PROMIS Versus Legacy Patient-Reported Outcome Measures in Patients Undergoing Surgical Treatment for Symptomatic Acetabular Dysplasia

Deborah J. Li,* BA, John C. Clohisy,†‡ MD, Maria T. Schwabe,† BS, Elizabeth L. Yanik,† PhD, ScM, and Cecilia Pascual-Garrido,† MD, PhD
Investigation performed at Department of Orthopaedic Surgery, Washington University School of Medicine, Washington University in St Louis, St Louis, Missouri, USA

Background: No previous study has investigated how the Patient-Reported Outcomes Measurement Information System (PROMIS) performs compared with legacy patient-reported outcome measures in patients with symptomatic acetabular dysplasia treated with periacetabular osteotomy (PAO).

Purpose: To (1) measure the strength of correlation between the PROMIS and legacy outcome measures and (2) assess floor and ceiling effects of the PROMIS and legacy outcome measures in patients treated with PAO for symptomatic acetabular dysplasia.

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

Methods: This study included 220 patients who underwent PAO for the treatment of symptomatic acetabular dysplasia. Outcome measures included the Hip disability and Osteoarthritis Outcome Score (HOOS) pain, HOOS activities of daily living (ADL), modified Harris Hip Score (mHHS), PROMIS pain, and PROMIS physical function subsets, with scores collected preoperatively and/or postoperatively at a minimum 12-month follow-up. The change in mean scores from preoperatively to postoperatively was calculated only in a subgroup of 57 patients with scores at both time points. Distributions of the PROMIS and legacy scores were compared to evaluate floor and ceiling effects, and Pearson correlation coefficients were calculated to evaluate agreement.

Results: The mean age at the time of surgery was 27.7 years, and 83.6% were female. The mean follow-up time was 1.5 years. Preoperatively, neither the PROMIS nor the legacy measures showed significant floor or ceiling effects. Postoperatively, all legacy measures showed significant ceiling effects, with 15% of patients with a maximum HOOS pain score of 100, 29% with a HOOS ADL score of 100, and 21% with an mHHS score of 100. The PROMIS and legacy instruments showed good agreement preoperatively and postoperatively. The PROMIS pain had a moderate to strong negative correlation with the HOOS pain ($r = -0.66$; $P < .0001$) and mHHS ($r = -0.60$; $P < .0001$) preoperatively and the HOOS pain ($r = -0.64$; $P < .0001$) and mHHS ($r = -0.64$; $P < .0001$) postoperatively. The PROMIS physical function had a moderate positive correlation with the HOOS ADL ($r = 0.51$; $P < .0001$) and mHHS ($r = 0.49$; $P < .0001$) preoperatively and a stronger correlation postoperatively with the HOOS ADL ($r = 0.56$; $P < .0001$) and mHHS ($r = 0.56$; $P < .0001$).

Conclusion: We found good agreement between PROMIS and legacy scores preoperatively and postoperatively. PROMIS scores were largely normally distributed, demonstrating an expanded ability to capture variability in patients with improved outcomes after treatment.

Keywords: PROMIS; legacy; acetabular dysplasia; periacetabular osteotomy

Periacetabular osteotomy (PAO) is an established surgical procedure to treat patients with symptomatic acetabular dysplasia.[10,17,37,44,45] These patients are commonly young and healthy with active lifestyles, which can introduce challenges in measuring physical function and activity before and after treatment. Currently, patient-reported outcomes (PROs) in patients undergoing PAO are primarily measured using legacy scores (historic PRO measures developed by clinicians with limited patient input, not intended for measuring provider performance[36]), which include the Hip disability and Osteoarthritis Outcome Score (HOOS), Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC), and modified Harris Hip Score (mHHS).[10,11,17,22,26,37] However, these scores have some limitations, and multiple studies have shown legacy PRO measures to be inconsistently administered and reported across the literature in upper and lower extremity orthopaedics.[27,28,46]

Despite legacy scores being validated for use across multiple hip osteoarthritis abnormalities,[18,23,32] the mHHS and a significant portion of the HOOS were not developed using modern psychometric approaches. Additionally, several legacy measures including the WOMAC,[35] 36-Item Short Form Health Survey (SF-36),[35] HHS,[41] and several HOOS subdomains[16,32] have been shown to have significant floor or ceiling effects after total hip arthroplasty, thus failing to distinguish between patients after treatment for hip osteoarthritis. This may lead to shortcomings in the discriminative ability of a test to detect clinically relevant changes in populations with very high or very low outcome scores and can negatively affect sample size requirements.[32,33,41] Given the changing landscape of the US health care system as it moves away from a volume-driven payment model to a quality-based reimbursement plan, identifying a standardized, validated, and reliable PRO measure becomes important for future utilization and reimbursement.

In an effort to address the longstanding shortcomings of PRO measures, the National Institutes of Health (NIH) initiated a multicenter group in 2004 called the Patient-Reported Outcomes Measurement Information System (PROMIS), with the purpose of developing and validating a standardized PRO item bank that could be used across a range of common medical conditions.[8] The resulting PROMIS instrument utilizes computer adaptive testing and item response theory to dynamically adapt to and choose the most appropriate questions based on the patient's previous responses, resulting in an instrument that can capture PROs more precisely and efficiently.[12,14] The ability to eliminate some questionnaire burden on patients so that they are not required to answer all questions on a survey makes it an attractive measure for assessing patient outcomes in clinical care.[7] Additionally, the PROMIS has been shown to require significantly smaller sample sizes compared with legacy instruments when detecting differences within patient populations, potentially lowering research costs for clinical trials.[13,14] Of particular interest in the context of musculoskeletal disease are the PROMIS pain and physical function domains. Both domains have been validated for use in various orthopaedic populations including total shoulder arthroplasty,[9] upper and lower extremity,[20] nonshoulder upper extremity,[4] and osteoarthritic[6] populations. However, the utility of these PROMIS domains for measuring outcomes in patients undergoing PAO still is not fully understood.

The main goal of this study was to compare the PROMIS to existing legacy measures in a population of patients who had undergone PAO for the treatment of symptomatic acetabular dysplasia. Our objectives were to (1) measure the strength of correlation between the PROMIS and legacy measures at preoperative and postoperative time points and (2) assess floor and ceiling effects of the PROMIS and legacy measures. We hypothesized (1) that PROMIS scores would show a strong correlation with legacy scores when assessing similar parameters and (2) that the PROMIS would not show floor or ceiling effects.

METHODS

Patient Population

As part of a prospective longitudinal cohort, one of the senior authors (J.C.C.) prospectively enrolled patients undergoing PAO for the treatment of symptomatic acetabular dysplasia from January 2008 to December 2017 who did not respond to nonoperative treatment. The collection of PROMIS scores in this cohort occurred between 2015 and 2017; there were 335 patients who completed the PROMIS instrument at ≥1 time points. The questionnaires were given to all patients on an iPad (Apple) during preoperative and postoperative visits by a medical assistant. Patients younger than 18 years (n = 77) and patients with a history of ipsilateral hip surgery (n = 38) were excluded for the purposes of this study to remove additional morbidity from past surgery on the ipsilateral hip and to exclude patients who filled out the pediatric PROMIS form (all patients aged <18 years at our institution), producing a final study population of 220 patients. A subset of 57 patients had complete data on both preoperative and minimum 1-year postoperative PROMIS and legacy scores. The PROMIS was not implemented until June 2015 at our institution; thus, only a subset of patients were able to provide preoperative and postoperative PROMIS scores. All other patients treated before June 2015 were only able to provide postoperative PROMIS scores. This study was approved by our institution's institutional review board.

Data Collection

After providing informed consent for participation in this study, patients were given a series of questionnaires, which included demographics, medical comorbidities, social history, previous hip surgery, and PRO scores. These

‡Address correspondence to John C. Clohisy, MD, Department of Orthopaedic Surgery, Washington University School of Medicine, 660 South Euclid Avenue, Campus Box 8233, St Louis, MO 63110, USA (email: jclohisy@wustl.edu).

*University of Miami Miller School of Medicine, Miami, Florida, USA.

†Department of Orthopaedic Surgery, Washington University School of Medicine, St Louis, Missouri, USA.

outcome scores were collected preoperatively and/or at a minimum of 1 year postoperatively. Legacy measures included the mHHS to assess pain and function[3]; the HOOS including 5 subscores to evaluate lower extremity symptoms, pain, sports/recreation, activities of daily living (ADL), and quality of life[5]; the University of California, Los Angeles (UCLA) score to assess activity level[1]; and the SF-12 Health Survey (short form of SF-36) to measure overall physical and mental health (version 2).[42] PROMIS scores, which included pain and physical function subsets, were collected concurrently starting in 2015.

Statistical Analysis

This study focused on how PROMIS pain and physical function scores correlate and compare with legacy scores relating to pain and function, which include the HOOS pain, HOOS ADL, and mHHS. The PROMIS is scored on a 100-point scale, with higher scores indicating more pain (for PROMIS pain) and better physical function (for PROMIS physical function). Each subscale of the HOOS is scored on a scale of 1 to 100, with a score of 100 indicating no problems (best outcome). The mHHS is scored on a scale of 0 to 100, with a score of 100 denoting the least dysfunction (best outcome).[31]

Demographic, clinical, and PRO findings were described for the included study population. Mean scores for the PROMIS and legacy measures are reported. Time to follow-up was calculated from the difference between the date of postoperative PROMIS/legacy survey completion and the date of surgery in the 57-patient subgroup. Distributions of PROMIS and legacy scores were compared using histograms to evaluate floor and ceiling effects. A floor or ceiling effect of 15% was considered the maximum acceptable value.[39] Changes in PROMIS and legacy scores from preoperatively to postoperatively in the 57-patient subgroup were visualized using scatterplots. To assess the relationship between related legacy and PROMIS scores preoperatively and postoperatively, Pearson correlation coefficients were calculated between the PROMIS pain and HOOS pain, PROMIS physical function and HOOS ADL, and PROMIS physical function and mHHS. A correlation of 0.9 to 1.0 was considered to be very strong, 0.7 to 0.9 to be strong, 0.5 to 0.7 to be moderate, 0.3 to 0.5 to be weak, and 0.0 to 0.3 to be negligible.[30] A P value <.05 was considered to be statistically significant for all statistical measures. All statistical analyses were performed using SAS 9.4 (SAS Institute).

RESULTS

Of the 220 patients included in our study, the mean age at the time of surgery was 27.7 ± 7.7 years, the mean body mass index was 24.5 ± 3.6 kg/m^2, and 83.6% of the cohort were female. Hip abnormalities manifested on the right side in 54.1% of patients. Over half of patients had pain from their hip abnormality for up to 3 years (66.9%), and 28.1% of patients had a Tönnis grade of 1, while the rest had a Tönnis grade of 0. The mean lateral center edge angle was 14.8° ± 8.1°, the mean acetabular inclination angle was 15.7° ± 6.3°, the mean anterior-posterior alpha angle was 55.9° ± 23.7°, and the mean anterior center edge angle was 16.6° ± 10.4°. A total of 114 (51.8%) patients had preoperative PROMIS scores, and 191 (86.8%) patients had postoperative PROMIS scores. Additionally, 57 patients had PROMIS and legacy scores both preoperatively and postoperatively, and the mean time to follow-up was 1.5 years (Table 1).

Floor and Ceiling Effects

Preoperatively, neither the PROMIS nor the legacy scores showed significant floor or ceiling effects, although the HOOS ADL score was skewed left, with over 10% of patients clustered with a score ≥90. All other scores preoperatively were normally distributed (Figure 1). Postoperatively, all legacy scores were heavily skewed left and showed significant ceiling effects, with 15% of patients with a maximum HOOS pain score of 100, 29% with a HOOS ADL score of 100, and 21% with an mHHS score of 100. None of the legacy scores demonstrated a floor effect (Figure 2). PROMIS physical function and pain scores were normally distributed, although 13% of patients were clustered with a PROMIS pain score of 38.7 (Figure 2A). The PROMIS instrument had a mean score closer to 50 postoperatively (PROMIS pain: 52.9 ± 8.1; PROMIS physical function: 48.1 ± 9.0), while the legacy instruments had a mean score closer to the ceiling postoperatively (HOOS pain: 84.6 ± 16.4; HOOS ADL: 92.4 ± 11.2; mHHS: 85.5 ± 15.3) (Table 2).

PROMIS and Legacy Correlations

The PROMIS and legacy instruments showed good agreement between pain and function scores preoperatively and postoperatively in the total population (Figure 3). Preoperatively, the PROMIS pain had a moderate to strong negative correlation with the HOOS pain ($r = -0.66$; $P < .0001$) and mHHS ($r = -0.60$; $P < .0001$), indicating that a high PROMIS pain score (severe pain) is correlated with a low HOOS pain score (severe pain) and low mHHS score (low function). Postoperatively, the correlation stayed about the same with the HOOS pain ($r = -0.64$; $P < .0001$) and became slightly more robust with the mHHS ($r = -0.64$; $P < .0001$). Preoperatively, the PROMIS physical function had a moderate positive correlation with the HOOS ADL ($r = 0.51$; $P < .0001$) and mHHS ($r = 0.49$; $P < .0001$), indicating that a high PROMIS physical function score (high function) is correlated with a high HOOS ADL score (high function) and high mHHS score (high function). Postoperatively, the correlations became stronger, with a moderate positive correlation between the PROMIS physical function and HOOS ADL ($r = 0.56$; $P < .0001$) and a moderate positive correlation between the PROMIS physical function and mHHS ($r = 0.56$; $P < .0001$).

TABLE 1
General Patient Characteristics[a]

	Total (n = 220)	Subgroup[b] (n = 57)
Sex, n (%)		
Male	36 (16.4)	7 (12.3)
Female	184 (83.6)	50 (87.7)
Age at surgery, mean ± SD, y	27.7 ± 7.67	28.1 ± 6.66
Body mass index, mean ± SD, kg/m^2	24.5 ± 3.60	23.6 ± 3.23
Side, n (%)		
Left	101 (45.9)	24 (42.1)
Right	119 (54.1)	33 (57.9)
Pain chronicity, n (%)		
<1 y	56 (25.5)	9 (15.8)
1-3 y	91 (41.4)	31 (54.4)
3-5 y	32 (14.6)	8 (14.0)
>5 y	41 (18.6)	9 (15.8)
Tönnis grade, n (%)		
Grade 0	156 (71.9)	45 (80.4)
Grade 1	61 (28.1)	11 (19.6)
Radiographic finding, mean ± SD, deg		
Lateral center edge angle	14.8 ± 8.11	15.1 ± 8.46
Acetabular inclination angle	15.7 ± 6.34	15.5 ± 6.42
Anterior-posterior alpha angle	55.9 ± 23.7	58.1 ± 26.1
Anterior center edge angle	16.6 ± 10.4	17.3 ± 10.8
Time to follow-up, mean (range), y		
PROMIS pain	—	1.46 (1.00-2.87)
PROMIS physical function	—	1.46 (1.00-2.87)
HOOS pain	—	1.48 (1.00-2.87)
HOOS ADL	—	1.48 (1.00-2.87)
mHHS	—	1.48 (1.00-2.87)

[a]Dashes indicate value not calculated. ADL, activities of daily living; HOOS, Hip disability and Osteoarthritis Outcome Score; mHHS, modified Harris Hip Score; PROMIS, Patient-Reported Outcome Measurement Information System.
[b]Subgroup of patients who have both PROMIS and legacy scores preoperatively and postoperatively.

Change in PROMIS and Legacy Scores From Preoperatively to Postoperatively

Overall, among the 57 patients with scores both preoperatively and postoperatively, the PROMIS and legacy scores showed significant improvement at postoperative time points (Table 2). For postoperative PROs collected between 1.0 and 2.9 years after surgery, scores did not differ significantly over time since surgery for the PROMIS physical function, HOOS pain, and HOOS ADL. Scores were improved with increasing time since surgery for the PROMIS pain (6.00-point difference per year; P = .003) and mHHS (9.41-point difference per year; P = .03) (Figures 4 and 5). The PROMIS and legacy scores preoperatively and postoperatively are reported in Table 2. The mean change in the UCLA score from preoperatively to postoperatively was 0.8 ± 3.2. The mean change in the mHHS score was 26.0 ± 18.7. The HOOS was divided into 5 subsets, with a mean change in the HOOS symptoms score of 23.6 ± 20.7, HOOS pain score of 30.9 ± 22.4, HOOS ADL score of 27.2 ± 20.0, HOOS sports/recreation score of 37.7 ± 24.4, and HOOS quality of life score of 35.4 ± 26.8. The WOMAC was divided into 3 subsets, with a mean change in the WOMAC pain score of 27.9 ± 19.6, WOMAC physical function score of 31.8 ± 22.0, WOMAC stiffness score of 27.2 ± 20.0, and WOMAC total score of 23.7 ± 24.5. The mean

change in the SF-12 physical component score was 11.7 ± 12.8, and the mean change in the SF-12 mental component score was −1.2 ± 10.4. The PROMIS was divided into 2 subsets, with a mean change in the PROMIS pain score of −8.0 ± 9.1 and a mean change in the PROMIS physical function score of 7.6 ± 10.5 (Table 2).

DISCUSSION

In our study, we found good agreement between PROMIS and legacy pain and function scores preoperatively and postoperatively, with correlations becoming slightly stronger postoperatively. Legacy scores also demonstrated substantial ceiling effects postoperatively, which may limit the ability of legacy instruments to capture differences in patients who have improved pain and function scores after surgical treatment for acetabular dysplasia. These data indicate that the PROMIS could offer complementary information for measuring PROs in patients undergoing PAO for the treatment of symptomatic acetabular dysplasia.

Although the PROMIS and legacy scores were both normally distributed preoperatively, there were significant and visible differences in the distribution of scores postoperatively: while PROMIS scores showed a generally
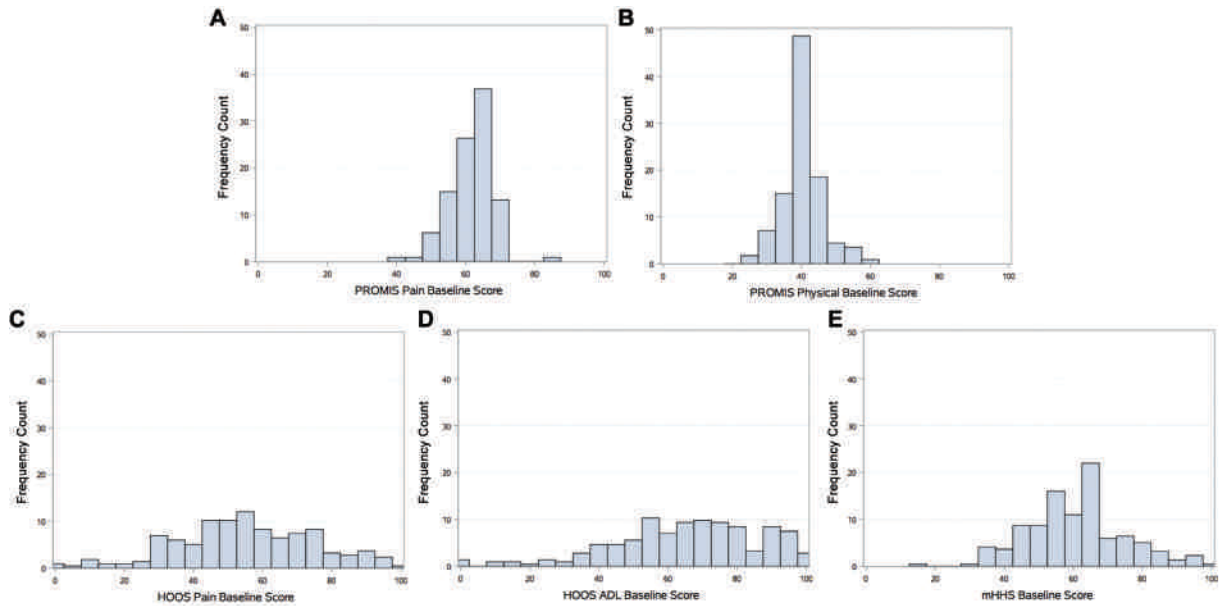
Figure 1. Histograms comparing the distribution of preoperative Patient-Reported Outcomes Measurement Information System (PROMIS) scores to preoperative legacy scores in the total patient population (n = 220). (A) Preoperative PROMIS pain score, (B) preoperative PROMIS physical function score, (C) preoperative Hip disability and Osteoarthritis Outcome Score (HOOS) pain score, (D) preoperative HOOS activities of daily living (ADL) score, and (E) preoperative modified Harris Hip Score (mHHS).
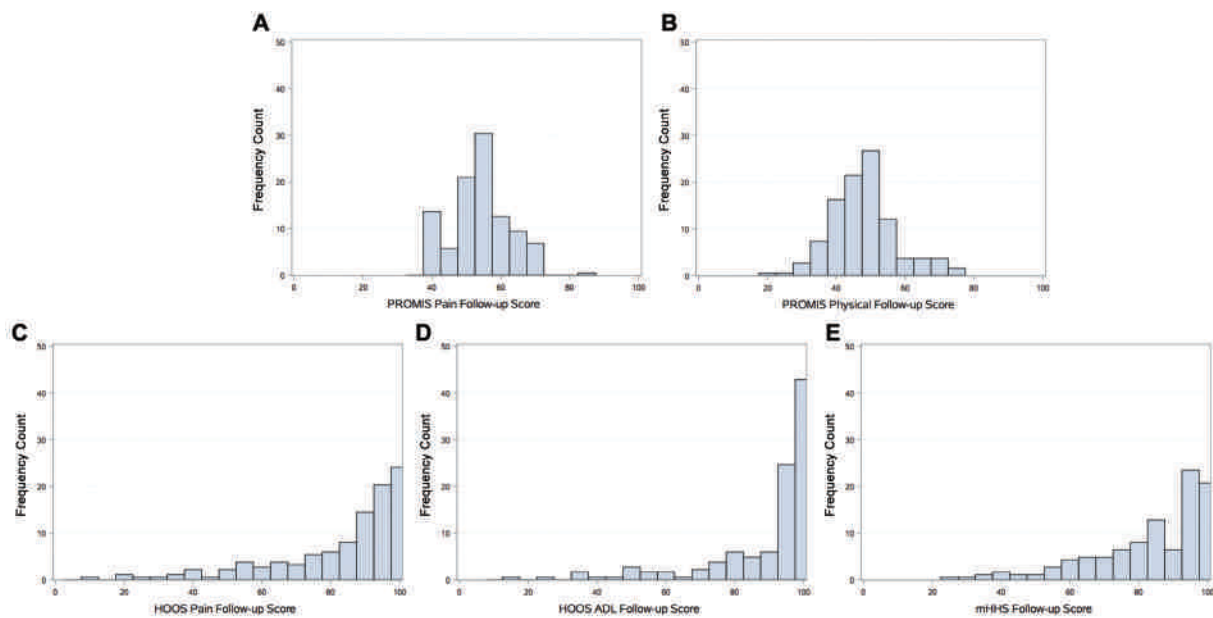
Figure 2. Histograms comparing the distribution of postoperative Patient-Reported Outcomes Measurement Information System (PROMIS) scores to postoperative legacy scores in the total patient population (n = 220). (A) Postoperative PROMIS pain score, (B) postoperative PROMIS physical function score, (C) postoperative Hip disability and Osteoarthritis Outcome Score (HOOS) pain score, (D) postoperative HOOS activities of daily living (ADL) score, and (E) postoperative modified Harris Hip Score (mHHS).

normal distribution, the HOOS pain, HOOS ADL, and mHHS scores were all heavily skewed left and demonstrated a visible ceiling effect. Our findings are consistent with previous studies that have similarly found ceiling effects postoperatively in legacy measures when reporting on patients with hip osteoarthritis undergoing total hip arthroplasty.[16,32,41] These ceiling effects may limit the ability of legacy instruments to capture differences in patients

TABLE 2
Preoperative and Postoperative PROMIS and Legacy Scores in Patient Subgroup (n = 57)[a]

	Preoperative Score	Postoperative Score	Change in Score From Preoperatively to Postoperatively	
			Mean ± SD	P Value
PROMIS				
Pain	60.9 ± 6.15	52.9 ± 8.10	−8.03 ± 9.05	<.0001
Physical function	40.4 ± 6.26	48.1 ± 9.03	7.62 ± 10.50	<.0001
Legacy measures				
UCLA score[b]	6.30 ± 2.74	7.23 ± 2.44	0.82 ± 3.23	.07
mHHS	59.5 ± 12.7	85.5 ± 15.3	26.0 ± 18.7	<.0001
HOOS				
Symptoms[b]	53.8 ± 17.9	77.3 ± 18.2	23.6 ± 20.7	<.0001
Pain	53.8 ± 18.3	84.6 ± 16.4	30.9 ± 22.4	<.0001
ADL	65.1 ± 19.0	92.4 ± 11.2	27.2 ± 20.0	<.0001
Sports/recreation[b]	40.8 ± 19.6	77.7 ± 19.8	37.7 ± 24.4	<.0001
Quality of life	30.8 ± 17.2	66.2 ± 26.3	35.4 ± 26.8	<.0001
WOMAC				
Pain	56.6 ± 19.8	88.4 ± 15.5	31.8 ± 22.0	<.0001
Physical function	65.1 ± 19.0	92.4 ± 11.2	27.2 ± 20.0	<.0001
Stiffness	53.7 ± 20.3	77.4 ± 21.2	23.7 ± 24.5	<.0001
Total	62.4 ± 18.2	90.3 ± 12.3	27.9 ± 19.6	<.0001
SF-12				
Physical component	35.9 ± 10.1	47.5 ± 10.0	11.7 ± 12.8	<.0001
Mental component	54.8 ± 8.93	53.7 ± 8.31	−1.2 ± 10.4	.40

[a]Subgroup of patients who have both PROMIS and legacy scores preoperatively and postoperatively. Data are shown as mean ± SD. ADL, activities of daily living; HOOS, Hip disability and Osteoarthritis Outcome Score; mHHS, modified Harris Hip Score; PROMIS, Patient-Reported Outcome Measurement Information System; SF-12, 12-Item Short Form Health Survey; UCLA, University of California, Los Angeles; WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index.
[b]One patient was missing data on the UCLA score preoperatively, 1 patient was missing data on the HOOS symptoms preoperatively, and 2 patients were missing data on the HOOS sports/recreation postoperatively.

who have improved pain and function scores after surgical treatment for acetabular dysplasia.

A study recently published by Wasko et al,[43] which includes members of our current group, found legacy instruments to have adequate psychometric properties for measuring PROs in patients undergoing PAO for the treatment of symptomatic acetabular dysplasia. Specifically, investigators measured the internal consistency, validity, minimal clinically important difference, responsiveness, and floor and ceiling effects of the HOOS, WOMAC, and mHHS instruments. The study only reported floor/ceiling effects at preoperative time points and found no floor or ceiling effects >15% in any of their measures. However, given that the interquartile ranges for the postoperative HOOS pain, HOOS ADL, and mHHS scores were 75.0 to 97.5, 83.8 to 100.0, and 71.5 to 95.7, respectively, 25% of patients were either at or close to the ceiling for all 3 instruments. These results are in agreement with our findings of ceiling effects in legacy instruments postoperatively.

Future surgical refinements will likely target active patients with a preserved hip but mild to moderate symptoms and suboptimal activity levels.[45] Outcome metrics that are more sensitive to the spectrum of patient-reported pain and function will enable more precise treatment refinements and eventually improved clinical outcomes. Given our results in the context of the recent study addressing the appropriateness of using legacy measures
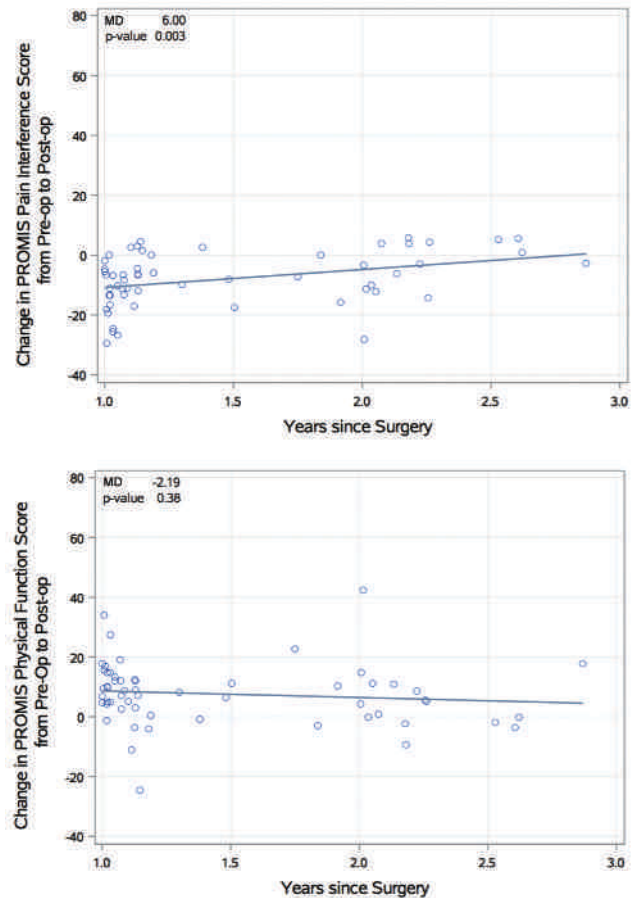
to assess PROs in patients undergoing PAO, our findings provide evidence that the PROMIS can capture variability in patients with improved PROs that would not be captured by using legacy measures alone. Our study also demonstrates that the PROMIS was able to adequately capture improvements in pain and function scores after PAO. From preoperatively to postoperatively, patients showed a mean improvement of 8.0 points on the PROMIS pain and 7.6 points on the PROMIS physical function, demonstrating clinically significant improvement in PROs based on previously determined minimal clinically important differences calculated for joint reconstruction[19] and spinal[2,21] orthopaedic populations. It should be noted that the PROMIS is not a hip-specific instrument, and increased variability postoperatively may in part indicate the capture of other musculoskeletal problems. This may be the case less frequently in a young population such as patients with acetabular dysplasia[15] (mean age, 27.7 years). Regardless, overall functional improvement and pain resolution are important long-term outcomes in this population.

There are several limitations to our study that should be noted. First, we have incomplete data on PROMIS scores, with 57 of the 220 patients having complete data on both preoperative and postoperative scores; this is because the PROMIS was only implemented within the orthopaedics department of Washington University in June 2015, and patients are continuing to come back to
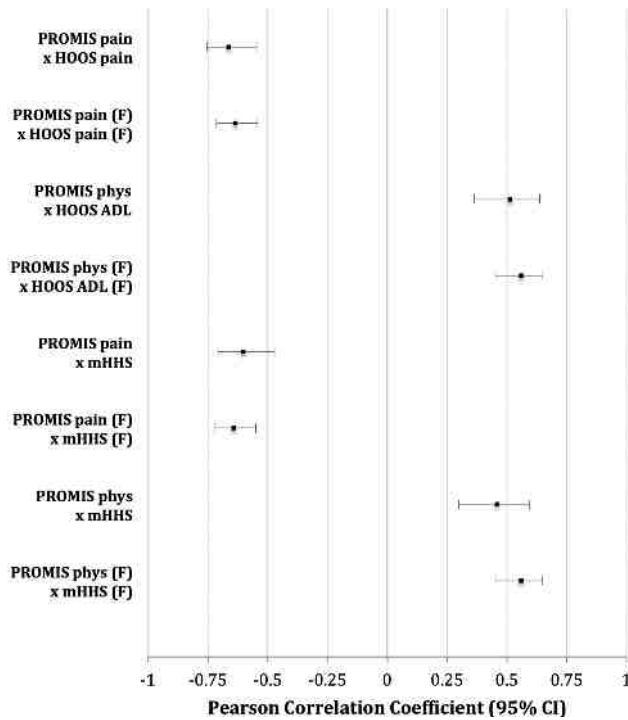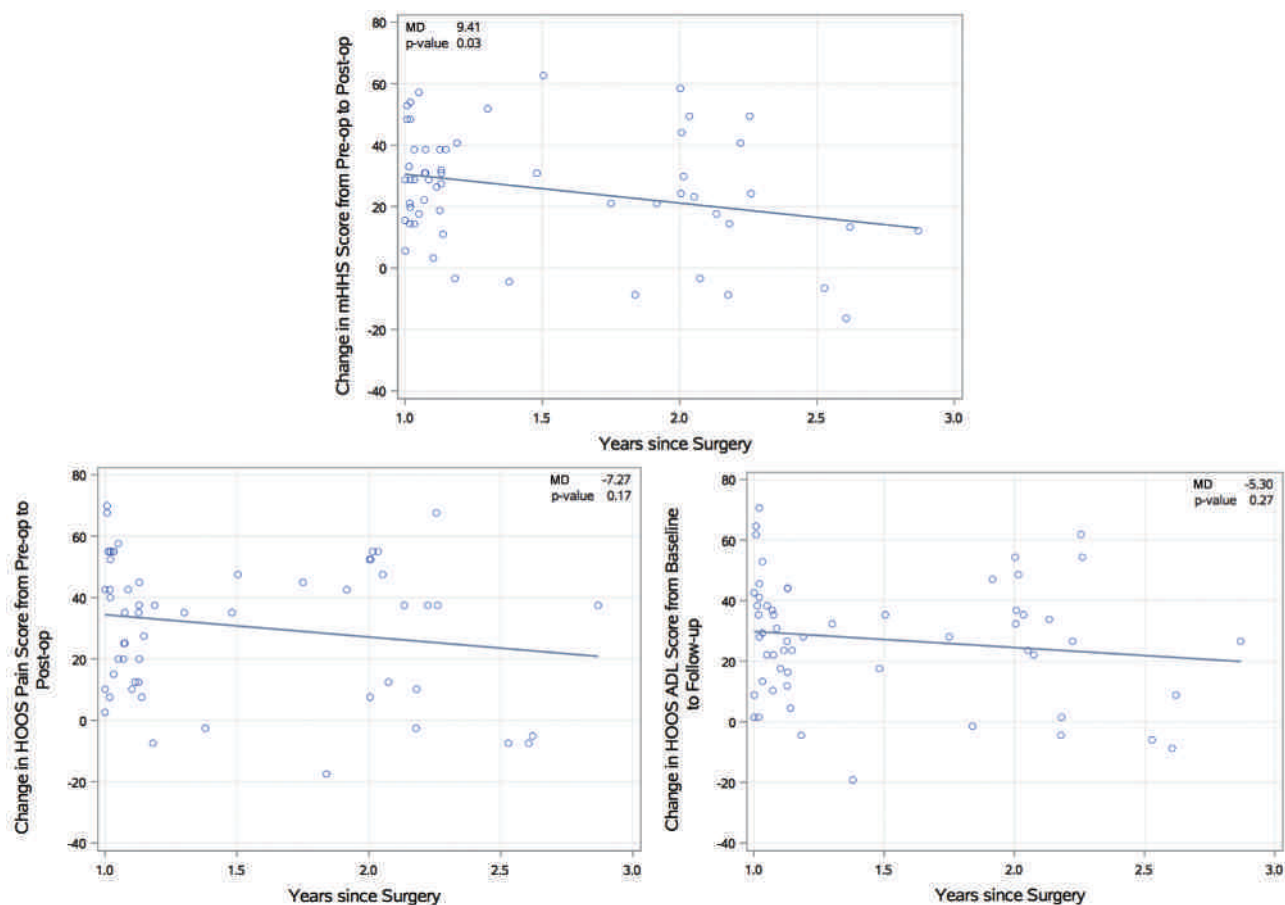
Figure 3. Forest plot of Pearson correlation coefficients calculated between Patient-Reported Outcomes Measurement Information System (PROMIS) and legacy scores preoperatively and postoperatively in the total patient population (n = 220). A PROMIS pain score of 0 indicates no pain (best outcome). A PROMIS physical function score of 100 indicates highest physical function (best outcome). ADL, activities of daily living; HHS, Harris Hip Score; HOOS, Hip disability and Osteoarthritis Outcome Score.

Figure 4. Scatterplots of the change in Patient-Reported Outcomes Measurement Information System (PROMIS) pain and physical function scores from preoperatively to postoperatively of the subgroup patient population (subgroup of patients who have both PROMIS and legacy scores preoperatively and postoperatively [n = 57]). Regression trend lines with absolute difference and associated P value are reported. A PROMIS pain score of 0 indicates no pain (best outcome). A PROMIS physical function score of 100 indicates highest physical function (best outcome). MD, mean difference in PROMIS score per year since surgery.

the clinic for postoperative follow-up. Therefore, there are fewer patients with preoperative PROMIS scores compared with preoperative legacy scores, and postoperative PROMIS scores are being recorded on a continuous basis as patients return for follow-up. Second, the mean time to follow-up was 1.5 years, with a maximum follow-up of 2.9 years. Because PAO is a procedure in which long-term pain and functional outcomes are particularly important, a comparison of PROMIS and legacy scores at longer periods of follow-up will be valuable to investigate in the future. It should be noted that there was no significant difference between postoperative PROMIS physical function and HOOS scores collected at different time points between 1.0 and 2.9 years postoperatively and statistically significant but clinically small absolute differences between postoperative PROMIS pain and mHHS scores at different postoperative time points (Figures 4 and 5). As a result, we chose to analyze all PROMIS and legacy scores occurring between 1.0 and 2.9 years postoperatively together. Third, although all correlation analyses showed statistically significant correlations between the PROMIS and legacy measures at preoperative and postoperative time points, some correlations were only weak to moderate

in strength. A recent study conducted by Kollmorgen et al[24] investigated agreement between the PROMIS and legacy measures in patients presenting to a tertiary care hip preservation center using correlations and found moderate to strong correlations in all relationships measured. However, this study did not differentiate between preoperative/postoperative or surgical/nonsurgical patients; rather we grouped all preoperative, postoperative, and nonsurgical patients together into a single group and measured correlations as if at a single time point. Similarly, another recently published study by Padilla et al[34] also reported moderate to strong correlations between legacy (Knee injury and Osteoarthritis Outcome Score joint replacement and HOOS–joint replacement) and PROMIS scores in patients presenting to an adult reconstructive outpatient clinic with hip or knee

Figure 5. Scatterplots of the change in modified Harris Hip Score (mHHS), Hip disability and Osteoarthritis Outcome Score (HOOS) pain, and HOOS activities of daily living (ADL) scores from preoperatively to postoperatively of the subgroup patient population (subgroup of patients who have both Patient-Reported Outcomes Measurement Information System [PROMIS] and legacy scores preoperatively and postoperatively [n = 57]). Regression trend lines with absolute difference and associated *P* value are reported. MD, mean difference in PROMIS score per year since surgery.

complaints. Although this study did separate patients who were treated operatively with total joint replacement versus those who were managed nonoperatively, they did not differentiate PRO scores between preoperative and postoperative time points. Although our study did not produce correlations as strong as the Kollmorgen et al[24] or Padilla et al[34] studies, we believe that our study offers a more valid estimate of agreement between PROMIS and legacy scores, given that we looked distinctly at preoperative and postoperative time points and that surgical and nonoperative patients have very different clinical courses. Fourth, there are inherent limitations in measuring subjective outcomes, particularly when it comes to function. The creation of the PROMIS, a standardized and validated PRO measuring system, helps to minimize these potential weaknesses but cannot completely eliminate them. For example, anxiety and depression are subjective contributors to patient outcomes that can significantly affect pain and function reporting. As these are both existing PROMIS domains available for use, we have started collecting PRO data on these results as well but did not focus on these domains in the current

study, as the main aim of our study was to measure agreement between legacy measures and the PROMIS. Finally, we did not include in our data collection several other legacy PRO measures also traditionally used to assess pain and function in orthopaedic patient populations with hip dysfunction, including the Copenhagen Hip and Groin Outcome Score and the International Hip Outcome Tool (iHOT), both of which have been shown to be appropriate and validated for use in young to middle-aged patients with symptomatic hip or groin abnormalities.[29,40] As of the beginning of 2019, our institution has implemented the use of new contemporary PRO measures, which include the iHOT-12.

Additional potential benefits to using the PROMIS, which our study did not investigate, include the establishment of a standardized common language when reporting PROs in orthopaedic populations. Given the changes that the US health care system is currently undergoing as it shifts away from the traditional volume-driven fee-for-service reimbursement model to a value-based payment plan, it is likely that the payers in the health care system may soon more broadly start requiring some sort of PRO

measure to be used in clinical practice.[36] As such, a PRO measure that is standardized, is validated, and reliably assesses the value of physician care from the patient's perspective needs to be identified. Currently, there are a large number of joint-specific PRO instruments used in orthopaedic surgery, with high levels of variability in PROs across studies reporting on patients with the same diagnosis, limiting the translatability of these studies.[25,28,38] These shortcomings in older PRO instruments are part of what prompted the NIH to create the PROMIS, a validated, standardized item bank applicable to the general population that can be used across a wide variety of morbidities. Also, because the PROMIS utilizes computer adaptive testing and item response theory to produce a concise and targeted survey, averaging only 7 questions per domain in our study, the time constraints that come with a high number of survey questions should not be a barrier to implementing the PROMIS in the clinic alongside existing legacy instruments. Implemented alongside existing legacy PRO instruments that are specific to hip abnormalities, the PROMIS could provide for a better comparison of results across different clinical trials and improve the applicability of research results in clinical practice. In our study, we demonstrated that the PROMIS pain and physical function domains can provide valuable additional insight into patients with improved pain and function scores undergoing PAO for the treatment of symptomatic acetabular dysplasia when used alongside existing legacy instruments.

REFERENCES

1. Amstutz HC, Thomas BJ, Jinnah R, Kim W, Grogan T, Yale C. Treatment of primary osteoarthritis of the hip: a comparison of total joint and surface replacement arthroplasty. *J Bone Joint Surg Am*. 1984;66(2):228-241.
2. 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.
3. Banaszkiewicz PA. Traumatic arthritis of the hip after dislocation and acetabular fractures: treatment by mold arthroplasty. An end-result study using a new method of result evaluation. In: Banaszkiewicz PA, Kader DF, eds. *Classic Papers in Orthopaedics*. London: Springer; 2014:13-17.
4. Beckmann JT, Hung M, Voss MW, Crum AB, Bounsanga J, Tyser AR. Evaluation of the Patient-Reported Outcomes Measurement Information System upper extremity computer adaptive test. *J Hand Surg Am*. 2016;41(7):739-744.e4.
5. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *J Rheumatol*. 1988;15(12):1833-1840.
6. Broderick JE, Schneider S, Junghaenel DU, Schwartz JE, Stone AA. Validity and reliability of Patient-Reported Outcomes Measurement Information System instruments in osteoarthritis. *Arthritis Care Res (Hoboken)*. 2013;65(10):1625-1633.
7. Cella D, Riley W, Stone A, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005-2008. *J Clin Epidemiol*. 2010;63(11):1179-1194.
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. Chen RE, Papuga MO, Nicandri GT, Miller RJ, Voloshin I. Preoperative Patient-Reported Outcomes Measurement Information System (PROMIS) scores predict postoperative outcome in total shoulder arthroplasty patients. *J Shoulder Elbow Surg*. 2019;28(3):547-554.
10. Clohisy JC, Ackerman J, Baca G, et al. Patient-reported outcomes of periacetabular osteotomy from the prospective ANCHOR cohort study. *J Bone Joint Surg Am*. 2017;99(1):33-41.
11. d'Aubigné RM, Postel M. The classic: functional results of hip arthroplasty with acrylic prosthesis. *Clin Orthop Relat Res*. 2009;467:7-27.
12. Fries JF, Bruce B, Cella D. The promise of PROMIS: using item response theory to improve assessment of patient-reported outcomes. *Clin Exp Rheumatol*. 2005;23(5)(suppl 39):S53-S57.
13. Fries JF, Krishnan E, Rose M, Lingala B, Bruce B. Improved responsiveness and reduced sample size requirements of PROMIS physical function scales with item response theory. *Arthritis Res Ther*. 2011;13(5):R147.
14. Fries JF, Witter J, Rose M, Cella D, Khanna D, Morgan-DeWitt E. Item response theory, computerized adaptive testing, and PROMIS: assessment of physical function. *J Rheumatol*. 2014;41(1):153-158.
15. Gala L, Clohisy JC, Beaule PE. Hip dysplasia in the young adult. *J Bone Joint Surg Am*. 2016;98(1):63-73.
16. Gandek B, Roos EM, Franklin PD, Ware JE Jr. A 12-item short form of the Hip disability and Osteoarthritis Outcome Score (HOOS-12): tests of reliability, validity and responsiveness. *Osteoarthritis Cartilage*. 2019;27(5):762-770.
17. Ganz R, Klaue K, Vinh TS, Mast JW. A new periacetabular osteotomy for the treatment of hip dysplasias: technique and preliminary results. *Clin Orthop Relat Res*. 1988;232:26-36.
18. Hoeksma HL, Van Den Ende CH, Ronday HK, Heering A, Breedveld FC. Comparison of the responsiveness of the Harris Hip Score with generic measures for hip function in osteoarthritis of the hip. *Ann Rheum Dis*. 2003;62(10):935-938.
19. Hung M, Bounsanga J, Voss MW, Saltzman CL. Establishing minimum clinically important difference values for the Patient-Reported Outcomes Measurement Information System physical function, Hip disability and Osteoarthritis Outcome Score for joint reconstruction, and Knee injury and Osteoarthritis Outcome Score for joint reconstruction in orthopaedics. *World J Orthop*. 2018;9(3):41-49.
20. Hung M, Clegg DO, Greene T, Saltzman CL. Evaluation of the PROMIS physical function item bank in orthopaedic patients. *J Orthop Res*. 2011;29(6):947-953.
21. Hung M, Saltzman CL, Kendall R, et al. What are the MCIDs for PROMIS, NDI, and ODI instruments among patients with spinal conditions? *Clin Orthop Relat Res*. 2018;476(10):2027-2036.
22. Jacobs CA, Peabody MR, Duncan ST, et al. Development of the HOOS global to assess patient-reported outcomes in patients undergoing hip preservation procedures. *Am J Sports Med*. 2018;46(4):940-946.
23. Klassbo M, Larsson E, Mannevik E. Hip disability and Osteoarthritis Outcome Score: an extension of the Western Ontario and McMaster Universities Osteoarthritis Index. *Scand J Rheumatol*. 2003;32(1):46-51.
24. Kollmorgen RC, Hutyra CA, Green C, Lewis B, Olson SA, Mather RC 3rd. Relationship between PROMIS computer adaptive tests and legacy hip measures among patients presenting to a tertiary care hip preservation center. *Am J Sports Med*. 2019;47(4):876-884.
25. Lukenchuk J, Sims LA, Shin JJ. Variability in outcome reporting for operatively managed anterior glenohumeral instability: a systematic review. *Arthroscopy*. 2017;33(2):477-483.
26. Mahomed NN, Arndt DC, McGrory BJ, Harris WH. The Harris Hip Score: comparison of patient self-report with surgeon assessment. *J Arthroplasty*. 2001;16(5):575-580.
27. Makhni EC, Meadows M, Hamamoto JT, Higgins JD, Romeo AA, Verma NN. Patient Reported Outcomes Measurement Information System (PROMIS) in the upper extremity: the future of outcomes reporting? *J Shoulder Elbow Surg*. 2017;26(2):352-357.
28. Makhni EC, Padaki AS, Petridis PD, et al. High variability in outcome reporting patterns in high-impact ACL literature. *J Bone Joint Surg Am*. 2015;97(18):1529-1542.

29. Mohtadi NG, Griffin DR, Pedersen ME, et al. The development and validation of a self-administered quality-of-life outcome measure for young, active patients with symptomatic hip disease: the International Hip Outcome Tool (iHOT-33). *Arthroscopy*. 2012;28(5):595-605.

30. Mukaka MM. Statistics corner: a guide to appropriate use of correlation coefficient in medical research. *Malawi Med J*. 2012;24(3):69-71.

31. Nilsdotter A, Bremander A. Measures of hip function and symptoms: Harris Hip Score (HHS), Hip disability and Osteoarthritis Outcome Score (HOOS), Oxford Hip Score (OHS), Lequesne Index of Severity for Osteoarthritis of the Hip (LISOH), and American Academy of Orthopedic Surgeons (AAOS) Hip and Knee Questionnaire. *Arthritis Care Res (Hoboken)*. 2011;63(suppl 11):S200-S207.

32. Nilsdotter AK, Lohmander LS, Klassbo M, Roos EM. Hip disability and Osteoarthritis Outcome Score (HOOS): validity and responsiveness in total hip replacement. *BMC Musculoskelet Disord*. 2003;4:10.

33. Ornetti P, Parratte S, Gossec L, et al. Cross-cultural adaptation and validation of the French version of the Hip disability and Osteoarthritis Outcome Score (HOOS) in hip osteoarthritis patients. *Osteoarthritis Cartilage*. 2010;18(4):522-529.

34. Padilla JA, Rudy HL, Gabor JA, et al. Relationship between the Patient-Reported Outcome Measurement Information System and traditional patient-reported outcomes for osteoarthritis. *J Arthroplasty*. 2019;34(2):265-272.

35. Rampazo-Lacativa MK, Santos AA, Coimbra AM, D'Elboux MJ. WOMAC and SF-36: instruments for evaluating the health-related quality of life of elderly people with total hip arthroplasty. A descriptive study. *Sao Paulo Med J*. 2015;133(4):290-297.

36. Squitieri L, Bozic KJ, Pusic AL. The role of patient-reported outcome measures in value-based payment reform. *Value Health*. 2017;20(6):834-836.

37. Steppacher SD, Tannast M, Ganz R, Siebenrock KA. Mean 20-year followup of Bernese periacetabular osteotomy. *Clin Orthop Relat Res*. 2008;466:1633-1644.

38. Stone AV, Jacobs CA, Luo TD, et al. High degree of variability in reporting of clinical and patient-reported outcomes after hip arthroscopy. *Am J Sports Med*. 2018;46(12):3040-3046.

39. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;60(1):34-42.

40. Thorborg K, Holmich P, Christensen R, Petersen J, Roos EM. The Copenhagen Hip and Groin Outcome Score (HAGOS): development and validation according to the COSMIN checklist. *Br J Sports Med*. 2011;45(6):478-491.

41. Wamper KE, Sierevelt IN, Poolman RW, Bhandari M, Haverkamp D. The Harris Hip Score: do ceiling effects limit its usefulness in orthopedics? *Acta Orthop*. 2010;81(6):703-707.

42. Ware JE. *The SF-12v2™: How to Score Version 2 of the SF-12® Health Survey (With a Supplement Documenting Version 1)*. Lincoln, Rhode Island: Quality Metric; 2002.

43. Wasko MK, Yanik EL, Pascual-Garrido C, Clohisy JC. Psychometric properties of patient-reported outcome measures for periacetabular osteotomy. *J Bone Joint Surg Am*. 2019;101(6):e21.

44. Wells J, Millis M, Kim YJ, Bulat E, Miller P, Matheney T. Survivorship of the Bernese periacetabular osteotomy: what factors are associated with long-term failure? *Clin Orthop Relat Res*. 2017;475(2):396-405.

45. Wells J, Schoenecker P, Duncan S, Goss CW, Thomason K, Clohisy JC. Intermediate-term hip survivorship and patient-reported outcomes of periacetabular osteotomy: the Washington University experience. *J Bone Joint Surg Am*. 2018;100(3):218-225.

46. Yeung M, Khan M, Schreiber VM, et al. Global discrepancies in the diagnosis, surgical management, and investigation of femoroacetabular impingement. *Arthroscopy*. 2014;30(12):1625-1633.