Diagnostic Accuracy of 3 Physical Examination Tests in the Assessment of Hip Microinstability

Daniel J. Hoppe,* MD, MEd, FRCSC, Jeremy N. Truntzer,* MD, Lauren M. Shapiro,* MD, Geoffrey D. Abrams,* MD, and Marc R. Safran,*† MD

Investigation performed at the Stanford Sports Medicine Clinic, Redwood City, California, USA

Background: Hip microinstability is a diagnosis gaining increasing interest. Physical examination tests to identify microinstability have not been objectively investigated using intraoperative confirmation of instability as a reference standard.

Purpose: To determine the test characteristics and diagnostic accuracy of 3 physical examination maneuvers in the detection of hip microinstability.

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

Methods: A review was conducted of 194 consecutive hip arthroscopic procedures performed by a sports medicine surgeon at a tertiary-care academic center. Physical examination findings of interest, including the abduction–hyperextension–external rotation (AB-HEER) test, the prone instability test, and the hyperextension–external rotation (HEER) test, were obtained from prospectively collected data. The reference standard was intraoperative identification of instability based on previously published objective criteria. Test characteristics, including sensitivity, specificity, positive and negative predictive values, and accuracy, were calculated for each test as well as for combinations of tests.

Results: A total of 109 patients were included in the analysis. The AB-HEER test was most accurate, with a sensitivity of 80.6% (95% CI, 70.8%-90.5%) and a specificity of 89.4% (95% CI, 80.5%-98.2%). The prone instability test had a low sensitivity (33.9%) but a very high specificity (97.9%). The HEER test performed second in both sensitivity (71.0%) and specificity (85.1%). The combination of multiple tests with positive findings did not yield significantly greater accuracy. All tests had high positive predictive values (range, 86.3%-95.5%) and moderate negative predictive values (range, 52.9%-77.8%). When all 3 tests had positive findings, there was a 95.0% (95% CI, 90.1%-99.9%) chance that the patient had microinstability.

Conclusion: The AB-HEER test most accurately predicted hip instability, followed by the HEER test and the prone instability test. However, the high specificity of the prone instability test makes it a useful test to “rule in” abnormalities. A positive result from any test predicted hip instability in 86.3% to 90.9% of patients, but a negative test result did not conclusively rule out hip instability, and other measures should be considered in making the diagnosis. The use of these tests may aid the clinician in diagnosing hip instability, which has been considered a difficult diagnosis to make because of its dynamic nature.

Keywords: hip; groin pain; hip instability; hip arthroscopic surgery; clinical assessment

Atraumatic hip instability, or microinstability, is a relatively new clinical entity that has been increasingly recognized as a cause of disability and pain in young patients and athletes.[17] It can be defined as extraphysiological hip motion or pathological laxity that leads to symptomatic abnormal mechanics of the hip.[27] The cause can be either traumatic or atraumatic; however, the pathophysiological mechanism is related to repetitive axial or rotational loading of the hip joint in the setting of subtle anatomic abnormalities of the structures that typically provide stability to the joint, leading to labral and articular cartilage damage.[5,27] Contributing factors include ligamentous or capsular laxity, muscular weakness of the hip and pelvic girdle, and repetitive hip joint loading with sporting activity; however, a direct cause may not always be identified.[5,21,27]

Symptomatic microinstability remains an elusive diagnosis in part because the pertinent history, physical examination, and radiographic findings have not been clearly

†Address correspondence to Marc R. Safran, MD, Department of Orthopaedic Surgery, Stanford University, 450 Broadway Street, MC6342, Redwood City, CA 94603, USA (email: msafran@stanford.edu).

*Department of Orthopaedic Surgery, Stanford University, Redwood City, California, USA.

The authors declared that they have no conflicts of interest in the authorship and publication of this contribution.

Ethical approval for this study was obtained from Stanford University (protocol No. 36570).

The Orthopaedic Journal of Sports Medicine, 5(11), 2325967117740121
DOI: 10.1177/2325967117740121

defined. Complicating the diagnosis is the depth of the hip joint, surrounded by a complex soft tissue envelope, and its constrained nature of a ball within a socket. Moreover, the dynamic nature of the disorder makes traditional diagnostic techniques such as radiography and magnetic resonance imaging less helpful. Many clinicians find hip microinstability analogous to shoulder microinstability 30 years ago: a deep joint with no objective clinical measures of microinstability.

While physical examination maneuvers afford the assessment of dynamic changes, the literature contains few descriptions of the tests used to evaluate for hip microinstability, and there are no reports on the diagnostic utility of these tests. Three tests for eliciting instability have been proposed: the abduction–hyperextension–external rotation (AB-HEER) test,[11] the prone instability test,[12] and the hyperextension–external rotation (HEER) test.[17] Knowing the usefulness of these tests will become important as the understanding of hip instability improves and awareness of this entity continues to grow.

The purpose of this study was to evaluate these 3 tests and determine which test, or combinations of tests, would provide the most accurate diagnosis of hip instability by determining their sensitivity, specificity, positive and negative predictive values, and overall accuracy using arthroscopically identified hip instability as a reference standard.

METHODS

We retrospectively reviewed the charts of 194 consecutive patients who underwent hip arthroscopic surgery performed by a sports medicine fellowship–trained, high-volume hip arthroscopic surgeon (M.R.S.) at a tertiary-care academic center between December 2014 and March 2016. The study was approved by an institutional review board (Stanford University, protocol No. 36570).

Preoperatively, all patients were evaluated for gait, hip range of motion, and hip strength, and special tests were performed, including the flexion/adduction/internal rotation (FADIR) and labral stress tests, Patrick test, Trendelenburg sign, and ligamentous laxity using the Beighton criteria.[3] Instability tests were performed by the senior author (M.R.S.) for all patients. Patient demographics, postoperative diagnosis, and procedure performed were also recorded.

Patients were excluded from the study if they had any radiographic features of moderate-to-severe acetabular dysplasia, including a Tönnis angle of >10° and a center-edge angle of Wiberg of <20°, or if there were any missing physical examination data from the preoperative evaluation.

The AB-HEER test has been described by Domb et al.[11] This test is performed with the patient in the lateral decubitus position with the affected hip placed upward. The hip is abducted to 30° to 45°, extended, and externally rotated, while an anteriorly directed force is applied to the posterior greater trochanter. The test finding is positive when this results in anterior hip pain (Figure 1).

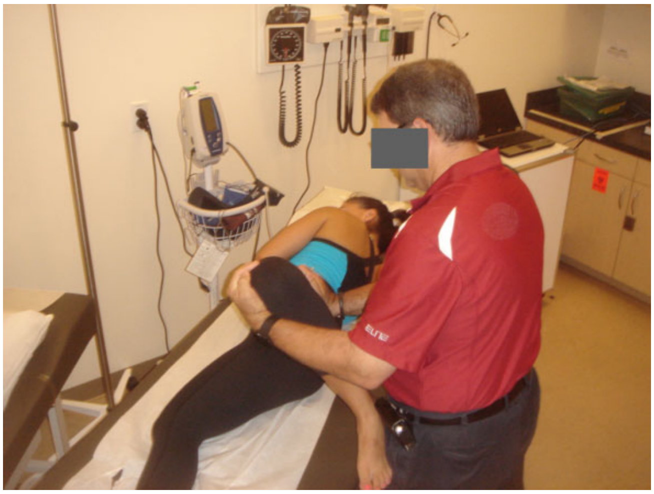

Figure 1. The abduction–hyperextension–external rotation (AB-HEER) test.

Figure 2. The prone instability test.

The prone instability test as described by Domb et al[12] is performed with the patient in the prone position. The hip is externally rotated, while the examiner applies a downward force on the posterior greater trochanter. The reproduction of anterior hip pain is consistent with a positive test result for instability (Figure 2).

The anterior apprehension test, or HEER test,[17,25] is performed with the patient supine at the foot of the table with the legs dangling free. Applying an anteriorly directed force at the hip, the contralateral hip is flexed, while the ipsilateral hip is hyperextended and externally rotated. A positive test result reproduces the patient's anterior hip pain (Figure 3).

For all 3 tests, the hip is externally rotated, which has been shown to result in anterior translation of the femoral head. The prone instability test evaluates the hip in neutral flexion-extension and neutral abduction-adduction, while the examiner applies an anteriorly directed force. The

Figure 3. The hyperextension–external rotation (HEER) test.

HEER test evaluates the hip in hyperextension with neutral abduction-adduction, with no additional force by the examiner. Finally, the AB-HEER test evaluates the hip in abduction and hyperextension with the examiner applying an anteriorly directed force.

The reference standard used to determine the diagnostic accuracy of all 3 tests was a diagnosis of hip microinstability made at the time of surgery (Table 1). Instability was confirmed if 1 or more of the following criteria were met, in accordance with the recent literature[1,16,26]: (1) distraction of the hip under general anesthesia with body weight alone; (2) adequate distraction of the hip joint with less than 11 turns of fine traction, equivalent to 44 mm of screw traction (MIS Hip Interventions table; Maquet); (3) inability of the hip to fully reduce the joint after negative intra-articular pressure is released and traction is removed; and (4) arthroscopic confirmation of microinstability, including tearing of the ligamentum teres, straight anterior labral tears, and an anterior inside-out chondral wear pattern.

For patients with identified instability, arthroscopic plication was performed by passing three No. 2 nonabsorbable sutures (FiberWire and TigerWire; Arthrex) using a crescent suture passer (ACCU-PASS Suture Shuttle; Smith & Nephew) through both limbs of the partial capsulectomy and tying the sutures outside the hip joint.[16]

Statistical Analysis

Statistics were tabulated using SPSS (Version 24; IBM). A 2-by-2 table was created to determine the sensitivity and specificity, with 95% CIs, for each test (Table 2). Positive and negative likelihood ratios were then calculated from these data. The positive predictive value, negative predictive value, and accuracy were also calculated.

The characteristics of combinations of tests were calculated to check for any possible improvement in the ability to diagnose microinstability. This was done for at least 1 test with positive results, at least 2 tests with positive results, and all tests with positive results.

TABLE 1
Criteria for Intraoperative Diagnosis of Hip Instability

Full distraction with body weight traction alone
Adequate distraction with <11 turns of fine traction
Inability to fully reduce hip after hip is vented
Arthroscopic findings:
Extensive tearing of ligamentum teres
Straight anterior labral tears (4 to 2 o'clock)
Anterior inside-out chondral wear pattern

TABLE 2
2 × 2 Table Used to Calculate Diagnostic Accuracy of Each Test[a]

	Disease	
Test	Positive	Negative
Positive	True positive (TP)	False positive (FP)
Negative	False negative (FN)	True negative (TN)

[a]Values were calculated with the following formulas: sensitivity = TP/(TP + FN); specificity = TN/(TN + FP); positive predictive value = TP/(TP + FP); negative predictive value = TN/(FN + TN); accuracy = (TP + TN)/(TP + FP + FN + TN); and likelihood ratio = sensitivity/(1 − specificity).

RESULTS

Of the 194 consecutive patients who underwent hip arthroscopic surgery, 85 patients did not undergo instability testing because a diagnosis of microinstability was not suspected. The remaining 109 patients underwent instability testing and were included in the analysis; there were 44 men and 65 women. The mean age at the time of surgery was 27.8 years (range, 13-58 years). Ten of the patients had previously undergone hip arthroscopy and were being evaluated for recurrent symptoms. Sixty-two patients (56.9%) were found to have intraoperative instability, all of whom underwent capsular plication. Table 3 shows the demographics, preoperative diagnosis, and distribution of procedures performed on the study sample.

The diagnostic measures for each of the 3 individual tests are summarized in Tables 4 and 5. The AB-HEER test was overall the most accurate test, having the highest sensitivity (80.6%), negative predictive value (77.8%), overall accuracy (84.4%), and second highest specificity (89.4%). The most specific test was the prone instability test (97.9%), which also had the highest positive predictive value (95.5%) but a very low sensitivity of 33.9%. The HEER test performed second best in both sensitivity (71.0%) and overall accuracy (77.1%). The positive predictive value was high for all 3 tests in isolation (range, 86.3-95.5%). However, only the AB-HEER test performed above 75% for the negative predictive value.

As shown in Table 6, when combining tests, an increase in specificity was observed with an increasing number of positive test findings, which coincided with an expected increase in the likelihood ratio. Sensitivity decreased as the number of positive test findings increased. When all 3 tests

TABLE 3
Demographics and Distribution of Procedures Performed for Both Groups

	Stable Group (n = 47)	Microinstability Group (n = 62)
Mean age, y	29.9	26.3
Female, %	17.0	91.9
Traumatic injury, %	68.1	48.4
Ligamentous laxity, %	25.0	97.2
Revision surgery, %	10.6	8.1
Type of impingement, %		
Cam	23.4	43.5
Pincer	17.0	19.4
Mixed	59.6	33.9
Type of surgery, %		
Labral debridement	51.1	27.4
Labral repair	42.6	56.4
No labral surgery	6.4	16.1
Chondroplasty	97.9	83.9
Femoral osteochondroplasty	83.0	64.5
Acetabuloplasty	76.6	46.8
Microfracture	21.3	14.5
Capsular plication	0.0	100.0
Mean No. of turns on traction table	18.4	8.7

had positive results, there was a 95.0% chance (95% CI, 90.1%-99.9%) that a patient had an intraoperative diagnosis of microinstability.

DISCUSSION

The purpose of our study was to evaluate the 3 most commonly published physical examination maneuvers in detecting hip microinstability; we used arthroscopically identified hip microinstability as a reference standard and assessed the characteristics of each test alone and in combination with one another.

Based on our findings, the AB-HEER test is the most accurate predictor of hip instability because of its combination of high sensitivity and specificity. However, all 3 tests showed specificities and positive predictive values above 85%. Relatedly, the high likelihood ratios, an indicator of a test's utility and how likely a patient has a disease or condition, further supported the validity of the 3 tests.[4,15] Overall, a positive prone instability test result may lead a clinician to have a high suspicion that a patient has microinstability, although a negative test result should not rule this out because of the test's low sensitivity. Conversely, a negative AB-HEER test finding should

TABLE 4
Sensitivities, Specificities, and Likelihood Ratios for Hip Instability Testing[a]

Test	Sensitivity (95% CI), %	Specificity (95% CI), %	Likelihood Ratio (95% CI)
AB-HEER test	80.6 (70.8-90.5)	89.4 (80.5-98.2)	7.6 (3.3-17.5)
Prone instability test	33.9 (22.1-45.7)	97.9 (93.7-100.0)	15.9 (2.2-114.2)
HEER test	71.0 (59.7-82.3)	85.1 (74.9-95.3)	4.8 (2.4-9.6)

[a]AB-HEER, abduction–hyperextension–external rotation; HEER, hyperextension–external rotation.

TABLE 5
Positive Predictive Values, Negative Predictive Values, and Accuracy for Hip Instability Testing[a]

Test	Positive Predictive Value (95% CI), %	Negative Predictive Value (95% CI), %	Accuracy (95% CI)
AB-HEER test	90.9 (87.0-94.8)	77.8 (72.1-83.4)	84.4 (78.2-90.6)
Prone instability test	95.5 (91.0-99.9)	52.9 (47.5-58.2)	61.5 (54.5-68.4)
HEER test	86.3 (81.5-91.1)	69.0 (62.9-75.0)	77.1 (72.9-86.8)

[a]AB-HEER, abduction–hyperextension–external rotation; HEER, hyperextension–external rotation.

TABLE 6
Diagnostic Values for Combinations of Tests for Hip Instability

	Sensitivity (95% CI), %	Specificity (95% CI), %	Likelihood Ratio (95% CI)
≥1 test with positive results	87.1 (78.8-95.4)	78.7 (67.0-90.4)	4.1 (2.2-7.7)
≥2 tests with positive results	67.7 (56.1-79.4)	95.7 (91.7-99.8)	15.9 (4.1-62.5)
All 3 tests with positive results	30.6 (19.2-42.1)	97.9 (94.7-100.0)	14.4 (2.0-104.8)

effectively rule out most patients who do not have instability. In comparison to shoulder instability, Farber et al[13] found that the anterior apprehension test had a sensitivity and specificity of 81% and 92%, respectively, and that the anterior drawer test had a sensitivity and specificity of 53% and 85%, respectively.

We also observed that a combination of positive test results increased the specificity and likelihood ratio above that of any single test. Although all 3 tests are used to reproduce anterior instability, this is not unexpected, given that each test stresses the hip in a slightly different manner. As each test is relatively quick to perform, the improvement in diagnostic utility with a combination of tests should help the physician to become familiar with all 3 tests. This is not dissimilar to a physical examination of other joints including the shoulder and knee, where a "toolbox" of tests is employed to examine subtle variations of a single problem (ie, shoulder impingement, posterolateral instability of the elbow, anterior cruciate ligament insufficiency).[19,23-25,28]

Although hip arthroscopic surgery has been demonstrated to be successful for specific conditions, such as femoroacetabular impingement (FAI) in the absence of significant arthritis,[2,6-9,18,22] there is a subset of patients that have continued symptoms and require revision surgery.[10,14] Residual FAI has previously been the most common reason for this, but revision surgery often addresses instability in the form of capsular lesions, laxity, and attenuation, which may be iatrogenic because of performing T-capsulotomy during surgery and/or incomplete capsular closure.[10,20,30] Additionally, Wenger et al[29] demonstrated that 87% of those presenting with symptomatic labral tears have bony dysmorphology (developmental dysplasia of the hip, FAI). It is quite possible that the other 13% have labral tears as a result of microinstability. The ability to reliably evaluate patients requiring primary hip arthroscopic surgery for instability or to identify instability in revision hip arthroscopic surgery as a source of pain is critical to improving outcomes and limiting secondary operative procedures. In our patient cohort, 10 patients had previously undergone hip arthroscopic surgery. The 5 patients who required plication all had positive instability rest results. Therefore, these tests appear to be reliable in evaluating for instability in patients who had previously undergone surgery; however, this finding warrants further study because of the low number of revision patients in our series.

Selection bias and study design represent 2 limitations of this study. The sample included consecutive patients who were preoperatively evaluated for hip arthroscopic surgery for a variety of reasons. Because of the referral pattern of our center, a higher percentage of patients may present with instability than observed in other practices. This is reflected in the high rate of capsular plications performed.

Because our study sample likely had a higher rate of microinstability than the general population, this may have increased the positive predictive values and decreased the negative predictive values, as the calculation of these tests is affected by the prevalence of the disease. The prevalence does not, however, affect the accuracy, sensitivity, and specificity of the calculations. We felt that this was justified by the fact that, in clinical practice, every physical examination test is not performed on every patient; rather, the tests for instability will primarily be used when there exists pretest probability that a patient may have a component of hip instability. An additional limitation includes the lack of blinding with regard to the physical examination results by the surgeon. In this study, the decision to perform capsular plication was based on strict objective criteria, and the physical examination results did not influence this decision. Lastly, because of the nature of the study design, with only the senior author performing the physical examination maneuvers, we cannot comment on the interreliability of these specific tests. However, we feel that a single evaluator for the initial analysis is a strength of the study because it eliminated the variability that may exist between multiple evaluators.

Although no studies exist that currently validate a reference standard with which to universally define instability intraoperatively, we did systematically define intraoperative instability upon which surgical decisions were made. These physical examination maneuvers, when used together, were reliable at predicting which patients had instability and required capsular plication. Further studies may include the ability of these tests to differentiate between iatrogenic, traumatic, and atraumatic instability.

CONCLUSION

Hip arthroscopic surgery continues to evolve as our understanding of the sources of hip pain improves. Microinstability has more recently emerged as a source of disability and pain. Unfortunately, the diagnosis of microinstability based on a physical examination has not been well defined. In this study, 3 physical examination tests were shown to have utility in diagnosing hip instability when evaluated against the reference standard of microinstability diagnosed at the time of arthroscopic surgery. Future studies are required to further validate these tests.

REFERENCES

1. Abrams GD, Luria A, Sampson J, et al. Decreased synovial inflammation in atraumatic hip microinstability compared with femoroacetabular impingement. *Arthroscopy*. 2017;33(3):553-558.
2. Alradwan H, Philippon MJ, Farrokhyar F, et al. Return to preinjury activity levels after surgical management of femoroacetabular impingement in athletes. *Arthroscopy*. 2012;28(10):1567-1576.
3. Beighton P. Hypermobility scoring. *Br J Rheumatol*. 1988;27(2):163.
4. Bhandari M, Guyatt GH. How to appraise a diagnostic test. *World J Surg*. 2005;29(5):561-566.
5. Boykin RE, Anz AW, Bushnell BD, Kocher MS, Stubbs AJ, Philippon MJ. Hip instability. *J Am Acad Orthop Surg*. 2011;19(6):340-349.
6. Boykin RE, Patterson D, Briggs KK, Dee A, Philippon MJ. Results of arthroscopic labral reconstruction of the hip in elite athletes. *Am J Sports Med*. 2013;41(10):2296-2301.
7. Byrd JW. Femoroacetabular impingement in athletes: current concepts. *Am J Sports Med*. 2014;42(3):737-751.
8. Byrd JW, Jones KS. Hip arthroscopy for labral pathology: prospective analysis with 10-year follow up. *Arthroscopy*. 2009;25(4):365-368.

9. Byrd JW, Jones KS. Prospective analysis of hip arthroscopy with 10-year follow up. *Clin Orthop Relat Res*. 2010;468(3):741-746.
10. Cvetanovich GL, Harris JD, Erickson BJ, Bach BR Jr, Bush-Joseph CA, Nho SJ. Revision hip arthroscopy: a systematic review of diagnoses, operative findings, and outcomes. *Arthroscopy*. 2015;31(7):1382-1390.
11. Domb BG, Brooks AG, Guanche CA. Physical examination of the hip. In: Guanche CA, ed. *Hip and Pelvis Injuries in Sports Medicine*. Philadelphia: Wolters Kluwer/Lippincott Williams and Wilkins; 2010:62-70.
12. Domb BG, Stake CE, Lindner D, El-Bitar Y, Jackson TJ. Arthroscopic capsular plication and labral preservation in borderline hip dysplasia: two-year clinical outcomes of a surgical approach to a challenging problem. *Am J Sports Med*. 2013;41(11):2591-2598.
13. Farber AJ, Castillo R, Clough M, Bahk M, McFarland EG. Clinical assessment of three common tests for traumatic anterior shoulder instability. *J Bone Joint Surg Am*. 2006;88(7):1467-1474.
14. Harris JD, McCormick FM, Abrams GD, et al. Complications and reoperations during and after hip arthroscopy: a systematic review of 92 studies and more than 6,000 patients. *Arthroscopy*. 2013;29(3):589-595.
15. Jaeschke R, Guyatt GH, Sackett DL. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? *JAMA*. 1994;271(9):703-707.
16. Kalisvaart MM, Safran MR. Hip instability treated with arthroscopic capsular plication. *Knee Surg Sports Traumatol Arthrosc*. 2017;25(1):24-30.
17. Kalisvaart MM, Safran MR. Microinstability of the hip: it does exist. Etiology, diagnosis and treatment. *J Hip Preserv Surg.* 2015;2(2):123-135.
18. Larson CM, Giveans MR, Taylor M. Does arthroscopic FAI correction improve function with radiographic arthritis? *Clin Orthop Rel Res*. 2011;469(6):1667-1676.
19. Lebland MC, Kowalczuk M, Andruszkiewicz N, et al. Diagnostic accuracy of physical examination for anterior knee instability: a systematic review. *Knee Surg Sports Traumatol Arthrosc*. 2015;23(10):2805-2813.
20. McCormick F, Slikker W 3rd, Harris JD, et al. Evidence of capsular defect following hip arthroscopy. *Knee Surg Sports Traumatol Arthrosc*. 2014;22(4):902-905.
21. Moorman C, Warren R, Hershman E, et al. Traumatic posterior hip subluxation in American football. *J Bone Joint Surg Am*. 2003;85:1190-1196.
22. Nepple JJ, Byrd JW, Siebenrock KA, Prather H, Clohisy JC. Overview of treatment options, clinical results, and controversies in the management of femeroacetabular impingement. *J Am Acad Orthop Surg*. 2013;21(suppl 1):S53-S58.
23. Park HB, Yokota A, Gill HS, El Rassi G, MFarland EG. Diagnostic accuracy of clinical tests for the different degrees of subacromial impingement syndrome. *J Bone Joint Surg Am*. 2005;87(7):1446-1455.
24. Regan W, Lapner PC. Prospective evaluation of two diagnostic apprehension signs for posterolateral instability of the elbow. *J Shoulder Elbow Surg*. 2006;15(3):344-346.
25. Safran M. Evaluation of the painful hip in tennis players. *Aspetar Sports Med J*. 2014;3:516-523.
26. Shibata KR, Matsuda S, Safran MR. Is there a distinct pattern to the acetabular labrum and articular cartilage damage in the non-dysplastic hip with instability? *Knee Surg Sports Traumatol Arthrosc*. 2017;25(1):84-93.
27. Shu B, Safran MR. Hip instability: anatomic and clinical considerations of traumatic and atraumatic instability. *Clin Sports Med*. 2011;30(2):349-367.
28. Somerville LE, Willits K, Johnson AM, et al. Clinical assessment of physical examination maneuvers for rotator cuff lesions. *Am J Sports Med*. 2014;42(8):1911-1919.
29. Wenger DE, Kendell KR, Miner MR, Trousdale RT. Acetabular labral tears rarely occur in the absence of bony abnormalities. *Clin Orthop Relat Res*. 2004;(426):145-150.
30. Wylie JD, Beckmann JT, Maak TG, Aoki SK. Arthroscopic capsular repair for symptomatic hip instability after previous hip arthroscopic surgery. *Am J Sports Med*. 2016;44(1):39-45.