

# Failure to Achieve Threshold Scores on Patient-Reported Outcome Measures Within 1 Year Has a Predictive Risk of Subsequent Hip Surgery Within 5 Years of Primary Hip Arthroscopy

## A Case-Control Study

Jacob D. Feingold,<sup>\*†</sup> BS, Erica L. Swartwout,<sup>†</sup> BA, Sacha A. Roberts,<sup>†</sup> BS, Benedict U. Nwachukwu,<sup>†</sup> MD, and Anil S. Ranawat,<sup>†</sup> MD

*Investigation performed at Hospital for Special Surgery, New York, New York, USA*

**Background:** Research has identified factors that influence achievement rates of threshold scores on patient-reported outcome measures (PROMs) after hip arthroscopy. However, little data exist on whether failure to achieve a threshold score (minimal clinically important difference [MCID] or substantial clinical benefit [SCB]) in the short term after hip arthroscopy predicts the risk of future hip surgery.

**Purpose/Hypothesis:** The purpose of this study was to determine if failure to achieve the MCID or SCB on PROMs within 1 year of hip arthroscopy can be considered a risk factor for repeat surgery within 5 years of primary hip arthroscopy. It was hypothesized that failure to achieve threshold scores would increase the risk of subsequent hip surgery.

**Study Design:** Case-control study; Level of evidence, 3.

**Methods:** A retrospective review of prospectively collected data was completed. Four PROMs were collected preoperatively and within 1 year of hip arthroscopy: modified Harris Hip Score, Hip Outcome Score–Activities of Daily Living, Hip Outcome Score–Sports, and the 33-Item International Hip Outcome Tool.

**Results:** Two cohorts were formed: (1) a study cohort ( $n = 88$ ) composed of patients who underwent repeat hip surgery within 5 years of hip arthroscopy and (2) a control cohort ( $n = 288$ ) composed of patients who did not require repeat hip surgery. The study cohort had significantly ( $P < .001$ ) lower scores on all postoperative PROMs, and a significantly ( $P < .001$ ) smaller percentage of the study cohort met the MCID and SCB. Multivariable regression analysis demonstrated that not achieving the MCID or SCB on each of the PROMs is an independent risk factor for repeat hip surgery. For every PROM in which a patient failed to achieve the MCID, the odds of subsequent surgery increased by 1.68 (95% CI, 1.42-1.98;  $P < .001$ ). For every PROM in which a patient failed to achieve the SCB, the odds of subsequent surgery increased by 1.63 (95% CI, 1.35-1.97;  $P < .001$ ).

**Conclusion:** Failure to meet threshold scores on PROMs after hip arthroscopy was an independent risk factor for subsequent hip surgery. This study establishes a novel utility of PROMs and confirms the importance of these metrics in the orthopaedic literature.

**Keywords:** hip arthroscopy; femoroacetabular impingement; hip preservation; sports medicine; PROMs

The number of patients undergoing hip arthroscopy for symptomatic femoroacetabular impingement (FAI) with or without a concomitant acetabular labral tear continues to increase annually.<sup>17,23,37</sup> This can be attributed to the documented success of the operation in randomized control

trials, prospective clinical studies, and retrospective cohort studies<sup>22</sup> combined with lower rates of complications compared with open surgical approaches.<sup>21</sup>

The metrics used to define the outcome of hip arthroscopy as a success or failure are numerous and include revision surgery, physical examination, and return to sport.<sup>3,4,11,35,39</sup> With the ongoing shift in medical care toward a patient-centered approach, patient-reported outcome measures (PROMs) have become a focal point of hip

The Orthopaedic Journal of Sports Medicine, 9(11), 23259671211053012  
DOI: 10.1177/23259671211053012  
© The Author(s) 2021

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>.

arthroscopy literature.<sup>1,15,19,24</sup> These outcome measures are completed by the patient, instead of the clinician, and therefore represent an accurate patient-centered perspective on the benefits of medical care. However, a clinical interpretation of PROMs can be a unique challenge because statistically significant changes in PROMs are sometimes associated with clinically insignificant differences.<sup>16,31</sup> This can make it difficult for a practicing physician to extrapolate data on PROMs to everyday practice. However, the challenge is somewhat overcome via utilization of the minimal clinically important difference (MCID) and substantial clinical benefit (SCB). While the value of the MCID and SCB can be dependent on the particular technique utilized for calculation, these are numeric changes in PROMs that indicate a significant improvement in clinical function.<sup>10,14,27</sup>

The MCID was initially introduced to determine the significance of any change in PROMs<sup>10</sup> and is now defined as the smallest change in outcome score a patient considers meaningful.<sup>27</sup> The SCB is a similar value that ascertains the upper threshold of change in outcome score and is defined as improvements perceived as clinically considerable.<sup>14</sup> Achievement of these threshold scores after hip arthroscopy allows a clinician to state that the operation was a success. As a result, a central component of hip-preservation research has focused on the identification of preoperative factors that increase or decrease the likelihood of achieving the MCID or SCB. However, little data exist on the clinical progression of FAI for patients that fail to meet a threshold outcome score such as the MCID and/or SCB after hip arthroscopy. This concept has been investigated in the field of total hip arthroplasty (THA), with the results demonstrating that failure to meet a threshold score is predictive of revision risk status post-primary THA.<sup>8,33,38</sup> Based on these findings, it is reasonable to infer a similar trend in hip-preservation surgery.

The purpose of this study was to determine if failure to meet the MCID and/or SCB within 1 year of hip arthroscopy can be considered an independent risk factor for subsequent hip surgery within 5 years of primary hip arthroscopy. Our principal hypothesis was that failure to achieve the MCID and/or SCB within 1 year of hip arthroscopy would significantly increase the risk of subsequent hip surgery within 5 years of the primary procedure. In addition, we hypothesized that the risk of subsequent surgery would increase for each additional outcome measure in which a patient failed to achieve the MCID and/or SCB.

## METHODS

This study was approved by our organization's institutional review board. PROM data utilized in the present study were collected from 2006 to 2018 via a separate institutional review board-approved hip preservation registry. The registry is a standard-of-care, multisurgeon, hospital-wide registry that tracks the clinical course and PROMs of patients diagnosed with symptomatic FAI that is managed medically and/or surgically. PROMs are collected prospectively for all patients preoperatively and at set time intervals after hip arthroscopy. In the present study, patients were included from four high-volume, fellowship-trained hip arthroscopy surgeons. A retrospective chart review was then completed to identify patients who could be included in the analyses.

All patients were assessed before and after hip arthroscopy with 4 hip specific PROMs: the modified Harris Hip Score (mHHS),<sup>2</sup> Hip Outcome Score—Activities of Daily Living (HOS-ADL),<sup>18</sup> Hip Outcome Score—Sports (HOS-Sports),<sup>20</sup> and the 33-Item International Hip Outcome Tool (iHOT-33).<sup>25</sup> All postoperative PROMs were collected between 5 and 11 months after primary hip arthroscopy. Previously established values for the MCID and SCB using outcomes for patients 1 year post-hip arthroscopy<sup>12,19,25,26</sup> were applicable for the present study and therefore utilized for statistical analyses. MCID values of 8.2, 8.3, 14.5, and 12.1<sup>27</sup> and SCB values of 19.8, 10.0, 29.9, and 24.5<sup>26</sup> were used on mHHS, HOS-ADL, HOS-Sports, and iHOT-33, respectively.

Exclusion criteria consisted of (1) prior ipsilateral hip surgery, (2) history of traumatic ipsilateral hip injury, or (3) radiographically confirmed hip osteoarthritis (Tönnis grade >1). Further patients were excluded from the study if they had not completed at least one of the outcome measures before surgery and the same outcome measure within one year of their primary procedure.

After application of the aforementioned criteria, 1,509 patients who met the criteria were included in initial chart review. Chart review was completed to identify patients (1) who underwent subsequent ipsilateral hip surgery within 5 years of their primary procedure and (2) who were greater than 5 years removed from hip arthroscopy and did not require additional hip surgery. We initially identified 101 patients who underwent subsequent surgery; however, 13 of these patients underwent surgery longer than 5 years after their initial hip arthroscopy and were therefore excluded from the study. Of the 1408 patients who did not undergo additional surgery after their primary procedure,

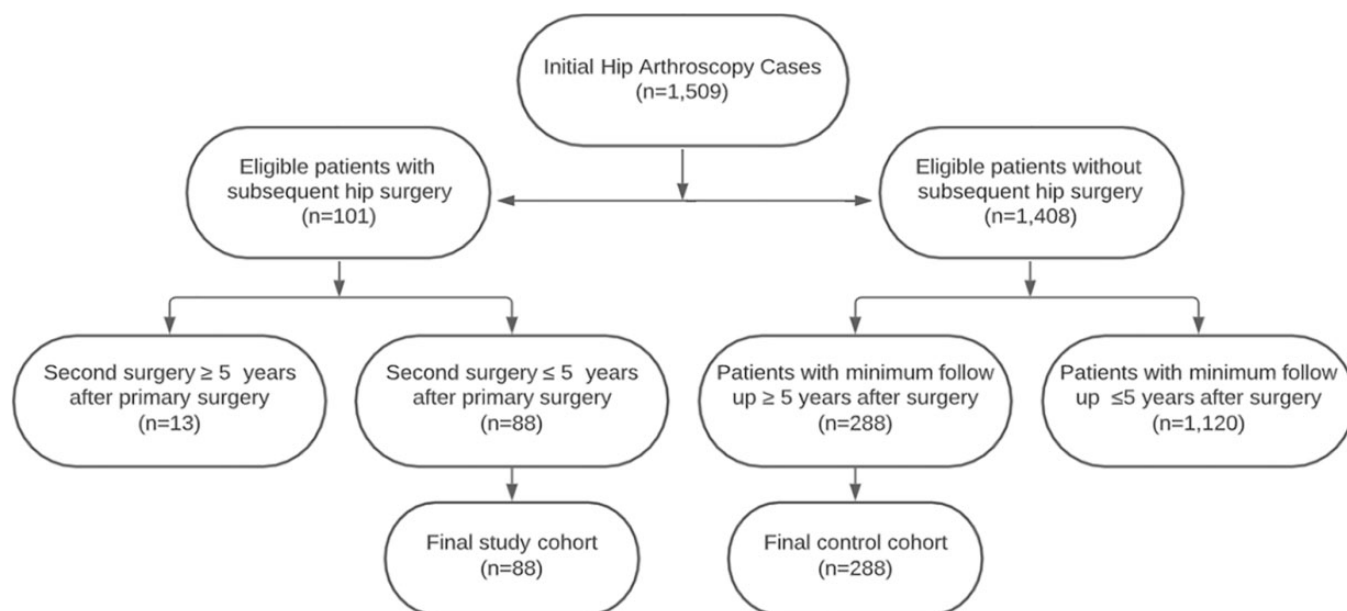
\*Address correspondence to Jacob D. Feingold, BS, Hospital for Special Surgery, 541 E 71st Street, 3rd Floor, New York, NY 10021, USA (email: Jfeingold56@gmail.com).

<sup>†</sup>Hospital for Special Surgery, New York, New York, USA.

Final revision submitted July 3, 2021; accepted August 4, 2021.

One or more of the authors has declared the following potential conflict of interest or source of funding: B.U.N. has received grant support from Arthrex, education payments from Arthrex and Smith & Nephew, and hospitality payments from Stryker and Wright Medical, and is the founder of and has an ownership interest in BICMID. A.S.R. has received consulting fees from Arthrex, Anika, Bodycad, Enhatch, Marrow Cellution, NewClip, Ranfac, Smith & Nephew, and Xiros; and speaking fees from Pfizer; and has an ownership interest in Conformis and Enhatch. AOSSM checks author disclosures against the Open Payments Database (OPD). AOSSM has not conducted an independent investigation on the OPD and disclaims any liability or responsibility relating thereto.

Ethical approval for this study was obtained from the institutional review board of Hospital for Special Surgery (study No. 2020-1840).



**Figure 1.** Flowchart of patient inclusion and exclusion process for this study.

1120 of these patients either were less than 5 years removed from primary surgery ( $n = 585$ ) or had not returned for an office visit greater than 5 years after their primary procedure ( $n = 535$ ) and were thus excluded from the study.

After application of the above parameters, the study cohort consisted of 88 patients, whereas the control cohort consisted of 288 patients. The flowchart documenting patient selection can be seen in Figure 1. Of the 88 patients in the study cohort, 86 completed mHHS and HOS-ADL both pre- and postoperatively, 84 completed HOS-Sports both pre- and postoperatively, and 64 completed iHOT-33 both pre- and postoperatively. Of the 288 patients in the control cohort, 284 completed HOS-ADL both pre- and postoperatively, 278 completed mHHS both pre- and postoperatively, 274 completed HOS-Sports both pre- and postoperatively, and 222 completed iHOT-33 both pre- and postoperatively.

### Statistical Analysis

In each of the cohorts, patient characteristics and outcome measures were described. Continuous variables were reported as means with ranges, while categorical variables were reported as counts and percentages. Chi-square and  $t$  tests were used to compare categorical and continuous variables between the 2 cohorts, respectively. Multivariable mixed-effects logistic regression was used to determine if failure to meet the MCID or SCB for each PROM within 1 year of surgery is associated with undergoing a subsequent surgery. A final multivariable model was run to assess the impact of failing to meet the MCID or SCB on more than one outcome measure. Age, sex, and baseline PROM score were controlled for in each multivariable

analysis. Differences with a  $P < 0.05$  were considered statistically significant in the present study.

The sensitivity, specificity, positive likelihood ratio (LR+), and negative likelihood ratio (LR-) were calculated, accompanied by 95% confidence intervals (CIs), for each outcome measure. Of note, because MCID and SCB values are calculated as a change from baseline, a small number of patients with a high enough baseline score were unable to achieve the MCID/SCB. All statistical analyses were performed using RStudio Version 1.2.1335.

### RESULTS

Patient characteristics, baseline PROMs, surgical indications, and surgical procedures for both cohorts are reported in Table 1. No significant differences were apparent between cohorts concerning age, sex, baseline PROMs, and surgical indications. When comparing procedures performed during arthroscopy, the control cohort did have a significantly greater percentage of patients undergo labral repair (83.7% vs 68.2%;  $P = .005$ ) and cam decompression (92% vs 77.3%,  $P = .003$ ). There were no significant differences between cohorts concerning the remaining surgical procedures.

The average time to subsequent surgery in the study cohort was 22.3 months (range, 5.5-59.3 months). Twenty-four patients (27.3%) in the study cohort underwent a subsequent THA, with the average time to THA being 26.4 months (range, 5.5-59.3 months). The remaining 64 patients (72.7%) underwent subsequent surgery at an average of 20.8 months (range, 6.0-51.7 months). Of these 64 patients, the most common subsequent surgery was a revision hip arthroscopy (50 patients). The remaining procedures included surgical hip dislocation (6 patients),

TABLE 1  
Patient Characteristics, Baseline PROMs, Surgical Indications, and Surgical Procedures<sup>a</sup>

	Study Cohort (n = 88)	Control Cohort (n = 288)	P
Sex			
Male	37 (42.0)	138 (47.9)	.399
Female	51 (58.0)	150 (52.1)	
Age, y	33.1 (18-59)	32.8 (18-57)	.804
Baseline PROMs			
HHS	59.6 (12.1-95.7)	63.1 (19.8-100)	.062
HOS-ADL	69.7 (13.2-97.1)	72.7 (20.6-100)	.206
HOS-Sports	48.7 (0.00-100)	48.9 (0.00-100)	.937
iHOT-33	37.7 (2.2-78.6)	38.5 (1.68-90.1)	.745
Preoperative surgical indications			
Labral tear	73 (83.0)	246 (85.4)	.588
Pincer lesion	25 (28.4)	91 (31.6)	.567
Cam lesion	65 (73.9)	231 (80.2)	.23
Arthroscopic procedures			
Labral repair	60 (68.2)	241 (83.7)	<b>.005</b>
Labral debridement	16 (18.2)	43 (14.9)	.485
Cam decompression	68 (77.3)	265 (92.0)	<b>.003</b>
Rim decompression	28 (31.8)	117 (40.6)	.129

<sup>a</sup>Data are reported as n (%) or mean (range). Bolded P values indicate a statistically significant difference between groups ( $P < .05$ ). HOS-ADL, Hip Outcome Score–Activities of Daily Living; HOS-Sports, Hip Outcome Score–Sports; iHOT-33, 33-Item International Hip Outcome Tool; mHHS, modified Harris Hip Score; PROMs, patient-reported outcome measures.

TABLE 2  
Type and Time to Subsequent Surgery for the Study Cohort

Type of Subsequent Surgery	No. of Patients (%)	Mean Months to Subsequent Surgery (range)
Overall study cohort	88 (100)	22.3 (5.5-59.3)
Total hip arthroplasty	24 (27.3)	26.4 (5.5-59.3)
Other <sup>a</sup>	64 (72.7)	20.8 (6.0-51.7)

<sup>a</sup>Other subsequent surgeries included hip arthroscopy (n = 50), surgical hip dislocation (n = 6), periacetabular osteotomy (n = 5), excision of heterotopic ossification (n = 2), and gluteus medius repair (n = 1).

periacetabular osteotomy (5 patients), open excision of heterotopic ossification (2 patients), and 1 patient underwent repair of the gluteus medius. This information is presented in Table 2.

Postoperative PROMs and the percentage of each cohort to achieve the MCID and SCB are presented in Table 3. The study cohort reported significantly lower results across all PROMs ( $P < .001$ ). Additionally, on all PROMs, a significantly smaller percentage of patients in the study cohort achieved the MCID and SCB when compared with those in the control cohort ( $P < .001$ ).

Multivariable logistic regression demonstrated that failure to achieve the MCID or SCB on any PROM significantly

TABLE 3  
Postoperative Score and Percentage of Cohort to Achieve the MCID and SCB<sup>a</sup>

	Study Cohort (n = 88)	Control Cohort (n = 288)	P
Postoperative PROM score			
mHHS	62.1 (14.3-95.7)	80.1 (29.7-100.1)	<.001
HOS-ADL	72.7 (17.7-100)	87.7 (23.2-100)	<.001
HOS-Sports	46.6 (0.00-100)	66.8 (0.00-100)	<.001
iHOT-33	46.9 (0.43-94.7)	68.5 (9.15-99.9)	<.001
Achieved MCID <sup>b</sup>			
mHHS	34.9 (30/86)	70.5 (196/278)	<.001
HOS-ADL	36.0 (31/86)	64.8 (184/284)	<.001
HOS-Sports	25.0 (21/84)	53.3 (146/274)	<.001
iHOT-33	40.6 (26/64)	74.8 (166/222)	<.001
Achieved SCB <sup>b</sup>			
mHHS	17.4 (15/86)	41.4 (115/278)	<.001
HOS-ADL	34.9 (30/86)	60.2 (171/284)	<.001
HOS-Sports	11.9 (10/84)	33.6 (92/274)	<.001
iHOT-33	20.3 (13/64)	55.4 (123/222)	<.001

<sup>a</sup>Data are reported as mean (range) or % (n/total). All differences between groups were statistically significant ( $P < .05$ ). iHOT-33, 33-Item International Hip Outcome Tool; HOS-ADL, Hip Outcome Score–Activities of Daily Living; HOS-Sports, Hip Outcome Score–Sports; MCID, minimal clinically important difference; mHHS, modified Harris Hip Score; PROM, patient-reported outcome measure; SCB, substantial clinical benefit.

<sup>b</sup>Because MCID and SCB values are calculated as a change from baseline, a small number of patients with a high enough baseline score were unable to achieve the MCID/SCB.

increased the risk of a subsequent surgery when controlling for age, sex, and baseline PROMs. Failure to achieve the MCID on any single PROM significantly increased the odds of subsequent surgery, with odds ratios of 5.39 for mHHS, 5.20 for HOS-ADL, 4.78 for iHOT-33, and 4.02 for the HOS-Sports ( $P < .001$  for all). Failure to meet the SCB on any single PROM also significantly increased the odds of revision surgery (odds ratios, 5.2 for iHOT-33, 4.69 for HOS-ADL, 4.57 for HOS-Sports, and 4.11 for mHHS;  $P < .001$  for all). This information is presented in Table 4.

A final multivariable model showed that each additional PROM in which a patient did not achieve a threshold score significantly increased the risk of subsequent surgery when adjusting for age, sex, and baseline scores. For each additional PROM, failure to meet the MCID increased the odds of subsequent surgery by 1.68 ( $P < .001$ ), and failure to meet the SCB increased the odds of subsequent surgery by 1.63 ( $P < .001$ ). This model is presented in Table 4.

Sensitivity, specificity, and likelihood ratios are reported in Table 5. The mHHS had the highest sensitivity (0.41; 95% CI, 0.33-0.49) and LR+ (3.06; 95% CI, 2.12-4.67) for the MCID. The mHHS (0.87; 95% CI, 0.82-0.91) and HOS-Sports (0.87; 95% CI, 0.81-0.92) had the highest specificity for the MCID. When the SCB was examined, iHOT-33 had the highest sensitivity (0.34; 95% CI, 0.27-0.42) and LR+ (3.55; 95% CI, 2.13-6.86). The highest specificity for the

TABLE 4

Multivariable Logistic Regression for Subsequent Surgery Risk Controlling for Age, Sex, and Baseline PROMs <sup>a</sup>

(A) Subsequent Surgery Risk Based Off Not Achieving the MCID or SCB		
	OR (95% CI)	P
MCID		
mHHS	5.39 (3.14-9.25)	<.001
HOS-ADL	5.20 (2.89-9.36)	<.001
HOS-Sports	4.02 (2.23-7.24)	<.001
iHOT-33	4.78 (2.60-8.77)	<.001
SCB		
mHHS	4.11 (2.19-7.72)	<.001
HOS-ADL	4.69 (2.58-8.51)	<.001
HOS-Sports	4.57 (2.14-9.74)	<.001
iHOT-33	5.20 (2.65-10.2)	<.001
(B) Cumulative Subsequent Surgery Risk for Not Achieving the MCID or SCB on Multiple PROMs		
	OR (95% CI)	P
MCID	1.68 (1.42-1.98)	<.001
SCB	1.63 (1.35-1.97)	<.001

<sup>a</sup>Part A shows the risk of revision surgery associated with failure to achieve the MCID or SCB on any PROM. Part B shows the risk of revision surgery associated with the cumulative effect of failure to achieve the MCID or SCB on multiple PROMs. All results were statistically significant ( $P < .05$ ). iHOT-33, 33-Item International Hip Outcome Tool; HOS-ADL, Hip Outcome Score—Activities of Daily Living; HOS-Sports, Hip Outcome Score—Sports; MCID, minimal clinically important difference; mHHS, modified Harris Hip Score; PROM, patient-reported outcome measure; SCB, substantial clinical benefit.

SCB was seen on HOS-Sports (0.91; 95% CI, 0.83-0.95). The LR<sup>-</sup> was <1.0 for all outcome measures for both the MCID and SCB.

## DISCUSSION

The results of this study support our initial hypothesis that failure to meet the MCID and/or SCB on the mHHS, HOS-ADL, HOS-Sports, or iHOT-33 within 1 year of hip arthroscopy significantly increases the risk of a subsequent hip surgery within 5 years of the primary procedure. Patients who underwent subsequent surgery had significantly lower scores on all postoperative PROMs compared with those who did not undergo a subsequent surgery. Additionally, a significantly smaller percentage of patients who underwent subsequent ipsilateral hip surgery reported a change in outcome score that met the MCID and/or SCB compared with those who did not undergo a subsequent hip surgery. Our principal hypothesis was then substantiated by multivariable logistic regression controlling for age and sex. The results demonstrated that failure to meet the MCID or SCB on any of the 4 included PROMs is an independent risk

factor for subsequent hip surgery. A final multivariable logistic regression also examined the effect of failing to meet the MCID or SCB on more than one PROM. This analysis of the cumulative effect of failure to meet a threshold score revealed that for each additional PROM in which a patient failed to achieve a threshold score, their risk of subsequent surgery significantly increased. Overall, failure to meet the threshold score for the SCB on iHOT-33 was associated with the largest increase in risk for subsequent surgery. When examining only the MCID, failure to meet the threshold score for iHOT-33 was again associated with the greatest risk of subsequent surgery.

The outcome measure with the highest reported sensitivity for subsequent surgery was mHHS (0.41) and iHOT-33 (0.34) for the MCID and SCB, respectively. Regarding specificity, the highest values were reported on both mHHS and HOS-Sports (0.87) for the MCID and on HOS-Sports (0.91) for the SCB. It is important to recall that sensitivity, defined as the true-positive rate, and specificity, defined as the true-negative rate, function in an inverse relationship, and as one value increases, the other will decrease.<sup>36</sup> Therefore, our data, upon initial review, suggest that the utilization of threshold values after hip arthroscopy can act as a highly specific measure of the risk of subsequent surgery. Despite this initial assessment, further analysis of the data reveals that the LR<sup>+</sup> and LR<sup>-</sup> capture the utility of these measures with greater accuracy.

Likelihood ratios function as a means to understand how probable it is that a patient with a condition is expected to have a result compared with patients without the condition.<sup>6,29</sup> In regard to the present study, it is the ratio of the probability of subsequent hip surgery in people who met the MCID and/or SCB compared with people who did not. A likelihood ratio increases in strength as it gets farther away from a value of 1.0, and in this analysis, all LR<sup>+</sup> values were >1.0 and all LR<sup>-</sup> values were <1.0. This finding is in line with our initial hypothesis, as it suggests that failure to achieve threshold scores for the MCID or SCB increases the risk of undergoing a subsequent surgery and that patients who do achieve these threshold scores conversely have a smaller risk of subsequent surgery.

However, for diagnostic and prognostic purposes, an LR<sup>+</sup> value >10 is desired for a test to be considered highly specific, and an LR<sup>-</sup> value <0.1 is desired for a test to be considered highly sensitive.<sup>28,29</sup> While all LR<sup>+</sup> values in this study were >1.0, the highest LR<sup>+</sup> value was 3.55. Similarly, while all LR<sup>-</sup> values were <1.0, the smallest LR<sup>-</sup> was 0.69. One can therefore conclude that failing to achieve a threshold score after hip arthroscopy is indeed associated with increased risk of subsequent surgery. However, when reviewing our results and the information from this discussion on the utility of likelihood ratios, it is apparent that the actual prognostic value of threshold scores from a clinical perspective may be limited. Therefore, we recommend these measures be used for prognostic purposes only when combined with other objective criteria. A threshold score can provide a comprehensive risk assessment for future hip surgery when combined with certain criteria such as physical examination findings, radiographic measures, and the patient's desired level of physical activity. This can serve as

TABLE 5  
Sensitivity, Specificity, Positive Likelihood Ratio, and Negative Likelihood Ratio for the PROMs Evaluated<sup>a</sup>

	mHHS	HOS-ADL	HOS-Sport	iHOT-33
<b>MCID</b>				
Sensitivity	0.41 (0.33-0.49)	0.35 (0.28-0.43)	0.33 (0.26-0.40)	0.40 (0.31-0.51)
Specificity	0.87 (0.82-0.91)	0.86 (0.81-0.90)	0.87 (0.81-0.92)	0.86 (0.81-0.91)
LR+	3.06 (2.12-4.67)	2.46 (1.71-3.75)	2.62 (1.66-4.07)	2.99 (1.99-4.80)
LR-	0.69 (0.59-0.79)	0.75 (0.66-0.85)	0.77 (0.69-0.86)	0.69 (0.58-0.82)
<b>SCB</b>				
Sensitivity	0.30 (0.25-0.37)	0.33 (0.26-0.40)	0.29 (0.23-0.35)	0.34 (0.27-0.42)
Specificity	0.88 (0.82-0.94)	0.85 (0.79-0.89)	0.91 (0.83-0.95)	0.90 (0.85-0.95)
LR+	2.62 (1.64-4.76)	2.22 (1.48-3.26)	2.95 (1.57-5.43)	3.55 (2.13-6.86)
LR-	0.79 (0.70-0.87)	0.79 (0.70-0.89)	0.79 (0.72-0.88)	0.73 (0.64-0.82)

<sup>a</sup>Data in parentheses are 95% CIs. Sensitivity, specificity, LR+, and LR- were calculated as follows:

$$\text{sensitivity} = \frac{TP}{TP + FN} \quad \text{specificity} = \frac{TN}{TN + FP} \quad LR+ = \frac{\text{sensitivity}}{1 - \text{specificity}} \quad LR- = \frac{1 - \text{sensitivity}}{\text{specificity}}$$

$$\text{True positive (TP)} = \frac{\text{Failure to achieve threshold score}}{\text{underwent subsequent surgery}} \quad \text{False positive (FP)} = \frac{\text{Achieved threshold score}}{\text{underwent subsequent surgery}}$$

$$\text{False negative (FN)} = \frac{\text{Failure to achieve threshold score}}{\text{Did not undergo subsequent surgery}} \quad \text{True Negative (TN)} = \frac{\text{Achieved threshold score}}{\text{Did not undergo subsequent surgery}}$$

HOS-ADL, Hip Outcome Score-Activities of Daily Living; HOS-Sports, Hip Outcome Score-Sports; iHOT-33, 33-Item International Hip Outcome Tool; MCID, minimal clinically important difference; mHHS, modified Harris Hip Score; LR+, positive likelihood ratio; LR-, negative likelihood ratio; PROM, patient-reported outcome measure; SCB, substantial clinical benefit.

valuable information to the hip-preservation surgeon for patient counseling throughout the postoperative course.

Several studies<sup>7,9,30</sup> in the field of hip-preservation surgery have demonstrated that preoperative PROMs carry prognostic value for postoperative outcome because of an established correlation with postoperative outcome scores as well as subsequent surgeries, including conversion to THA. Despite this, there is a paucity of data on the predictive risk of subsequent hip surgery with regard to achievement of the MCID and/or SCB after arthroscopic treatment of FAI. A recent retrospective study by Rosinsky et al,<sup>32</sup> to our knowledge, was the first to publish data supporting the novel concept that a patient's failure or achievement of a threshold score correlates with rates of subsequent hip surgery after hip arthroscopy. This, however, was a single-surgeon study that investigated 2 PROMs (mHHS and HOS-ADL) and only established a correlation between threshold scores and subsequent hip surgery. A later study by Shapira et al<sup>34</sup> did investigate this relationship in a causal manner and found that the risk for secondary hip surgery may be evaluated with mHHS scores at 1 year after hip arthroscopy.

The results of our study substantiate the findings of the aforementioned studies. Further, the present study has expanded the discussion to include additional PROMs and included a regression analysis with parameters that controlled for confounding variables. Our findings indicate that failure to reach a threshold score can serve as an independent risk factor for subsequent hip surgery. As such, close monitoring should be implemented for patients who fail to achieve a threshold score in the short-term after hip arthroscopy as they are at a statistically significantly

increased risk of subsequent hip surgery. Our results underscore the importance of the continued utilization of PROMs and threshold scores in orthopaedic literature. The inclusion of PROMs, in the form of both absolute score and comparisons to threshold values, allows the reader to gain a more robust understanding of the generalizability of study results and therefore provides the clinician the capacity to extrapolate the results to everyday clinical practice.

The strength of this study includes the large sample size, the use of established metrics to define meaningful clinical improvement, and the robust statistical analysis of our data on short-term PROMs. Further, the follow-up window requiring patients to have a clinical visit at a minimum of 5 years after surgery provides insight into how short-term achievement of threshold scores predicts mid- to long-term results after hip arthroscopy.

The present study has several limitations. This was a single-center retrospective study based on hospital-wide registry data collected in a prospective manner, which may limit the study's generalizability. Future research that distinguishes between the causes of secondary surgery may be useful to determine where threshold scores have the greatest predictive quality. Additionally, this study did not differentiate patients based off the type of subsequent hip surgery or by specific procedures performed during arthroscopy. However, a significantly greater percentage of the control cohort underwent a cam decompression and a labral repair, which presents a possible confounding variable. A study wherein patients were divided on the basis of procedures during hip arthroscopy and the type of subsequent hip surgery they underwent

would be of further use in investigating this topic. Certain objective preoperative clinical information, such as preoperative body mass index, joint parameters, and residual impingement, was not available for all patients and therefore not included in this analysis. Our study may underestimate the benefit of hip arthroscopy in the small number of patients who were unable to achieve the MCID/SCB owing to a high baseline PROM. Lastly, this study used previously established values for the MCID and SCB from other studies as reference values. While substantial heterogeneity does exist in MCID quantification methods, the MCID can be calculated as a score specific to a study population,<sup>5,13</sup> which reduces inconsistency. Thus, our decision to use reference values inherently limits the validity of the study.

## CONCLUSION

The results of the current study show that the MCID and SCB thresholds for the mHHS, HOS-Sports, HOS-ADL, and iHOT-33 are all predictive measures of subsequent surgery risk for patients after primary hip arthroscopy. This study establishes a novel utility of PROMs and underscores the importance of these metrics in the orthopaedic literature.

## REFERENCES

- Bernstein DN, Nwachukwu BU, Bozic KJ. Value-based health care: moving beyond “minimum clinically important difference” to a tiered system of evaluating successful clinical outcomes. *Clin Orthop Relat Res*. 2019;477(5):945-947. doi:10.1097/CORR.0000000000000741
- Byrd JWT. Hip arthroscopy: patient assessment and indications. *Instr Course Lect*. 2003;52:711-719. Accessed April 18, 2021. <https://europepmc.org/article/med/12690896>
- Chambers CC, Zhang AL. Outcomes for surgical treatment of femoroacetabular impingement in adults. *Curr Rev Musculoskelet Med*. 2019;12(3):271-280. doi:10.1007/s12178-019-09567-1
- Chona DV, Bonano JC, Ayeni OR, Safran MR. Definitions of return to sport after hip arthroscopy: are we speaking the same language and are we measuring the right outcome? *Orthop J Sports Med*. 2020;8(9):2325967120952990.
- Copay AG, Eyberg B, Chung AS, Zurcher KS, Chutkan N, Spangehl MJ. Minimum clinically important difference: current trends in the orthopaedic literature, Part II: Lower extremity. *JBJS Rev*. 2018;6(9):e2. doi:10.2106/JBJS.RVW.17.00160
- Deeks JJ, Altman DG. Diagnostic tests 4: Likelihood ratios. *BMJ*. 2004;329(7458):168-169. doi:10.1136/bmj.329.7458.168
- Domb BG, Gui C, Hutchinson MR, Nho SJ, Terry MA, Lodhia P. Clinical outcomes of hip arthroscopic surgery: a prospective survival analysis of primary and revision surgeries in a large mixed cohort. *Am J Sports Med*. 2016;44(10):2505-2517. doi:10.1177/0363546516663463
- Eneqvist T, Nemes S, Brisby H, Fritzell P, Garellick G, Rolfsen O. Lumbar surgery prior to total hip arthroplasty is associated with worse patient-reported outcomes. *Bone Joint J*. 2017;99(6):759-765. doi:10.1302/0301-620X.99B6.BJJ-2016-0577.R2
- Griffin DW, Kinnard MJ, Formby PM, McCabe MP, Anderson TD. Outcomes of hip arthroscopy in the older adult: a systematic review of the literature. *Am J Sports Med*. 2017;45(8):1928-1936. doi:10.1177/0363546516667915
- Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertain the minimal clinically important difference. *Control Clin Trials*. 1989;10(4):407-415. doi:10.1016/0197-2456(89)90005-6
- Kemp JL, Collins NJ, Makdissi M, Schache AG, Machotka Z, Crossley K. Hip arthroscopy for intra-articular pathology: a systematic review of outcomes with and without femoral osteoplasty. *Br J Sports Med*. 2012;46(9):632-643. doi:10.1136/bjsports-2011-090428
- Kemp JL, Collins NJ, Roos EM, Crossley KM. Psychometric properties of patient-reported outcome measures for hip arthroscopic surgery. *Am J Sports Med*. 2013;41(9):2065-2073. doi:10.1177/0363546513494173
- Kunze KN, Bart JA, Ahmad M, Nho SJ, Chahla J. Large heterogeneity among minimal clinically important differences for hip arthroscopy outcomes: a systematic review of reporting trends and quantification methods. *Arthroscopy*. 2021;37(3):1028-1037.e6. doi:10.1016/j.arthro.2020.10.050
- Lum ZC, Coury JG, Cohen JL, Dorr LD. The current knowledge on spinopelvic mobility. *J Arthroplasty*. 2018;33(1):291-296. doi:10.1016/j.arthro.2017.08.013
- Maempel JF, Ting JZ, Gaston P. Assessing the outcome of hip arthroscopy for labral tears in femoroacetabular impingement using the minimum dataset of the British Non-arthroplasty Hip Register: a single-surgeon experience. *Arthroscopy*. 2018;34(7):2131-2139. doi:10.1016/j.arthro.2018.02.038
- Maltenfort M, Díaz-Ledezma C. Statistics in brief: minimum clinically important difference—availability of reliable estimates. *Clin Orthop Relat Res*. 2017;475(4):933-946. doi:10.1007/s11999-016-5204-6
- Maradit Kremers H, Schilz SR, Van Houten HK, et al. Trends in utilization and outcomes of hip arthroscopy in the United States between 2005 and 2013. *J Arthroplasty*. 2017;32(3):750-755. doi:10.1016/j.arthro.2016.09.004
- Martin RRL, Kelly BT, Philippon MJ. Evidence of validity for the Hip Outcome Score. *Arthroscopy*. 2006;22(12):1304-1311. doi:10.1016/j.arthro.2006.07.027
- Martin RRL, Kivlan BR, Christoforetti JJ, et al. Defining variations in outcomes of hip arthroscopy for femoroacetabular impingement using the 12-Item International Hip Outcome Tool (iHOT-12). *Am J Sports Med*. 2020;48(5):1175-1180. doi:10.1177/0363546520906408
- Martin RRL, Philippon MJ. Evidence of validity for the Hip Outcome Score in hip arthroscopy. *Arthroscopy*. 2007;23(8):822-826. doi:10.1016/j.arthro.2007.02.004
- Matsuda DK, Carlisle JC, Arthurs SC, Wierks CH, Philippon MJ. Comparative systematic review of the open dislocation, mini-open, and arthroscopic surgeries for femoroacetabular impingement. *Arthroscopy*. 2011;27(2):252-269. doi:10.1016/j.arthro.2010.09.011
- Minkara AA, Westermann RW, Rosneck J, Lynch TS. Systematic review and meta-analysis of outcomes after hip arthroscopy in femoroacetabular impingement. *Am J Sports Med*. 2019;47(2):488-500. doi:10.1177/0363546517749475
- Montgomery SR, Ngo SS, Hobson T, et al. Trends and demographics in hip arthroscopy in the United States. *Arthroscopy*. 2013;29(4):661-665. doi:10.1016/j.arthro.2012.11.005
- 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(11):2601-2606. doi:10.1177/0363546518786480
- Nwachukwu BU, Chang B, Beck EC, et al. How should we define clinically significant outcome improvement on the iHOT-12? *HSS J*. 2019;15(2):103-108. doi:10.1007/s11420-018-9646-0
- Nwachukwu BU, Chang B, Fields K, et al. Defining the “substantial clinical benefit” after arthroscopic treatment of femoroacetabular impingement. *Am J Sports Med*. 2017;45(6):1297-1303. doi:10.1177/0363546516687541
- Nwachukwu BU, Fields K, Chang B, Nawabi DH, Kelly BT, Ranawat AS. Preoperative outcome scores are predictive of achieving the minimal clinically important difference after arthroscopic treatment of femoroacetabular impingement. *Am J Sports Med*. 2017;45(3):612-619. doi:10.1177/0363546516669325
- Parikh R, Mathai A, Parikh S, Sekhar GC, Thomas R. Understanding and using sensitivity, specificity and predictive values. *Indian J Ophthalmol*. 2008;56(1):45-50. doi:10.4103/0301-4738.37595

29. Ranganathan P, Aggarwal R. Understanding the properties of diagnostic tests—Part 2: Likelihood ratios. *Perspect Clin Res*. 2018;9(2):99. doi:10.4103/picr.picr\_41\_18
30. Redmond JM, Gupta A, Dunne K, Humayun A, Yuen LC, Domb BG. What factors predict conversion to THA after arthroscopy? *Clin Orthop Relat Res*. 2017;475(10):2538-2545. doi:10.1007/s11999-017-5437-z
31. Rojanasopondist P, Galea VP, Connelly JW, et al. What preoperative factors are associated with not achieving a minimum clinically important difference after THA? Findings from an international multicenter study. *Clin Orthop Relat Res*. 2019;477(6):1301-1312. doi:10.1097/CORR.0000000000000667
32. Rosinsky PJ, Chen JW, Yelton MJ, et al. Does failure to meet threshold scores for mHHS and iHOT-12 correlate to secondary operations following hip arthroscopy? *J Hip Preserv Surg*. 2020;7(2):272-280. doi:10.1093/jhps/hnaa015
33. Rothwell AG, Hooper GJ, Hobbs A, Frampton CM. An analysis of the Oxford hip and knee scores and their relationship to early joint revision in the New Zealand Joint Registry. *J Bone Joint Surg Br*. 2010;92(3):413-418. doi:10.1302/0301-620X.92B3.22913
34. Shapira J, Bheem R, Kyin C, et al. Can patient-reported outcomes predict the need for secondary surgeries after hip arthroscopy? *Am J Sports Med*. 2021;49(1):97-103. doi:10.1177/0363546520974374
35. Shin JJ, de Sa DL, Burnham JM, Mauro CS. Refractory pain following hip arthroscopy: evaluation and management. *J Hip Preserv Surg*. 2018;5(1):3-14. doi:10.1093/jhps/hnx047
36. Shreffler J, Huecker MR. *Diagnostic Testing Accuracy: Sensitivity, Specificity, Predictive Values and Likelihood Ratios*. StatPearls Publishing; 2020. Accessed December 28, 2020. <http://www.ncbi.nlm.nih.gov/pubmed/32491423>
37. Sing DC, Feeley BT, Tay B, Vail TP, Zhang AL. Age-related trends in hip arthroscopy: a large cross-sectional analysis. *Arthroscopy*. 2015;31(12):2307-2313.e2. doi:10.1016/j.arthro.2015.06.008
38. Singh JA, Schleck C, Harmsen S, Lewallen D. Clinically important improvement thresholds for Harris Hip Score and its ability to predict revision risk after primary total hip arthroplasty. *BMC Musculoskelet Disord*. 2016;17(1):256. doi:10.1186/s12891-016-1106-8
39. Stephan P, Röling MA, Mathijssen NMC, Hannink G, Bloem RM. Developing a risk prediction model for the functional outcome after hip arthroscopy. *BMC Musculoskelet Disord*. 2018;19(1):122. doi:10.1186/s12891-018-2030-x