Assessment of Disability Related to Hip Dysplasia Using Objective Measures of Physical Performance

Elizabeth J. Scott,*† MD, Michael C. Willey,† MD, Arthur Mercado,‡ BS, John Davison,† BS, and Jason M. Wilken,† PT, PhD

Investigation performed at the University of Iowa, Iowa City, Iowa, USA

Background: Lower extremity physical performance measures (PPMs), which can objectively quantify functional ability, are an attractive adjuvant to patient-reported outcome (PRO) instruments. However, few tests have been validated for use in hip instability.

Purpose: To evaluate 4 different PPMs for their ability to differentiate between young adults with hip dysplasia indicated for treatment with periacetabular osteotomy (PAO) and asymptomatic controls and to test inter- and intratest reliability and relationship with popular hip PRO instruments.

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

Methods: A total of 24 symptomatic patients aged 15 to 39 years (100% female) with hip dysplasia (lateral center-edge angle $<25^{\circ}$) indicated for treatment with PAO completed the visual analog scale (VAS) for pain, Hip disability and Osteoarthritis Outcome (HOOS) Pain subscale, HOOS Short Version (HOOS PS), International Hip Outcome Tool Short Version (iHOT-12), modified Harris Hip Score (mHHS), Patient Reported Outcome Measurement Information System (PROMIS) physical function (PF) and pain interference (PI), and 4 physical function tests: (1) self-selected walking speed (SSWS), sit-to-stand 5 times (STS5), (3) 4-square step test (FSST), and (4) timed stair ascent (TSA). A further 21 young, asymptomatic adults aged 18 to 39 years (91% female) also underwent testing. Between-group comparisons were made with unpaired *t* test with Bonferroni-Holm correction. Inter- and intrarater reliability was assessed in 38 participants by repeating PPMs at a second visit and using 2 raters. Spearman rank correlation coefficients were used to determine associations between PPMs and PRO instruments.

Results: Significant differences between patients with hip dysplasia and controls were observed for all PRO instruments (HOOS Pain, 47.8 vs 99.2; HOOS PS, 61.9 vs 99.2; iHOT-12, 32.2 vs 99.2; mHHS, 54.5 vs 90.6; PROMIS PF, 41.4 vs 65.6; and PROMIS PI, 62.0 vs 39.1 [all $P < .001$]), and all PPMs (SSWS, 1.21 vs 1.53 m/s; STS5, 10.85 vs 5.95 s; FSST, 6.59 vs 4.03 s [all $P < .001$]; and TSA, 4.58 vs 3.29 s [$P = .002$]). All 4 PPMs demonstrated excellent intra- and intertest reliability (intraclass correlation coefficient, 0.83-0.99). STS5, FSST, and TSA were correlated highly ($r > 0.5$) with physical function PRO instruments, including PROMIS PF, mHHS, and iHOT-12.

Conclusion: Patients with symptomatic hip dysplasia demonstrated significant impairment on functional testing compared with asymptomatic controls, and performance measure testing demonstrated excellent test-retest reliability. Timed stair ascent and sit-to-stand testing in particular were correlated strongly with physical function PRO instruments. PPMs may be a viable and well-received adjuvant to PRO instrument administration for patients with nonarthritic hip conditions, and investigation of the ability of PPMs to assess surgical outcomes for hip dysplasia is warranted.

Keywords: clinical outcomes; performance measures; hip; dysplasia; hip arthroscopy

Hip dysplasia is a common cause of nonarthritic hip pain in the adolescent and young adult, whereby capsular laxity and asymmetry between the femoral head and a nonhemispherical or shallow acetabulum lead to relative increases in joint contact forces with a decrease in weightbearing surface area.[10,29] Common clinical manifestations include generalized hip or groin pain, subjective feelings of instability, and altered gait mechanics. Over time, cartilage degeneration is accelerated from excessive mechanical stress, and ultimately, dysplasia is likely attributable to 20% to 40% of hip osteoarthritis.[8,12,24,34]

Although the effects of femoroacetabular impingement (FAI) on hip range of motion and physical function have been fairly well quantified, functional deficit from symptomatic hip dysplasia in the skeletally mature patient has been relatively less explored. Previous studies have observed that these patients demonstrate decreased walking speed[33] or an abductor lurch or limp[27] and report pain and impairment in physical function including both activities of daily living and sport-specific testing on patient-

The Orthopaedic Journal of Sports Medicine, 8(2), 2325967120903290
DOI: 10.1177/2325967120903290
© The Author(s) 2020

This open-access article is published and distributed under the Creative Commons Attribution - NonCommercial - No Derivatives License (https://creativecommons.org/licenses/by-nc-nd/4.0/), which permits the noncommercial use, distribution, and reproduction of the article in any medium, provided the original author and source are credited. You may not alter, transform, or build upon this article without the permission of the Author(s). For article reuse guidelines, please visit SAGE's website at http://www.sagepub.com/journals-permissions.

reported outcome (PRO) instruments.[37] It is unknown how these perceived deficits objectively limit patients' ability to perform more physically demanding tasks requiring hip flexion, balance, and strength; the literature on hip dysplasia is particularly sparse compared with anterior cruciate ligament deficiency or FAI, for which excellent data on a variety of athletic and squatting maneuvers exist.[1,3,26] As we continue to explore ways to maximize results from treatment for dysplasia with periacetabular osteotomy (PAO),[9] arthroscopy, and/or other osteotomies, means of quantifying physical function both before and after surgery are critically important.

Wilken at al[38] described a series of simple performance measures that can be used without specialized equipment or procedures to objectively quantify physical function and that have been used for patients with a variety of lower extremity injuries. These tests include self-selected walking speed (SSWS), timed stair ascent (TSA), sit-to-stand 5 times (STS5), and 4-square step test (FSST). These tests were recently validated in patients with FAI, and all 4 tests were found to be responsive to FAI-associated disability, revealing detriments in walking speed, agility, balance, and mobility compared with normative controls.[32] Given that these tests require little to no specialized equipment, can be conducted in a relatively small space, and require minimal instruction from trained administrators, they represent a new avenue to improve our ability to track changes in patient function and may be able to improve our understanding of the adult dysplastic hip.

The purpose of this study was (1) to evaluate the ability of these 4 performance measures to differentiate between young adults with hip dysplasia indicated for treatment with PAO and asymptomatic controls, (2) to evaluate the test-retest and interobserver reliability of these physical performance measures (PPMs), and (3) to measure the relationship with PRO instruments well-validated in this population. We hypothesized that patients would demonstrate significant decrements on all 4 tests, that the tests would demonstrate excellent intra- and interobserver reliability, and that excellent correlation would be observed with self-reported measures of patient physical function.

METHODS

Patients

This prospective study was approved by the institutional review board. We enrolled 21 asymptomatic volunteers (controls) and 24 participants aged 15 to 39 years with hip dysplasia scheduled for treatment with PAO. Control participants were recruited from available orthopaedic and emergency medicine residents and staff and members of our institution's orthopaedic biomechanics laboratory. Participants with symptomatic hip dysplasia were enrolled from a hip preservation clinic staffed by M.C.W. and a second fellowship-trained orthopaedic surgeon. Dysplasia was defined as a lateral center-edge angle (LCEA) of $<25^\circ$ measured at the lateral margin of the sourcil on standing anteroposterior (AP) pelvic radiograph; patients undergoing PAO exclusively for acetabular retroversion were, therefore, excluded. All measurements and clinical diagnoses were reviewed and confirmed by both surgeons. Patients with a diagnosed neuromuscular condition, history of Perthes disease, radiographic evidence of osteoarthritis (Tönnis grade >1), or previous open hip surgery were excluded. Participants in the control group did not undergo radiography or other advanced imaging; however, they were instructed to confirm that they did not currently have lower extremity musculoskeletal pain and that they had never received a diagnosis of dysplasia, FAI, or osteoarthritis or undergone hip or femur surgery.

History, clinical examination, radiographs, and advanced imaging were collected from chart review. All patients at our clinic undergo diagnostic imaging consisting of standing AP pelvis, false profile, 45° Dunn lateral, and frog leg radiographs, as well as a low-dose noncontrast pelvic computed tomography scan with selected cuts through the distal femur and noncontrast 3.0-T magnetic resonance imaging. From radiographic series, the LCEA, anterior center-edge angle, Tönnis angle, extrusion index, posterior wall sign, crossover sign, and acetabular index were measured on radiographs by a fellowship-trained surgeon (M.C.W.). Head-neck offset of the femur (alpha angle) was measured from Dunn lateral and frog leg lateral views. Femoral version was measured from low-dose pelvic computed tomography, and magnetic resonance imaging scans were reviewed for presence of labral and cartilage abnormalities. Descriptive and treatment data, including age, sex, laterality, and history of treatment including injection, physical therapy, nonsteroidal medication or previous arthroscopic surgery, were collected from the electronic medical record.

Outcomes Assessment

All participants completed the modified Harris Hip Score (mHHS),[13] International Hip Outcome Tool Short Version (iHOT-12),[11] Hip disability and Osteoarthritis Outcome Score (HOOS) Short Version (HOOS PS)[21] and Pain subscale (HOOS Pain),[18] the Patient Reported Outcome Measurement Information System (PROMIS) physical function

*Address correspondence to Elizabeth J. Scott, MD, Department of Orthopedics and Rehabilitation, University of Iowa, 200 Hawkins Drive, Iowa City, IA 52242, USA (email: elizabeth-j-scott@uiowa.edu) (Twitter: LizScottMD).
†Department of Orthopedics and Rehabilitation, University of Iowa, Iowa City, Iowa, USA.
‡Carver College of Medicine, University of Iowa, Iowa City, Iowa, USA.
Final revision submitted November 3, 2019; accepted November 8, 2019.
One or more of the authors has declared the following potential conflict of interest or source of funding: E.J.S. has received a research grant from the Arthroscopy Association of North America. M.C.W. has received research support from Biomet and hospitality payments from Smith & Nephew 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 received from the University of Iowa Institutional Review Board (ID No. 201802848).

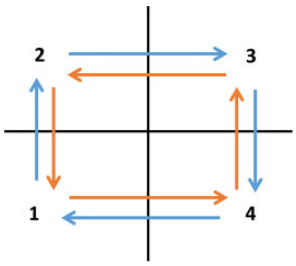

Figure 1. Four-square step test. Starting with both feet in quadrant 1, the participant first steps clockwise over the poles (blue arrows) through quadrants 2, 3, and 4 until arriving with both feet back in quadrant 1. The participant then reverses direction (orange arrows), proceeding back through quadrants 4, 3, and 2 and then returning to the starting place. Timing is stopped when both feet touch the ground in the final position.

and pain interference adaptive tests (PROMIS PF and PROMIS PI),[2] and a visual analog scale (VAS) for pain. PRO instruments were administered in a randomized order through use of a hand-held tablet computer. Participants were also queried on current or recent (within 30 days) use of opioid medication. Immediately after administration of PRO tests and screening questions, the participants underwent functional testing by 1 of 4 trained examiners using 4 previously validated physical performance measures (SSWS, STS5, FSST, and TSA).[32,38] Examiners were not blinded to the grouping of participants. In the SSWS, the participant walked 20 m at a comfortable pace, and the time taken to walk the middle 10 m was recorded. The STS5 was performed with the participant standing up and sitting down 5 times as quickly as possible while keeping the arms crossed over the chest. In the FSST, the participant was required to sequentially step over four 2.5 cm–diameter poles that were laid flat on the floor in the shape of a cross. Participants began with both feet in the left rear square and completed the following pattern of steps as fast as possible: (1) forward, (2) sideways to the right, (3) backward, (4) sideways to the left, (5) sideways to the right, (6) forward, (7) sideways to the left, and (8) backward. At all times, the participant was charged with keeping 1 foot in contact with the ground; timing began once the participant placed a foot into the box in front of him or her and finished once the patient had placed both feet in the final box (Figure 1). In the TSA, the participant ascended 12 stairs as quickly as possible without using a handrail or skipping a step, and timing ended once the participant had both feet on the top step. Participants were not shown the tests until after enrollment; before each task, they were given visual demonstration by the administrator along with verbal instruction, and then they practiced each test once before data were collected. Tests were administered in order of what we hypothesized to be increasing difficulty (walking, standing-sitting, lateral movements, and reciprocal stair climbing) to minimize risk of symptomatic patients stopping before test completion because of pain or fatigue. All testing was performed with the participant in athletic footwear. Guidance and coaching for the participant were standardized by use of a detailed written protocol available to administrators during the performance testing. The time to completion for each trial was recorded, with 2 or 3 trials completed per performance test as previously described by Sheean et al.[32] Repeated trials were averaged for each participant to produce a single final value for each test. At the conclusion of testing, participants completed an electronic exit survey querying them on the perceived difficulty, comfort, and interest in the different tests.

To evaluate test-retest reliability, participants returned for a second performance evaluation at the time of the participant's choosing at a minimum of 24 hours after the first collection. Participants repeated the 4 performance tests in the same order with the same test administrator but did not repeat PRO testing at that time. Additionally, to evaluate interrater reliability, the first 19 participants in both groups (n = 38) were timed simultaneously by 2 administrators (A.M. and J.D.).

Statistical Analysis

Unpaired t tests (alpha = 0.05) with Bonferroni-Holm correction for multiple comparisons were used to compare PRO and performance measure means between dysplasia and control groups. Pearson correlation coefficients were used to describe relationships between physical performance measures and PRO instruments. Correlations were defined as low ($r < 0.03$), moderate ($r = 0.3$-0.5), or high ($r > 0.5$). For physical performance measures and PRO instruments, Cliff delta,[6] median, minimum-maximum, and interquartile range were calculated because data were not normally distributed. Test-retest variability for performance measures was evaluated by use of intraclass correlation coefficients (ICC, 2,1) with the Shrout-Fleiss method. SPSS (Version 25, IBM Corp) was used for these calculations. Before these statistical analyses were completed, all variables were evaluated for normality, and nonparametric methods (Wilcoxon rank sum test instead of t test, Spearman rank instead of Pearson correlation coefficients) were used when indicated.

A priori power analysis indicated that 16 participants in each group would need to complete testing in order to detect an effect size equal to 1 standard deviation for each variable of interest with a power of 80% and an alpha level of .05. Statistical analysis was performed by a trained statistician; SAS Version 9.4 statistical software (SAS Institute) was used for analysis, and a P value $<.05$ was considered statistically significant.

RESULTS

Participant characteristics are shown in Tables 1 and 2. Age, height, weight, and sex were not significantly different between asymptomatic volunteers and the patients with symptomatic hip dysplasia. The mean LCEA for participants with dysplasia was 14°. The average time between successive tests for intrarater reliability was 11 days (median, 1 day). Given the long distance required for travel,

TABLE 1
Characteristics of Participants With Dysplasia (n = 24) and Asymptomatic Controls (n = 21)

	Control Group	Dysplasia Group	*P* Value
Age, y, mean ± SD	25 ± 6.32	24 ± 8.97	.27
Body mass index, mean ± SD	24.38 ± 3.31	23.61 ± 3.92	.59
Female sex, n (%)	19 (90.5)	24 (100)	.21
Opioid use, n (%)			.02
None	21 (100)	19 (79.1)	
Intermittent	0 (0)	4 (16.6)	
Daily	0 (0)	1 (4.1)	

TABLE 2
Additional Characteristics of Participants With Dysplasia (n = 24)[a]

Measurements	Mean	Median	IQR	Range
LCEA, deg	14	16	3.5	–55 to 25
Extrusion index	0.3	0.33	0.07	0.23 to 0.90
Tönnis angle, deg	14	13	5.5	5 to 46
	n (%)			
Tönnis grade >1	4 (16.7)			
Posterior wall sign	4 (16.7)			
Laterality				
Left	8 (33.3)			
Right	16 (66.6)			
Bilateral symptoms	7 (29.2)			

[a]The LCEA was measured on coronal computed tomography. The extrusion index, Tönnis angle, Tönnis grade, and posterior wall sign were measured on anteroposterior standing radiograph. EI, extrusion index; IQR, interquartile range; LCEA, lateral center-edge angle.

6 participants with dysplasia waited until their day of surgery for repeat testing, which was between 40 and 116 days after enrollment and initial testing. Excellent agreement was found during testing of both groups between the 2 independent reviewers, with interrater reliability ranging from 0.977 to 0.999 for the 4 tests (Table 3). Test results were also consistent between the 2 testing sessions for most participants, with intrarater reliability ranging from 0.83 to 0.93 (Table 4).

Of 45 participants included in the study, 1 patient with dysplasia and 2 control participants did not fully complete PRO instruments. Significant differences were found in all PRO measures between groups (Table 5). Patients with dysplasia demonstrated significant decrements in performance for all 4 tested physical performance measures compared with asymptomatic controls (Table 6, Figure 2). Cliff delta was large (*R* value, –0.83 to 1.00) for both PRO and performance measures.

The performance tests involving hip flexion (STS5, TSA) were most strongly associated with lower physical function and sport-specific domains of PROMIS (PF), mHHS (Gait, Activities of Daily Living, Function), and iHOT-12. PROMIS PI was also strongly associated with these tests. FSST, which tests agility and balance, also demonstrated strong correlation with physical domains of PROMIS and mHHS (Table 7). SSWS, which requires the least hip flexion and mobility, demonstrated little correlation with all PRO measures ($r = -0.141$ to 0.262). A separate subanalysis evaluating the Spearman rank correlations for both groups of participants (dysplasia and controls; n = 42) is available in the Appendix.

Subjectively, patients reported high satisfaction (5/5) with performance testing. A total of 3 patients preferred performance testing to PRO instrument administration, with the remainder (n = 20) reporting that both were equally useful. No patients preferred PRO instrument administration. Patients were universally interested (n = 23) in knowing their scores on testing and how these scores compared with peers who had similar conditions.

TABLE 3
Interrater Reliability of Physical Performance Measures (n = 38)[a]

Measure	ICC	95% CI	SEM True	MDC_{95}
SSWS	0.99	0.99-0.99	0.03	0.09
STS5	0.99	0.99-0.99	0.13	0.36
FSST	0.99	0.98-0.99	0.19	0.53
TSA	0.97	0.95-0.98	0.31	0.87

[a]FSST, 4-square step test; ICC, intraclass correlation coefficient; MDC_{95}, minimal detectable change 95% CI; SSWS, self-selected walking speed; STS5, sit-to-stand 5 times; TSA, timed stair ascent.

TABLE 4
Intrarater (Test-Retest) Reliability of Physical Performance Measures (n = 34)[a]

Measure	ICC	95% CI	SEM	MDC_{95}
SSWS	0.93	0.87-0.96	0.13	0.35
STS5	0.83	0.69-0.91	1.70	4.71
FSST	0.93	0.87-0.96	0.60	1.66
TSA	0.9	0.81-0.95	0.65	1.81

[a]FSST, 4-square step test; ICC, intraclass correlation coefficient; MDC_{95}, minimal detectable change 95% CI; SSWS, self-selected walking speed; STS5, sit-to-stand 5 times; TSA, timed stair ascent.

DISCUSSION

Four functional performance tests (SSWS, STS5, FSST, TSA) designed for assessment of lower extremity injuries[34] were evaluated for use with hip dysplasia; we demonstrated statistically significant differences between asymptomatic controls and patients with symptomatic dysplasia ($P < .05$); patients with dysplasia showed substantial decrements on all 4 physical tests. We found that 3 of the 4 tests (STS5, FSST, and TSA) correlated strongly

TABLE 5
Patient-Reported Outcomes for Participants With Dysplasia (n = 23) and Asymptomatic Controls (n = 19)[a]

Outcome	Control Group	Dysplasia Group	*P* Value[b]
VAS pain	0.7 ± 1.5 0-4 (0)	55.0 ± 21.1 10-82 (21)	<.0001
HOOS			
Pain subscale	99.2 ± 3.4 85-100 (0)	47.8 ± 13.7 10-82 (21)	<.0001
Short Version	100.0 ± 0.0 100-100 (0)	61.9 ± 16.1 17.6-87.3 (14.8)	<.0001
iHOT-12	99.2 ± 2.0 91.8-100 (0.18)	32.2 ± 13.0 12.8-58.0 (2.73)	<.0001
PROMIS			
Physical function, *t* score	65.6 ± 7.1 50.0-73.3 (8.1)	41.4 ± 7.1 33.2-61.7 (5.7)	<.0001
Pain interference, *t* score	39.1 ± 1.1 85.0-100 (0)	62.0 ± 5.0 51.9-70.3 (9.1)	<.0001
mHHS, maximum 100			
Pain subscale	43.7 ± 0.9 40.0-44.0 (0)	19.5 ± 6.3 10.0-30.0 (0)	<.0001
ADL subscale	13.8 ± 0.7 11.0-14.0 (0)	9.5 ± 3.0 4.0-14.0 (5)	<.0001
Gait subscale	33.0 ± 0.0 33.0-33.0 (0)	25.3 ± 6.1 6.0-33.0 (6)	<.0001
Function subscale	46.8 ± 0.7 44.0-47.0 (0)	34.9 ± 8.0 13.0-47.0 (10)	<.0001
Total mHHS	90.6 ± 1.6 84.0-91.0 (0)	54.5 ± 13.1 23.0-72.0 (10)	<.0001

[a]Values are expressed as mean ± SD, minimum-maximum (interquartile range). ADL, activities of daily living; HOOS, Hip disability and Osteoarthritis Outcome Score; iHOT-12, International Hip Outcome Tool, Short Form; mHHS, modified Harris Hip Score; PROMIS, Patient Reported Outcome Information System; VAS, visual analog scale.
[b]Indicates significance ($P \leq .05$) determined by Wilcoxon rank sum test.

TABLE 6
Physical Performance Measures for Participants With Dysplasia (n = 24) and Asymptomatic Controls (n = 21)[a]

Measure	Control Group	Dysplasia Group	*P* Value[b]
SSWS, m/s	1.53 ± 0.22 1.23-2.02 (1.54)	1.21 ± 0.21 0.79-1.59 (1.18)	<.0001
STS5, s	5.95 ± 1.15 4.28-9.22 (0.99)	10.85 ± 4.43 4.86-22.2 (5.08)	<.0001
FSST, s	4.03 ± 0.66 3.13-5.56 (0.99)	6.59 ± 2.51 3.95-13.91 (2.94)	<.0001
TSA, s	3.29 ± 0.33 2.64-3.88 (0.33)	4.58 ± 2.69 2.53-16.23 (3.83)	.0027

[a]Values are expressed as mean ± SD, minimum-maximum (interquartile range). FSST, 4-square step test; SSWS, self-selected walking speed; STS5, sit-to-stand 5 times; TSA, timed stair ascent.
[b]Indicates significance ($P \leq .05$) determined by an unpaired *t* test.

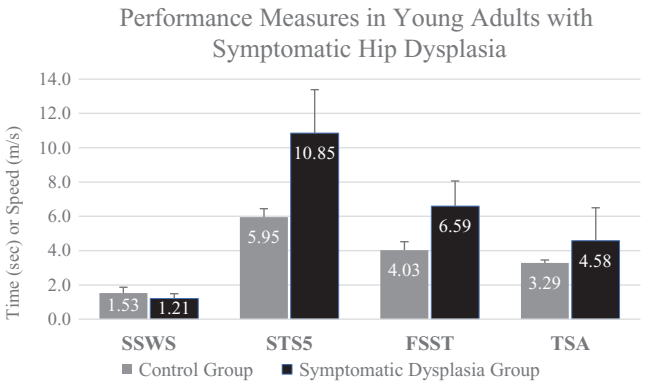

Figure 2. Performance measures in young adults with symptomatic hip dysplasia and controls. SSWS is reported in meters per second (m/s); STS5, FSST, and TSA are reported in seconds (s). FSST, 4-square step test; SSWS, self-selected walking speed; STS5, sit-to-stand 5 times; TSA, timed stair ascent.

TABLE 7
Spearman Rank Correlations for Participants With Hip Dysplasia (n = 23)[a]

	SSWS	STS5	FFST	TSA
VAS pain	0.061	0.364	0.211	0.262
HOOS Pain	–0.141	–0.418	–0.196	–0.253
HOOS PS	–0.251	–0.329	–0.120	–0.205
iHOT-12	0.001	**–0.632**	–0.413	**–0.538**
PROMIS PF	0.302	**–0.656**	**–0.575**	**–0.644**
PROMIS PI	–0.163	**0.598**	0.387	**0.544**
mHHS Pain	0.138	–0.499	–0.272	–0.298
mHHS ADL	0.069	**–0.528**	–0.439	**–0.506**
mHHS Gait	0.262	**–0.654**	**–0.652**	**–0.711**
mHHS Function	0.189	**–0.650**	**–0.618**	**–0.684**
mHHS	0.190	**–0.626**	**–0.519**	**–0.576**

[a]Values indicating high correlation ($r > 0.5$) are presented in boldface. Rank correlations for all participants (n = 42) are available in the Appendix. ADL, activities of daily living; FSST, 4-square step test; HOOS, Hip disability and Osteoarthritis Outcome Score; iHOT-12, International Hip Outcome Tool, Short Form; mHHS, modified Harris Hip Score; PF, physical function; PI, pain interference; PROMIS, Patient Reported Outcome Information System; PS, Short Version; SSWS, self-selected walking speed; STS5, sit-to-stand 5 times; TSA, timed stair ascent; VAS, visual analog scale.

with PRO instruments associated with physical function, including PROMIS PF (–0.656, –0.575, –0.644), mHHS Gait (–0.654, –0.652, –0.711), mHHS Function (–0.650, –0.618, –0.684), and total mHHS (–0.626, –0.519, –0.576). SSWS demonstrated the weakest association with PRO measures, which may be attributable to the test requiring little in the way of hip flexion or dynamic movement. Our data support the hypothesis that SSWS, STS5, FSST, and TSA are reliable measures that serve as a quantitative means of assessing functional decrements in patients with dysplasia. Additionally, subjective data

from patients indicated that these tests were well-tolerated and patients were interested in understanding how their scores compared with those of their peers.

Our findings are consistent with previous studies demonstrating that adult participants with hip dysplasia walk with decreased gait speed and strike.[16] Other studies also have reported decreased hip extension, increased hip abduction, and reduced hip flexor movement during walking, which corresponds with the impairment we observed in hip flexion–based tests.[15,28,30,33]

We observed excellent interrater reliability (ICC, 0.977-0.999) and reliability between days (ICC, 0.83-0.93), similar to the findings reported by Wilken et al[38] in a study of asymptomatic military recruits (interrater reliability, 0.97-0.99; test-retest reliability, 0.86-0.93). Notably, the standard deviation for our cohort was substantially larger than that reported by Wilken et al; given that our population included both healthy controls and symptomatic patients with varying functional deficits, this is unsurprising.

Our participants' scores on PRO testing align with previously reported preoperative hip pain and function measures in patients with dysplasia before PAO. Wasko et al[37] observed a preoperative HOOS Pain score of 54.4 ± 20.7 and mHHS of 60.0 ± 14.8 in 294 patients. In 336 patients (420 hips), Livermore et al[19] reported PROMIS PF adjusted mean scores of 52.3, 47.3, and 51.0 for mild, moderate, and severe dysplasia, respectively, before surgery. A baseline value for iHOT-12 has not previously been reported, to our knowledge; however, the test's psychometric properties for patients with hip dysplasia are similar to those of the original iHOT-33.[35] Of note, 3 of our participants did not fully complete PRO instruments. Of these, 2 instances were the result of technical failures with testing software; in the third case, the participant left the testing area before completing PRO instruments. Use of newer alternative delivery methods for PRO instruments, such as text-messaging or smartphone application, may have prevented these events.

Interestingly, despite reporting substantial pain on pain-related PRO instruments, such as HOOS Pain, VAS, and PROMIS PI, many of the patients with dysplasia were able (and willing) to complete physical function testing at a relatively high level; although their times were slower than those of control participants, the patients with dysplasia were not as slow as we had anticipated in someone about to undergo a large pelvic osteotomy. In other words, even though the patients with dysplasia reported significant functional pain, they universally demonstrated ability and willingness to push through their reported discomfort for the purpose of testing when asked, and they responded that they found testing equal to or more useful than standard PRO instrument administration (23/23; 100%). We did not collect data on grit, self-efficacy, or other measures of mental fortitude, but this may be worth considering in future studies when identifying patient populations who may respond well to physical performance testing. We did not track the number of patients who declined enrollment; however, this may also be worth evaluating in future studies.

Regarding physical performance, overall our patients experienced disability comparable with that described by Wilken et al[38] in adult patients with hip impingement (FAI). On the STS5 test, which involves deep hip flexion and hip flexor strength, our patients with dysplasia performed at a level equal to that of the patients with FAI (10.85 vs 10.75 s), which was worse than young veterans with transtibial amputation (7.8 s, 8.4 s),[22,38] similar to older control participants (9.9-13.4 s),[7] and superior to older participants with knee osteoarthritis (13.3 s).[5] On the FSST, which tests balance and lateral movements, our patients' impairment was similar to that of transtibial amputees (6.59 vs 6.0 s), slightly better than that of FAI patients (8.89 s), and superior to that of older adult controls (8.7).[5] On the TSA, which tests agility and power, patients with dysplasia performed in a manner similar to patients with transtibial amputation (4.8 vs 4.58 s); this was faster than the FAI cohort described by Wilken et al[38] (5.92 s). Because impingement is typically associated with hip flexion, this finding may reflect a distinct functional difference between these 2 populations. An unstable hip without impingement may be comparatively less aggravated by high hip-flexion tasks, such as stair climbing, than a hip with FAI. Although more study would be required, it is possible that this test could even help distinguish between abnormalities. Self-selected walking speed was slower in our dysplasia cohort than in the patients with FAI (1.21 vs 1.32 m/s); we hypothesize that this is more a reflection of age and sex, although our controls performed equal to theirs for this particular test (1.53 vs 1.51 m/s). Normative data indicates women typically walk slower than men due to differences in limb and stride length. Our study population was entirely female, with an average age 24 years; in the Wilken et al[38] study of military recruits, 25% were female, with a mean age of 32.2 years.

Our study found that physical function PRO instruments including PROMIS PF and mHHS were correlated strongly with physical function testing ($r > 0.5$; Table 6). This high correlation suggests that patients are able to adequately rate their function using PRO instruments (ie, not grossly over- or underestimating their perceived difficulty). In comparison, pain-based metrics including HOOS Pain, VAS, and PROMIS PI were correlated poorly, with patients scoring very high on some pain metrics but demonstrating comparatively small deficits in function. Pain-based PRO instruments appeared to evaluate domains not represented with physical function testing for this population. Accordingly, PRO instruments that combine pain and physical function domains together (ie, iHOT) were correlated relatively poorly, as well. Although this may be viewed as a limitation by some, we believe that this finding supports the objective nature of PPMs. PPMs could be used, for instance, as a benchmark to show patients that they are improving in function over time despite pain. Because PPMs can be completed safely even by those with significant disability, these measures may also serve as a better "report card" to share with patients with dysplasia during treatment or rehabilitation than more difficult tests, such as the single-leg squat or Star Excursion Balance Test proposed for detecting FAI-related disability.[17,25] PPMs also may be useful in some situations as a way of corroborating results of a functional PRO instrument. Like

PRO instruments, however, PPMs cannot be used to distinguish between specific causes of hip pain or disfunction. Providers interested in using PPMs should be aware of how these tests are correlated with pain-related PRO instruments and continue to use a variety of pain metrics in their practice.

Limitations

This study has a number of limitations. Foremost, the small number of participants in both control and dysplasia groups limited the strength of our results. Based on a priori power analysis of a similar publication in patients with FAI,[32] however, an adequate number of participants were enrolled in each group to detect a significant difference in our functional measures. Greater numbers may perhaps result in stronger correlation with HOOS and its subscales or with pain-related metrics. We selected the PRO instruments used in this study (HOOS Pain and PS, PROMIS PF and PI, mHHS, VAS) based on current literature regarding the psychometric properties of PRO instruments for patients undergoing PAO[1] and hip arthroscopy.[4,14,20,37] We elected to omit other well-validated PRO instruments in order to limit participant fatigue with testing. Inclusion of other hip metrics, including the Western Ontario and McMaster Universities Osteoarthritis Index and the Hip Outcome Score,[23] could be considered in future study. Further evaluation of the relationship between physical function scores and quality-of-life metrics assessing the emotional state of patients, such as the 12-item Short Form Health Survey[36] or other scales for anxiety, depression, and pain catastrophizing, would also be worthwhile considering the substantial physical effort patients had to put forth for accurate collection of performance measures. Additionally, the restrictions innate to performance measures and their accurate collection should be noted and include the need for dedicated and unobstructed space, an individual trained in collection, appropriate footwear for testing, and cooperation on the part of the patient. Without any one of these features, testing results may be inaccurate. Additionally, environmental considerations, such as lack of a staircase may limit the ability for some clinics to collect performance measures.

Finally, when our data are applied to future studies in dysplasia, variations in sex, height, and body mass should be considered. Our participants were primarily females with normal body mass index; this reflects the epidemiological patterns of hip dysplasia (6:1 female to male)[31]; however, normative values may need to be established for mixed-sex populations or patients with elevated body mass index in the future in order to directly compare data with populations in other regions. We also chose to include 5 hips with borderline dysplasia (center-edge angles 19°-25°) in our study; subanalysis in a larger population may reveal a relationship between specific dysplastic features and functional ability, as well. Comparison with patients with isolated acetabular retroversion or with combined dysplastic and impingement features may also be worth considering in future study. Finally, it is unknown how individuals with asymptomatic or minimally symptomatic dysplasia might perform on PPMs; comparison with such a group (perhaps by administering a lidocaine injection or adding a third cohort) could help delineate between pain-related and mechanical (anatomic) causes of dysplasia-related functional impairment.

CONCLUSION

This study describes 4 physical performance measures that may be of interest to the clinician in quantifying disability in patients with hip dysplasia. Patients with dysplasia demonstrated marked decrements in functional ability when evaluated with simple tests requiring balance, agility, hip strength, and walking speed. These decrements were correlated well with self-reported limitations in physical function on PRO testing; alternatively, looser correlation with pain-based PRO instruments demonstrated that patient perception of pain remains a distinct domain less reflected by functional ability. We propose incorporation of physical function measures into future prospective studies to more precisely define the effects of treatment for dysplasia. Additional research could determine the ability for these measures to quantify postoperative change in physical function.

ACKNOWLEDGMENT

Natalie Glass, PhD, assisted in statistical analysis for this project. Shannon Ortiz provided research assistance.

REFERENCES

1. Beckmann JT, Havrilak EE, Millis MB, Wylie JD. Functional outcome assessment in hip preservation surgery. *JBJS Rev*. 2018;6(7):e6.
2. 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.
3. Cheatham SW, Stull KR, Fantigrassi M, Montel I. Hip musculoskeletal conditions and associated factors that influence squat performance: a systematic review. *J Sport Rehabil*. 2018;27(3):263-273.
4. Childs S, Canham C, Kenney RJ, Silas DR, Adler K, Giordano BD. Correlation of PROMIS CAT with validated hip outcome scores in patients undergoing hip arthroscopy. *Orthop J Sports Med*. 2017; 5(7)(suppl 6):2325967117S2325900451.
5. Christiansen CL, Stevens-Lapsley JE. Weight-bearing asymmetry in relation to measures of impairment and functional mobility for people with knee osteoarthritis. *Arch Phys Med Rehabil*. 2010;91(10): 1524-1528.
6. Cliff N. Ordinal analysis of behavioral data. In: Weiner B, ed. *Handbook of Psychology*. 6th ed. New York, NY: Wiley; 2003:635-661.
7. Dite W, Temple VA. A clinical test of stepping and change of direction to identify multiple falling older adults. *Arch Phys Med Rehabil*. 2002; 83(11):1566-1571.
8. Gala L, Clohisy JC, Beaule PE. Hip dysplasia in the young adult. *J Bone Joint Surg Am*. 2016;98(1):63-73.
9. Ganz R, Klaue K, Mast J. Peri-acetabular reorientation osteotomy [in French]. *Acta Orthop Belg*. 1990;56(1, pt B):357-369.
10. Ganz R, Leunig M, Leunig-Ganz K, Harris WH. The etiology of osteoarthritis of the hip: an integrated mechanical concept. *Clin Orthop Relat Res*. 2008;466(2):264-272.

11. Griffin DR, Parsons N, Mohtadi NG, Safran MR. A short version of the International Hip Outcome Tool (iHOT-12) for use in routine clinical practice. *Arthroscopy*. 2012;28(5):611-616.
12. Harris NH, Lloyd-Roberts GC, Gallien R. Acetabular development in congenital dislocation of the hip: with special reference to the indications for acetabuloplasty and pelvic or femoral realignment osteotomy. *J Bone Joint Surg Br*. 1975;57(1):46-52.
13. Harris WH. 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. *J Bone Joint Surg Am*. 1969;51(4):737-755.
14. Hinman RS, Dobson F, Takla A, O'Donnell J, Bennell KL. Which is the most useful patient-reported outcome in femoroacetabular impingement? Test-retest reliability of six questionnaires. *Br J Sports Med*. 2014;48(6):458-463.
15. Jacobsen JS, Nielsen DB, Sørensen H, Søballe K, Mechlenburg I. Changes in walking and running in patients with hip dysplasia. *Acta Orthop*. 2013;84(3):265-270.
16. Karam MD, Gao Y, McKinley T. Assessment of walking pattern pre and post peri-acetabular osteotomy. *Iowa Orthop J*. 2011;31:83-89.
17. Kivlan BR, Martin RL. Functional performance testing of the hip in athletes: a systematic review for reliability and validity. *Int J Sports Phys Ther*. 2012;7(4):402-412.
18. 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.
19. Livermore AT, Anderson LA, Anderson MB, Erickson JA, Peters CL. Correction of mildly dysplastic hips with periacetabular osteotomy demonstrates promising outcomes, achievement of correction goals, and excellent five-year survivorship. *Bone Joint J*. 2019;101-B(6)(suppl B):16-22.
20. Lodhia P, Slobogean GP, Noonan VK, Gilbart MK. Patient-reported outcome instruments for femoroacetabular impingement and hip labral pathology: a systematic review of the clinimetric evidence. *Arthroscopy*. 2011;27(2):279-286.
21. Lyman S, Lee YY, Franklin PD, Li W, Mayman DJ, Padgett DE. Validation of the HOOS, JR: a short-form hip replacement survey. *Clin Orthop Relat Res*. 2016;474(6):1472-1482.
22. Lythgo N, Marmaras B, Connor H. Physical function, gait, and dynamic balance of transfemoral amputees using two mechanical passive prosthetic knee devices. *Arch Phys Med Rehabil*. 2010;91(10):1565-1570.
23. Martin RL, Kelly BT, Philippon MJ. Evidence of validity for the hip outcome score. *Arthroscopy*. 2006;22(12):1304-1311.
24. Mavcic B, Iglic A, Kralj-Iglic V, Brand RA, Vengust R. Cumulative hip contact stress predicts osteoarthritis in DDH. *Clin Orthop Relat Res*. 2008;466(4):884-891.
25. McGovern RP, Christoforetti JJ, Martin RL, Phelps AL, Kivlan BR. Evidence for reliability and validity of functional performance testing in the evaluation of nonarthritic hip pain. *J Athl Train*. 2019;54(3):276-282.
26. Mullins K, Hanlon M, Carton P. Differences in athletic performance between sportsmen with symptomatic femoroacetabular impingement and healthy controls. *Clin J Sport Med*. 2018;28(4):370-376.
27. Nunley RM, Prather H, Hunt D, Schoenecker PL, Clohisy JC. Clinical presentation of symptomatic acetabular dysplasia in skeletally mature patients. *J Bone Joint Surg Am*. 2011;93(suppl 2):17-21.
28. Pedersen EN, Simonsen EB, Alkjaer T, Soballe K. Walking pattern in adults with congenital hip dysplasia: 14 women examined by inverse dynamics. *Acta Orthop Scand*. 2004;75(1):2-9.
29. Pompe B, Daniel M, Sochor M, Vengust R, Kralj-Iglic V, Iglic A. Gradient of contact stress in normal and dysplastic human hips. *Med Eng Phys*. 2003;25(5):379-385.
30. Romano CL, Frigo C, Randelli G, Pedotti A. Analysis of the gait of adults who had residua of congenital dysplasia of the hip. *J Bone Joint Surg Am*. 1996;78(10):1468-1479.
31. Sankar WN, Duncan ST, Baca GR, et al. Descriptive epidemiology of acetabular dysplasia: the Academic Network of Conservational Hip Outcomes Research (ANCHOR) periacetabular osteotomy. *J Am Acad Orthop Surg*. 2017;25(2):150-159.
32. 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(11):2476-2482.
33. Skalshoi O, Iversen CH, Nielsen DB, et al. Walking patterns and hip contact forces in patients with hip dysplasia. *Gait Posture*. 2015;42(4):529-533.
34. Solomon L. Patterns of osteoarthritis of the hip. *J Bone Joint Surg Br*. 1976;58(2):176-183.
35. Thorborg K, Tijssen M, Habets B, et al. Patient-reported outcome (PRO) questionnaires for young to middle-aged adults with hip and groin disability: a systematic review of the clinimetric evidence. *Br J Sports Med*. 2015;49(12):812.
36. Ware J Jr, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med Care*. 1996;34(3):220-233.
37. 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.
38. Wilken JM, Darter BJ, Goffar SL, et al. Physical performance assessment in military service members. *J Am Acad Orthop Surg*. 2012;20(suppl 1):S42-S47.

APPENDIX

Spearman Rank Correlations for All Patients (Combined Controls and Patients With Dysplasia, n = 42)[a]

	STS5	FSST	SSWS	TSA
VAS pain	0.66381	0.63787	-0.53233	0.47502
HOOS Pain	–0.70655	–0.65941	0.51196	–0.51549
P	<.0001	<.0001	.0005	.0005
n	42	42	42	42
HOOS PS	–0.66008	–0.64945	0.47246	–0.48293
P	<.0001	<.0001	.0016	.0012
n	42	42	42	42
iHOT-12	–0.71379	–0.73123	0.52567	–0.58414
P	<.0001	<.0001	.0004	<.0001
n	42	42	42	42
PROMIS PF	–0.71039	–0.72504	0.58167	–0.61973
P	<.0001	<.0001	<.0001	<.0001
n	42	42	42	42
PROMIS PI	0.7272	0.7234	–0.58174	0.57804
P	<.0001	<.0001	<.0001	<.0001
n	42	42	42	42
mHHS Pain	–0.68844	–0.65498	0.59821	–0.47768
P	<.0001	<.0001	<.0001	.0016
n	41	41	41	41
mHHS ADL	–0.71012	–0.66512	0.5312	–0.60341
P	<.0001	<.0001	.0004	<.0001
n	41	41	41	41
mHHS Gait	–0.69205	–0.75671	0.59089	–0.63786
P	<.0001	<.0001	<.0001	<.0001
n	41	41	41	41
mHHS Function	–0.7382	–0.75311	0.60167	–0.65042
P	<.0001	<.0001	<.0001	<.0001
n	41	41	41	41
mHHS total	–0.72848	–0.71664	0.59795	–0.57771
P	<.0001	<.0001	<.0001	<.0001
n	41	41	41	41

[a]ADL, activities of daily living; FSST, 4-square step test; HOOS, Hip disability and Osteoarthritis Outcome Score; iHOT-12, International Hip Outcome Tool, Short Form; mHHS, modified Harris Hip Score; PF, physical function; PI, pain interference; PROMIS, Patient Reported Outcome Information System; PS, Short Version; SSWS, self-selected walking speed; STS5, sit-to-stand 5 times; TSA, timed stair ascent; VAS, visual analog scale.