.1) and a mean lower function (<50 for 2-year PROMIS-PF; 44.6 ± 8.8) than the average population, our calculated cohort-specific PASS threshold scores (PROMIS-PI, 56.0; PROMIS-PF, 42.7) and SCB threshold scores (PROMIS-PI, 52.6; PROMIS-PF, 43.7) were also inferior to the average population score of 50. This finding indicates that the primary GM repair patients observed in this study were often satisfied with their postoperative state of pain and function, despite their levels of pain and function being worse than the reference population used to establish normalized PROMIS scores.

We found moderate to strong correlations between the PROMIS and included legacy hip-specific PROs. In the context of hip arthroscopy literature showing correlations between PROMIS CATs and legacy hip-specific PROs ranging from poor to moderate to strong,[3,4,19,28] the correlations evidenced in this study support the efficacy of PROMIS CATs in evaluating patient outcomes after GM repair.

PROMIS effect size and responsiveness compared with legacy questionnaires have been assessed in prior literature on hip preservation procedures. A recent systematic review by Jan et al[19] reported on 15 studies assessing outcomes after hip arthroscopy for FAIS and/or periacetabular osteotomy. They found that the PROMIS was not as responsive when compared with historically used and validated PROs in quantitatively assessing function and pain after hip preservation. On the contrary, Nwachukwu et al[28] evaluated the effect size of PROMIS CATs and legacy hip-specific PROs after hip arthroscopy and demonstrated that, while HOS-ADL (1.29), HOS-SS (1.06), and iHOT-12 (1.87) showed qualitatively larger effect sizes that PROMIS-PI (0.96) and PROMIS-PF (0.94), all effect sizes were considered large (>0.80). Our present study expands these findings to the primary GM repair population by showing large and comparable effect sizes for both PROMIS CATs and legacy hip-specific PROs.

PROMIS tools were designed to mitigate the floor and ceiling effects naturally found in legacy PROs.[5] However, this is not always the case, as the present study showed significant postoperative floor effects for PROMIS-PI. While the legacy PROs demonstrated some significant floor and ceiling effects, PROMIS-PI showed the second-highest floor effect, behind VAS pain. In contrast to prior studies on PROMIS CAT utility in hip arthroscopy for FAIS patients,[4,19] our study showed a greater prevalence of the postoperative floor effect for PROMIS-PI. Nevertheless, the postoperative floor effect of PROMIS-PI remained lower than VAS pain, indicating that the PROMIS CATs remained effective in limiting the floor and ceiling effects commonly associated with legacy hip-specific PROs.

Limitations

The limitations of the present study include its retrospective design, heterogenicity in tear pathology and procedures performed, and potentially limited external validity. Despite the prospective collection of PROs, this study's retrospective design is at risk for selection bias based on the patients available for follow-up, although an 81.7% compliance rate at 2-year follow-up helped mitigate some potential selection bias. Furthermore, patients in this study were grouped for analysis regardless of gluteal tendons involved, tear thickness, gluteal repair approach used, and labral tear treatment provided. We sought to provide a holistic view of gluteal repair outcomes and 2-year CSO thresholds, although this was done at the risk of underappreciating the diverse patient population evaluated. The nonsignificant difference in 2-year outcomes in our subgroup analysis comparing open and endoscopic gluteal repair supports the utility of combing these groups in defining 2-year PROMIS CSO thresholds, although future studies are needed to further evaluate demographic and outcome differences between these groups. While the inclusion of patients who underwent GM repair with concomitant labral procedures may have confounded our findings, a 2024 study[21] demonstrated that patients who underwent gluteal repair with versus without concomitant labral treatment achieved comparable 2-year outcomes, which may help mitigate some concern of this possible confounder, although the lack of randomization of patient groups in the study precludes further understanding of how labral preservation in the setting of GM repair influences outcomes. A further limitation of our study is that all surgical procedures were performed by a single sports medicine fellowship–trained surgeon who treats a high volume of gluteal tears. The patient outcomes evaluated in this study may not reflect those of lower volume surgeons, potentially limiting the external validity of the study findings.

CONCLUSION

In this study, we defined the MCID, PASS, and SCB thresholds for the PROMIS-PI and PROMIS-PF after primary GM repair at minimum 2-year follow-up. Moderate to strong correlations were shown between the PROMIS CATs and legacy hip-specific PROs, and a significant postoperative floor effect was evidenced for the PROMIS-PI. Findings from this study indicate that PROMIS CATs are effective for use in GM repair patients, given the moderate to strong correlations between the PROMIS CATs and the legacy hip-specific PROs, mostly limited floor and ceiling effects, and large effect sizes. Use of PROMIS CATs instead of legacy PROs may aid in limiting survey burden.

REFERENCES

1. Allahabadi S, Chapman RS, Fenn TW, Browning RB, Nho SJ. The Gluteus-Score-7 predicts the likelihood of both clinical success and failure following surgical repair of the hip gluteus medius and/or minimus. *Arthroscopy*. 2024;40(2):343-351.
2. Anthony CA, Glass NA, Hancock K, Bollier M, Wolf BR, Hettrich CM. Performance of PROMIS instruments in patients with shoulder instability. *Am J Sports Med*. 2017;45(2):449-453.
3. Bido J, Sullivan SW, Dooley MS, et al. PROMIS Global-10 poorly correlates with legacy outcomes for patients undergoing hip arthroscopy. *J Hip Preserv Surg*. 2021;8(1):67-74.
4. Bodendorfer BM, Clapp IM, Browning RB, Alter TD, Nwachukwu BU, Nho SJ. Patient-Reported Outcomes Measurement Information System test is less responsive than legacy hip-specific patient-reported outcome measures in patients undergoing arthroscopy for femoroacetabular impingement syndrome. *Arthrosc Sports Med Rehabil*. 2021;3(6):e1645-e1650.
5. Bodendorfer BM, DeFroda SF, Clapp IM, Newhouse A, Nwachukwu BU, Nho SJ. Defining clinically significant improvement on the Patient-Reported Outcomes Measurement Information System test at 1-year follow-up for patients undergoing hip arthroscopy for the treatment of femoroacetabular impingement syndrome. *Am J Sports Med*. 2021;49(9):2457-2465.

6. 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):S3-S11.
7. Chandrasekaran S, Gui C, Hutchinson MR, Lodhia P, Suarez-Ahedo C, Domb BG. Outcomes of endoscopic gluteus medius repair: study of thirty-four patients with minimum two-year follow-up. *J Bone Joint Surg Am*. 2015;97(16):1340-1347.
8. Clapp I, DeFroda S, Newhouse A, et al. Paper 26: defining clinically significant improvement on Patient-Reported Outcomes Measurement Information System test for patients undergoing hip arthroscopy for the treatment of femoroacetabular impingement syndrome at 1-year follow-up. *Orthop J Sports Med*. 2022;10(5)(suppl 3):2325967121S00564.
9. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Psychology Press; 2009.
10. Day MA, Hancock KJ, Selley RS, et al. Patient-Reported Outcomes Measurement Information System mobility computerized adaptive testing maintains high correlation and low test burden compared with legacy hip-specific instruments in patients undergoing hip arthroscopy for femoroacetabular impingement. *Arthroscopy*. 2022;38(11):3023-3029.
11. Domb BG, Botser I, Giordano BD. Outcomes of endoscopic gluteus medius repair with minimum 2-year follow-up. *Am J Sports Med*. 2013;41(5):988-997.
12. Dunlap WP, Cortina JM, Vaslow JB, Burke MJ. Meta-analysis of experiments with matched groups or repeated measures designs. *Psychol Methods*. 1996;1(2):170-177.
13. Franovic S, Gulledge CM, Kuhlmann NA, Williford TH, Chen C, Makhni EC. Establishing ''normal'' Patient-Reported Outcomes Measurement Information System Physical Function and Pain Interference scores. *JBJS Open Access*. 2019;4(4):e0019.
14. Gulledge CM, Smith DG, Ziedas A, Muh SJ, Moutzouros V, Makhni EC. Floor and ceiling effects, time to completion, and question burden of PROMIS CAT domains among shoulder and knee patients undergoing nonoperative and operative treatment. *JBJS Open Access*. 2019;4(4):e0015.1-7.
15. Hajian-Tilaki K. Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. *Caspian J Intern Med*. 2013;4(2):627-635.
16. Horner NS, Chapman RS, Larson JH, Nho SJ. Results of endoscopic labral repair with concomitant gluteus medius and/or minimus repair compared with outcomes of labral repair alone: a matched comparative cohort analysis at minimum 2-year follow-up. *Am J Sports Med*. 2023;51(7):1818-1825.
17. Hung M, Baumhauer JF, Licari FW, Voss MW, Bounsanga J, Saltzman CL. PROMIS and FAAM minimal clinically important differences in foot and ankle orthopedics. *Foot Ankle Int*. 2019;40(1):65-73.
18. Jaeschke R, Singer J, Guyatt GH. Measurement of health status: ascertaining the minimal clinically important difference. *Control Clin Trials*. 1989;10(4):407-415.
19. Jan K, Krivicich L, Hevesi M, Nho S. Patient-Reported Outcomes Measurement Information System is not as responsive as legacy scores in detecting patient outcomes in hip preservation: a systematic review. *Arthroscopy*. 2023;39(3):838-850.
20. Kucharik MP, Abraham PF, Nazal MR, et al. Arthroscopic acetabular labral repair versus labral debridement: long-term survivorship and functional outcomes. *Orthop J Sports Med*. 2022;10(7):232596 71221109012.
21. Larson JH, Brusalis CM, Allahabadi S, et al. Outcomes of isolated endoscopic gluteal tendon repair compared with concomitant endoscopic gluteal tendon repair and arthroscopic hip labral repair: a propensity-matched analysis with minimum 2-year follow-up. *Orthop J Sports Med*. 2024;12(2):23259671231215340.
22. Levy DM, Bogunovic L, Grzybowski JS, Kuhns BD, Bush-Joseph CA, Nho SJ. All-endoscopic single-row repair of full-thickness gluteus medius tears. *Arthrosc Tech*. 2016;5(1):e1-e6.
23. Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. *J Thorac Oncol*. 2010;5(9):1315-1316.
24. Menge TJ, Briggs KK, Dornan GJ, McNamara SC, Philippon MJ. Survivorship and outcomes 10 years following hip arthroscopy for femoroacetabular impingement: labral debridement compared with labral repair. *J Bone Joint Surg Am*. 2017;99(12):997-1004.
25. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med Care*. 2003;41(5):582-592.
26. Nwachukwu BU, Beck EC, Kunze KN, Chahla J, Rasio J, Nho SJ. Defining the clinically meaningful outcomes for arthroscopic treatment of femoroacetabular impingement syndrome at minimum 5-year follow-up. *Am J Sports Med*. 2020;48(4):901-907.
27. Nwachukwu BU, Chang B, Beck EC, et al. How should we define clinically significant outcome improvement on the iHOT-12? *HSS J*. 2019;15(2):103-108.
28. Nwachukwu BU, Rasio J, Beck EC, et al. Patient-Reported Outcomes Measurement Information System Physical Function has a lower effect size and is less responsive than legacy hip specific patient reported outcome measures following arthroscopic hip surgery. *Arthroscopy*. 2020;36(12):2992-2997.
29. Okoroha KR, Beck EC, Nwachukwu BU, Kunze KN, Nho SJ. Defining minimal clinically important difference and patient acceptable symptom state after isolated endoscopic gluteus medius repair. *Am J Sports Med*. 2019;47(13):3141-3147.
30. Rai SK, Yazdany J, Fortin PR, Aviña-Zubieta JA. Approaches for estimating minimal clinically important differences in systemic lupus erythematosus. *Arthritis Res Ther*. 2015;17(1):143.
31. Rice MW, Browning RB, Nho SJ. Surgical treatment of gluteus medius tears. *Arthroscopy*. 2022;38(7):2115-2117.
32. Rice MW, Sivasundaram L, Hevesi M, et al. Defining the minimal clinically important difference and Patient Acceptable Symptom State after endoscopic gluteus medius or minimus repair with or without labral treatment and routine capsular closure at minimum 5-year follow-up. *Am J Sports Med*. 2022;50(10):2629-2636.
33. Saltzman BM, Louie PK, Clapp IM, et al. Assessment of association between spino-pelvic parameters and outcomes following gluteus medius repair. *Arthroscopy*. 2019;35(4):1092-1098.
34. Terwee CB, Bot SDM, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;60(1):34-42.
35. Unger RZ, Burnham JM, Gammon L, Malempati CS, Jacobs CA, Makhni EC. The responsiveness of patient-reported outcome tools in shoulder surgery is dependent on the underlying pathological condition. *Am J Sports Med*. 2019;47(1):241-247.
36. Uppstrom TJ, Sullivan SW, Burger JA, Ranawat AS, Kelly BT, Nwachukwu BU. Defining minimal clinically important difference after open hip abductor repair. *Orthop J Sports Med*. 2021;9(4):23259 671211007740.
37. Vogel MJ, Wright-Chisem J, Alvero AB, Chan JJ, Chapman RS, Nho SJ. The promising 2-year performance of the Patient-Reported Outcomes Measurement Information System in primary hip arthroscopy for femoroacetabular impingement syndrome. *Am J Sports Med*. 2024;52(4):998-1004.
38. Williams BS, Cohen SP. Greater trochanteric pain syndrome: a review of anatomy, diagnosis and treatment. *Anesth Analg*. 2009;108(5): 1662-1670.
39. Yanke AB, Hart MA, McCormick F, Nho SJ. Endoscopic repair of a gluteus medius tear at the musculotendinous junction. *Arthrosc Tech*. 2013;2(2):e69-e72.