7-14-2020

PROMIS Physical Function Has a Lower Effect Size and is Less Responsive than Legacy Hip Specific Patient Reported Outcome Measures Following Arthroscopic Hip Surgery

Benedict U. Nwachuwu

Jonathan Rasio

Edward C. Beck

Kelechi R. Okoroha

Spencer W. Sullivan

See next page for additional authors

Authors

Benedict U. Nwachuwu, Jonathan Rasio, Edward C. Beck, Kelechi R. Okoroha, Spencer W. Sullivan, Eric C. Makhni, and Shane J. Nho

PROMIS Physical Function Has a Lower Effect Size and is Less Responsive than Legacy Hip Specific Patient Reported Outcome Measures Following Arthroscopic Hip Surgery

Benedict U. Nwachuwu, M.D., M.B.A., Jonathan Rasio, B.S., Edward C. Beck, M.D., M.B.A., Kelechi R. Okoroha, M.D., Spencer W. Sullivan, B.S., Eric C. Makhni, M.D., and Shane J. Nho, M.D., M.S.

Purpose: To compare the use and responsiveness of Patient Reported Outcomes Measurement Information System (PROMIS) to legacy patient-reported outcome measures (PROMs) in patients undergoing hip arthroscopy for femoroacetabular impingement syndrome (FAIS) at 6-month follow-up. **Methods:** Data from patients who underwent primary hip arthroscopy with routine capsular closure between August 2018 and January 2019 for the treatment of FAIS were analyzed. Preoperative outcomes, 6-month postoperative outcomes, and demographics were recorded. Primary outcome measures included PROMIS Physical Function (PROMIS-PF), PROMIS Pain Interference (PROMIS-PI), and PROMIS Depression. The legacy PROMs included Hip Outcome Score Activities of Daily Living (HOS-ADL), Hip Outcome Score Sport Subscale (HOS-SS), and the international hip outcome tool 12 questions (iHOT-12). Floor and ceiling effects along with the responsiveness and Cohen's *d* effect size of each PROM tool were calculated. **Results:** Ninety-six patients with an average age and body mass index of 32.4 ± 11.9 years and 25.9 ± 6.1 kg/m^2, respectively, were included in the final analysis. All outcomes were significantly higher at 6 months compared with the preoperative level ($P < .001$) except for PROMIS Depression ($P = .873$). PROMIS-PF demonstrated excellent correlation with HOS-SS ($r = 0.81$; $P < .001$), very good correlation with HOS-ADL ($r = 0.73$; $P < .001$), and good correlation with iHOT-12 ($r = 0.68$; $P < .001$). No floor was observed for any measure. The effect size was large for all outcomes, except PROMIS Depression ($d = 0.04$), but largest for iHOT12 ($d = 1.87$) followed by HOS-ADL ($d = 1.29$). The iHOT-12 was more responsive than PROMIS-PI (relative efficiency [RE] = 3.95), PROMIS-PF (RE = 4.13), HOS-ADL (RE = 2.26), and HOS-SS (RE = 3.84). HOS-SS was similarly responsive to PROMIS-PI (RE=1.03) and PROMIS-PF (RE=1.08). However, PROMIS-PF was overall the least responsive. **Conclusions:** In patients at 6 months postoperatively from hip arthroscopy for FAIS, iHOT-12 was the most responsive and had the largest effect size. In contrast, PROMIS-PF had a lower effect size compared with legacy hip-specific PROMs. Additionally, PROMIS-PF did not correlate as well with iHOT-12 compared with HOS-SS. **Level of Evidence:** Level IV, case series.

There is increasing attention to streamlining methods for patient-reported outcome collection and delivery. The Patient Reported Outcomes Measurement Information System (PROMIS) was developed by the National Institutes of Health with the goal of providing a single, generalizable, and validated outcome measure that can be administered across a number of conditions.[1] More recently, computer adaptive testing

From the Department of Orthopaedic Surgery, Hospital for Special Surgery (B.U.N., E.C.B., S.W.S.), New York, New York, U.S.A.; Division of Sports Medicine, Midwest Orthopaedics at Rush, Rush University Medical Center (J.R., S.J.N.), Chicago, Illinois, U.S.A.; and Department of Orthopaedic Surgery, Henry Ford Health Center (K.R.O., E.C.M.), Detroit, Michigan, U.S.A.

The authors report the following potential conflicts of interest or sources of funding: S.J.N. reports personal fees from Stryker, American Journal of Orthopedics, and Ossur. Full ICMJE author disclosure forms are available for this article online, as supplementary material.

Received January 31, 2020; accepted July 3, 2020.

Address correspondence to Benedict U. Nwachukwu, M.D., M.B.A., Department of Orthopaedic Surgery, Hospital for Special Surgery, 610 W. 58th St., 3rd Floor, New York, NY 10019. E-mail: nwachukwub@hss.edu

0749-8063/2091/$36.00

https://doi.org/10.1016/j.arthro.2020.07.008

(CAT) versions of PROMIS instruments have been developed that use item selection to reduce the number of responses required to provide an outcome score assessment.[1-4]

There is increasing interest in understanding the performance of PROMIS for evaluating orthopaedic conditions.[5] The psychometric properties of PROMIS have been studied for knee meniscal surgery, rotator cuff abnormalities, shoulder arthritis, shoulder instability, femoroacetabular impingement syndrome (FAIS), and anterior cruciate ligament injuries.[6-12] Although the majority of prior studies have suggested good correlation between PROMIS and legacy outcome measures, some authors have suggested that PROMIS is inherently anchored to the preferences of a United States-based population and may lack the generalizability required of a "gold standard."[13] Specific to hip pain, 2 prior studies have evaluated the preoperative psychometric properties of PROMIS for patients presenting at a hip preservation center. Both studies demonstrated that PROMIS-PF demonstrated excellent to good correlation preoperatively with legacy hip-specific instruments as well as with health-related quality of life measures.[11,12] However, a study specific to FAIS noted that PROMIS Physical Function (PROMIS-PF) demonstrated the weakest correlation with the Hip Outcome Score Sports Specific subscale.[12] The authors theorized that there is "a component of hip-specific disability that is not captured in the PROMIS-PF."

In the hip preservation evidence base, there is a paucity of longitudinal and postoperative PROMIS data. As such, there is an opportunity to evaluate the early postoperative performance of PROMIS-PF CAT in patients undergoing arthroscopic FAIS surgery. The purpose of this study was to compare the use and responsiveness, which refers to the ability of an outcome measure to detect change over time, of PROMIS to legacy patient-reported outcome measures (PROMs) in patients undergoing hip arthroscopy for FAIS at 6-month follow-up. The authors hypothesize that PROMIS will demonstrate good correlation; however, the authors predict that there may be limited responsiveness for PROMIS compared with legacy PROMs.

Methods

Patient Selection

The current study was approved by the Rush University Medical Center institutional review board. Data were prospectively collected on all patients undergoing hip arthroscopy for the treatment of FAIS by a single, fellowship-trained surgeon (S.J.N.). Data were retrospectively analyzed in a clinical repository. Consecutive patients undergoing primary hip arthroscopy for the treatment of FAIS between August 2018 and January 2019 were included in this study. Inclusion criteria consisted of clinical and radiographic diagnosis of symptomatic FAIS, failure of conservative management (including physical therapy, activity modification, oral anti-inflammatories, and/or intra-articular cortisone injection), and surgical treatment with hip arthroscopy for FAIS.[14] Exclusion criteria consisted of prior history of bilateral hip surgery, hip arthroscopy for an indication other than FAIS, signs of osteoarthritis (Tonnis grade >1), hip dysplasia (Lateral Center Edge Angle $<20°$),[15] a history of pediatric hip disorders (eg, slipped capital femoral epiphysis, hip dysplasia, Perthes disease, etc.), concomitant trochanteric bursectomy, and concomitant gluteus medius/minimus repair.

Functional Outcome Evaluation

Preoperatively, demographic data were collected from all patients, including sex, age, operative extremity, body mass index, sports participation, duration of symptoms, and comorbidities. All patients completed preoperative and minimum 6 months of legacy hip-specific PROMs including the Hip Outcome Score Activities of Daily Living Subscale (HOS-ADL),[16] HOS-Sports Subscale (HOS-SS), and the International Hip Outcome Tool-12 (iHOT-12).[17] Nwachukwu et al[18] previously demonstrated that a high proportion of patients achieved significant clinical outcomes in legacy PROMs, iHOT-33, HOS-ADL, and HOS-SS at the 6-month postoperative follow-up. In addition, primary outcome measures, PROMIS-PF, PROMIS Pain Interference (PROMIS-PI),[19,20] and PROMIS Depression,[21] were obtained via CAT. Questionnaires were completed using Outcome Based Electronic Research Database (Universal Research Solutions, Columbia, MO).

Statistical Analysis

Analysis was performed to identify any existing floor and ceiling effects for the legacy hip-specific PROMs, PROMIS-PF, PROMIS-PI, and PROMIS Depression. For the legacy PROMs, which are scaled to 100, any percentage greater or equal to 15% of the study population in the top or bottom 5% of the score range was deemed as a significant ceiling or floor effect.[6,22] For PROMIS, a significant floor or ceiling effect for PROMIS outcomes was determined as any percentage greater or equal to 15% of the study population in the top or bottom 5th percentile of the population was deemed as significant ceiling or floor effect.[6,22]

All data were screened to determine the achievement of all parametric statistical assumptions before analysis. Pearson coefficient analysis was used to identify correlations between PROMIS scores and the legacy PROMs. Correlation coefficients were classified by the strength of the correlation, which were defined as follows: excellent (>0.80), very good (0.71-0.80), good (0.61-0.70), fair (0.41-0.60), and poor (0.21-0.40).[17]

Table 1. Patient Demographics

	Mean (±SD)/No. (%)
N	96
Age (**y**)	32.4 ± 11.9
Females	76 (79.2)
BMI	25.9 ± 6.1

BMI, body mass index; SD, standard deviation.

To directly compare the responsiveness between PROMs, the effect size and relative efficiency (RE) were calculated for each PROM tool.[23-25] The effect size, or Cohen's *d*, is a measure of the magnitude of the preoperative to postoperative change in relation to the amount of variability in the scores.[25-27] Cohen's *d* was calculated by dividing the absolute difference in the mean change score for each PROM by the pooled standard deviation for that PROM tool.[25-27] Consequently, a *d* of 1 means that the pre- to postoperative change in a PROM differs by 1 standard deviation. Effect sizes between 0.20 and 0.49 are considered small, between 0.50 and 0.79 are considered moderate, and ≥0.80 are considered large and are clearly visible to the naked eye.[25-27] The RE was calculated to directly compare responsiveness between PROM tools.[23-25] The RE is derived from the resultant *t* score for paired *t* tests comparing the preoperative score for a PROM tool with the postoperative score. The *t* score from 1 PROM tool is then divided by the *t* score of another PROM tool and the resulting value is squared to obtain the RE between the 2 PROMs.[23-25] If the RE is <1, the first PROM tool would be considered "less responsive" than the second PROM tool. Conversely, if the RE is >1, the first PROM would be considered "more responsive" than the second PROM tool.[23-25]

Descriptive statistics for all continuous variables are reported as means with standard deviations, and frequency statistics are reported for all noncontinuous variables. Paired samples *t* tests were used to compare preoperative and 6-month postoperative PROMs in FAIS patients. All continuous data were screened for normality using the Shapiro-Wilk test and all data were found to be drawn from a normal distribution. This allowed us to use parametric tests such as Pearson's correlation and Student *t* test. In addition, Levene's test was used to test for homogeneity of variance in cases where the pooled standard deviation had to be calculated. Statistical significance for all analyses was set at an $\alpha \leq 0.05$. All statistical analyses were performed using JMP PROM (version 14.2, SAS Institute, Cary, NC).

Results

Of 140 eligible patients, 13 underwent revision surgery and 7 underwent concomitant gluteus medius repair and were excluded from the study. A total of 96 patients (80%) had 6-month follow-up and were included in the study, with an average age and body mass index of 32.4 ± 11.9 years and 25.9 ± 6.1 kg/m^2, respectively. Patient demographics are summarized in Table 1. The majority of patients were female (76%). There were no workers compensation patients in this cohort. All patients underwent primary labral repair, acetabular rim trimming, debridement, synovectomy, femoral osteochondroplasty, and capsular closure.

Clinical Outcomes Analysis

Using PROMIS CAT, patients answered between 4 and 12 questions with an average of 5.0, 5.7, and 4.4 questions for PI, Depression, and PF, respectively. Paired *t* test analysis of pre- and minimum 6-months postoperative PROMs are summarized in Table 2. There was a statistically significant increase in mean HOS-ADL (62.7 ± 16.1 vs 83.1 ± 15.3; $P < .001$), HOS-SS (39.1 ± 21.0 vs 64.7 ± 26.5; $P < .001$), iHOT-12, (29.2 ± 18.3 vs 68.0 ± 23.0; $P < .001$), and PROMIS-PF (40.3 ± 6.4 vs 47.1 ± 7.9; $P = .003$). In addition, there was a significant reduction in reported mean PROMIS-PI score (61.3 ± 5.8 vs 54.6 ± 8.0; $P = .004$). There was no statistical difference in reported mean PROMIS depression scores (48.8 ± 8.6 vs 49.2 ± 8.2; $P < .87$).

Floor and ceiling effects at 6-months postoperatively for PROM tools are summarized in Table 3. There was a ceiling effect observed for HOS-ADL, with 24.44% in the top 5% of possible scores, and for PROMIS-PI, with 15.22% in the top 5th percentile. No floor effect was observed for any measure.

Correlation Analysis

Pearson's coefficient analysis between PROMIS outcome and legacy PROMs score are summarized in Table 4. PROMIS-PF scores demonstrated excellent correlation with HOS-SS ($r = 0.81$; $P < .001$), very good correlation with HOS-ADL ($r = 0.73$; $P < .001$), and good correlation with iHOT-12 ($r = 0.68$; $P < .001$). PROMIS-PI demonstrated very good correlation with HOS-ADL ($r = -0.78$; $P < .001$), HOS-SS ($r = -0.74$; $P < .001$), and iHOT12 ($r = -0.77$; $P < .001$). PROMIS Depression demonstrated fair correlation with HOS-ADL ($r = -0.46$;

Table 2. Analysis of Pre- and Postoperative Reported Outcomes

	Preoperative	6 mo	*P* Value
HOS-ADL (± SD)	62.7 ± 16.1	83.1 ± 15.3	<.001*
HOS-SS (± SD)	39.1 ± 21.0	64.7 ± 26.5	<.001*
iHOT-12 (± SD)	29.2 ± 18.3	68.0 ± 23.0	<.001*
PROMIS depression (± SD)	48.8 ± 8.6	49.2 ± 8.2	.873
PROMIS Pain interference (± SD)	61.3 ± 5.8	54.6 ± 8.0	.004*
PROMIS Physical function (± SD)	40.3 ± 6.4	47.1 ± 7.9	0.003*

HOS-ADL, Hip Outcome Score-Activities of Daily Living; HOS-SS, Hip Outcome Score-Sports Subscale; iHOT-12, International Hip Outcome Tool 12; PROMIS, Patient Reported Outcomes Measurement Information System; SD, standard deviation.

*Statistical significance with $P < .01$.

Table 3. Postoperative Outcomes Floor and Ceiling Effects

	Floor (%)	Ceiling (%)
HOS-ADL	0	**24.44**
HOS-SS	2.67	8.00
iHOT-12	0	4.26
PROMIS Depression	10.34	4.60
PROMIS Pain Interference	0	**15.22**
PROMIS Physical Function	8.79	2.20

Boldface type indicates >15% of patient outcomes in the bottom or top 5th percentile for floor and ceiling, respectively.

HOS-ADL, Hip Outcome Score-Activities of Daily Living; HOS-SS, Hip Outcome Score-Sports Subscale; iHOT-12, International Hip Outcome Tool 12; PROMIS, Patient Reported Outcomes Measurement Information System.

$P < .001$), HOS-SS ($r = -0.46$; $P < .001$), and iHOT12 ($r = -0.43$; $P < .001$).

Responsiveness Analysis

The pooled effect sizes for each PROM are summarized in Table 5. Overall, the PROMIS outcomes had smaller differences between mean pre- and postoperative outcomes compared with the legacy PROMs, but were also more precise with standard deviations 2 to 3 times smaller than the legacy PROMs. The pooled effect size was considered large and visible to the naked eye for all PROMs except PROMIS Depression ($d = 0.04$). However, the effect size was largest for iHOT12 ($d = 1.87$) followed by HOS-ADL ($d = 1.29$). PROMIS-PI, PROMIS-PF, and HOS-SS had very similar effect sizes of $d = 0.96$, $d = 0.94$, and $d = 1.06$, respectively.

Based on the results of the correlation analysis and effect size, PROMIS Depression was not included in the comparative responsiveness analysis. The RE between outcomes are summarized in Table 6. If the value is >1, the PROM tool on the left column is defined as more responsive than the respective PROM tool on the top row. The iHOT-12 was more responsive than PROMIS-PI (RE = 3.95), PROMIS-PF (RE = 4.13), HOS-ADL (RE = 2.26), and HOS-SS (RE = 3.84). In contrast, HOS-SS was similarly responsive to PROMIS-PI (RE = 1.03) and PROMIS-PF (RE = 1.08); however, PROMIS-PF was overall the least responsive across all outcome comparisons.

Discussion

In this study, we found that both PROMIS-PF and PROMIS-PI demonstrate good to excellent correlation with legacy hip-specific PROMs. We also found that at minimum 6-month follow-up there were no floor effects on any PROM; however, HOS-ADL and PROMIS-PI demonstrated significant ceiling effects. When comparing responsiveness, we found that the effect size was large for all outcomes, except PROMIS Depression, but largest for the iHOT-12. PROMIS-PF was overall the least responsive and the iHOT-12 was more responsive than PROMIS-PI, PROMIS-PF, HOS-ADL, and HOS-SS. In contrast, HOS-SS was similarly responsive to PROMIS-PF and PROMIS-PI. These findings confirm our hypotheses.

Few studies have previously evaluated the psychometric properties of PROMIS in patients undergoing arthroscopic FAIS. Sheean et al[28] first reported PROMIS scores in 20 patients with FAIS, compared with 22 asymptomatic patients. The authors demonstrated that PROMIS scores were significantly different between FAIS patients and asymptomatic controls. The authors therefore proposed that PROMIS was useful for capturing disability associated with FAIS.[27] More recently, Kollmorgen et al[11] investigated the psychometrics of PROMIS in a mixed population of 125 patients presenting to a tertiary care hip preservation center. The authors reported strong positive Spearman correlations between PROMIS-PF scores and hip-specific legacy scores, iHOT-12, HOS, modified Harris Hip Score, and Veterans RAND-6D (VR-6D). However only 55% of patients were post-surgical (mean follow-up of 5.2 months) and the compositions of surgical procedures was quite heterogeneous.[11] Nwachukwu et al[12] evaluated the preoperative performance of PROMIS-PF in 197 nonoperative patients with FAIS. The authors found good to excellent Pearson correlations between PROMIS and the following hip-specific legacy scores: iHOT-12, HOS-ADL, HOS-SS, and modified Harris Hip Score. Each PROM was measured preoperatively in a cohort with 100% diagnosis of FAIS indicated for hip arthroscopic surgery.[12] The study, however, was limited by a lack of postoperative follow-up and analysis.

In this study, we found that PROMIS-PF correlated well with legacy hip-specific PROMs with the best correlation demonstrated for HOS-SS and HOS-ADL. Interestingly, PROMIS-PF did not correlate as well with iHOT-12. These findings are somewhat in contrast to the prior findings reported by Nwachukwu et al[12]

Table 4. Correlation Analysis of Postoperative Reported Outcomes

	PROMIS Physical Function (*r*)	*P* Value	PROMIS Pain Interference (*r*)	*P* Value	PROMIS Depression (*r*)	*P* Value
HOS-ADL	0.732	<.001*	-0.781	<.001*	-0.457	<.001*
HOS-SS	0.810	<.001*	-0.741	<.001*	-0.455	<.001*
iHOT-12	0.679	<.001*	-0.770	<.001*	-0.426	<.001*

HOS-ADL, Hip Outcome Score-Activities of Daily Living; HOS-SS, Hip Outcome Score-Sports Subscale; iHOT-12, International Hip Outcome Tool 12; PROMIS, Patient Reported Outcomes Measurement Information System.

*Statistical significance with $P < .01$.

Table 5. Pre- vs Postoperative Outcome Score Difference and Pooled Effect Size

	Mean Score Difference	Pooled SD	Pooled Effect Size (*d*)
PROMIS Pain Interference	-6.67	6.99	0.96
PROMIS Depression	0.37	8.39	0.04
PROMIS Physical Function	6.79	7.21	0.94
HOS-ADL	20.37	15.75	1.29
HOS-SS	25.29	23.93	1.06
iHOT-12	38.82	20.77	1.87

HOS-ADL, Hip Outcome Score-Activities of Daily Living; HOS-SS, Hip Outcome Score-Sports Subscale; iHOT-12, International Hip Outcome Tool 12; PROMIS, Patient Reported Outcomes Measurement Information System; SD, standard deviation.

because, in the present study, HOS-SS demonstrates the best correlation with PROMIS. This suggests that although the HOS-SS may only demonstrate good correlation with PROMIS-PF preoperatively, by the postoperative time point HOS-SS and PROMIS-PF may demonstrate excellent correlation. In contrast however, iHOT-12, as demonstrated in the prior Nwachukwu et al[12] study, has excellent preoperative correlation with PROMIS-PF but by the postoperative time point iHOT-12 demonstrated only good correlation, representing a substantial drop. These findings are important as we consider how PROMIS-PF can replace legacy measures. Specifically, the described trend may lend credence to the notion that PROMIS-PF may not fully capture all aspects of hip-specific disability compared with other joint specific forms. Additionally, there may also be a dynamic nature to symptomatology of FAIS as captured by PROMIS and legacy PROMs.

Beyond correlational analysis, we also evaluated the responsiveness of PROMIS and legacy PROMs. Responsiveness is a key factor that describes how the PROM is able to track clinically important changes longitudinally.[22,29] The validity of iHOT-12 is supported because it is able to detect minimally important change in a clinical setting.[30] In this study, we found that the iHOT-12 is the most responsive of hip-specific PROMs and is more responsive than PROMIS-PF. In contrast, however, we found that PROMIS-PF is the least responsive. These findings suggest that among legacy hip-specific PROMs, the iHOT-12 should be preferentially administered alongside PROMIS because of its large effect size and strong responsiveness. Although our study findings highlight that PROMIS is lacking as a disease-specific outcome measure in the hip preservation population, we do believe that there is a role for the collection of PROMIS in these patients. Ultimately, PROMIS instruments can provide comparability across disease states and procedures, thereby allowing for collection of normative data. Widely applicable instruments such as PROMIS can provide benefits for patient counseling and implementation of health policy initiatives. As such, we counsel against replacing legacy hip-specific measures but rather administering PROMIS measures as an adjunct.

Limitations

The current study has certain limitations worth noting. The major limitation in this study is the limited follow-up of 6 months postoperatively. We can support the use of PROMIS and how it correlates with legacy PROMs over a 6-month period, but we are unable to predict long-term outcomes. However, a prior study found that more than one-half of FAIS patients that underwent hip arthroscopy achieved minimally important clinical difference at the 6-month follow-up.[18] In addition, our study population consisted of primarily surgical patients; as such, we are unable to evaluate the longitudinal correlation of PROMIS with legacy hip measures in patients not being treated with hip arthroscopy. Another limitation is that this is a single surgeon series with procedures performed by a high-volume hip arthroscopist, limiting us somewhat in our generalizability. Finally, this study carries the common limitations associated with a nonrandomized retrospective study design, as well as from a small sample size.

Conclusions

In patients at 6 months postoperatively from hip arthroscopy for FAIS, iHOT-12 was the most responsive and had the largest effect size. In contrast, PROMIS-PF had a lower effect size compared with legacy hip-specific PROMs. Additionally, PROMIS-PF did not correlate as well with iHOT-12 compared with HOS-SS.

Table 6. Relative Efficiency Between Pre- and 6-Months Postoperative Outcome Measures

	PROMIS Pain Interference	PROMIS Physical Function	HOS-ADL	HOS-SS	iHOT-12
PROMIS Pain Interference	-	1.05	0.57	0.97	0.25
PROMIS Physical Function	0.96	-	0.55	0.93	0.24
HOS-ADL	1.75	1.83	-	1.70	0.44
HOS-SS	1.03	1.08	0.59	-	0.26
iHOT-12	3.95	4.13	2.26	3.84	-

HOS-ADL, Hip Outcome Score-Activities of Daily Living; HOS-SS, Hip Outcome Score-Sports Subscale; iHOT-12, International Hip Outcome Tool 12; PROMIS, Patient Reported Outcomes Measurement Information System.

References

1. Patient-Reported Outcomes Measurement Information System. Physical Function: A brief guide to the PROMIS Physical Function Instruments. Available from. http://www.healthmeasures.net/images/PROMIS/manuals/PROMIS_Physical_Function_Scoring_Manual.pdf. Accessed January 6, 2020.
2. Cella D, Gershon R, Lai J-S, Choi S. The future of outcomes measurement: item banking, tailored short-forms, and computerized adaptive assessment. *Qual Life Res* 2007;16:133-141 (Suppl 1).
3. 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;41:153-158.
4. Wylie JD, Beckmann JT, Granger E, Tashjian RZ. Functional outcomes assessment in shoulder surgery. *World J Orthop* 2014;5:623-633.
5. Brodke DJ, Saltzman CL, Brodke DS. PROMIS for orthopaedic outcomes measurement. *J Am Acad Orthop Surg* 2016;24:744-749.
6. 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;45:449-453.
7. Anthony CA, Glass N, Hancock K, 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.
8. Dowdle SB, Glass N, Anthony CA, Hettrich CM. Use of PROMIS for patients undergoing primary total shoulder arthroplasty. *Orthop J Sports Med* 2017;5:2325967117726044.
9. Hancock KJ, Glass N, Anthony CA, et al. PROMIS: a valid and efficient outcomes instrument for patients with ACL tears. *Knee Surg Sports Traumatol Arthrosc* 2019;27:100-104.
10. Hancock KJ, Glass N, Anthony CA, et al. Performance of PROMIS for healthy patients undergoing meniscal surgery. *J Bone Joint Surg Am* 2017;99:954-958.
11. Kollmorgen RC, Hutyra CA, Green C, Lewis B, Olson SA, Mather RC. 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;47:876-884.
12. Nwachukwu BU, Beck EC, Chapman R, Chahla J, Okoroha K, Nho SJ. Preoperative performance of the PROMIS in patients undergoing hip arthroscopic surgery for femoroacetabular impingement syndrome. *Orthop J Sports Med* 2019;7:2325967119860079.
13. Dowie J, Kaltoft MK. Why a global PROMIS® can't be kept. *Stud Health Technol Inform* 2019;262:114-117.
14. Griffin DR, Dickenson EJ, O'Donnell J, et al. The Warwick Agreement on femoroacetabular impingement syndrome (FAI syndrome): an international consensus statement. *Br J Sports Med* 2016;50:1169-1176.
15. Byrd JWT, Jones KS. Hip arthroscopy in the presence of dysplasia. *Arthroscopy* 2003;19:1055-1060.
16. Martin RL, Kelly BT, Philippon MJ. Evidence of validity for the hip outcome score. *Arthroscopy* 2006;22:1304-1311.
17. Nwachukwu BU, Chang B, Beck EC, et al. How should we define clinically significant outcome improvement on the iHOT-12? *HSS J* 2019;15:103-108.
18. Nwachukwu BU, Chang B, Adjei J, et al. Time required to achieve minimal clinically important difference and substantial clinical benefit after arthroscopic treatment of femoroacetabular impingement. *Am J Sports Med* 2018;46:2601-2606.
19. Chen CX, Kroenke K, Stump TE, et al. Estimating minimally important differences for the PROMIS pain interference scales: results from 3 randomized clinical trials. *Pain* 2018;159:775-782.
20. Kendall R, Wagner B, Brodke D, et al. The relationship of PROMIS pain interference and physical function scales. *Pain Med* 2018;19:1720-1724.
21. Clover K, Lambert SD, Oldmeadow C, et al. PROMIS depression measures perform similarly to legacy measures relative to a structured diagnostic interview for depression in cancer patients. *Qual Life Res* 2018;27:1357-1367.
22. Terwee CB, Bot SDM, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007;60:34-42.
23. Liang MH, Larson MG, Cullen KE, Schwartz JA. Comparative measurement efficiency and sensitivity of five health status instruments for arthritis research. *Arthritis Rheum* 1985;28:542-547.
24. Nilsdotter AK, Roos EM, Westerlund JP, Roos HP, Lohmander LS. Comparative responsiveness of measures of pain and function after total hip replacement. *Arthritis Rheum* 2001;45:258-262.
25. 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;47:241-247.
26. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. Mahwah, NJ: Lawrence Erlbaum Associates Inc, 1988.
27. Dunlap W, Cortina J, Vaslow J, Burke M. Meta-analysis of experiments with matched groups or repeated measures designs. *Psychol Methods* 1996;1:170-177.
28. Sheean AJ, Schmitz MR, Ward CL, et al. Assessment of disability related to femoroacetabular impingement syndrome by use of the Patient-Reported Outcome Measure Information System (PROMIS) and objective measures of physical performance. *Am J Sports Med* 2017;45:2476-2482.
29. 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;41:582-592.
30. Baumann F, Popp D, Müller K, et al. Validation of a German version of the International Hip Outcome Tool 12 (iHOT12) according to the COSMIN checklist. *Health Qual Life Outcomes* 2016;14:3.