Minimal Clinically Important Difference and Substantial Clinical Benefit Values for the 12-Item International Hip Outcome Tool

RobRoy L. Martin, Ph.D., P.T., Benjamin R. Kivlan, Ph.D., P.T., John J. Christoforetti, M.D., Andrew B. Wolff, M.D., Shane J. Nho, M.D., M.S., John P. Salvo Jr., M.D., Thomas J. Ellis, M.D., Geoff Van Thiel, M.D., M.B.A., Dean K. Matsuda, M.D., and Dominic S. Carreira, M.D.

Purpose: To define minimal clinically important difference (MCID) and substantial clinical benefit (SCB) values for the 12-item International Hip Outcome Tool (iHOT-12) in patients undergoing hip arthroscopy for intra-articular pathology. **Methods:** This was a retrospective review of prospectively collected data on patients who underwent hip arthroscopy. On initial assessment and follow-up between 335 and 395 days after surgery, subjects completed the iHOT-12 and a categorical self-rating of function (severely abnormal, abnormal, nearly normal, or normal). One-half the standard deviation (SD) of the change in 1-year iHOT-12 scores was used to calculate the MCID. Receiver operator characteristic analysis was performed to determine SCB values. A change in SCB value was determined based on an improvement in the categorical rating of function. Absolute postoperative SCB scores were calculated to determine scores that would be associated with normal function ratings or with abnormal or severely abnormal function ratings. **Results:** Of 1,034 eligible patients, 733 (71%) met the inclusion criteria. The subjects consisted of 537 female patients (73%) and 196 male patients (27%), with a mean age of 35.3 years (SD, 13 years). At a mean of 352 days (SD, 21 days) after surgery, 536 patients (73%) were in the "improved" group and 197 (27%) were in the "not improved" group. The MCID was 13 points. An SCB change score of 28 points was able to identify patients who improved with high sensitivity (0.79) and specificity (0.72). Scores of 86 points or greater and 56 points or less were the cutoff values found to identify subjects who rated their function as normal and abnormal, respectively, with high sensitivity (0.74 and 0.90, respectively) and specificity (0.82 and 0.86, respectively). **Conclusions:** This study provides information to help interpret iHOT-12 scores for a follow-up period ranging between 335 and 395 days with MCID and SCB values of 13 and 28 points, respectively. In addition, a vpatient who scored 86 points or better was likely to have a normal rating of function, whereas a patient with a score of 56 points or less was likely to have an abnormal rating of function. **Level of Evidence:** Level III, retrospective comparative study.

See commentary on page 417

From the Department of Physical Therapy, Rangos School of Health Sciences, Duquesne University (R.L.M., B.R.K.), Pittsburgh, Pennsylvania; University of Pittsburgh Medical Center, Center for Sports Medicine (R.L.M.), Pittsburgh, Pennsylvania; Center for Athletic Hip Injury, Allegheny Health Network (J.J.C.), Pittsburgh, Pennsylvania; Department of Orthopaedic Surgery, Drexel University School of Medicine (J.J.C.), Philadelphia, Pennsylvania; American Hip Institute (J.J.C.), Pittsburgh, Pennsylvania; Hip Preservation and Sports Medicine, Washington Orthopaedics and Sports Medicine (A.B.W.), Washington, DC; Department of Orthopedic Surgery, Division of Sports Medicine, Hip Preservation Center, Rush University Medical Center (S.J.N.), Chicago, Illinois; Orthopaedic Surgery, Sidney Kimmel Medical College, Thomas Jefferson University Hospital (J.P.S.), Philadelphia, Pennsylvania; Hip Arthroscopy Program, Rothman Institute (J.P.S.), Philadelphia, Pennsylvania; Orthopedic One, Ohio Orthopedic Surgery Institute, Dublin Methodist Hospital (T.J.E.), Columbus, Ohio; OrthoIllinois (G.V.T.), Chicago, Illinois; Rush University Medical Center (G.V.T.), Chicago, Illinois; US National Soccer Teams (G.V.T.), Chicago, Illinois; Chicago Blackhawks Medical Network (G.V.T.), Chicago, Illinois; Hip Arthroscopy DISC Sports and Spine Center (D.M.), Marina del Rey, California; and Peachtree Orthopedics (D.S.C.), Atlanta, Georgia, U.S.A.

The authors report the following potential conflicts of interest or sources of funding: J.J.C. is a board member of the International Society for Hip Arthroscopy and paid consultant for Arthrex and has patents with and receives royalties from Arthrex and Breg. A.B.W. is a paid consultant for Stryker. S.J.N. is a board member of the American Orthopaedic Society for Sports Medicine and Arthroscopy Association of North America; a paid consultant for Stryker, Pizot Medical, and Ossur; receives IP royalties from Ossur; receives publishing royalties, financial or material support from Springer; is on the editorial or governing board of American Journal of Orthopedics; and receives research support from Allosource, Arthrex, Athletico, DJ Orthopaedics, Linvatec, Miomed, Smith & Nephew, and Stryker. J.P.S. is a paid consultant for Stryker. T.J.E. receives grant support from Medacta and is a paid consultant for Biomet Sports Medicine. G.V.T. is a consultant for Smith & Nephew, Trainer Rx, Vericel, and Zimmer; receives payments for lectures from Smith & Nephew and Zimmer; receives royalties from Zimmer; and owns stock or stock options in Trainer Rx. D.M. is an editorial board member of Orthopedics Today *but receives no money from that source, is a paid consultant for Zimmer Biomet, and receives royalties for intellectual property from Smith & Nephew and Zimmer Biomet. D.S.C. is a board member of the American Orthopaedic Foot and Ankle Society and is a paid consultant for Biomet. Full ICMJE author disclosure forms are available for this article online, as supplementary material.*

Received June 4, 2018; accepted September 7, 2018.

Address correspondence to RobRoy L. Martin, Ph.D., P.T., Department of Physical Therapy, Rangos School of Health Sciences, Duquesne University, 600 Forbes Ave, Pittsburgh, PA 15282, U.S.A. E-mail: martinr280@duq.edu

0749-8063/18666/$36.00

https://doi.org/10.1016/j.arthro.2018.09.028

Patient-reported outcome measures (PROMs) continue to be emphasized when investigating the outcomes of treatment intervention for musculoskeletal conditions.[1,2] To interpret score changes on a PROM when given periodically over time, there must be appropriate psychometric evidence of responsiveness. A number of PROMs have been studied in hip arthroscopy, including the 33-item International Hip Outcome Tool (iHOT-33) and 12-item International Hip Outcome Tool (iHOT-12).[3,4] The iHOT-33 was developed for research purposes and has evidence to support the interpretation of scores and changes in scores over time.[4-8] The iHOT-12 was developed to have similar psychometric properties to the iHOT-33 but a lower response burden and, therefore, to be more appropriate for clinical use.[3] Although there is evidence of responsiveness for the iHOT-33, information is needed to meaningfully interpret changes in scores for the iHOT-12.

The iHOT-33 and iHOT-12 were designed to measure the impact of hip disease and effect of treatment in young, active individuals.[3,4] One of the advantages of the International Hip Outcome Tool instruments was an extensive item development phase, which included patient involvement in both item generation and assessment of item importance. Initial testing supported use of the iHOT-33 as a valid, reliable, and responsive instrument to assess the effect of treatment in young, active patients with nonarthritic hip problems.[4] The iHOT-33 was recommended for use in a research setting to assess a broad range of symptoms and problems.[4] The iHOT-12 was developed to have similar characteristics of validity, reliability, and responsiveness to the iHOT-33 but with a smaller number of items to reduce patient burden and administrative effort in a clinical setting.[3] It was suggested that the iHOT-12 be used for initial and postoperative follow-up assessments in routine clinical practice. Despite the iHOT-12 being only about one-third the length, Griffin et al.[3] found it to be psychometrically similar to the iHOT-33. Recent studies have continued to support the use of the iHOT-33 with evidence of responsiveness and specific information that allows for interpreting scores and changes in scores.[4-8] Similar evidence is not available for the iHOT-12.

Responsiveness is defined as the ability of an instrument to detect change when a change in status has occurred; it can be defined with values for the minimal clinically important difference (MCID), patient acceptable symptom state (PASS), and substantial clinical benefit (SCB).[8-12] The MCID is defined as the smallest change in outcome that the patient is able to appreciate, whereas the PASS is defined as a satisfactory outcome status.[11] SCB is considered the improvement in outcome or absolute postoperative health state that the patient considers a substantial improvement.[9,10] Nwachukwu et al.[8] proposed that the MCID, PASS, and SCB should be considered the lower, intermediate, and upper thresholds, respectively, for clinically significant outcome improvement. Studies have defined MCID and SCB values for the iHOT-33 but not the iHOT-12.[5-8] The purpose of this study was to define MCID and SCB values for the iHOT-12 in patients undergoing hip arthroscopy for intra-articular pathology. It was hypothesized that the iHOT-12 would be responsive to change in functional status over a 1-year follow-up period for subjects who underwent hip arthroscopy.

Methods

This was a retrospective review of prospectively collected data maintained in a secure electronic registry. The registry consisted of patients who consented to undergo hip arthroscopy at 1 of 7 centers between January 2014 and April 2017. The inclusion criteria specific to this study included subjects who underwent surgery for intra-articular hip pathology with preoperative and follow-up iHOT-12 scores available. The follow-up data were required to be collected between 335 and 395 days after surgery. The exclusion criteria for the registry included patients with primary lumbopelvic pathology, advanced hip arthrosis (Tönnis grade > 1), or other conditions contraindicated for arthroscopic hip surgery. An inability to read or understand English was also an exclusion criterion for the registry. The pre hoc collection and storage of agreed-on clinical data points were granted according to individual institutional requirements, and institutional review board approval was granted to review the deidentified registry of patient data.

On initial assessment, subjects were given the iHOT-12 and a categorical self-rating of function to complete. For follow-up data collection, the iHOT-12 and self-rating of current function were e-mailed to subjects to complete between 335 and 395 days after surgery. Each item on the iHOT-12 was scored using a 100-mm horizontal line (visual analog scale); thus, the mean value of all questions amounts to a score result ranging from 0 to 100, with a higher score meaning better function, fewer symptoms, and higher quality of life.[3] The self-rating of function consisted of the following question: "How would you rate your current level of function?" The subject chose from the following categorical responses: severely abnormal, abnormal, nearly normal, or normal. Demographic information was recorded from the electronic registry.

Psychometric Analysis

The MCID and SCB were calculated using distribution-based and anchor-based methods, respectively, in a

similar fashion to that previously described for the iHOT-33.[6-8] One-half the standard deviation (SD) of the change in 1-year iHOT-12 scores was used to calculate the MCID.[13] An anchor-based method was used to calculate SCB values. The SCB change score and absolute postoperative SCB values were calculated with receiver operator characteristic (ROC) analysis calculating the area under the curve (AUC) at a 95% confidence interval (CI).[14,15] For the SCB change score, an improvement in the categorical rating of function was used to group patients into those who improved and those who did not improve. Absolute postoperative SCB scores were calculated to determine a score that would be associated with a patient response of having normal function and a score that would be associated with having abnormal or severely abnormal function. The AUC of the ROC analysis defines the strength of association and the accuracy of the instrument in distinguishing between groups.[15] An AUC value greater than 0.8 and a 95% CI that does not contain 0.5 are considered excellent properties of responsiveness.[15-17] The Youden index was used to optimize sensitivity and specificity values and identify the best cutoff values for SCB that represent the following: (1) a score change that is most likely to be associated with reporting being in the improved group, (2) an absolute score that is likely to represent a patient who reports a normal function rating, and (3) an absolute score that is likely to represent a patient who reports a function rating of abnormal or severely abnormal.[18] Statistical analysis was performed using the SPSS software package (version 24; IBM, Armonk, NY).

Results

Participants

Of 1,034 eligible patients, 733 (71%) met the inclusion criteria for this study and had outcome data available for analysis. The average follow-up time was 352 days (SD, 21 days). Demographic information, including, age, sex, diagnosis, and procedure performed, is presented in Table 1. It should be noted that 577 subjects (79%) underwent multiple procedures.

Psychometric Results

The mean preoperative iHOT-12, postoperative iHOT-12, and change in iHOT-12 scores were 34 points (SD, 17 points), 70 points (SD, 25 points), and 37 points (SD, 26 points), respectively. Preoperative and postoperative ratings of function are provided in Table 2. At follow-up, 536 patients (73%) were in the "improved" group and 197 (27%) were in the "not improved" group. The MCID for the change score was 13 points. The results of the ROC analysis for the change in SCB for the iHOT-12 score and absolute scores associated with a normal rating of function and with an abnormal rating of function are presented in Table 3. A change score of 28 points was able to identify patients who had a change in function versus those who did not with high sensitivity and specificity. Values of 86 points and 56 points were the cutoff scores found to be able to identify subjects who rated their function as normal and abnormal, respectively, with high sensitivity and specificity.

Table 1. Subject Demographic Characteristics

	Data
Mean age, yr	35.3 (SD, 13)
Sex, n	
Female	537 (73%)
Male	196 (27%)
Diagnosis, n	
Femoroacetabular impingement	677 (92%)
Labral pathology	659 (90%)
Procedures performed, n	
Femoroplasty	608 (83%)
Synovectomy	586 (80%)
Labral repair	557 (76%)
Acetabuloplasty	425 (58%)
Acetabular chondroplasty	410 (56%)
Femoral chondroplasty	169 (23%)
Labral reconstruction	139 (19%)

SD, standard deviation.

Discussion

This study found the iHOT-12 MCID and SCB values to be 13 and 28 points, respectively, with a score of 86 points or better identifying a normal rating of function and a score of 56 points or less identifying an abnormal rating of function. These results offer evidence of responsiveness for the iHOT-12 for a follow-up period ranging between 335 and 395 days after surgery in patients who underwent hip arthroscopy for intra-articular hip pathology. The iHOT-12 was found to have acceptable properties of responsiveness to change because the AUC value was greater than 0.80 with a 95% CI not containing 0.5. This information can be used to assess postsurgical outcomes.

Patient-centered assessments using PROMs have been increasingly used to report outcomes after orthopaedic procedures.[12] The information obtained from a PROM is only valuable if there is evidence of reliability, validity, and responsiveness to support its use.[19] Responsiveness may be the most relevant psychometric property to interpret changes in scores over time on a PROM.[19] There are a number of ways to determine responsiveness.[10,12,14] Obtaining values for the MCID and SCB as part of evidence of responsiveness has the advantage of providing clinicians with information that can directly help to interpret patient scores obtained during follow-up assessments in a clinical setting. The MCID is commonly determined using statistical

Table 2. Preoperative and Postoperative Rating of Function

Rating of Function	Preoperatively, n	Follow-Up (335-395 d), n
Normal	8 (1%)	224 (30%)
Nearly normal	138 (19%)	350 (48%)
Abnormal	461 (63%)	137 (19%)
Severely abnormal	126 (17%)	22 (3%)

methods and represents only the minimal requirement for improvement.[10] The SCB value is an alternative psychometric value that can define improvement based on patient-centered criteria using an anchor-based question and response.[6,9] When an anchor-based method is used to determine SCB values, the assessment of change is from the patient's perspective.[12] In our study, the patient-centered criterion was patients' self-rating of current function as severely abnormal, abnormal, nearly normal, or normal. The SCB values obtained for the iHOT-12 based on a change in function rating and reported current level of function can be used as part of patient-centered assessment and may be combined with standard clinical assessment for a comprehensive evaluation of a patient's status.

The iHOT-33 and iHOT-12 were developed as PROMs for use in young, active patients with nonarthritic hip problems.[3,4] For the iHOT-33, rigorous psychometric methodology in development and patient input were used to create a comprehensive list of relevant items that assess symptoms; quality of life (social, emotional, and lifestyle concerns); and functional limitations in activities of daily living, sports or recreational activities, and work.[4] The iHOT-33 has evidence of validity, reliability, and responsiveness.[4-8,20] However, its length was thought to be a potential concern for clinicians who want to use a PROM for all patients at all clinical visits.[3] This concern led to the development of a shorter version of the International Hip Outcome Tool. The iHOT-12 was developed by selecting 12 items from the original iHOT-33 based on psychometric evidence. Regression analysis found these 12 items accounted for greater than 96% of the total variation in the iHOT-33 score.[3] In addition, the iHOT-12 scores showed excellent agreement with the iHOT-33 and almost identical sensitivity to change after treatment.[3] Therefore, despite being 60% shorter, the iHOT-12 collects information similar to the original 33-item questionnaire.

A number of PROMs have been evaluated in systematic reviews for outcome assessment of patients who undergo hip arthroscopy.[20-22] Kemp et al.[5] evaluated the psychometric properties of the Modified Harris Hip Score, Hip Outcome Score, Hip Disability and Osteoarthritis Outcome Score, Copenhagen Hip and Groin Outcome Score, and iHOT-33 in a general hip arthroscopy population. Their study found these 5 instruments to have appropriate test-retest reliability, validity, and responsiveness, with the Hip Disability and Osteoarthritis Outcome Score and iHOT-33 performing best.[5] The iHOT-12, however, was not included in the study. At a 1-year follow-up interval, studies have determined the SCB values for the iHOT-33 using ROC analysis with the anchor-based question "How would you rate your overall physical ability?" and the response choices as follows: no change, slightly worse, worse, slightly improved, or much improved.[6,8] These studies have found the change score for SCB in identifying patients who were "unchanged" from "much improved" to be 24.5 points in those who underwent hip arthroscopic surgery for femoroacetabular impingement[6] and 25.2 points after revision hip arthroscopy.[8] These numbers are comparable with the change score of 28 points identified in our study of the iHOT-12 in a general population of patients undergoing hip arthroscopy. An absolute SCB score of 63.5 points on the iHOT-33 was able to identify patients who were much improved versus those who were unchanged among patients who underwent hip arthroscopic surgery for femoroacetabular impingement.[6] This score was similar to the absolute SCB value of 56 points or less to identify an abnormal rating of function in our study. Any difference in these SCB values could be attributed to the difference in the anchor-based question and responses as well as item differences between the iHOT-12 and iHOT-33.

The iHOT-12 has been internationally recognized and translated into Dutch, German, Portuguese, Spanish, and Swedish with evidence of reliability, validity, and responsiveness to support its use.[23-26] The German version of the iHOT-12 identified an important change in score of 14 points over an approximately 1-month period.[23] This value was determined by asking the patients to rate the change in their hip condition over approximately 1 month as much better, somewhat

Table 3. Receiver Operator Characteristic Analysis for iHOT-12 Substantial Clinical Benefit Change Score and Absolute Scores Associated With Normal and Abnormal Ratings of Function

	Score	Sensitivity	Specificity	AUC (95% CI)
Change score to identify improvement in function	≥28 points	0.79	0.72	0.82 (0.79-0.86)
Score to identify normal rating of function	≥86 points	0.74	0.82	0.87 (0.85-0.90)
Score to identify abnormal rating of function	≤56 points	0.90	0.86	0.95 (0.93-0.96)

AUC, area under curve; CI, confidence interval.

better, unchanged, somewhat worse, or much worse compared with their condition at initial evaluation.[23] The change in score over approximately 1 month was predictably lower than the SCB change in score of 28 points over approximately 1 year in our study.

Limitations

Although this study provides evidence of responsiveness with MCID and SCB values for the iHOT-12, there are a number of limitations. The study did not define a PASS value for the iHOT-12. The findings of our study are only applicable to patients with a follow-up period ranging between 335 and 395 days. Further studies are needed to provide specific information to help interpret scores for other evaluation periods that clinicians may typically encounter. This study is also limited to the anchor-based question and responses used herein. One of the main limitations of using anchor-based methods when determining MCID and SCB values is that the question and responses are determined by researchers and therefore do not necessarily represent what the patient considers worthwhile. Other methods to determine SCB values may provide slightly different results. This study included subjects with intra-articular hip pathology; therefore, the results cannot be used for other hip conditions. However, in most patients who undergo hip arthroscopy, multiple procedures are performed at the same time. The patients used in this data set, consisting of over 700 patients from 7 different centers across the United States, represent a fairly typical population that surgeons will encounter when using the iHOT-12. It should be noted that there was less than 80% follow-up and most subjects were female patients, which could possibly introduce bias. In relation to sex bias, Nwachukwu et al.[6] did not show that sex influenced outcome. In addition, the values for SCB change scores may vary depending on the baseline level of function of the subjects.[7]

Conclusions

This study provides information to help interpret iHOT-12 scores for a follow-up period ranging between 335 and 395 days with MCID and SCB values of 13 and 28 points, respectively. In addition, a patient who scored 86 points or better was likely to have a normal rating of function, whereas a patient with a score of 56 points or less was likely to have an abnormal rating of function.

References

1. Mercieca-Bebber R, Rouette J, Calvert M, et al. Preliminary evidence on the uptake, use and benefits of the CONSORT-PRO extension. *Qual Life Res* 2017;26:1427-1437.
2. Calvert M, Kyte D, Duffy H, et al. Patient-reported outcome (PRO) assessment in clinical trials: A systematic review of guidance for trial protocol writers. *PloS One* 2014;9:e110216.
3. Griffin DR, Parsons N, Mohtadi NG, Safran MR, Multicenter Arthroscopy of the Hip Outcomes Research Network. A short version of the International Hip Outcome Tool (iHOT-12) for use in routine clinical practice. *Arthroscopy* 2012;28:611-616. quiz 616-618.
4. 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;28:595-610. quiz 606-610.e1.
5. Kemp JL, Collins NJ, Roos EM, Crossley KM. Psychometric properties of patient-reported outcome measures for hip arthroscopic surgery. *Am J Sports Med* 2013;41:2065-2073.
6. Nwachukwu BU, Chang B, Fields K, et al. Defining the "substantial clinical benefit" after arthroscopic treatment of femoroacetabular impingement. *Am J Sports Med* 2017;45:1297-1303.
7. 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;45:612-619.
8. Nwachukwu BU, Chang B, Rotter BZ, Kelly BT, Ranawat AS, Nawabi DH. Minimal clinically important difference and substantial clinical benefit after revision hip arthroscopy. *Arthroscopy* 2018;34:1862-1868.
9. Glassman SD, Copay AG, Berven SH, Polly DW, Subach BR, Carreon LY. Defining substantial clinical benefit following lumbar spine arthrodesis. *J Bone Joint Surg Am* 2008;90:1839-1847.
10. Katz NP, Paillard FC, Ekman E. Determining the clinical importance of treatment benefits for interventions for painful orthopedic conditions. *J Orthop Surg Res* 2015;10:24.
11. Kvien TK, Heiberg T, Hagen KB. Minimal clinically important improvement/difference (MCII/MCID) and patient acceptable symptom state (PASS): What do these concepts mean? *Ann Rheum Dis* 2007;66:iii40-iii41 (suppl 3).
12. Harris JD, Brand JC, Cote MP, Faucett SC, Dhawan A. Research pearls: The significance of statistics and perils of pooling. part 1: Clinical versus statistical significance. *Arthroscopy* 2017;33:1102-1112.
13. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: The remarkable universality of half a standard deviation. *Med Care* 2003;41:582-592.
14. Beaton DE. Understanding the relevance of measured change through studies of responsiveness. *Spine (Phila Pa 1976)* 2000;25:3192-3199.
15. Lasko TA, Bhagwat JG, Zou KH, Ohno-Machado L. The use of receiver operating characteristic curves in biomedical informatics. *J Biomed Inform* 2005;38:404-415.
16. Park SH, Goo JM, Jo CH. Receiver operating characteristic (ROC) curve: Practical review for radiologists. *Korean J Radiol* 2004;5:11-18.
17. Gortmaker SL, Hosmer DW, Lemeshow S. Applied logistic regression. *Contemp Sociol* 1994;23:159.

18. Schisterman EF, Perkins NJ, Liu A, Bondell H. Optimal cut-point and its corresponding Youden Index to discriminate individuals using pooled blood samples. *Epidemiology* 2005;16:73-81.
19. Guyatt G, Walter S, Norman G. Measuring change over time: Assessing the usefulness of evaluative instruments. *J Chronic Dis* 1987;40:171-178.
20. Thorborg K, Roos EM, Bartels EM, Petersen J, Holmich P. Validity, reliability and responsiveness of patient-reported outcome questionnaires when assessing hip and groin disability: A systematic review. *Br J Sports Med* 2010;44:1186-1196.
21. 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;46:632-643.
22. Tijssen M, van Cingel R, van Melick N, de Visser E. Patient-reported outcome questionnaires for hip arthroscopy: A systematic review of the psychometric evidence. *BMC Musculoskelet Disord* 2011;12:117.
23. Baumann F, Popp D, Muller K, et al. Validation of a German version of the International Hip Outcome Tool 12 (iHOT12) according to the COSMIN checklist. *Health Qual Life Outcomes* 2016;14:3.
24. Jonasson P, Baranto A, Karlsson J, et al. A standardised outcome measure of pain, symptoms and physical function in patients with hip and groin disability due to femoro-acetabular impingement: Cross-cultural adaptation and validation of the international Hip Outcome Tool (iHOT12) in Swedish. *Knee Surg Sports Traumatol Arthrosc* 2014;22:826-834.
25. Stevens M, van den Akker-Scheek I, ten Have B, Adema M, Giezen H, Reininga IH. Validity and reliability of the Dutch version of the International Hip Outcome Tool (iHOT-12NL) in patients with disorders of the hip. *J Orthop Sports Phys Ther* 2015;45:1026-1034, A1021-A1022.
26. Polesello GC, Godoy GF, Trindade CA, de Queiroz MC, Honda E, Ono NK. Translation and cross-cultural adaptation of the International Hip Outcome Tool (iHOT) into Portuguese. *Acta Ortop Bras* 2012;20:88-92B.