



Systematic review

The validity and accuracy of clinical diagnostic tests used to detect labral pathology of the hip: A systematic review

Roanna M. Burgess^{a,*}, Alison Rushton^b, Chris Wright^b, Cathryn Daborn^a^aRegional Medical Centre, Royal Air Force Cosford, Wolverhampton WV7 3EX, UK^bSchool of Health and Population Sciences, College of Medical and Dental Sciences, The University of Birmingham, Birmingham, UK.

ARTICLE INFO

Article history:

Received 27 October 2010

Received in revised form

4 January 2011

Accepted 7 January 2011

Keywords:

Acetabular

Labral

Tears

Diagnosis

ABSTRACT

Acetabular labral tears are an area of increasing interest to clinicians involved in the diagnosis of musculoskeletal complaints of the hip. This review systematically evaluated the evidence for the diagnostic accuracy and validity of reported symptoms, physical examination and imaging in this complex population. Studies published in English prior to May 2010 were included. One reviewer searched information sources to identify relevant articles. Two reviewers independently assessed studies for inclusion, extracted data and evaluated quality using the Quality Assessment of Diagnostic Studies Tool.

Twenty one studies were included. Meta-analysis was limited owing to heterogeneity between studies. Results showed Magnetic Resonance Arthrography to consistently outperform Magnetic Resonance Imaging. Computerised Tomography also showed high accuracy levels for the few studies identified. Studies investigating physical tests were of poor quality demonstrating a need for further research in this area. Symptoms likely to be present in patients presenting with acetabular labral tears were found to be anterior groin pain and mechanical hip symptoms; however, additional good quality studies are needed to consolidate findings.

Crown Copyright © 2011 Published by Elsevier Ltd. All rights reserved.

1. Introduction

Intra-articular lesions of the hip have increasingly been identified as the source of unexplained hip and groin pain (Kelly et al., 2005). In a recent study evaluating diagnosis of longstanding groin pain, the most prevalent condition was hip joint pathology, with acute labral tears and impingement syndromes accounting for a large proportion of diagnoses (Bradshaw et al., 2008). Prevalence of acetabular labral tears in patients presenting with hip or groin pain has been reported to be between 22% (Narvani et al., 2003) and 55% (McCarthy et al., 2001). As a diagnosis, this is further supported by reports of positive outcome from arthroscopic resection and excision of the torn acetabular labra (Fitzgerald, 1995; McCarthy and Busconi, 1995; Farjo et al., 1999; Byrd and Jones, 2001).

The diagnosis of hip injuries is complex due to the challenging anatomy and biomechanics around the hip and groin (Feeley et al., 2008). Many factors predispose patients to acetabular labral tears, including femoral acetabular impingement, Legg-Calve-Perthes Disease (LCP), slipped capital epiphysis, and repetitive or acute hip trauma (McCarthy et al., 2003; Martin et al., 2006). Clinicians,

particularly musculoskeletal physiotherapists working in extended roles, require a thorough knowledge of presenting symptoms, physical findings and imaging techniques, and their diagnostic accuracy in relation to acetabular labral tears, to recognise the need for early investigations and to guide appropriate management. To be useful to clinicians, a diagnostic test must possess high sensitivity to rule in a condition and high specificity to rule out a condition (Deeks, 2001). Alongside other summary statistics, sensitivity and specificity provide measures of diagnostic accuracy (Deeks, 2001).

History taking was found to be the most important part of the diagnostic process by Peterson et al. (1992) in a study on medical outpatients, with examination and laboratory testing improving diagnostic ability by a further 12% and 11% respectively. McCarthy and Busconi (1995) judged history taking and physical examination to be the best predictors of intra-articular hip pathology, in their cohort study of patients with refractory hip pain ($n = 59$). The patient history and physical findings are therefore important entities to explore in this patient population alongside diagnostic imaging.

Symptoms associated with acetabular labral tears include pain, locking, catching, instability, giving way, and stiffness (Martin et al., 2006). These symptoms, however, are also highly prevalent in other patient populations including those with LCP, loose bodies, degenerative joint disease and osteonecrosis (O'Leary et al., 2001). A number of physical tests are used to assess the hip complex.

* Corresponding author. Tel.: +44 7738814782.

E-mail address: physio5@cosfordfus.afpaa.mod.uk (R.M. Burgess).

Those routinely used to test for acetabular labral tears include the Modified Thomas Test (Narvani et al., 2003), hip impingement test (Narvani et al., 2003; Burnett et al., 2006), hip quadrant test (Narvani et al., 2003), Scour test (Martin et al., 2008) and Fitzgerald's provocative test (Fitzgerald, 1995). Most of the physical tests that are used to aid diagnosis of labral pathology have not been formally validated (Martin et al., 2006).

Current imaging techniques include magnetic resonance imaging (MRI), magnetic resonance arthrography (MRA) (Czerny et al., 1996) and computerised tomography (CT) (Yamamoto et al., 2007). Hip arthroscopic examination is time consuming, costly and invasive (Keeney et al., 2004) and can only be performed by a small number of specialist orthopaedic surgeons. Consequently, most clinicians are reliant on conventional methods of diagnosis using data from the patient history, physical examination and accessible forms of imaging to guide decision making in clinical practice.

The objective of the study was to determine diagnostic accuracy and validity of the patient history, physical examination and imaging for the diagnosis of acetabular labral tears in patients presenting with hip pain. Current literature on the diagnosis of labral pathology was identified and presented to support clinical decision making when considering labral pathology as a differential diagnosis.

2. Methodology

2.1. Methods

A systematic review was conducted according to a pre-defined protocol following the Cochrane handbook (Reitsma et al., 2009 (Cochrane Collaboration)).

2.2. Search strategy

Information sources were searched from 1987 to May 2010. Electronic databases included Medline, Cinahl, Embase, Cochrane, Sport Discus, Web of Science and Institute of Health and Life Sciences. Electronic searches were carried out combining terms; hip labrum, labral tear, hip, acetabular, hip pain, groin pain, diagnosis, physical examination, Thomas test, Impingement test, Quadrant test, Scour test, Fitzgerald test, resisted straight leg raise test, log roll test, MRI, MRA, CT, investigation, imaging, sensitivity, specificity, accuracy and validity. Studies need sets of terms to both identify the index tests and identify the target population (de Vet et al., 2008, see Fig. 1).

The search was augmented using reference lists from retrieved articles, expert opinion, and hand searching of key journals including *Manual Therapy* and the *Journal of Orthopaedic and Sports Physical Therapy*.

2.3. Inclusion and exclusion criteria

Titles and abstracts of articles were screened by one reviewer (RB), using specific pre-determined inclusion criteria (Deville et al., 2002). Identified full text articles were further screened by two reviewers (RB, CD) for inclusion into the study, using the inclusion criteria. Inclusion criteria were: published in English; any study design; patients presenting with hip or groin pain; clinical or diagnostic test used to assess for acetabular labral tears; comparison against a reference standard; reporting of sensitivity, specificity, likelihood ratios or raw data; and 7 or more 'yes' answers generated from the Quality Assessment of Diagnostic Studies Tool (QUADAS). Exclusion criteria were: other pathologies at outset leading to hip pain; and studies not reporting separate findings for participants with or without a labral tear.

2.4. Quality assessment

Two reviewers (RB, CD) independently conducted the quality assessment using the QUADAS tool (Whiting et al., 2004), which was designed explicitly to assess the quality of diagnostic accuracy studies. Whiting et al. (2006) reported inter-rater agreement of 90% between two reviewers and 85% between three reviewers using the QUADAS tool. Hollingworth et al. (2006) however found that the mean agreement between individual QUADAS items was only fair, with a kappa value of 0.22. Lack of clarity on the meaning and scoring of items were the main sources of confusion. All items were therefore discussed in detail by the reviewers prior to commencement of the study. A pilot using the QUADAS to score 2 articles was conducted to assess agreement between the reviewers. Agreement was high, achieving an overall agreement of 93%.

A study evaluating the scoring of diagnostic studies using QUADAS found that different ways of weighting individual items could produce different quality scores (Whiting et al., 2005). Hence, for the present systematic review, information from the QUADAS tool was presented in tabular format so that items representing sources of bias could be considered on an individual basis. In addition, contrary to recommendations of Whiting et al. (2005) but in line with several other studies, a score of 7 or more yes answers out of 14 was used to indicate a higher quality diagnostic study (de Graaf et al., 2006; Sehgal et al., 2005; Shah et al., 2005). Studies with less than 7 yes answers were excluded from the review.

2.5. Data extraction and data analysis

Data extraction was carried out independently by two reviewers (RB, CD) using a pre-designed data extraction sheet.

A meta-analysis of diagnostic accuracy studies has two main aims: to obtain a pooled measure of diagnostic accuracy and to explore the heterogeneity amongst studies (CRD, 2009). Specific inclusion criteria for the meta-analysis were homogeneity of studies with regards to similarity of patient populations, comparable reference tests, no differences in diagnostic thresholds, inclusion of raw data and homogeneity of results (considered graphically from a receiver operating characteristic (ROC) plot of sensitivities and specificities (Deeks, 2001)).

3. Results

3.1. Study selection

A Flow Diagram of the search history is presented in Fig. 2. Six articles were excluded from screening the full text. Twenty one studies were identified for inclusion in the review.

3.2. Patient symptoms: quality assessment

A major limitation in a large proportion of studies reporting patient symptoms is subjects only being included if they had an identified tear at arthroscopy (Fitzgerald, 1995; Farjo et al., 1999; Burnett et al., 2006). Retrospective studies do not allow for comparison of findings against patients presenting with hip pain without labral pathology, and do not allow for calculation of specificity in relation to the symptom.

3.2.1. Patient symptom data

O'Leary et al., 2001 and McCarthy and Busconi (1995) both reported 100% of patients to present with anterior groin pain (sensitivity 100%), and the study by McCarthy and Busconi (1995) also allowed calculation of the specificity (4%). Clicking was widely reported across studies (Table 2). Farjo et al. (1999) reported

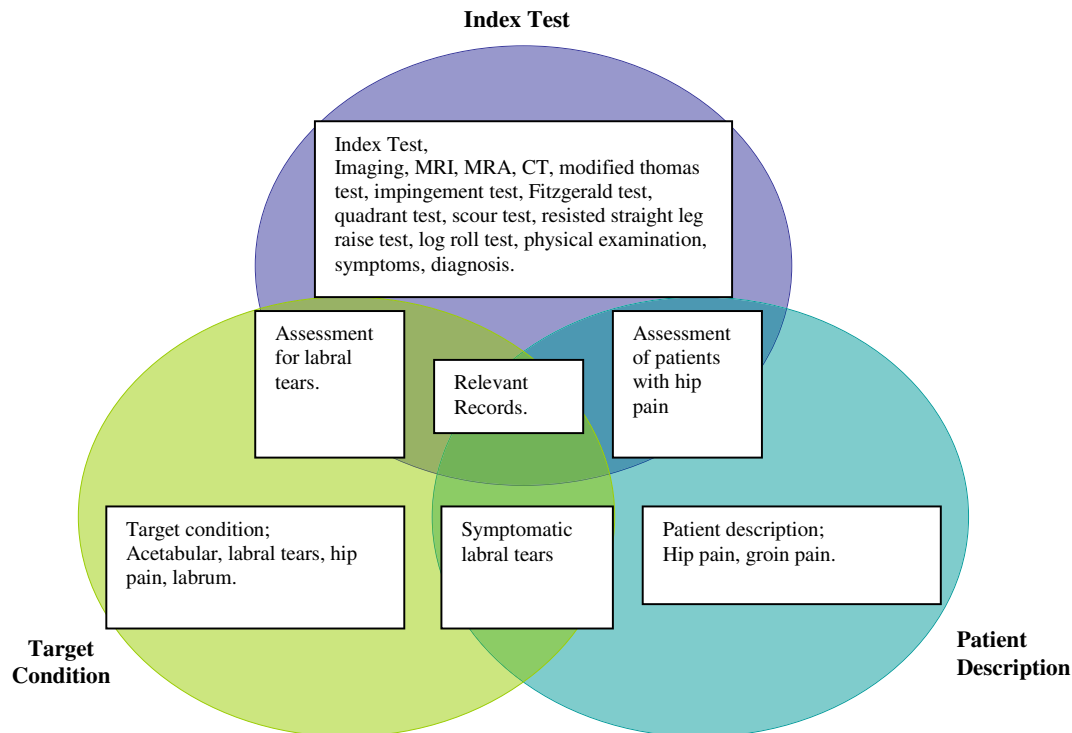


Fig. 1. Combining Concepts as Search Sets (de Vet et al., 2008).

16/28 patients with a labral tear reported clicking (sensitivity, 57%) whereas Narvani et al. (2003) reported 100% sensitivity for clicking (however, only 4/18 patients had a tear). Mechanical symptoms including clicking, locking, popping or giving way were grouped together in three studies (Farjo et al., 1999; O'Leary et al., 2001; Burnett et al., 2006). Sensitivity of mechanical symptoms ranged from 53% (Burnett et al., 2006) to 100% (O'Leary et al., 2001).

3.3. Physical examination: quality assessment

Only five studies provided data on physical examination findings in relation to subjects with labral tears (see Table 2, see Table 3 for test descriptions). The main limitation of all five studies was poor reporting of the physical tests. All subjects in two of these studies had a labral tear (Fitzgerald, 1995; Burnett et al., 2006), leading to an inability to compare results of physical tests against hip pain patients without labral pathology and, hence, calculation of specificities. Another significant weakness was that MRA was adopted as the reference standard in three studies because no subject underwent arthroscopy (Burnett et al., 2006; Narvani et al., 2003; Troelsen et al., 2009). Arthroscopy is however the gold standard investigation, for the identification of acetabular labral tears (Byrd and Jones, 2004).

3.3.1. Physical examination tests

Physical examination sensitivities ranged from 75% for the impingement test (Narvani et al., 2003) to 98% for the Fitzgerald test (Fitzgerald, 1995). Specificity was calculable for two of the tests (impingement test and modified Thomas test), with specificities of 43–100% (Narvani et al., 2003; Troelsen et al., 2009) and 92% (McCarthy and Busconi, 1995), respectively.

3.4. Imaging studies: quality assessment

There were a number of quality issues across the imaging studies (Table 1). Selection bias was a main source of bias. Only

subjects with positive findings from imaging were listed for arthroscopy (Neumann et al., 2007); only complex cases or those failing conservative management were listed for surgery (Byrd and Jones et al., 2004); and, only those found at arthroscopy to have a labral tear were included in retrospective analyses (Burnett et al., 2006). As a consequence, simple tears that responded to conservative management were not included and subjects with disabling hip symptoms but negative imaging results were, then, often excluded (Czerny et al., 1996; Chan et al., 2005; Neumann et al., 2007). Hence, there were very few reports of false negative images. Blinding was another quality issue in the majority of studies, with only 2/14 studies reporting that the surgeon was blinded to imaging results (Nishi et al., 2007; Yamamoto et al., 2007). Grading of tears relating to diagnostic threshold was also an issue. A number of studies used grading systems where both the radiologist and surgeon had to agree on the grade (extent) of tear (Czerny et al., 1996, 1999; Chan et al., 2005; Freedman et al., 2006). This greatly affects estimates of accuracy when comparing against those studies that only reported scans as positive or negative for a tear.

3.4.1. MRA

Twelve studies investigated the accuracy of MRA (Table 4). MRA sensitivity in studies ranged from 24% (Mitchell et al., 2003) to 100% (Chan et al., 2005), but the majority of studies (11/12) reported a sensitivity between 63% and 100%. Specificity ranged from 44% (Kenney et al., 2004) to 100% (Toomayan et al., 2006) with the majority of studies (11/12) reporting a specificity between 71% and 100%, although 4 studies did not report specificity values (Chan et al., 2005; Burnett et al., 2006; Freedman et al., 2006; Ziegert et al., 2009). Sensitivity might have been lower in Mitchell et al.'s study (2003) because, regardless of imaging findings, all patients with signs and symptoms indicative of hip pathology went on for arthroscopy. This study, therefore, contributed a much higher number of false negative MRA results.

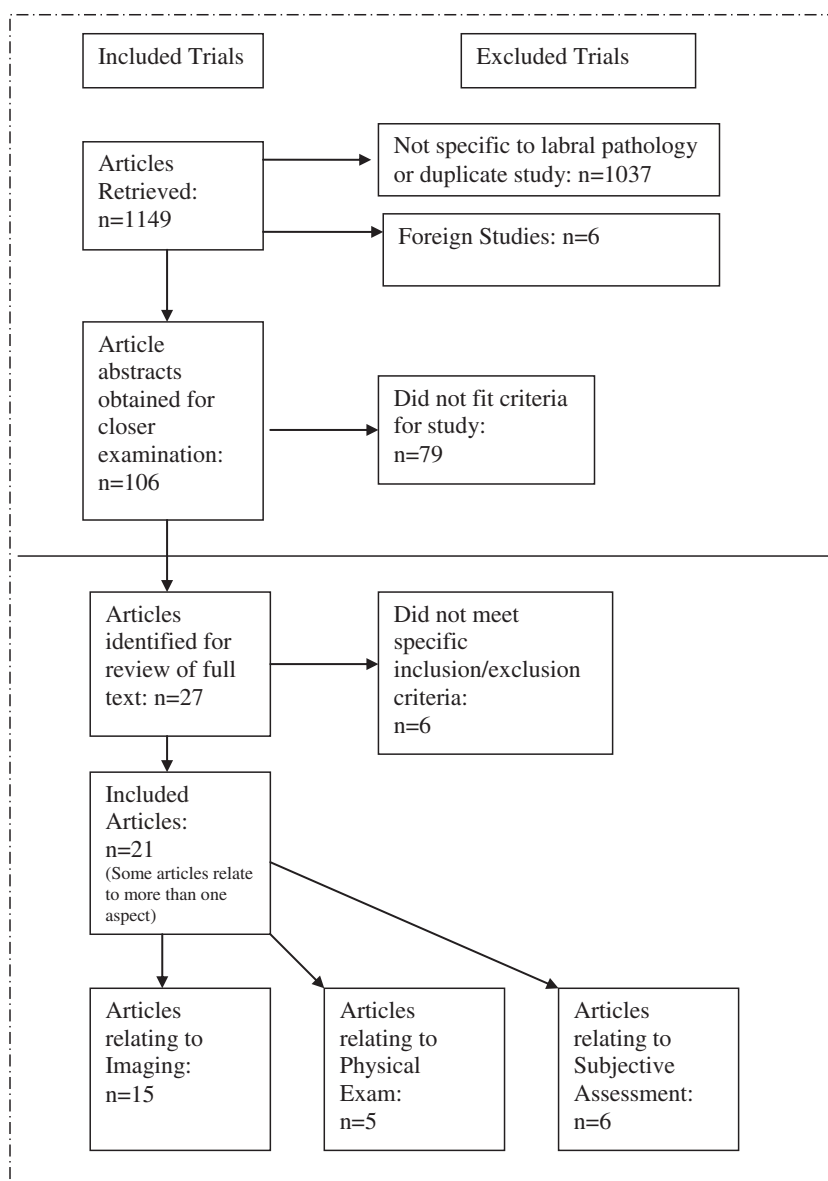


Fig. 2. Flow Diagram of Searches.

3.4.2. MRI and MRA

Three studies compared MRI and MRA findings (Byrd et al., 2004; Czerny et al., 1996; Toomayan et al., 2006) (Table 4). The studies by Czerny et al. (1996) and Byrd et al. (2004) were strengthened by the same subjects undergoing both imaging techniques; allowing direct comparison between the two. Czerny et al. (1996) found that MRI had a sensitivity of 30% and MRA 90%, and an accuracy of 36% and 91%, respectively. Byrd et al. (2004) also reported statistically significant differences between the sensitivity of MRI and MRA (25% and 66% respectively). Individual studies further demonstrated that MRA outperforms MRI in the detection of acetabular labral tears (see Table 4).

3.4.3. CT

Two studies investigated the accuracy of CT arthrography (Nishi et al., 2007; Yamamoto et al., 2007). Nishi et al. (2007) reported a sensitivity of 97%, specificity of 87% and accuracy of 92%; and Yamamoto et al. (2007) a sensitivity of 92%, specificity of 100% and

accuracy of 95%. These values demonstrate reasonable agreement between the two studies, and a higher level of accuracy compared to the MRI studies, and to a large proportion of the MRA studies.

3.5. Meta-analysis: sources of heterogeneity

The main limitation for meta-analysis was the degree of heterogeneity across studies. Sources of heterogeneity included:

- Type of study prospective/retrospective case/cohort study
- Grading of tear between surgeon and radiologist (different diagnostic threshold across studies)
- Blinding (non-blinded studies have been found to overestimate diagnostic performance (Deeks, 2001))
- Population: orthopaedic or sports
- Population: presence/absence of hip dysplasia and osteoarthritis
- Physical tests utilised
- Reference standard utilised

Table 1
Results of Quality Assessment Using QUADAS.

Study First Author	Appropriate Spectrum	Inclusion Exclusion Criteria	Appropriate Reference Test	Disease Progression Bias Avoided	Partial Verification Bias Avoided	Differential Verification Bias Avoided	Incorporation Bias Avoided	Index Test Details	Reference Test Details	Test Review Bias Avoided	Diagnostic Test Review Bias Avoided	Clinical information Available	Uninterpretable Results Explained	Withdrawal Bias Avoided
Burnett et al. (2006)*	Y	U	U	N	Y	Y	Y	N	Y	Y	N	Y	U	Y
Byrd and Jones (2004)*	U	U	Y	U	Y	Y	Y	Y	Y	Y	N	Y	U	Y
Chan et al. (2005)	Y	U	Y	Y	N	N	Y	Y	Y	Y	U	Y	Y	Y
Czerny et al. (1996)*	Y	N	Y	U	N	N	Y	Y	N	Y	N	U	Y	Y
Czerny et al. (1999)*	Y	N	Y	U	Y	Y	Y	Y	N	Y	U	U	Y	Y
Farjo et al. (1999)	Y	U	Y	U	Y	Y	Y	N	Y	Y	N	Y	Y	Y
Fitzgerald (1995)*	Y	U	Y	U	N	N	Y	Y	N	Y	N	Y	Y	Y
Freedman et al. (2006)*	Y	U	Y	U	Y	Y	Y	Y	N	Y	N	N	Y	Y
Keeney et al. (2004)*	Y	U	Y	U	Y	Y	Y	Y	Y	Y	U	U	U	Y
Leunig et al. (1997)*	Y	N	Y	N	Y	U	Y	Y	N	Y	N	Y	U	Y
McCarthy and Busconi (1995)	Y	U	Y	U	Y	Y	Y	N	Y	Y	N	Y	Y	Y
Mintz et al. (2005)	Y	N	Y	N	Y	Y	Y	Y	Y	Y	U	Y	Y	Y
Mitchell et al. (2003)*	N	N	Y	U	Y	Y	Y	N	Y	Y	U	Y	Y	Y
Narvani et al. (2003)*	N	N	U	N	Y	Y	Y	N	Y	Y	N	Y	Y	Y
Neumann et al. (2007)*	Y	N	Y	U	N	N	Y	Y	N	Y	N	Y	Y	Y
Nishi et al. (2007)*	N	Y	Y	Y	N	N	Y	Y	N	Y	Y	N	Y	N
O'Leary et al., (2001)*	Y	U	Y	U	N	N	Y	Y	N	Y	N	Y	Y	Y
Toomayan et al. (2006)*	Y	N	Y	U	Y	Y	Y	Y	N	Y	U	U	U	Y
Troelsen et al. (2009)	N	Y	U	U	Y	Y	Y	Y	Y	Y	Y	N	Y	Y
Yamamoto et al. (2007)	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	N	Y	Y
Ziegert et al. (2009)*	Y	N	Y	N	Y	Y	Y	Y	N	N	U	Y	Y	Y

(Y = yes, N = no, U = unclear, *studies had less than 10 yes answers to QUADAS assessment).

Table 2
Reported symptoms and physical examination findings.

First Author (Year)	Number of Subjects	Mean Age (years)	Setting	Physical Test			Symptom		
				Test	Sensitivity	Specificity	Symptom	Sensitivity	Specificity
Burnett et al. (2006)	66 with tear	38	O	Impingement Test (63/66)	95%	All had tear	Groin pain 61/66 Buttock pain 25/66 Anterior thigh/knee pain 34/66	92% 38% 53%	
				Audible palpable click (44/55)	80%		Mechanical Symptoms 35/66	53%	
Fitzgerald (1995)	55 with tear	NR	O	Click/pain with Fitzgerald test (54/55)	98%	All had tear	Groin pain 49/55 Clicking 34/55 Trochanteric pain 2/55 Buttock pain 1/55	88% 61% 4% 2%	
Farjo et al. (1999) ^a	28 with tear	41	O				Mechanical Symptoms 18/28 Clicking 16/28 Locking 5/28 Giving Way 4/28	64% 57% 18% 14%	
McCarthy and Busconi (1995)	59, 35 with tear	37	O	Thomas Test (33/59, 31/35 with tear)	89%	92%	Anterior groin pain 35/35 Clicking 31/35 Giving Way 20/35 Locking 4/35	100% 89% 57% 11%	4% 92% 83% 33%
Narvani et al. (2003)	18, 4 with tear	30.5	S	Pain on Hip Quadrant (11/18, 3 with tear)	75%	43%	Clicking 6/18	100%	85%
O'Leary et al., (2001) ^a	86 hips in 85 subjects, 22 with tear	33.6	O				Thigh/groin pain 22/22 Mechanical Symptoms 22/22	100% 100%	
Troelson et al. (2009)	18, 17 with tear	43	O	Impingement Test	59%	100%			

O = Orthopaedic S = Sports NR = Not Reported.

Numbers such as 33/59 denote number of subjects with positive test/symptom out of total number of subject.

^a Studies included patients with known hip dysplasia/osteoarthritic changes.

3.6. Synthesis of results

Sensitivities and specificities from the studies were plotted on a ROC plot for MRA imaging studies (Fig. 3). This was the only group evaluated as possessing potential for meta-analysis based on the inclusion/exclusion criteria and sources of heterogeneity. However, even within this subgroup, four studies could not be plotted owing

to the lack of specificity values (Chan et al., 2005; Burnett et al., 2006; Freedman et al., 2006; Ziegert et al., 2009).

The three studies in the top left corner of the ROC plot (Czerny et al., 1996; Toomayan et al., 2006; Neumann et al., 2007) demonstrate high accuracy, with high sensitivity and specificity (Sackett, 1992). The overall spread of points in Fig. 3 illustrates the considerable heterogeneity between studies. It was therefore not appropriate to pool summary statistics across these studies (Deeks, 2001). This was demonstrated further by the use of paired forest plots (Fig. 4) that demonstrate that sensitivities are reasonably homogeneous with the exception of Mitchell et al. (2003) and Leunig et al. (2005); but specificities are largely heterogeneous.

Table 3
Physical Tests.

Test	Description and Reliability
Impingement Test	- Combination of hip flexion, internal rotation and adduction to engage the femoral head neck junction into the anterior superior labrum and acetabular rim (Leunig et al., 2005) - Inter-rater reliability (Martin et al., 2008); kappa value of 0.58 (raters needed to get the same result for pain location).
Quadrant Test	- Manoeuvre of internal rotation, flexion and axial compression (Martin et al., 2006). - No reliability data were available for this test.
Modified Thomas Test	- Extension of the hip from the fully flexed position in either internal or external rotation (DeAngelis and Busconi, 2003). The test is deemed positive where pain and clicking are reported. - No reliability data were available for this test.
Fitzgerald Test	- Manoeuvre of acute flexion, external rotation and full abduction followed by extension, internal rotation and adduction (deemed to be positive if it precipitates pain with or without an associated click in patients with an anterior labral tear). Extension with abduction and external rotation from the fully flexed, internally rotated and adducted position is used to assess for a posterior labral tear (deemed positive with provocation of pain with or without an associated click) (Fitzgerald, 1995). - No reliability data were available for this test.

4. Discussion

Following analysis of the results regarding use of data from the patient history and physical testing, it was not possible to make recommendations for practice. In relation to reported symptoms in the patient history, anterior groin pain and mechanical symptoms (particularly clicking) are useful to rule out a diagnosis of acetabular labral tear if negative, but are of limited use in isolation to rule in a diagnosis if positive. Similarly, the research supports the use of physical tests to help to rule out the presence of a labral tear only in response to a negative test result. In isolation, the physical tests are also of limited use in clinical practice.

A number of key findings and recommendations for practice can be drawn following analysis of the results regarding imaging.

- The first recommendation is for MRA to be the preferred imaging method for patients presenting with hip pain with signs and symptoms consistent of a tear. Ziegert et al. (2009) found the use of oblique imaging improved the overall detection rate of labral tears, and found a greater than 95%

Table 4
Results of imaging studies.

First Author	Subject No	Grading of Tear (surgeon & rad)	Mean Age/Range (years)	Setting	Test	Sensitivity	Specificity	Accuracy
Burnett et al. (2006)	66 (S66)	No	38	O	MRA	79%		
Byrd and Jones (2004)	40(S40)	No	NR	O	MRA (40)	66%	75%	
					MRI (40)	25%	67%	
Chan et al. (2005)	30 (S17) (S22)	Yes; Czerny et al. (1996)	41	O	MRA	100%		94%
					MRA (22)	90%	100%	91%
Czerny et al. (1996)	(S22)	Yes; Czerny et al. (1996)	39	O	MRI (22)	30%	100%	36%
					MRA	91%	71%	88%
Freedman et al. (2006)	24 (S24)	Yes; Czerny et al. (1996)	37.1	O	MRA	96%		96%
Keeney et al. (2004)	101 (102 hips/S)	No	NR	O	MRA	71%	44%	69%
Leunig et al. (1997) ^a	23 (S23)	No	40+/- 2	O	MRA	63%	71%	65%
Mintz et al. (2005)	92 (S92)	No	38.5	O	MRI (Rad A)(92)	97%	33%	95%
					MRI (Rad B) (92)	96%	33%	93%
Mitchell et al. (2003)	25 (S25)	No	30.9	Sp	MRA (24)	24%	100	46%
Neumann et al. (2007)	100 (S23)	No	38	O	MRA	89%	95%	
Nishi et al. (2007) ^a	29 (41 hips/20S)	No	33	O	CT Arth	97%	87%	92%
Toomayan et al. 2006 ^a	48 (51 hips/S)	No	35	O	MRA (30)	92%	100%	93%
					MRI (SFOV)(7)	25%	100%	57%
					MRI (LFOV)(14)	8%	100%	21%
Yamamoto et al. 2007 ^a	189 (S22)	No	43	O	CT	92%	100%	95%
Ziegert et al. 2009 ^a	144 (S144)	No	40.6	O	MRA	97%		

O = orthopaedic, Sp = sports clinic, S = surgery, LFOV = large field of view, SFOV = small field of view, NR = not reported, rad = radiologist.
^a Studies included patients with known hip dysplasia and/or osteoarthritic change.

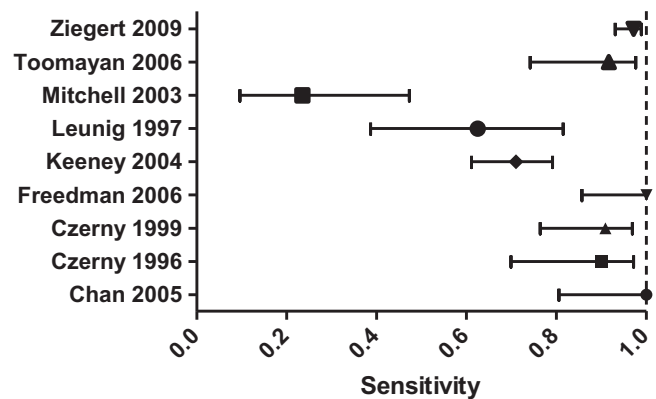
detection rate was achieved with 3 imaging sequences (axial oblique T1 weighted, sagittal T1 weighted, and coronal T2 weighted).

- II. The second recommendation is for patients where MRA is contraindicated (patients implanted with certain metallic devices and claustrophobic patients (Yamamoto et al., 2007)), for CT to be the imaging technique of choice. CT was shown to be a good alternative to MRA from the small number of available studies. CT is also less expensive and less time consuming than MRA, but possesses a major disadvantage of high levels of radiation exposure (Yamamoto et al., 2007).
- III. The third recommendation is in a situation where MRA is available, that MRI is not considered appropriate as the first line imaging technique. However, patients who are allergic to gadolinium (can cause breathing difficulty, throat swelling and convulsions (Yamamoto et al., 2007)) or are anxious regarding the invasive nature of MRA may still benefit from MRI investigation.

Further research with regards to diagnosis of labral pathology is needed, therefore, before decisions can be made solely on reported symptoms, physical examination, and imaging without the use of diagnostic arthroscopy. In clinical practice, a composite of tests

could provide higher levels of diagnostic accuracy when aiming to identify acetabular labral tears. Composite testing has previously been used in the diagnosis of sacro-iliac joint dysfunction to obtain higher levels of diagnostic accuracy (Van Der Wurff et al., 2006).

Sensitivities with 95% Confidence Intervals



Specificity with 95% Confidence Intervals

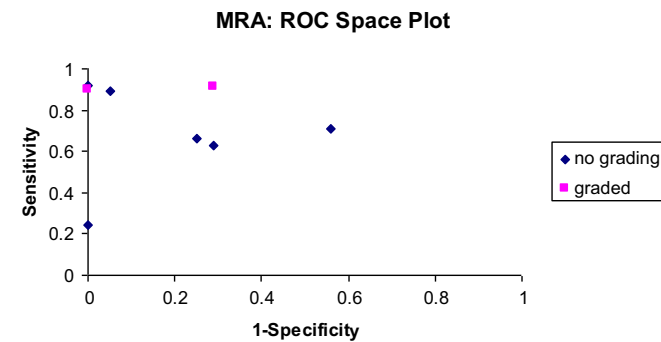
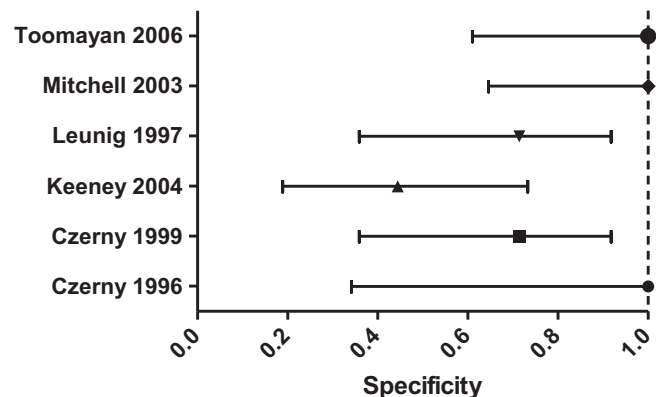


Fig. 3. ROC space plot of results of individual MRA studies (n = 8). (Those plotted in squares/pink graded tears according to Czerny et al., 1996). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Fig. 4. Forest Plots of Sensitivity and Specificity (excludes Neumann et al., 2007 and Burnett et al., 2006 due to lack of raw data for these studies).

In terms of a composite of tests, further research looking at patients presenting with anterior groin pain, a mechanical hip symptom (clicking, locking, giving way), a positive physical test (Hip Quadrant test, Impingement test, Fitzgerald test, Modified Thomas test) and a positive finding on MRA or CT, could be more accurate in identifying those patients with a symptomatic labral tear. This would then aid decision making in regard to referral for arthroscopic treatment in this complex patient population.

4.1. Limitations

Results of this study should however be interpreted with consideration to a number of limitations. Firstly foreign studies were not included which may have introduced bias (Song et al., 2002). Only one reviewer searched information sources and performed the initial screening of abstracts which again could introduce bias (Edwards et al., 2002). A limited amount of journals were available for hand searching at the research institution hence limiting the scope of the study. Finally the lead researcher was also a reviewer and, therefore, potentially included studies likely to produce anticipated results (Sim and Wright, 2000).

4.2. Conclusion

Quality of data across the studies was low. Due to the inherent limitations of the case and cohort studies presented, health professionals must be cautious in interpreting the study results for use in clinical practice. The key finding of the review was that MRA consistently outperformed MRI in the detection of acetabular labral tears. CT was also found to be a useful technique. Physical examination studies were largely heterogeneous, generally of low quality, and are of limited use at present to guide clinical practice. Symptoms likely to be present in patients with a symptomatic labral tear were found to include anterior groin pain and mechanical hip symptoms.

Due to the heterogeneity and poor quality of studies evaluated in the review, meta-analysis was limited, and a largely descriptive analysis of study results was necessary. Further good quality research would allow specific meta-analysis and pooling of study results, to give a more precise estimate of diagnostic accuracy (Egger et al., 2001). Future research into composite testing could also provide a more useful guideline to be utilised in clinical practice.

References

- Bradshaw C, Bundy M, Falvey E. The diagnosis of longstanding groin pain: a prospective clinical cohort study. *British Journal of Sports Medicine* 2008;(42):851–4.
- Byrd J, Jones K. Hip arthroscopy in athletes. *Clinical Sports Medicine* 2001;(20):749–61.
- Byrd J, Jones K. Diagnostic accuracy of clinical assessment, magnetic resonance imaging, magnetic resonance arthrography and intra articular injection in hip arthroscopy patients. *American Journal of Sports Medicine* 2004;(32):1668–74.
- Burnett R, Della Rocca G, Prather H, Curry M, Maloney W, Clohisy J. Clinical presentation of patients with tears of the acetabular labrum. *Journal of Bone and Joint Surgery* 2006;88(7):1448–57.
- Centre for Reviews and Dissemination (CRD). *Systematic Reviews: CRD's guidance for undertaking reviews in healthcare*, Chapter 2, 2009.
- Chan Y, Lien L, Hsu H, Wan Y, Lee M, Hsu K, et al. Evaluating hip labral tears using MRA: a prospective study comparing hip arthroscopy and MRA diagnosis. *The Journal of Arthroscopic and Related Surgery* 2005;21(10):1250–6.
- Czerny C, Hofman S, Neuhold A, Tschauner C, Engel A, Recht M, et al. Lesions of the acetabular labrum: accuracy of MR imaging and MR arthrography in detection and staging. *Radiology* 1996;(200):225–30.
- Czerny C, Hofmann S, Urban M, Tschauner C, Neuhold A, Pretterklieber M, et al. MR arthrography of the adult acetabular capsular-labral complex: correlation with surgery and anatomy. *American Journal of Rheumatology* 1999;(173):345–9.
- DeAngelis N, Busconi B. Assessment and differential diagnosis of the painful hip. *Clinical Orthopaedics and Related Research* 2003;(406):11–8.
- Deeks J. Systematic reviews of evaluations of diagnostic and screening tests. *BMJ* 2001;(323):157–62.
- de Vet H, Eisinga A, Riphagen I, Aertgeerts B, Pewsner D, Mitchell R. Chapter 7: searching for studies. In: Deeks J, Bossuyt P, Gatsonis C, editors. *Cochrane handbook for systematic reviews of diagnostic test accuracy version 1.1*. The Cochrane Collaboration. Available from, srdta.cochrane.org; 2008.
- de Graaf I, Prak A, Bierma-Zeinstra S, Thomas S, Peul W, Koes B. Diagnosis of lumbar spinal stenosis: a systematic review of the accuracy of diagnostic tests. *Spine* 2006;(31):1168–76.
- Deville W, Buntinx F, Bouter L, Montori V, de Vet H, van der Windt D, Bezemer D. Conducting systematic reviews of diagnostic studies: didactic guidelines. *2002*;1:9.
- Edwards P, Clarke M, DiGiuseppi C, Prata S, Roberts I, Wentz R. Identification of randomised controlled trials in systematic reviews: accuracy and reliability of screening records. *Statistics in Medicine* 2002;(21):1635–40.
- Egger M, Smith G, Altman D. *Systematic reviews in health care: meta analysis in context*. Bodmin: BMJ Publishing Group; 2001. 248–281.
- Farjo L, Glick J, Sampson T. Hip arthroscopy for acetabular labral tears. *Arthroscopy* 1999;(15):132–7.
- Feeley B, Powell J, Muller M, Barnes R, Warren R, Kelly B. Hip injuries and labral tears in the national football league. *The American Journal of Sports Medicine* 2008;(36):2187–95.
- Fitzgerald R. Acetabular labrum tears: diagnosis and treatment. *Clinical Orthopaedics and Related Research* 1995;(311):60–8.
- Freedman B, Dinauer P, Giuliani J, Kuklo T, Murphy K. Prognostic value of MRA for Czerny stage II and III acetabular labral tears. *Journal of Arthroscopic and Related Surgery* 2006;22(7):742–7.
- Hollingsworth W, Medina L, Lenkinski R, Shibata D, Bernal B, Zurakowski D, et al. Interrater reliability in assessing quality of diagnostic accuracy studies using the QUADAS tool. A preliminary assessment. *Academic Radiology* 2006;13(7):803–10.
- Kelly B, Weiland D, Schenker M, Philippon M. Arthroscopic labral repair in the hip; surgical technique and review of the literature. *Arthroscopy* 2005;(21):1496–504.
- Kenney J, Peelle M, Jackson J, Rubin D, Maloney W, Clohisy J. Magnetic resonance arthrography versus arthroscopy in the evaluation of articular hip pathology. *Clinical Orthopaedics and Related Research*; 2004:163–9.
- Leunig M, Werlen S, Ungersbock A, Ito K, Ganz R. Evaluation of the Acetabular Labrum by MR Arthrography. *Journal of Bone and Joint Surgery* 1997;(79-B):230–4.
- Leunig M, Beck M, Dora C, Ganz R. Femoroacetabular impingement: aetiology and surgical concept. *Operative Techniques in Orthopaedics* 2005;(15):247–55.
- Martin R, Eneski K, Draovitch T, Trapuzzano T, Philippon M. Acetabular labral tears of the hip: examination and diagnostic challenges. *Journal of Orthopaedics and Sports Physical Therapy* 2006;36(7):503–14.
- Martin R, Irrgang J, Sekiya J. The diagnostic accuracy of a clinical examination in determining intra articular hip pain for potential hip arthroscopy candidates. *The Journal of Arthroscopic and Related Surgery* 2008;24(9):1013–8.
- McCarthy J, Busconi B. The role of hip arthroscopy in the diagnosis and treatment of hip disease. *Canadian Journal of Surgery* 1995;(38):S13–7.
- McCarthy J, Noble P, Schuck M, Lee J, The Otto E. Aufranc Award: the role of labral lesions to development of early degenerative hip disease. *Clinical Orthopaedics* 2001;(393):25–37.
- McCarthy J, Noble P, Aluisio F. Anatomy, pathologic features and treatment of acetabular labral tears. *Clinical Orthopaedics* 2003;(406):38–47.
- Mintz D, Hooper T, Connell D, Buly R, Padgett D, Potter H. Magnetic Resonance Imaging of the Hip: detection of labral and chondral abnormalities using noncontrast imaging. *The Journal of Arthroscopic and Related Surgery* 2005;21(4):385–93.
- 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. *Clinical Journal of Sports Medicine* 2003;(13):152–6.
- Narvani A, Tsiroidis E, Kendall S, Chaudhuri R, Thomas P. A preliminary report on prevalence of acetabular labrum tears in sports patients with groin pain. *Sports Medicine* 2003;(11):403–8.
- Neumann G, Mendicuti A, Zou K, Minas T, Coblyn J, Winalski C, et al. Prevalence of labral tears and cartilage loss in patients with mechanical symptoms of the hip: evaluation using MR arthrography. *Osteoarthritis and Cartilage* 2007;(15):909–17.
- Nishi T, Tanaka H, Sugano M, Miki H, Takao M, Yoshikawa H. Disorders of the acetabular labrum and articular cartilage in hip dysplasia; evaluation of using isotropic high-resolution CT arthrography with sequential radial reformation. *Osteoarthritis and Cartilage* 2007;(15):251–7.
- O'Leary J, Berend M, Parker T. The relationship between diagnosis and outcome in arthroscopy of the hip. *The Journal of Arthroscopic and Related Surgery* 2001;17(2):181–8.
- Peterson MC, Holbrook JH, Von Hales D, Smith NL, Staker LV. Contributions of the history, physical examination and laboratory investigation in making medical diagnoses. *Western Journal of Medicine* 1992;156(2):163–5.
- Reitsma JB, Rutjes AWS, Whiting P, Vlassov VV, Leeflang MMG, Deeks JJ. Chapter 9: assessing methodological quality. In: Deeks JJ, Bossuyt PM, Gatsonis C, editors. *Cochrane handbook for systematic reviews of diagnostic test accuracy Version 1.0.0*. The Cochrane Collaboration. Available from, <http://srdta.cochrane.org/>; 2009.
- Sackett D. A primer on the precision and accuracy of the clinical examination. *JAMA* 1992;267(19):2638–44.
- Sehgal N, Shah R, McKenzie-Brown A, Everett C. Diagnostic utility of facet (zygapophysial) joint injections in chronic spinal pain: a systematic review of evidence. *Pain Physician* 2005;(8):211–24.

- Shah R, Everett C, McKenzie-Brown A, Sehgal N. Discography as a diagnostic test for spinal pain: a systematic and narrative review. *Pain Physician* 2005;(8):187–209.
- Sim J, Wright C. *Research in healthcare, concepts, designs and methods*. Cheltenham: Stanley Thomes Ltd; 2000. p. 28.
- Song F, Khan K, Dinnes J, Sutton A. Asymmetrical funnel plots and publication bias in meta-analyses of diagnostic accuracy. *International Journal of Epidemiology* 2002;(31):88–95.
- Toomayan G, Russell Holman W, Major N, Kozlowicz S, Parker Vail T. Sensitivity of MR arthrography in the evaluation of acetabular labral tears. *American Journal of Roentgenology* 2006;(186):449–53.
- Troelsen A, Mechlenburg I, Gelineck J, Bolvig L, Jacobsen S, Soballe K. What is the role of clinical tests and ultrasound in acetabular labral tear diagnostics? *Acta Orthopaedica* 2009;80(3):314–8.
- Van Der Wurff P, Buijs E, Groen G. A multitest regimen of pain provocation tests as an aid to reduce unnecessary minimally invasive sacro iliac joint procedures. *Archives of Physical Medical Rehabilitation* 2006;(87):10–4.
- Whiting P, Rutjes A, Dinnes J, Reitsma J, Bossuyt P, Kleijnen J. Development and validation of methods for assessing the quality and reporting of diagnostic accuracy studies. *Health Technology Assessment* 2004;(8):1–234.
- Whiting P, Rutjes A, Dinnes J, Reitsma J, Bossuyt P, Kleijnen J. A systematic review finds that diagnostic reviews fail to incorporate quality despite available tools. *Journal of Clinical Epidemiology* 2005;(58):1–12.
- Whiting P, Weswood M, Rutjes A, Reitsma J, Bossuyt P, Kleijnen J. Evaluation of QUADAS, a tool for the quality assessment of diagnostic accuracy studies. *BMC Medical Research Methodology* 2006;6(9).
- Yamamoto Y, Tonotsuka H, Ueda T, Hamada Y. Usefulness of radial contrast-enhanced computed tomography for the diagnosis of acetabular labrum injury. *The Journal of Arthroscopic and Related Surgery* 2007;23(12):1290–4.
- Ziebert A, Blankenbaker D, De Smet A, Keena J, Shinki K, Fine J. Comparison of the standard hip MR Arthrographic imaging planes and sequences for detection of arthroscopically proven labral tears. *American Journal of Roentgenology* 2009;(192):1397–400.