Original Research

The Diagnostic Validity of Hip Provocation Maneuvers to Detect Intra-Articular Hip Pathology

Erin Maslowski, MD, William Sullivan, MD, Jeri Forster Harwood, PhD, Peter Gonzalez, MD, Marla Kaufman, MD, Armando Vidal, MD, Venu Akuthota, MD

Objective: To determine which hip provocation maneuvers best predict the presence of an intra-articular hip pathology.
Design: Prospective diagnostic study.
Setting: Musculoskeletal clinic at a university-based multispecialty group practice.
Participants: Fifty subjects referred for intra-articular hip injection under fluoroscopic guidance.
Interventions: Subjects were examined with 4 pain provocation maneuvers before and after anesthetic intra-articular hip injection administered under fluoroscopic guidance.
Main Outcome Measurements: Presence of intra-articular hip pain generator was confirmed by ≥80% improvement on visual analog scale after intra-articular hip injection.
Results: The most sensitive tests were flexion abduction external rotation (FABER) test and internal rotation over pressure (IROP) maneuver. For the FABER test, sensitivity was 0.82 (95% CI 0.57-0.96); sensitivity for the IROP maneuver was 0.91 (95% CI 0.68-0.99). The most specific test was the Stinchfield maneuver, with specificity at 0.32 (95% CI 0.14-0.55). FABER and IROP had the highest positive predictive value, with 0.46 (95% CI 0.28-0.65) and 0.47 (95% CI 0.29-0.64), respectively. IROP had the highest negative predictive value at 0.71 (95% CI 0.25-0.98).
Conclusions: IROP and FABER may be worthwhile components of the clinical evaluation of hip pain to determine intra-articular hip pathology. These tests are nonspecific and therefore not necessarily negative in the absence of intra-articular hip pathology. These hip provocation maneuvers are a useful part of an evaluation that includes history, further examination findings, and other diagnostic studies.

PM R 2010;2:174-181

INTRODUCTION

Hip pain is a common complaint among adults. The prevalence of hip osteoarthritis ranges from 7% to 25% in people ages 55 years and older [1]. Neumann et al found that 66% of patients with mechanical symptoms of the hip (eg, pain, locking, giving way) had a labral tear on magnetic resonance arthrography (MRA) [2]. Along with a precise history, physical examination maneuvers can help localize the source of pain. Often, clinicians need to determine if hip pain is from intra-articular sources, such as osteoarthritis or labral tears, or from extra-articular sources. Early detection of these intra-articular lesions may affect diagnostic and treatment decisions. For example, the presence of positive hip physical examination provocation maneuvers can guide clinicians in the need for diagnostic hip intra-articular injections and hip MRA. Moreover, early detection of hip intra-articular lesions may lead to treatment options such as hip arthroscopy and therapeutic hip joint injections.

Physical examination of the hip includes a variety of maneuvers. These include hip range of motion testing, hip flexibility maneuvers (such as the Thomas test), hip musculature strength testing, gait evaluation (including the Trendelenburg test), and hip provocation maneuvers [3]. Hip provocation maneuvers include FABER (hip positioned in flexion, abduction, external rotation), Stinchfield, Scour, and hip internal rotation over pressure

See also Statistical Companion to this article, p. 182 in this issue.

E.M. Department of Physical Medicine and Rehabilitation, University of Colorado Denver, Mail Stop F-493, 12631 East 17th Avenue, Academic Office 1, Room 2513, Aurora, CO 80045. Address correspondence to: E.M.; e-mail: elmaslow@hotmail.com
Disclosure: nothing to disclose

W.S. Department of Physical Medicine and Rehabilitation, University of Colorado Denver, Denver, CO
Disclosure: nothing to disclose

J.F.H. Department of Pediatrics, University of Colorado Denver, Denver, CO
Disclosure: nothing to disclose

P.G. Department of Physical Medicine and Rehabilitation, University of Colorado Denver, Denver, CO
Disclosure: nothing to disclose

M.K. Department of Rehabilitation Medicine, University of Washington, Seattle, WA
Disclosure: nothing to disclose

A.V. Department of Orthopaedics, University of Colorado Denver, Denver, CO
Disclosure: 5A

V.A. Department of Physical Medicine and Rehabilitation, University of Colorado Denver, Denver, CO
Disclosure: nothing to disclose

Disclosure Key can be found on the Table of Contents and at www.pmrjournal.org

Submitted for publication August 5, 2009; accepted January 22, 2010.

PM&R
1934-1482/10/$36.00
Printed in U.S.A.

Vol. 2, 174-181, March 2010
DOI: 10.1016/j.pmrj.2010.01.014

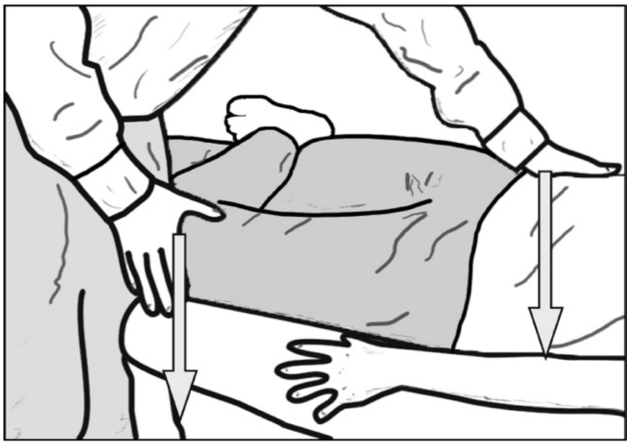

Figure 1. For the flexion abduction external rotation (FABER) maneuver, the subject lays supine with the foot of the tested leg on the knee of opposite leg. The examiner lowers the knee toward the table. Gentle, downward overpressure was exerted by the examiner at the subject's knee (of the tested limb) and anterior superior iliac spine (of the contralateral limb). The test is considered positive if the maneuver recreates the subject's pain.

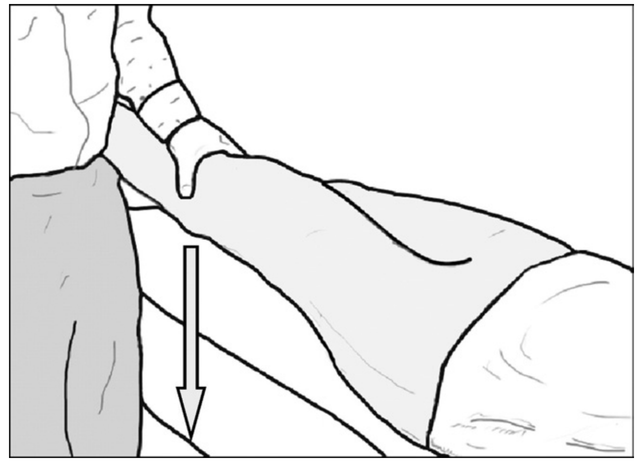

Figure 2. The Stinchfield maneuver is also performed with the subject supine. The tested leg is raised to 30° of hip flexion with the knee in full extension. The subject holds his or her leg in place while the examiner exerts downward force proximal to the knee. The test is considered positive if the maneuver recreates the subject's pain.

(IROP) tests (Figures 1-4). Many variations of these provocation tests exist. For example, IROP can be performed in maximal hip flexion (potentially more sensitive in detecting femoral acetabular impingement [FAI]) or at 0° of hip flexion/extension (also referred to as the "log roll" test often used to potentially eliminate confounding soft-tissue sources of hip pain).

The main purpose of this study is to validate the diagnostic utility of hip provocation maneuvers. The tests under investigation are FABER, Stinchfield, Scour, and IROP tests. A diagnostic intra-articular hip injection was used as the gold standard to confirm the presence of intra-articular hip pathology (IAHP), as a positive diagnostic block has been shown to be 90% accurate in detection of hip internal derangement and a good predictor of improvement after surgical intervention [4]. The validity of hip provocation tests was measured using sensitivity, specificity, and positive/negative predictive values.

METHODS

After institutional review board approval was obtained, a prospective trial examining the diagnostic validity of hip provocation maneuvers was initiated. After statistical power analysis, 50 subjects were deemed necessary for prospective recruitment. Subjects were recruited from a university medical center multispecialty musculoskeletal clinic between November 2007 and April 2009 when they were scheduled for anesthetic injection to the hip with one of the investigators. Subjects were referred for injection when they had typical symptoms, physical examination findings, and radiographic data (if available) suggesting that intra-articular hip pain was present. Inclusion criteria included: $\geq$18 years of age referred for injection after presenting with hip symptoms thought to be related to the presence of intra-articular hip pathology. Subjects were excluded if they had known oncological issues involving the hip, renal failure, allergy to contrast dye, inability to provide consent, or need for sedation. Subjects reported their baseline pain severity by indicating their pain level on a 10-cm visual analog scale (VAS).

The 4 provocation maneuvers were administered by physicians, who specialized in physical medicine and rehabilitation and interventional pain techniques, using the following techniques in the same order for each subject. For the FABER maneuver (also known as Patrick test) [5], the subject was asked to lie supine with the foot of the tested leg on the knee of opposite leg such that the tested extremity is in a hip flexion, abduction, and external rotation position. The examiner then lowered the ipsilateral knee toward the table. Gentle, downward overpressure was exerted by the examiner at the subject's ipsilateral knee. Downward pressure was applied at the contralateral anterior superior iliac spine to stabilize the pelvis (Figure 1). The test was considered positive if the maneuver recreated the subject's ipsilateral hip pain.

The Stinchfield maneuver was also performed with the subject supine. The tested leg was raised to 30° of hip flexion with the knee in full extension. The subject held his or her leg in place while the examiner exerted downward force proximal to the knee (Figure 2). The test was considered positive if the maneuver recreated the subject's pain.

The Scour maneuver (also known as the quadrant test) [5] was performed with the subject supine. The affected hip was

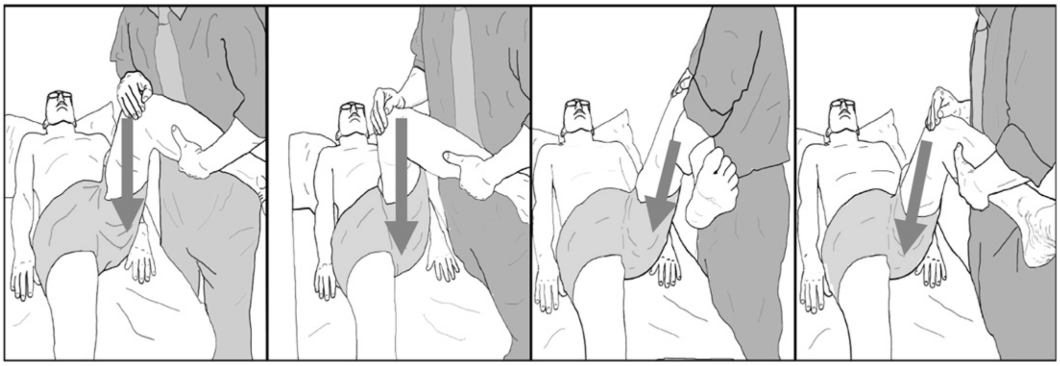

Figure 3. The Scour maneuver is performed with the subject supine. The hip is maximally flexed and adducted. Then, with a compressive force applied to the joint in the direction of the shaft of the femur, the examiner moves the femur through a circular arc of motion. The maneuver is considered positive if it recreates pain.

maximally flexed and adducted. Then, with a compressive force applied to the joint in the direction of the shaft of the femur, the examiner moved the femur through a circular arc of motion (Figure 3). The maneuver was considered positive if it recreated pain.

IROP testing was also performed with the subject supine. The affected hip was flexed to 90° and the knee flexed to 90°.

Figure 4. Internal rotation over pressure was performed with the subject supine. The affected hip is flexed 90° and the knee flexed 90°. The examiner internally rotates the hip by rotating the leg laterally while stabilizing the knee at the same time. The examiner's other hand is used to apply gentle downward pressure on the contralateral anterior superior iliac spine. The examiner applies gentle pressure to the tested leg in the lateral direction. The test is positive if it recreates the patient's hip pain.

The examiner internally rotated the hip by rotating the leg laterally while stabilizing the knee at the same time. Internal rotation overpressure was administered with further gentle rotation of the ipsilateral leg. The pelvis was stabilized, when necessary, by the examiner's other hand at the contralateral anterior superior iliac spine to reduce contralateral ilial rotation (Figure 4). The test was positive if it recreated the patient's hip pain.

Results of each test were recorded. Subjects then underwent concomitant diagnostic and therapeutic intra-articular hip injection. With use of sterile technique, 2 mL 1% lidocaine, 6 mL

Table 1. *Baseline Characteristics of Subjects and Available Radiographic Data*

Characteristic	n (%)
Age	60.2 (13)*
Male gender	20 (40%)
Right side	34 (68%)
Imaging study	
XR only	39 (78%)
MRI only	6 (12%)
MR arthrogram	3 (6%)
MRI + XR	1 (2%)
Radiographic findings†	
Normal/negative	6 (12%)
Mild OA	11 (22%)
Mild to moderate OA	5 (10%)
Moderate OA	4 (8%)
Moderate to severe OA	4 (8%)
Severe OA	8 (16%)
FAI	5 (10%)
Labral tear	6 (12%)
AVN	2 (4%)
Calcific tendonitis or bursitis	2 (4%)
Trochanteric bursitis	1 (2%)
Missing	1 (2%)

XR = x-ray; MRI = magnetic resonance imaging; OA = osteoarthritis; FAI = femoroacetabular impingement; AVN = avascular necrosis.

*Mean (standard deviation).

†One subject can appear in more than one category.

of 0.75% bupivacaine, and 2 mL of 6 mg/mL Betamethasone were injected into the hip joint capsule under fluoroscopic guidance. Injection of contrast material (Omnipaque; GE Healthcare, Princeton, NJ) confirmed correct needle placement within the hip joint before administration of injectate. Ten to 15 minutes after injection, patients rated their pain on a 10-cm VAS identical to the one used before the procedure. Finally, subjects estimated their pain relief as a percentage. Percentage of pain relief was also calculated as the difference between preprocedure VAS score and post-procedure VAS, divided by the pre-procedure VAS. An intra-articular pain generator was assumed when 80% pain relief was achieved by injection (calculated or estimated). This percentage relief level was chosen because it appears to increase the confidence that false-positive levels are mitigated [6].

Statistical Methods

For each maneuver and combination of maneuvers, sensitivity, specificity, positive predictive value (PPV) and negative

Table 2. *Subject Pain Scores Before and After Injection and Estimated Relief*

Pt. No	Radiographic Findings	Imaging Study	Gender	Side	Age	Pre-VAS	Pre-FABER	Pre-Stinch	Pre-Scour	Pre-IR	Post-VAS	Post-FABER	Post-Stinch	Post-Scour	Post-IR	Change Pre-Post VAS	Calculated Relief (Pre-VAS Post-VAS)/Pre VAS	Pt Estimated % Relief
1	Labral tear	MR Arthr	F	R	58	3.3	+	−	−	+	1	−	−	−	−	2.3	0.70	90
2	Mild to mod OA	XR	F	R	59	5.2	−	+	+	+	1.1	−	−	−	−	4.1	0.79	90
3	Mod to severe OA	XR	F	R	60	6.0	+	−	−	+	3.1	−	−	−	−	2.9	0.48	45
4	Mild to mod OA	XR	M	L	65	2.8	+	−	−	+	0	+	+	−	−	2.8	1.00	100
5	Negative	XR	F	L	65	4.4	+	+	+	−	2.6	−	+	+	−	1.8	0.41	40
6	Mild to Mod OA	MR Arthr	F	R	56	3.6	+	−	+	+	0	−	−	−	−	3.6	1.00	100
7	Negative	XR	F	R	50	3.0	−	−	−	+	2.9	−	−	−	+	0.1	0.03	0
8	Mod to severe OA	XR	F	L	55	6.3	+	+	+	+	1.3	−	−	−	−	5	0.79	100
9	Mild to mod OA	XR	F	L	69	9.4	+	+	−	+	0.5	−	−	−	−	8.9	0.95	100
10	Labral tear	MRI	M	R	66	5.6	+	+	+	+	6.5	+	+	+	+	0.9	0.16	5
11	Labral tear	MRI	F	R	22	8.9	+	+	+	−	8.7	+	+	+	−	0.2	0.02	0
12	Mod to severe OA	XR	F	L	69	3.7	+	+	+	+	1.5	−	+	−	−	2.2	0.59	90
13	FAI	XR, MRI	M	R	44	4.5	+	+	+	+	1.8	−	−	+	+	2.7	0.60	70
14	Negative	XR	F	R	55	4.2	+	−	+	−	1.5	+	−	−	−	2.7	0.64	90
15	Severe OA	XR	M	R	62	2.5	+	−	+	+	2.5	−	−	−	−	0	0.00	35
16	Severe OA	XR	F	R	84	2.0	−	+	+	+	0	−	−	−	−	2	1.00	90
17	FAI, mild OA	MRI	M	R	62	5.1	+	+	+	+	1.5	−	+	−	−	3.6	0.71	90
18	Mod OA	XR	M	L	59	5.7	+	+	+	+	0	−	−	−	−	5.7	1.00	100
19	AVN	XR	M	R	67	9.5	+	+	+	+	0.9	−	−	−	−	8.6	0.91	90
20	Calcific tendonitis over greater tuberosity	XR	F	R	63	7.1	+	−	+	+	6.9	+	−	+	+	0.2	0.03	0
21	Severe OA	XR	F	R	64	3.7	−	+	−	+	0.5	−	−	−	−	3.2	0.86	95
22	Severe OA	XR	M	R	72	7.2	+	+	−	+	0.2	−	−	−	−	7	0.97	100
23	FAI	XR	M	R	72	4.2	−	−	−	+	2.5	−	−	−	−	1.7	0.40	20
24	Mod OA	XR	F	L	72	6.6	+	+	+	+	3.8	+	+	+	+	3.2	0.48	50
25	Severe OA	XR	F	R	73	7.7	+	+	+	+	2.5	−	−	+	+	5.2	0.68	80
26	Normal	XR	M	R	42	1.5	+	+	+	+	0	−	+	−	−	1.5	1.00	100
27	Severe OA	XR	F	R	81	4.2	+	−	−	−	0	−	−	−	−	4.2	1.00	100
28	AVN	MRI	M	R	62	6.0	−	+	−	−	5.5	−	+	−	−	0.5	0.08	20
29	Mild to mod OA	XR	F	L	66	2.2	+	+	−	+	0.1	−	+	−	−	2.1	0.95	70
30	Labral tear, FAI	MR Arthr	M	L	38	1.0	−	−	+	+	0	−	−	+	+	1	1.00	100
31	Negative	XR	F	R	51	7.0	−	+	+	+	4.6	−	−	−	−	2.4	0.34	50
32	Moderate OA	XR	M	L	76	2.5	+	+	−	+	4.2	−	+	−	+	−1.7	−0.68	0
33	Mild OA	XR	M	R	72	3.0	+	+	−	−	0.3	−	−	−	−	2.7	0.90	95
34	Mod to severe OA	XR	F	L	50	5.6	+	+	+	+	0.2	−	+	−	−	5.4	0.96	99
35	Mild OA	XR	M	R	68	3.3	+	−	+	+	1.5	−	−	−	−	1.8	0.55	60
36	Severe OA	XR	F	L	63	6.9	+	+	+	+	5.7	+	+	+	+	1.2	0.17	30
37	Labral tear, trochanteric bursitis	MRI	F	L	56	8.7	+	+	+	+	8.2	+	+	−	+	0.5	0.06	0
38	Mild OA	XR	F	R	56	4.1	+	−	+	+	0.6	−	−	−	−	3.5	0.85	90
39	Mild OA	XR	F	R	49	4.4	+	−	−	+	0.6	−	−	−	−	3.8	0.86	100
40	Mild OA	XR	F	R	47	5.8	+	+	−	+	1.1	−	+	−	−	4.7	0.81	80
41	Severe OA	XR	M	R	73	7.0	−	+	−	2	1.8	−	+	−	−	5.2	0.74	80
42	Mild OA, calcific tendonitis or bursitis	XR	M	L	70	2.4	+	−	−	+	0	−	−	−	−	2.4	1.00	100
43	Normal	XR	F	L	26	8.5	+	−	+	+	0	−	+	−	+	8.5	1.00	100
44	Mild OA	XR	F	R	66	6.7	−	−	−	+	6.5	−	−	−	−	0.2	<0.01	0
45	Mild OA	XR	M	R	73	6.9	+	−	+	+	0.4	−	−	−	−	6.5	0.94	100
46	Labral tear, FAI	MRI	M	R	32	1.9	+	+	+	+	0.2	−	+	−	−	1.7	0.89	75
47	Mod OA	XR	M	R	72	2.9	−	+	−	+	0	−	−	−	−	2.9	1.00	100
48	OA	MRI	F	L	55	5.3	+	+	+	+	2.1	+	−	−	+	3.2	0.60	30
49	Mild OA	XR	F	R	66	4.7	+	+	+	−	3.8	+	+	+	−	0.9	0.19	0
50	Mild OA	XR	F	R	57	3.9	+	+	+	+	1.5	+	−	−	−	2.4	0.62	100

Pt = patient; VAS = visual analog scale; FABER = flexion abduction external rotation; Stinch = Stinchfield; IR = internal rotation over pressure; mod = moderate; XR = radiograph; MRI = magnetic resonance imaging; OA = osteoarthritis; FAI = femoroacetabular impingement; AVN = avascular necrosis.

Table 3. *Sensitivities and specificities using calculated ≥80% relief in VAS as confirmation of intra-articular hip pathology*

At Least 80% VAS Relief (n = 22/50)	FABER	Stinchfield	Scour	IROP	F+St	F+St+Sc	F+St+Sc+IR
Sensitivity (95% CI)	18/22 = 0.82 (0.57-0.96)	13/22 = 0.59 (0.34-0.82)	11/22 = 0.50 (0.26-0.74)	20/22 = 0.91 (0.68-0.99)	21/22 = 0.96 (0.74-1.0)	22/22 = 1.0 (0.82-1.0)	22/22 = 1.0 (0.82-1.0)
Specificity (95% CI)	7/28 = 0.25 (0.09-0.48)	9/28 = 0.32 (0.14-0.55)	8/28 = 0.29 (0.12-0.51)	5/28 = 0.18 (0.05-0.40)	3/28 = 0.11 (0.02-0.31)	3/28 = 0.11 (0.02-0.31)	0/28 = 0 (0-0.15)

VAS = visual analog scale; FABER = flexion abduction external rotation; IROP = internal rotation over pressure; F = FABER; St = Stinchfield; Sc = Scour; IR = IROP.

predictive value (NPV) were calculated considering 2 different "gold standards" (at least 80% improvement in VAS score and at least 80% estimated relief). Also, 95% rectangular confidence intervals (CI) were calculated for each maneuver and maneuver combination such that there is 95% confidence that the true sensitivity *and* the true specificity (or PPV *and* NPV) are contained in their respective intervals [7].

RESULTS

Fifty subjects were enrolled in the study. Twenty (40%) were male and 30 (60%) were female. Average age was 60 years old. Demographic and radiographic data are summarized in Tables 1 and 2.

Of the 20 subjects who had >80% decrease in pain with injection by VAS, there were 16 cases of osteoarthritis, one case of avascular necrosis (AVN), one case of a labral tear with FAI, and one with a normal radiograph. Of the 29 subjects with >80% decrease in pain by estimated VAS, 22 had osteoarthritis, one had a labral tear, one had osteoarthritis with FAI, one had AVN, one had a labral tear with FAI, and 3 had a negative radiograph. Of the 19 subjects with <80% improvement by either measure, 9 had osteoarthritis, 3 had negative radiographs, 2 had labral tears, 2 had FAI, one had calcific tendonitis over the greater trochanter, one had AVN, and one had a labral tear with trochanteric bursitis.

For results of sensitivities and specificities of each test and the tests in combination, see Tables 3 and 4. Using 80% percent improvement in VAS as the gold standard, the most sensitive test was IROP, with a sensitivity of 0.91 (95% CI 0.68-0.99). The next most sensitive test was FABER, with a sensitivity of 0.82 (95% CI 0.57-0.96). Sensitivity findings were similar when ≥80% perceived relief by the patient was used as the gold standard (see Table 4). The Stinchfield test was the most specific maneuver at 0.32 (95% CI 0.14-0.55) when using the VAS standard. When using ≥80% perceived relief as the standard, the most specific maneuver was the Scour test, with a specificity of 0.38 (95% CI 0.17-0.62).

A summary of PPV and NPV is outlined in Tables 5 and 6. The tests had similar PPV, ranging from 0.36 (95% CI 0.17-0.57) for the Scour test to 0.47 (95% CI 0.29-0.64) for IROP test when using the VAS standard. PPV were slightly lower when estimated relief was used as the gold standard. For the VAS standard, NPV was highest for IROP at 0.71 (95% CI 0.25-0.98), followed by FABER with 0.64 (95% CI 0.27-0.91). NPV were lower when ≥80% perceived relief was used as the standard.

When the same statistical analysis was performed excluding the subjects who had negative imaging (n = 2), sensitivities, specificities, PPV, and NPV did not change to an appreciable extent.

Prevalence of IAHP was 44% when ≥80% improvement in VAS score was used as the standard. Prevalence of IAHP was 52% when ≥80% perceived relief was considered significant.

DISCUSSION

This study is one of the first studies to look at the diagnostic validity of these tests. In particular, the hip IROP has not been studied before this investigation. The results from this study demonstrate that hip provocation maneuvers have relatively high sensitivity but relatively low specificity for detecting IAHP. The IROP and FABER appear to be the most sensitive of the hip provocation maneuvers. Because these tests are used in conjunction with other diagnostic activities, they do not call for the stringent sensitivity or specificity requirements of a screening test for a life-threatening illness. However, the overall sensitivities and specificities (and therefore PPV and NPV) are not robust. Although these tests are widely used by clinicians in the assessment of hip pain, hip provocation maneuvers have had limited diagnostic validity data.

Table 4. *Sensitivities and specificities using ≥80% perceived relief as confirmation of intra-articular hip pathology*

At Least 80% Perceived Relief (n = 26/50)	FABER	Stinchfield	Scour	IROP	F+St	F+St+Sc	F+St+Sc+IR
Sensitivity (95% CI)	21/26 = 0.81 (0.58-0.95)	15/26 = 0.58 (0.34-0.79)	16/26 = 0.62 (0.38-0.82)	23/26 = 0.88 (0.67-0.98)	25/26 = 0.96 (0.78-1.0)	26/26 = 1.0 (0.84-1.0)	26/26 = 1.0 (0.84-1.0)
Specificity (95% CI)	6/24 = 0.25 (0.08-0.50)	7/24 = 0.29 (0.11-0.54)	9/24 = 0.38 (0.17-0.62)	4/24 = 0.17 (0.04-0.40)	3/24 = 0.13 (0.02-0.35)	3/24 = 0.13 (0.02-0.35)	0/24 = 0 (0-0.17)

VAS = visual analog scale; FABER = flexion abduction external rotation; IROP = internal rotation over pressure; F = FABER; St = Stinchfield; Sc = Scour; IR = IROP.

Table 5. *PPV and NPV, using calculated ≥80% in VAS as confirmation of intra-articular hip pathology*

At Least 80% VAS Relief (n = 22/50)	FABER	Stinchfield	Scour	IROP	F+St	F+St+Sc	F+St+Sc+IR
PPV (95% CI)	18/39 = 0.46 (0.28-0.65)	13/32 = 0.41 (0.22-0.62)	11/31 = 0.36 (0.17-0.57)	20/43 = 0.47 (0.29-0.64)	21/46 = 0.46 (0.29-0.63)	22/47 = 0.47 (0.30-0.64)	22/50 = 0.44 (0.28-0.61)
NPV (95% CI)	7/11 = 0.64 (0.27-0.91)	9/18 = 0.50 (0.23-0.77)	8/19 = 0.42 (0.18-0.69)	5/7 = 0.71 (0.25-0.98)	3/4 = 0.75 (0.15-1.0)	3/3 = 1.0 (0.23-1.0)	0/0 = not estimable

VAS = visual analog scale; FABER = flexion abduction external rotation; IROP = internal rotation over pressure; F = FABER; St = Stinchfield; Sc = Scour; IR = IROP.

Diagnostic hip intra-articular injection as a gold standard has several advantages over imaging. Positive response to injection suggests intra-articular pathology as the source of pain. In contrast, plain radiographs, magnetic resonance imaging and MRA can show pathology in asymptomatic patients. Furthermore, radiographs are usually negative in the presence of labral tears or other non-bony, intra-articular abnormalities.

Intra-articular hip injection under fluoroscopic guidance has been suggested as a suitable gold standard in hip join investigations [8]. This is based on the premise that injections have been used in an analogous fashion in sacroiliac joint and lumbar zygapophyseal joint pain. Illgen found that of 21 patients with positive response to intra-articular hip injection, all had positive findings at the time of arthroplasty [4]. Byrd reported that response to intra-articular hip injection was 90% accurate compared with findings on arthroscopy [9]. Pateder et al found intra-articular hip injection was 100% sensitive and 88% specific in identification of an intra-articular hip pain generator [10].

The FABER test is the most widely used hip provocation test. The test appears to be nonspecific because the FABER position can cause pain emanating from the sacroiliac joint and lower lumbar facet joints [11]. Theiler et al reported a 0.54 correlation coefficient between FABER test results and findings of osteoarthritis on hip radiographs. Mitchell et al reviewed the records of 25 patients who underwent hip arthroscopy. Of these 25 patients, 17 patients had a documented FABER test before surgery. Of those 17, 15 patients had a positive FABER. All 17 of the patients had positive findings on arthroscopy. Thus, the authors concluded that test was 88% sensitive for identifying patients with a labral tear [12]. Martin et al studied 49 patients considered potential candidates for arthroscopic surgery for hip pain. They found that neither FABER nor flexion-internal rotation-adduction reliably predicted the presence of an intra-articular hip pain generator, as identified by a >50% improvement in hip pain after intra-articular diagnostic injection [13]. That authors' study demonstrates that FABER is approximately 81% sensitive and 25% specific.

Hip IROP testing has not been adequately tested for reliability and validity. In addition, IROP testing has various modifications. For instance, so called "hip impingement" maneuvers for FAI and labral tears are modifications of the IROP where maximal hip flexion as well as maximal internal rotation is applied [14]. Others employ the log-roll test where IROP is performed in 0° of hip extension/flexion [15]. In this study, IROP was performed with hip flexion of 90°. Simple hip internal rotation range of motion testing may also be considered a variation of IROP. Narvani et al found that an internal rotation-flexion-axial compression maneuver (hip impingement test) was 75% sensitive and 43% specific in predicting the presence of a labral tear on MRA in 18 athletes [16]. Brown reported that limited hip internal rotation range of motion increased the likelihood of a hip disorder or a hip and spine disorder 14 times greater than a spine only disorder [17]. In a study of 195 individuals in a primary care setting, Birrell found that restricted hip internal range of motion was 85% sensitive and 54% specific for detecting mild to moderate hip osteoarthritis, compared with findings on hip radiographs [18]. Altman described the characteristics of persons with hip osteoarthritis, including painful hip range of motion ≥15°, morning hip stiffness lasting ≥60 minutes, and age ≥50 years old. This constellation of findings identified patients with hip osteoarthritis with a sensitivity of 86% and specificity of 75%. The same group also reported that internal rotation ROM <15° and flexion ≤115° was 86% sensitive and 75% specific for hip osteoarthritis [19]. We found the IROP test was 88%-91% sensitive and 17%-18 % specific for intra-articular hip pathology.

Table 6. *PPV and NPV, using ≥80% perceived relief as confirmation of intra-articular hip pathology*

At Least 80% Perceived Relief (n = 26/50)	FABER	Stinchfield	Scour	IROP	F+St	F+St+Sc	F+St+Sc+IR
PPV (95% CI)	21/39 = 0.54 (0.35-0.72)	15/32 = 0.47 (0.27-0.68)	16/31 = 0.52 (0.31-0.72)	23/43 = 0.54 (0.36-0.71)	25/46 = 0.54 (0.37-0.71)	26/47 = 0.55 (0.38-0.72)	26/50 = 0.52 (0.36-0.68)
NPV (95% CI)	6/11 = 0.55 (0.20-0.86)	7/18 = 0.39 (0.15-0.67)	9/19 = 0.47 (0.22-0.74)	4/7 = 0.57 (0.15-0.92)	3/4 = 0.75 (0.15-1.0)	3/3 = 1.0 (0.23-1.0)	0/0 = not estimable

VAS = visual analog scale; FABER = flexion abduction external rotation; IROP = internal rotation over pressure; F = FABER; St = Stinchfield; Sc = Scour; IR = IROP.

Sutlive reported the Scour test was 62% sensitive and 75% specific for presence of hip osteoarthritis. In that study, presence of osteoarthritis was confirmed by finding Kellgren and Lawrence Grade 2 or higher on anteroposterior hip radiographs. We found the Scour test was 50%-62% sensitive and 29%-38% specific for intra-articular hip pathology. The authors found no previous validity data for the Stinchfield test. In this study, Stinchfield was 58%-59% sensitive and 29%-32% specific for intra-articular hip pathology.

Sutlive et al sought to determine a clinical prediction rule for diagnosis of hip osteoarthritis [20]. They considered multiple variables including range of motion in various planes and specific tests including Patrick (ie, FABER) test, the Scour test, and the squat test. Logistic regression analysis was used to identify 5 predictor variables, including squatting as an aggravating factor, active hip flexion causing lateral hip pain, Scour test with adduction causing lateral hip pain, active hip extension causing hip pain, and passive internal rotation of less than or equal to 25°. They reported positive likelihood ratio of 5.2 with 3 of variables present. The positive likelihood ratio was 24.3 with 4 of 5 variables present. In that study, hip osteoarthritis was deemed present when an anteroposterior radiograph of the pelvis showed Grade 2 or greater changes on the Kellgren and Lawrence Grading Scale for hip osteoarthritis.

In this present study, of the 4 tests, IROP was the most sensitive, followed by FABER, suggesting that if intra-articular hip pathology IAHP is the pain generator, the tests are more likely to be positive. However, given the sample size, the 95% confidence intervals do not rule out poorer performance (0.67 and 0.58, respectively). The specificity of the tests were low (all less than random chance), suggesting that if IAHP is not the pain generator, testing may not reliably yield a negative result. As expected, sensitivity of tests increased as tests were added. However, this is at the cost of specificity, which decreases dramatically as tests are added. NPV and PPV values were generally no better than chance alone. Though some estimates were greater than 0.5, the confidence intervals are wide, and therefore do not rule out poorer performance.

In a clinical setting, these maneuvers are used in conjunction with a medical history, other physical assessments, and imaging studies to diagnose and guide treatment decisions. The invasiveness and expense of the possible diagnostics and treatments vary widely. Typical radiographic evaluation of hip pain may include plain radiographs of hip and lumbar spine, diagnostic intra- and extra-articular injections, computed tomography, magnetic resonance imaging, and MRA. Arthroscopy is often the most definitive diagnostic measure, though also the most invasive. Treatments include rest, gait aid prescription, physical and manual therapies, intra- and extra-articular injections, medications, arthroscopic debridement, and total hip arthroplasty.

Strengths of this study include use of response to intra-articular injection as the gold standard outcome measure and stringent criteria (80% improvement in pain) for positive response to injection. There are several limitations in this study. The examiners were not blinded to the radiographic results, fluoroscopic findings, and response to injection. Also, there were multiple examiners without knowledge of the inter-rater reliability between the authors. Finally, the prevalence of significant IAHP in this group may be higher than that of the general population with hip pain, as these subjects had a clinical picture (history, exam, and radiographic findings) that warranted referral for intra-articular hip injection. A lower prevalence would lead to a lower PPV and a higher NPV, by definition.

Another potential source of error in this study is the volume of injectate used. The authors used a total volume of 10 mL of anesthetic and steroid. Theoretically, this volume could cause overflow to adjacent extra-articular soft tissue structures, thus decreasing the specificity of a positive test. On review of the literature, other authors have used 5 mL [21], 6 mL [8], 7 mL [4], 9.5 mL [9], 10 mL [10,22], and 14 mL [13]. Notably, the group using 9.5 mL found 3 false-negative responses and one false-positive response to injection compared with findings on arthroscopy, suggesting that at least in that study, a similar amount of injectate resulted in more false negatives than false positives.

Concomitant back pain might lead to another source of error in this study. Because the authors did not investigate low back pain generators as part of the study, subjects with referred pain from the back would have reduced pain relief despite a successful hip injection. Future investigations could minimize this error by excluding subjects with known spine pathology or including diagnostic data from the spine.

Localization of pain before and after injection (eg, lateral, groin, posterior) would be helpful to determine whether pain might be emanating from a site other than the hip. Furthermore, the authors did not collect other historical data from the patients including history of injury, presence of mechanical symptoms, or duration of symptoms. Therefore it is impossible to correlate these items with response to injection. This information could further define the clinical utility of diagnostic injections to the hip and would be a helpful area of further study.

Two of the authors' study subjects had negative imaging. Exclusion of these subjects in the calculations did not change the results of this statistical analysis to an appreciable extent, likely because of the small number of subjects involved. A larger study, involving a greater number of subjects with negative imaging, would be necessary to determine to what extent imaging findings impact sensitivity, specificity, PPV, and NPV of the tests.

In summary, these findings suggest that IROP and FABER are the most sensitive of the provocation maneuvers studied in identifying IAHP. Though they may be sensitive, none of these tests were specific for IAHP. In addition, PPVs and NPVs were not superior to chance. Therefore, the authors conclude that IROP and FABER may be worthwhile components of the clinical evaluation of hip pain. There is a high probability that these tests will be positive in patients with IAHP. However, the tests are nonspecific, and therefore not necessarily negative in the absence of IAHP. Furthermore, they cannot meaningfully predict a positive or negative response to intra-articular injection.

ACKNOWLEDGMENTS

The authors wish to acknowledge Blake Humphrey for his assistance with the figures used in this text.

REFERENCES

1. Felson DT. Epidemiology of hip and knee osteoarthritis. Epidemiol Rev 1988;10:1-28.
2. Neumann G, Mendicuti AD, Zou KH, et al. Prevalence of labral tears and cartilage loss in patients with mechanical symptoms of the hip: evaluation using MR arthrography. Osteoarthr Cartilage 2007;15:909-917.
3. Cibere J, Thorne A, Bellamy N, et al. Reliability of the hip examination in osteoarthritis: effect of standardization. Arthritis Rheum 2008;59: 373-381.
4. Illgen RL, 2nd, Honkamp NJ, Weisman MH, Hagenauer ME, Heiner JP, Anderson PA. The diagnostic and predictive value of hip anesthetic arthrograms in selected patients before total hip arthroplasty. J Arthroplasty 2006;21:724-730.
5. Magee D. Orthopaedic physical assessment. 4th ed. Philadelphia: WB Saunders; 2002.
6. Bogduk N, ed. Spinal diagnostic treatment procedures. San Francisco: International Spine Intervention Society; 2004.
7. Pepe M. The statistical evaluation of medical tests for classification and prediction. New York: Oxford University Press; 2003.
8. Lesher JM, Dreyfuss P, Hager N, Kaplan M, Furman M. Hip joint pain referral patterns: a descriptive study. Pain Med 2008;9:22-25.
9. Byrd JW, Jones KS. Diagnostic accuracy of clinical assessment, magnetic resonance imaging, magnetic resonance arthrography, and intra-articular injection in hip arthroscopy patients. Am J Sports Med 2004;32: 1668-1674.
10. Pateder DB, Hungerford MW. Use of fluoroscopically guided intra-articular hip injection in differentiating the pain source in concomitant hip and lumbar spine arthritis. Am J Orthop 2007;36:591-593.
11. Porterfield JA, DeRosa C. Mechanical low back pain: perspectives in functional anatomy. Philadelphia: W. B. Saunders; 1991.
12. Mitchell B, McCrory P, Brukner P, O'Donnell J, Colson E, Howells R. Hip joint pathology: clinical presentation and correlation between magnetic resonance arthrography, ultrasound, and arthroscopic findings in 25 consecutive cases. Clin J Sport Med 2003;13:152-156.
13. Martin RL, Irrgang JJ, Sekiya JK. The diagnostic accuracy of a clinical examination in determining intra-articular hip pain for potential hip arthroscopy candidates. Arthroscopy 2008;24:1013-1018.
14. Sink EL, Gralla J, Ryba A, Dayton M. Clinical presentation of femoroacetabular impingement in adolescents. J Pediatr Orthop 2008;28:806-811.
15. Byrd JWT. Hip arthroscopy in athletes. Oper Tech Sports Med 2005; 13:24-36.
16. Narvani AA, Tsiridis E, Kendall S, Chaudhuri R, Thomas P. A preliminary report on prevalence of acetabular labrum tears in sports patients with groin pain. Knee Surg Sports Traumatol Arthrosc 2003;11:403-408.
17. Brown MD, Gomez-Marin O, Brookfield KF, Li PS. Differential diagnosis of hip disease versus spine disease. Clin Orthop Relat Res 2004;419: 280-284.
18. Birrell F, Croft P, Cooper C, Hosie G, Macfarlane G, Silman A. Predicting radiographic hip osteoarthritis from range of movement. Rheumatology (Oxford) 2001;40:506-512.
19. Altman R, Alarcon G, Appelrouth D, et al. The American College of Rheumatology criteria for the classification and reporting of osteoarthritis of the hip. Arthritis Rheum 1991;34:505-514.
20. Sutlive TG, Lopez HP, Schnitker DE, et al. Development of a clinical prediction rule for diagnosing hip osteoarthritis in individuals with unilateral hip pain. J Orthop Sports Phys Ther 2008;38:542-550.
21. Robinson P, Keenan AM, Conaghan PG. Clinical effectiveness and dose response of image-guided intra-articular corticosteroid injection for hip osteoarthritis. Rheumatology (Oxford) 2007;46:285-291.
22. Crawford RW, Gie GA, Ling RS, Murray DW. Diagnostic value of intra-articular anaesthetic in primary osteoarthritis of the hip. J Bone Joint Surg Br 1998;80:279-281.