

Assessing outcomes following arthroscopic labral debridement—what can the IHOT-33 reveal?

Maureen K. Dwyer^{1,2*}, Marlon Green¹ and Joseph C. McCarthy^{1,2}

1. Kaplan Joint Center, Department of Orthopaedics, Newton Wellesley Hospital, Newton, MA 02462

2. Department of Orthopaedics, Massachusetts General Hospital, Boston, MA 02114, USA

*Correspondence to: M. K. Dwyer. E-mail: mkdwyer@partners.org

Submitted 17 November 2014; Revised 29 January 2015; revised version accepted 18 February 2015

ABSTRACT

As the complexity of arthroscopic procedures continues to increase, assessing the success of these procedures is of utmost importance to determine appropriate treatment plans for patients. This study compares outcome scores on the International Hip Outcome Tool (IHOT-33) using an analysis of its four subscales to scores on the Modified Harris Hip Score (MHHS) in patients following hip arthroscopy. Patients who had undergone hip arthroscopy between 1 and 5 years ago were assessed using both the IHOT-33 and MHHS at their most recent follow-up visit. Total scores for each outcome measure were calculated. A composite score for each of the IHOT subscales was calculated by averaging the total points in each section. Total score on the MHHS was compared with total score on the IHOT-33 and its four subscales using separate wilcoxon signed ranks tests. In total, 44 patients met our inclusion criteria with an average follow-up of 24.8 ± 18.8 months. MHHS total score was greater than IHOT-33 total score ($P = 0.04$). Comparisons between MHHS total score and the IHOT-33 subscales revealed reductions in sport ($P = 0.001$) and social ($P = 0.004$), but no difference in symptoms ($P = 0.74$) and job ($P = 0.84$). Our findings demonstrated that scores on the MHHS are inflated when compared with scores on the IHOT-33 for patients following hip arthroscopy. Subscale analysis revealed that the reductions in IHOT-33 scores exist in the sport and recreational domain and social, emotional, lifestyle domain compared with the MHHS. Our results suggest that analysing the subscales of the IHOT-33 would provide a more thorough understanding of functional limitations in patients undergoing hip arthroscopy.

INTRODUCTION

Hip arthroscopy is an emerging minimally invasive technique used to treat a variety of hip conditions, including labral tears, chondral lesions, loose bodies, synovitis and bony morphologic alterations. Arthroscopic treatment of both intra- and extra-articular hip conditions continues to expand through advancements in both our understanding of the pathogenesis of these conditions and emerging surgical techniques. As the complexity of arthroscopic procedures continues to increase, assessing the success of these procedures is of utmost importance to determine appropriate treatment plans for patients. Patient reported outcome measures continue to be the standard with which to examine symptoms and function in patients following surgery and have been used extensively in hip arthroscopy patients.

Numerous measurement tools have been utilized to assess outcomes in patients following hip arthroscopy [1–7]. The modified version of the Harris Hip Score (MHHS) is the most commonly reported measure of outcomes for this patient population [7], and has been evaluated in patients undergoing arthroscopic labral debridement [8–12], labral repair and reconstruction [13–15], and treatment for femoroacetabular impingement [16–20]. Other outcome measures, including the Hip Outcome Score (HOS) [1], Hip Disability and Osteoarthritis Outcome Score (HOOS) [21], and Non-arthritic Hip Score [6] have also been reported for hip arthroscopy patients, though in much lesser frequency. However, none of these tools was developed specifically for the assessment of outcomes in patients undergoing hip arthroscopy. The International Hip Outcome Tool (IHOT-33) was created to address this

need and has been validated for use with hip arthroscopy populations [2]. A recent systematic evaluation of the psychometric properties of all outcome measures assessed in this population revealed important limitations in the MHHS, HOOS and HOS, including the presence of ceiling effects and decreased responsiveness [22]. The authors recommended using the IHOT-33 [22]; however, they criticized its use of a single aggregate score, which limited its ability to discriminate between important functional domains.

Despite consistent reports of ceiling effects and poor psychometric properties [22, 23], the MHHS remains the outcome measure of choice to assess patients undergoing hip arthroscopy. The intent of the IHOT-33 was to address the limitations of the MHHS and expand our understanding of the functional limitations experienced by these patients; yet the use of an aggregate score restricts its interpretability. Therefore, we examined IHOT-33 scores in patients following hip arthroscopy using an analysis of its four subscales and compared these scores to the MHHS to determine which specific facets of function are uniquely captured with the IHOT-33.

MATERIALS AND METHODS

From 2008 to 2013, all patients who had undergone unilateral arthroscopic labral debridement by the senior author and had 1–5-year follow-up MHHS and IHOT-33 data were included in the study. Data were collected as part of our institutionally approved prospective hip registry, but were analysed retrospectively. Patients were excluded if they had undergone a repeat hip preservation surgery since the index arthroscopy or had converted to a total hip replacement. Indications for surgery included anterior, trochanteric, inguinal or buttock pain with or without mechanical symptoms, along with confirmation of intra-articular pathology via MRI or arthrography. Additional indications included patients who had residual pain and symptoms, which did not respond to activity modification, nonsteroidal anti-inflammatories, physical therapy or time. We identified 53 patients who met our inclusion criteria (14 males, 39 females), with an average age of 40.6 ± 13.1 years at time of surgery and an average follow-up of 24.9 ± 18.7 months (range: 12–93 months) following arthroscopy. Institutional approval was obtained prior to the initiation of this study.

All patients underwent arthroscopy in the lateral decubitus position with a dedicated hip distractor set at 7–10 mm of distraction. Arthroscopies were performed as outpatient procedures. All labral tears were treated with debridement. Full thickness chondral lesions were treated with microfracture surgery, while partial-thickness lesions

were minimally resected to a stable base. Loose bodies were noted and removed. The synovial lining was resected if impingement was noted. At the time of surgery, the presence and location of all labral tears were noted. In addition, the morphologic features and location of all articular cartilage lesions were identified and graded for severity using the Outerbridge classification by the senior author. Findings from each surgical procedure were recorded using audio and visual media and entered into our institutionally approved database.

Patients completed both the IHOT-33 and MHHS. The MHHS consists of 8 questions assessing pain (44 points) and function during daily activities (47 points). The total number of points is multiplied by a factor of 1.1 to provide a score out of a possible 100 points. All questions must be answered to calculate a score. The IHOT-33 consists of 33 questions, broken down into four domains—Symptoms and Functional Limitations (IHOT-symptoms—16 questions); Sports and Recreational Activities (IHOT-sport—6 questions); Job Related Concerns (IHOT-job—4 questions); and Social, Emotional and Lifestyle Concerns (IHOT-social—7 questions). If a patient is retired or not working for other reasons, they can opt out of the job-related questions. In addition, there are four questions to which a patient can check a box to indicate that the activity in question is not applicable to them, one in IHOT-sport inquiring about cutting activities, one in IHOT-job inquiring about lifting heavy objects, and two in IHOT-social inquiring about sexual activity and lifting children. However, unlike the MHHS, not answering a question on the IHOT-33 does not negate calculation of a total score. Each question is scored out of a possible 100 points, and the total score is calculated by averaging the total number of points by the total number of questions answered (0–100 points). Completed outcome scores were stored in our database.

The total score for each outcome measure was calculated (0–100 points). A composite score for each of the IHOT subscales (IHOT-symptoms, IHOT-sport, IHOT-job, IHOT-social) was then calculated by averaging the total number of points in each section by the number of questions answered in each section to get the four subscale scores (0–100 points). Data were assessed for normality using the Shapiro-Wilk test. Our data were not found to be normal distributed; therefore, nonparametric analyses were used. Total score on the MHHS was compared with total score on the IHOT using a wilcoxin signed ranks test. Total score on the MHHS was compared with the composite score of each of the four subscales of the IHOT (IHOT-symptoms, IHOT-sport, IHOT-job, IHOT-social) using separate wilcoxin signed ranks tests. All statistical

analyses were performed using IBM SPSS v.21.0 (IBM Corporation, Armonk, NY). Level of significance was set *a priori* at $P \leq 0.05$.

RESULTS

Upon further review, data for nine patients were excluded because of incomplete MHHS scores. Therefore, 44 patients were included in the final analysis (11 males, 33 females). Average age and time from surgery were similar between patients included in the analysis and the cohort of patients excluded (Table I). Observations made at the time of surgery revealed the presence of an anterior labral tear in 100% of patients in our study, with concomitant lesions to the posterior (6 hips) and lateral (8 hips) labrum in 12 patients. Four patients had grade I wear of the femoral head, located anteriorly in three and laterally in one, while three patients had grade II wear antero-superiorly. Four additional patients had grade IV wear of the femoral head, one anterior, one superior, and two posterior. On the articular cartilage of the acetabulum, abnormalities were noted anteriorly in 59% of patients, posteriorly in 16% of patients and laterally in 9% of patients. Specifics regarding all labral and cartilage lesions for both the included and excluded patients are presented in Table I.

Outcome scores for our cohort of patients at an average of 25 months following hip arthroscopy are presented in Fig. 1. Total score on the MHHS was >70 points, with 45% of patients rating a good to excellent outcome (scores >80 points). Total score on the MHHS (74.1 ± 19.2) was greater than total score on the IHOT-33 (66.9 ± 24.3 ; $P=0.04$). Separate comparisons between total score on the MHHS and scores on the subscales of the IHOT-33 revealed significant reductions in IHOT-sport (55.6 ± 32.5 ; $P=0.001$) and IHOT-social (60.6 ± 31.5 ; $P=0.004$), but no difference in IHOT-symptoms (72.6 ± 23.2 ; $P=0.74$) and IHOT-job (73.5 ± 26.6 ; $P=0.84$).

DISCUSSION

Assessing patient outcomes following hip arthroscopy is essential to determining the success of the procedure. Numerous outcome measures have been created for and implemented in patients undergoing hip preservation surgery; however, the most commonly cited outcome measure remains the MHHS, despite its high ceiling effect in this population. The IHOT-33 was created to address the limitations of the MHHS; yet it has been criticized for its use of a single aggregate score. As such, we sought to examine whether analysing the separate subscales of the IHOT-33 would better portray the functional abilities of patients undergoing hip arthroscopy. We demonstrated that

Table I. Demographic and surgical variables for the study cohort

Variable	Category	Included cohort	Excluded cohort
Age at arthroscopy		40.4 \pm 13.7	41.7 \pm 10.4
Gender	Male	11 (25%)	3 (33%)
	Female	33 (75%)	6 (67%)
Follow-up time (months)		24.8 \pm 18.8	25.0 \pm 19.3
Labral tear	Anterior	44 (100%)	9 (100%)
	Posterior	6 (14%)	0
	Lateral	8 (18%)	0
Femoral outerbridge grade			
Anterior	Normal	38 (86%)	9 (100%)
	Mild	3 (7%)	0
	Severe	3 (7%)	0
Posterior	Normal	42 (95%)	9(100%)
	Mild	0	0
	Severe	2 (5%)	0
Superior	Normal	41 (93%)	6 (67%)
	Mild	1 (2%)	2 (22%)
	Severe	2 (5%)	1 (11%)
Acetabular outerbridge grade			
Anterior	Normal	19 (43%)	0
	Mild	11 (25%)	6 (67%)
	Severe	15 (34%)	3 (33%)
Posterior	Normal	37 (84%)	9 (100%)
	Mild	1 (2%)	0
	Severe	6 (14%)	0
Superior	Normal	40 (91%)	6 (67%)
	Mild	1 (2%)	3 (33%)
	Severe	3 (7%)	0

assessing patients following hip arthroscopy using the total score resulted in inflated scores on the MHHS compared with the IHOT-33. Comparing scores on the four separate subscales of the IHOT-33 with the total score of the MHHS revealed discrepancies between with the domains of Sports and Recreational Activities and Social, Emotional

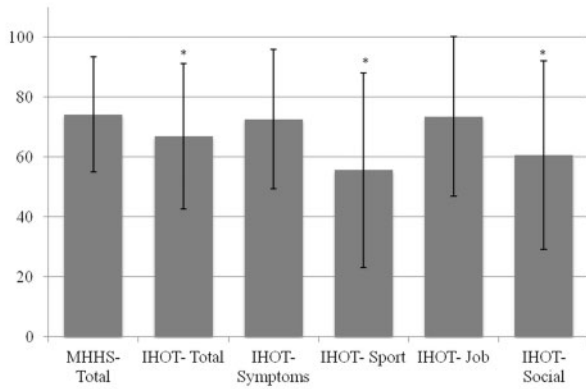


Fig. 1. Average score on the MHHS, IHOT-33 and the IHOT-33 subscales for patients at 1–5 years following hip arthroscopy.

and Lifestyle Concerns. Our results support the use of subscale analysis to garner a more comprehensive depiction of subjective pain and function in hip arthroscopy patients.

Outcome scores measured using the MHHS for our series of patients are in accordance with previous studies, which reported midterm results of hip arthroscopy. Average MHHS scores were reported to range from 70 to 85 points for patients who were 1–5 years post-arthroscopy [11, 12, 24–26]. Total MHHS score for our patients averaged 74 points at a mean of 2-years following surgery. Only one previous study has reported outcomes for hip arthroscopy patients using the IHOT-33 [27]. Total scores for their patients at 18 months following hip preservation surgery fell between 70 and 80 points, while our total score averaged 66 points. Therefore, while reporting midterm results of hip arthroscopy was not the primary aim of our study, based on the similarity of our findings with those of others, we are confident that the cohort of patients in whom we compared the MHHS and IHOT-33 is representative of hip arthroscopy patients as a whole.

Total MHHS score was greater than total IHOT-33 score, but did not differ from scores for two of the separate subscales of the IHOT-33—Symptoms and Functional Limitations and Job Related Concerns. The Symptoms and Functional Limitations domain of the IHOT-33 assesses similar constructs as the MHHS—pain, gait, climbing stairs, getting into/out of a car, sitting and putting on socks/shoes; therefore, one would not expect to see a difference between these two scores. The IHOT-symptoms subscale consists of 16 of the 33 total questions on the IHOT-33. Therefore, roughly 50% of the total score for the IHOT-33 comes from questions that mirror the MHHS. Only 25% of our patients answered every question on the IHOT-33, but all patients answered every question in the IHOT-symptoms domain. This increased the

influence of the IHOT-symptoms subscale on the total IHOT-33 score to >50% for three-quarters of our patients. Because the IHOT-symptoms subscale covers a wider range of daily activities than the MHHS, we feel assessment of these constructs in hip arthroscopy patients is best done by using a subscale analysis of the IHOT-33. The Job-Related Concerns domain consists of four questions, with only one assessing higher level function—lifting or carrying heavy objects—and 29% of our patients opted out of answering this question because it was not part of their job. The questions on the IHOT-job subscale do not address similar constructs as the MHHS; therefore, the lack of difference in score between the two may reflect that hip arthroscopy patients do not feel their hip symptoms affect their job performance any more than they affect their daily activities.

In contrast to the IHOT-symptoms and IHOT-job subscales, total score on the MHHS was greater than scores on the IHOT-sport and IHOT-social domains. Composite scores on these two subscales ranged from 15 to 25% lower than total MHHS score. In total, 18% of patients opted out of answering the cutting activities question of the IHOT-sport domain, while 9% of patients opted out of answering the sexual activity and lifting children questions each of the IHOT-social subscale. The remaining questions inquire about symptoms during and performance of sport-related skills as well as quality of life, which are not addressed in any question on the MHHS. We believe the reduction in scores on these two subscales accounted for the increased total score on the MHHS compared with the total score on the IHOT-33. Given the reduction in scores on these domains when compared with the MHHS, assessing the IHOT-social and IHOT-sport subscales separately will provide valuable novel information regarding potential specific impairments experienced by hip arthroscopy patients.

Our study is not without limitations. First, we chose to only assess the MHHS and IHOT-33 and did not include the other measures, which have been utilized to assess outcomes in patients following hip arthroscopy. However, given that the MHHS is the most commonly assessed outcome measure in this population, we felt it was the most appropriate measure to use for comparison to the IHOT-33. Second, given that the IHOT-33 was not published until 2012, we did not have pre-operative scores for our patients. However, our primary purpose was not to report outcomes for our patients, but to compare scores on different outcome measures collected at the same point in time. Third, our sample size was relatively small, and, thus, parametric tests were inappropriate. However, given that we did find statistically significant differences between the

total MHHS score and the composite IHOT-33 score and the total MHHS score and the IHOT-sport and IHOT-social domains, our study was powered enough to detect these differences. In addition, using our effect sizes, and an alpha level of 0.05, 1270 patients would be needed to obtain 80% power for the MHHS to IHOT-symptoms comparison and 353 409 patients would be needed to obtain 80% power for the MHHS to IHOT-job comparison. Therefore, we feel confident that if a meaningful difference existed, we would have detected it.

Given the importance of patient-reported outcome data in determining the success of a procedure, selection of the measurement tool that captures the most comprehensive information about a patient is critical. The IHOT-33 has been recommended for use in patients undergoing hip arthroscopy based on the strength of its psychometric properties; yet it has been criticized for its use of a composite score. Our results support this criticism and show that the total score of the IHOT-33 may be disproportionately influenced by the Symptoms and Functional Limitations subscale, providing comparative information as the MHHS. However, subscale analysis revealed that the Sports and Recreational Activities and Social, Emotional and Lifestyle Concerns domains capture unique information regarding functional impairments and quality of life in this population. Our results suggest that analysing the subscales of the IHOT-33 would provide a more thorough understanding of functional limitations in patients undergoing hip arthroscopy.

CONFLICT OF INTEREST STATEMENT

None declared.

REFERENCES

- Martin R, Kelly B, Phillippon M. Evidence of validity for the hip outcome score. *Arthroscopy* 2006; 22: 1304–11.
- Mohtadi NG, Griffin DR, Pedersen ME, et al. The development and validation of a self-administered quality-of-life outcome measure for young, active patients with symptomatic hip disease: the International Hip Outcome Tool (iHOT-33). *Arthroscopy* 2012; 28: 595–605.
- Griffin DR, Parsons N, Mohtadi NG, et al. A short version of the International Hip Outcome Tool (iHOT-12) for use in routine clinical practice. *Arthroscopy* 2012; 28: 611–6.
- Thorborg K, Holmich P, Christensen R, et al. The Copenhagen Hip and Groin Outcome Score (HAGOS): development and validation according to the COSMIN checklist. *Br J Sport Med* 2011; 45: 478–91.
- Nilsdotter A, Lohmander LS, Klassbo M, et al. Hip disability and osteoarthritis outcome score (HOOS) - validity and responsiveness in total hip replacement. *BMC Musculoskel Dis* 2003; 4: 10–8.
- Christensen C, Althausen P, Mittleman M, et al. The Nonarthritic Hip Score: Reliable and validated. *Clin Orthop Relat R* 2003; 406: 75–83.
- 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* 1969; 51-A: 737–55.
- Byrd JT, Jones K. Hip arthroscopy in athletes. *Clin Sport Med* 2001; 20: 749–60.
- Meftah M, Rodriguez JA, Panagopoulos G, et al. Long-term results of arthroscopic labral debridement: predictors of outcomes. *Orthopedics* 2011; 34: e588–92.
- Byrd JWT, Jones KS. Prospective analysis of hip arthroscopy with 10-year followup. *Clin Orthop Relat R* 2010; 468: 741–6.
- Santori N, Villar RN. Acetabular labral tears: result of arthroscopic partial limbectomy. *Arthroscopy* 2000; 16: 11–5.
- Kamath AF, Componovo R, Baldwin K, et al. Hip arthroscopy for labral tears: review of clinical outcomes with 4.8-year mean follow-up. *Am J Sport Med* 2009; 37: 1721–7.
- Larson CM, Giveans MR, Stone RM. Arthroscopic debridement versus refixation of the acetabular labrum associated with femoroacetabular impingement: mean 3.5-year follow-up. *Am J Sport Med* 2012; 40: 1015–21.
- Phillippon MJ, Briggs KK, Hayy CJ, et al. Arthroscopic labral reconstruction in the hip using iliotibial band autograft: technique and early outcomes. *Arthroscopy* 2010; 26: 750–6.
- Boykin RE, Patterson D, Briggs KK, et al. Results of arthroscopic labral reconstruction of the hip in elite athletes. *Am J Sport Med* 2013; 41: 2296–301.
- Haviv B, Singh PJ, Takla A, et al. Arthroscopic femoral osteochondroplasty for cam lesions with isolated acetabular chondral damage. *J Bone Joint Surg Br* 2010; 92: 629–33.
- Cooper A, Basheer S, Maheshwari R, et al. Outcomes of hip arthroscopy. *A prospective analysis and comparison between patients under 25 and over 25 years of age*. *Br J Sport Med* 2012; 47: 234–8.
- Phillippon MJ, Briggs KK, Yen YM, et al. Outcomes following hip arthroscopy for femoroacetabular impingement with associated chondrolabral dysfunction: minimum two-year follow-up. *J Bone Joint Surg Br* 2009; 91: 16–23.
- Phillippon MJ, Ejnisman L, Ellis HB, et al. Outcomes 2 to 5 years following hip arthroscopy for femoroacetabular impingement in the patient aged 11 to 16 years. *Arthroscopy* 2012; 28: 1255–61.
- Singh PJ, O'Donnell JM. The outcome of hip arthroscopy in Australian football league players: a review of 27 hips. *Arthroscopy* 2010; 26: 743–9.
- Nilsdotter AK, Lohmander LS, Klassbo M, et al. Hip disability and osteoarthritis outcome score (HOOS)—validity and responsiveness in total hip replacement. *BMC Musculoskel Dis* 2003; 4: 10.
- Kemp JL, Collins NJ, Roos EM, et al. Psychometric properties of patient-reported outcome measures for hip arthroscopic surgery. *Am J Sport Med* 2013; 41: 2065–73.
- Tijssen M, van Cingel R, van Melick N, et al. Patient-Reported Outcome questionnaires for hip arthroscopy: a systematic review

- of the psychometric evidence. *BMC Musculoskel Dis* 2011; 12: 117.
24. Streich NA, Gotterbarm T, Barié A, *et al.* Prognostic value of chondral defects on the outcome after arthroscopic treatment of acetabular labral tears. *Knee Surg Sport Trauma* 2009; 17: 1257–63.
 25. Byrd JT, Jones KS. Prospective analysis of hip arthroscopy with 2-year follow-up. *Arthroscopy* 2000; 16: 578–87.
 26. Burnett SJ, Della Rocca GJ, Prather H, *et al.* Clinical presentation of patients with tears of acetabular labrum. *J Bone Joint Surg* 2006; 88: 1448–57.
 27. Kemp JL, Makdissi M, Schache AG, *et al.* Hip chondropathy at arthroscopy: prevalence and relationship to labral pathology, femoroacetabular impingement and patient-reported outcomes. *Br J Sport Med* 2014; 48: 1102–7.