Reliability and Validity of Diagnosing Acetabular Labral Lesions with Magnetic Resonance Arthrography

Gustaaf Reurink, MD, MSc, Sebastiaan P.L. Jansen, MD, Jasper M. Bisselink, MD, Patrice W.J. Vincken, MD, Adam Weir, PhD, and Maarten H. Moen, MD

Investigation performed at Rijnland Hospital, Leiderdorp, the Netherlands

**Background:** Arthroscopic surgery of the hip is being increasingly used to diagnose and treat various abnormalities, including acetabular labral tears. Magnetic resonance arthrography has been suggested as the imaging test of choice for the evaluation of the acetabular labrum. There is substantial variability in the previously reported accuracy of magnetic resonance arthrography for diagnosing labral lesions. Interobserver reliability has not been established previously. The purpose of this study was to establish the interobserver reliability and the validity of magnetic resonance arthrography for detecting lesions of the acetabular labrum in a retrospective case series.

**Methods:** Two radiologists independently assessed the acetabular labrum on magnetic resonance arthrograms of ninety-five hips in ninety-three patients who underwent hip arthroscopy for a suspected acetabular labral lesion. Magnetic resonance arthrography findings were compared with the gold standard, which was defined as the assessment of the labrum during the hip arthroscopy.

**Results:** At arthroscopy, ninety-one labral lesions were identified in the ninety-five hips. The interobserver reliability of detecting labral lesions with magnetic resonance arthrography was fair (kappa = 0.268). Magnetic resonance arthrography, as interpreted by observers A and B, showed a sensitivity of 86% and 86%, specificity of 75% and 50%, negative predictive value of 19% and 13%, and positive predictive value of 99% and 98%, respectively.

**Conclusions:** Because of its limited reliability and the high prevalence of labral lesions, magnetic resonance arthrography provides a limited complementary benefit in the detection of labral lesions in patients with a high clinical suspicion of labral pathology. When there is a high clinical suspicion of a labral lesion, magnetic resonance arthrography has a poor negative predictive value and cannot be used to rule out a labral lesion. Physicians should critically consider whether the findings on a magnetic resonance arthrogram will alter the treatment strategy for an individual patient with a clinical suspicion of labral pathology.

Acetabular labral tears are a common cause of hip pain, especially in young and middle-aged adults. These tears can be associated with previous hip trauma or a history of hip disease such as developmental dysplasia, Legg-Calvé-Perthes disease, slipped capital femoral epiphysis, degenerative hip disorders, and osseous dysmorphism resulting in femoroacetabular impingement. Labral tears may affect the development and progression of osteoarthritis in the hip joint.

The major symptom of an acetabular labral tear is activity-related anterior groin pain, but the clinical presentation may vary considerably. Radiographs can identify osseous abnormalities but cannot be used to diagnose a labral lesion, which can result in...
Radiographs, changes on radiographs, and/or signs of femoroacetabular impingement on and locking, decreased hip motion, an absence of or minimal degenerative impingement test on physical examination, mechanical symptoms of snapping hip arthroscopy in the period from January 2007 to October 2010 and had had (ninety-five hips) with a clinical suspicion of a labral tear who had undergone it provides suboptimal visualization of the acetabular labrum.

Several studies comparing magnetic resonance arthrography of the hip with surgical findings showed sensitivity and specificity of magnetic resonance arthrography for detecting labral tears ranging from 60% to 100% and 0% to 100%, respectively10-17. However, these studies had some methodological flaws, including relatively small sample sizes10,11,13-17, a retrospective design11-17, and a risk of patient selection bias11-17.

Keeney et al.12 reported substantial variability in sensitivity and specificity between experienced observers for detecting acetabular labral lesions using magnetic resonance arthrography, raising concerns about the interobserver reliability of this imaging study for these lesions. Detection of cartilage abnormalities with use of magnetic resonance arthrography has been reported to have poor reliability (kappa [k] < 0.2)14. To our knowledge, the interobserver reliability of detecting labral lesions with magnetic resonance arthrography has not been established previously.

The purpose of this study was twofold: to determine the interobserver reliability and the validity of detecting acetabular labral tears with magnetic resonance arthrography in a retrospective case series.

Materials and Methods

Patient Selection

The patient cohort consisted of patients presenting with hip disorders to a referral center at a medium-sized regional hospital. Ninety-three patients (ninety-five hips) with a clinical suspicion of a labral tear who had undergone hip arthroscopy in the period from January 2007 to October 2010 and had preoperative magnetic resonance arthrography were included in this study. Clinical suspicion was based on a clinical presentation of groin pain, a positive impingement test on physical examination, mechanical symptoms of snapping and locking, decreased hip motion, an absence of or minimal degenerative changes on radiographs, and/or signs of femoroacetabular impingement on radiographs12-17. Hips on which magnetic resonance arthrography had been performed in other hospitals were excluded from the study. The patient selection procedure is presented in Figure 1. All of the hip arthroscopy procedures were performed by the same orthopaedic surgeon.

The preoperative diagnostic workup consisted of a history, physical examination, two hip radiographs (anteroposterior and Lauenstein projections), hip magnetic resonance arthrography, and in most patients a diagnostic injection containing a local anesthetic and steroid (Chirocaine [levobupivacaine] and Depo-Medrol [methylprednisolone acetate]) to differentiate between intra-articular and extra-articular pathology. Patient characteristics are presented in Table I.

Hip Magnetic Resonance Arthrography

Preoperatively, all patients underwent direct magnetic resonance arthrography of the hip to assess labral and other intra-articular pathology. Under fluoroscopic guidance and under sterile conditions, a 20-gauge spinal needle was inserted into the joint. After confirmation of the intra-articular position of the needle tip with injection of a small amount of iodinated contrast material (Hexabrix 320; Guerbet, Roissy, France), a solution of 0.1 mL of gadopentetate dimeglumine (Magnevist; Bayer, Leverkusen, Germany) in 20 mL of saline solution was injected. Following the injection, magnetic resonance images were obtained in three planes with a 1.5-T magnet system (Achieva 1.5T; Philips Medical Systems, Best, the Netherlands) with use of a dedicated transmit-receive coil. The magnetic resonance imaging protocol consisted of five sequences: a coronal T1-weighted spin-echo (SE) sequence (repetition time/echo time [TR/TE] of 628/18 ms and field of view [FOV] of 360 mm) and a coronal T2-weighted SE sequence with frequency-selective fat suppression (TR/TE of 3146/70 ms and FOV of 360 mm) of both hips and transverse-oblique, sagittal, and coronal T1-weighted SE sequences with frequency-selective fat suppression (TR/TE of 645 to 582/14 to 12 and FOV of 150 mm) of the affected hip. All sequences had a section thickness of 3 mm. The total time to perform the magnetic resonance imaging examination (including the initial survey sequence) was less than twenty-five minutes.

Magnetic resonance arthrograms were retrospectively assessed by two radiologists, with six years (observer A) and three years (observer B) of experience in musculoskeletal radiology. The radiologists were blinded to the findings at arthroscopy, and they assessed the magnetic resonance arthrograms independently. They were informed that the magnetic resonance arthrography had been performed in patients with a clinical suspicion of a labral lesion, but

![Patient cohort flow diagram.](image-url)
no information about whether a patient had hip arthroscopy was provided. A standard assessment form was used for scoring labral abnormalities. Criteria for labral abnormalities were derived from studies of imaging of the acetabular labrum.\textsuperscript{10,12,19} The appearance of the labrum was scored as “normal labrum,” “abnormal labral shape,” “abnormal signal within labrum,” “labral cyst,” “contrast fluid within labrum,” and/or “detachment from acetabulum.” The appearance of an abnormal labral shape, abnormal signal within the labrum, labral cyst, contrast fluid within the labrum, and/or detachment from the acetabulum was considered to be a labral lesion. Labral abnormalities were identified as being located in the anterior-superior, posterior-superior, anterior-inferior, and/or posterior-inferior quadrant in the sagittal plane.\textsuperscript{19} An abnormality could be located in more than one quadrant.

**Hip Arthroscopy**

Hip arthroscopy was performed with the patient under general or spinal anesthesia. The patient was positioned supine on a standard fracture table. The hip to be operated on was placed in slight flexion, abduction, and internal rotation to provide access to the central compartment. Sufficient force was applied with traction to achieve adequate distraction, which was confirmed with fluoroscopy.

To gain access to the peripheral compartment, the hip was flexed to about 60° and could be rotated freely. No traction was applied for access to the peripheral compartment. A 70° arthroscope, hip arthroscopy set (Hip Arthroscopy Master Instrument Set; Arthrex), shaver, and radiofrequency device were used. All operations were recorded on DVD and documented in the operative report.

The arthroscopic findings were evaluated retrospectively with use of the documentation in the operative report, in which the aspect of the acetabular labrum was systematically classified as “normal,” “degenerative,” “torn,” or “detached from acetabulum,” with the latter three considered to be labral lesions. In the operative report, the location of the labral lesion was described with use of a clock-face method as described by Blankenbaker et al.\textsuperscript{19}

**Source of Funding**

There was no external funding source for the study.

**Statistical Methods**

Statistical analysis of interobserver reliability was performed with use of SPSS software (version 17.0; SPSS, Chicago, Illinois). The interobserver agreement for detecting acetabular labral lesions was determined by calculating the kappa coefficient ($\kappa$). The $\kappa$ value was interpreted as poor ($\kappa = 0$), slight ($\kappa = 0.01$ to 0.200), fair ($\kappa = 0.21$ to 0.40), moderate ($\kappa = 0.41$ to 0.60), substantial ($\kappa = 0.61$ to 0.80), or almost perfect ($\kappa = 0.81$ to 1.00).\textsuperscript{20,21}

The sensitivity, specificity, negative predictive value, positive predictive value, and likelihood ratios for a negative test and a positive test were calculated for the detection of labral lesions with magnetic resonance arthrography. Arthroscopic assessment of the labrum was used as the reference standard. Confidence intervals (CIs) were set at 95%.

**Results**

At arthroscopy, ninety-one labral lesions were identified in ninety-five hips (Fig. 2). The prevalence of labral lesions in the studied population was 96%.

**Interobserver Reliability**

The interobserver reliability of the two radiologists for detecting labral lesions using magnetic resonance arthrography was fair ($\kappa = 0.268$, CI = 0.023 to 0.513). The agreement between the observers is presented as a cross-tabulation in Table II.

### Table II Interobserver Agreement for Diagnosing Labral Lesions with Magnetic Resonance Arthrography

<table>
<thead>
<tr>
<th>Observer B</th>
<th>Labral Lesion Present</th>
<th>No Labral Lesion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Labral lesion present</td>
<td>70 (74%)</td>
<td>9 (9%)</td>
</tr>
<tr>
<td>No labral lesion</td>
<td>10 (11%)</td>
<td>6 (6%)</td>
</tr>
</tbody>
</table>

### Table III Comparison of Hip Arthroscopy Findings with Magnetic Resonance Arthrography Findings by Observers A and B

<table>
<thead>
<tr>
<th>Magnetic resonance arthrography</th>
<th>Hip Arthroscopy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observer A</td>
<td>Labral lesion present</td>
</tr>
<tr>
<td>Labral lesion present</td>
<td>13 (14%)</td>
</tr>
<tr>
<td>Observer B</td>
<td>Labral lesion present</td>
</tr>
<tr>
<td>Labral lesion present</td>
<td>13 (14%)</td>
</tr>
</tbody>
</table>

---

**Fig. 2**

**Fig. 2-a** Sagittal T1-weighted magnetic resonance arthrogram showing an intact labrum (arrow): a triangle-shaped labrum of homogeneous low signal intensity. **Fig. 2-b** Coronal T1-weighted magnetic resonance arthrogram showing a linear intralabral contrast fluid collection, indicating a labral tear (arrow).
Magnetic Resonance Arthrography Compared with Arthroscopy

The results of both observers when they used magnetic resonance arthrography to detect labral lesions, compared with the results of hip arthroscopy, are presented in Tables III and IV. The sensitivity and specificity of detecting labral lesions with magnetic resonance arthrography were 86% (CI = 79% to 93%) and 75% (CI = 33% to 100%) when the studies were interpreted by observer A and 86% (CI = 79% to 93%) and 50% (CI = 1% to 99%) when they were interpreted by observer B.

Location of Labral Lesions

Of the ninety-one labral lesions identified at arthroscopy, 89% (eighty-one) were anterior-superior, 67% (sixty-one) were posterior-superior, 13% (twelve) were anterior-inferior, and 11% (ten) were posterior-inferior. Table V presents the locations of the lesions identified by arthroscopy and the sensitivity of detecting lesions with magnetic resonance arthrography according to location for both observers. There were no significant differences in sensitivity among the locations of the labral lesions.

Discussion

To our knowledge, this is the first study to establish the interobserver reliability of detecting labral lesions with use of magnetic resonance arthrography. Furthermore, in this study magnetic resonance arthrography was compared with hip arthroscopy in a large series of patients (ninety-five hips) with a clinical suspicion of labral pathology. The interobserver reliability of detecting labral tears was fair. Magnetic resonance arthrography had a high positive predictive value and a low negative predictive value in the studied population.

Interobserver Reliability

The interobserver reliability of magnetic resonance arthrography for patients who subsequently underwent arthroscopy for suspected labral pathology was fair. The magnitude of \( \kappa \) needs to be considered in relation to the studied population; in a homogeneous study population, the value of \( \kappa \) will be lower than it will be in a heterogeneous population evaluated with the same measurement tool\(^{20,21} \). The population in this study can be considered homogeneous regarding the selection of patients with clinical suspicion of labral pathology and the high prevalence of labral lesions. In a population with a lower prevalence of labral lesions, \( \kappa \) may be higher.

Keeney et al.\(^{12} \) found variability among four observers using magnetic resonance arthrography to assess labral lesions; sensitivity ranged from 54% to 100% and specificity, from 50% to 67%. However, because these radiologists were not evaluating the same magnetic resonance arthograms, interobserver reliability could not be determined with the data.

The authors of one prior study reported excellent intraobserver reliability of detecting labral tears with use of magnetic

---

### TABLE IV Diagnostic Performance of Magnetic Resonance Arthrography for Diagnosing Labral Lesions with Arthroscopy as Standard Reference*

<table>
<thead>
<tr>
<th>Observer</th>
<th>Sensitivity (**)</th>
<th>Specificity (**)</th>
<th>Prevalence (%)</th>
<th>Negative Predictive Value (%)</th>
<th>Positive Predictive Value (%)</th>
<th>For Negative Test</th>
<th>For Positive Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>86 (79-93)</td>
<td>75 (33-100)</td>
<td>96</td>
<td>19 (0-38)</td>
<td>99 (96-100)</td>
<td>0.19 (0.09-0.41)</td>
<td>3.34 (0.63-18.76)</td>
</tr>
<tr>
<td>B</td>
<td>86 (79-93)</td>
<td>50 (1-99)</td>
<td>96</td>
<td>13 (4-31)</td>
<td>98 (94-100)</td>
<td>0.29 (0.09-0.86)</td>
<td>1.71 (0.64-4.58)</td>
</tr>
</tbody>
</table>

*The 95% confidence intervals are presented in parentheses.

---

### TABLE V Location of Labral Lesions Identified by Arthroscopy in Ninety-five Hips and Sensitivity of Detecting Labral Lesions with Magnetic Resonance Arthrography Distributed by Location

<table>
<thead>
<tr>
<th>Location (Quadrant)</th>
<th>No. of Lesions Seen on Arthroscopy</th>
<th>Sensitivity* (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Observer A</td>
<td>Observer B</td>
</tr>
<tr>
<td>Anterior-superior</td>
<td>81</td>
<td>88 (80-95)</td>
</tr>
<tr>
<td>Posterior-superior</td>
<td>61</td>
<td>92 (85-99)</td>
</tr>
<tr>
<td>Anterior-inferior</td>
<td>12</td>
<td>83 (62-100)</td>
</tr>
<tr>
<td>Posterior-inferior</td>
<td>10</td>
<td>70 (42-98)</td>
</tr>
</tbody>
</table>

*The 95% confidence intervals are presented in parentheses.
resonance arthrography ($\kappa = 0.96$) in twenty-four hips. However, a $\kappa$ cannot be calculated with use of the formula for $\kappa^2$ on the basis of the data reported in that study. As it is unclear which data were used for the calculation of $\kappa$ in that study, this value for intraobserver reliability cannot be adequately interpreted.

**Magnetic Resonance Arthrography Compared with Arthroscopy**

The validity of detecting acetabular labral lesions with use of magnetic resonance arthrography was determined by comparison with arthroscopy as the gold standard. Comparison with arthroscopy resulted in a sensitivity of 86% and 86% and a specificity of 75% and 50% for observers A and B, respectively. Whereas a high positive predictive value was noted for both observers (99% and 98%), the negative predictive values were only 19% and 13%. The low number of patients without a labral lesion in the studied population limits the accuracy of the predictive values. However, considering the 95% confidence intervals of the negative predictive values (0% to 38% for observer A and 4% to 31% for observer B), they still indicate a substantial number of false-negative magnetic resonance arthograms for patients selected for arthroscopy for suspected labral lesions. Therefore, it can be concluded that magnetic resonance arthrography is not an appropriate diagnostic instrument for ruling out labral lesions in patients with a high clinical suspicion of labral pathology but can be used to confirm a suspected labral lesion prior to surgery.

We could not accurately determine the specificity of magnetic resonance arthrography for detecting acetabular labral lesions because of the small number of patients without a labral tear. This is confirmed by the large width of the confidence intervals for both observers (Table IV).

The high prevalence of labral lesions in the studied population was due to the selection of patients for arthroscopy on the basis of the clinical evaluation. Despite the acceptable sensitivity of 86%, this resulted in a low negative predictive value. It should be emphasized that these results reflect the studied population, in which the subjects had a high clinical suspicion of labral pathology. Magnetic resonance arthrography may be more useful in the diagnostic workup for labral lesions in less symptomatic patients. Physicians should consider whether the outcome of magnetic resonance arthrography will alter the treatment strategy in less symptomatic patients.

A limitation of this study is that the radiologists assessing the lesions were biased by the information that they were evaluating the magnetic resonance arthograms of patients selected for hip arthroscopy. This risk of bias may have resulted in an increased reporting of magnetic resonance arthograms being positive for labral lesions and a reduced reporting of them being negative for the lesions. Consequently, the true sensitivity may be lower and the specificity may be higher than reported in this study.

Another limitation of the present study is that the surgeon was not blinded to the magnetic resonance arthrography report at the time of surgery. Although the indication for arthroscopic surgery was based on the symptoms indicated by the history and physical examination, the fact that arthroscopic evaluation was not fully segregated from the assessment of the magnetic resonance arthrography introduced a potential source of bias. The consequence of this potential bias is that the surgeon could have been influenced to assess the labrum according to the report, thereby increasing the agreement between the surgical and magnetic resonance arthrography findings. This could have caused an overestimation of the true-positive and true-negative outcomes, resulting in an overestimation of the sensitivity, specificity, negative predictive value, and positive predictive value. In light of this potential source of bias, the true value of magnetic resonance arthrography for detecting labral lesions may be less than suggested by the results of the present study. As a result of this potential source of bias, the conclusions of the present study addressing the limitations of magnetic resonance arthrography might even be an underestimation.

Magnetic resonance arthrography and surgery have been compared in several previous studies. The results of these studies are presented in the Appendix; variables were calculated with use of cross-tabulations performed with data reported in the studies. Sensitivity ranged from 60% to 100%, and specificity ranged from 0% to 100%. These studies had some methodological flaws, including relatively small sample sizes and retrospective designs. Retrospective studies comparing the findings of magnetic resonance arthrography with those of arthroscopic surgery, including the present investigation, are susceptible to bias in the selection of patients, as patients with positive findings on magnetic resonance arthrography are probably more likely to undergo surgery. As a result of the possibility of such selection bias, sensitivity may be overestimated and specificity underestimated. One study, by Chan et al., had a prospective study design. However, Chan et al. compared magnetic resonance arthrography with arthroscopic surgery in seventeen patients with a positive finding on magnetic resonance arthrography, resulting in a substantial selection bias. Patients with negative magnetic resonance arthograms were excluded from the analysis; therefore, there were no true-negative or false-negative results with which to correctly calculate sensitivity and specificity.

In accordance with the present study, the prevalence of labral lesions in previous series has been high (75% to 96%), resulting in low negative predictive values. The reported positive predictive values seem to be relatively high (77% to 100%). However, compared with the high probability in studied series, the probability of a labral lesion is only marginally increased with a positive magnetic resonance arthrogram. Consequently, a negative magnetic resonance arthrogram will not rule out a labral lesion and a positive magnetic resonance arthrogram will only marginally increase an already high probability of a labral lesion.

**Location of Labral Lesions**

In accordance with previous published studies, the majority of lesions were located in the anterior-superior and posterior-superior sites of the labrum in our study. Magnetic resonance
arthrography may be less effective in identifying posterior and lateral tears. However, the results of the present study do not support this notion, as the sensitivities for detecting labral tears were similar among the various locations (Table V).

**Conclusions**

This study showed that magnetic resonance arthrography has only fair reliability ($\kappa = 0.268$) for detecting acetabular labral lesions in patients with a clinical suspicion of labral pathology. Because of this limited reliability and the high prevalence of labral lesions, the complementary benefit of magnetic resonance arthrography for the detection of labral lesions in patients with a high clinical suspicion of labral pathology is limited. Magnetic resonance arthrography has a poor negative predictive value and cannot be used to rule out a labral tear when there is a high clinical suspicion of such a tear. Physicians should critically consider whether the findings of a magnetic resonance arthrogram will alter the treatment strategy for an individual patient suspected of having labral pathology on clinical examination.

**Appendix**

A table showing a comparison of magnetic resonance arthrography with arthroscopic surgery for diagnosing labral lesions in previously published studies is available with the online version of this article as a data supplement at jbjs.org.

**References**


