Ultrasound Detection of Rotator Cuff Tears: Observer Agreement Related to Increasing Experience

Matthieu J. C. M. Rutten1
Gerrit J. Jager1
Lambertus A. L. M. Kiemeney2

OBJECTIVE. Ultrasound of the rotator cuff is considered to be operator-dependent with its accuracy being related to the operator’s level of experience. This study was conducted to test the hypothesis that ultrasound performed by operators with different levels of experience will give nonreproducible results.

SUBJECTS AND METHODS. Two radiologists, one general radiologist with no experience in musculoskeletal ultrasound and one experienced musculoskeletal radiologist, independently performed ultrasound on 200 shoulders in 183 consecutive patients. Agreement was assessed. Cohen’s kappa values with standard errors were calculated. In 71 patients the ultrasound findings could be related to surgical findings.

RESULTS. The diagnoses of full-thickness and partial-thickness rotator cuff tears were made with agreements of 98% (κ-value: 0.95 [standard error, 0.03]) and 90% (κ-value: 0.79 [0.05]), respectively. Agreement for full-thickness tears was constant; the agreement for partial-thickness tears improved from 80% to 98% in the last quarter of the study period. Based on the 71 patients who underwent shoulder surgery, sensitivity, specificity, and accuracy for detecting full-thickness tears by the experienced and general radiologists were 94%, 94%, and 94% and 89%, 91%, and 90%, respectively. Sensitivity, specificity, and accuracy for detecting partial-thickness tears by the experienced and general radiologists were 100%, 32%, and 57% and 84%, 35%, and 53%, respectively.

CONCLUSION. The hypothesis that ultrasound of the shoulder is operator-dependent and related to experience was refuted. In this study, there was excellent agreement for the detection of rotator cuff tears, which only slightly improved with the increasing experience of the general radiologist. Accuracy of rotator cuff tear detection was high and in accordance with the results in the literature.
a general radiologist with little musculoskeletal experience in an extensive series of 200 shoulders of 183 consecutive patients.

**Subjects and Methods**

**Patients**
During a time period of approximately 1 year, two radiologists prospectively performed 200 ultrasound examinations of one or both shoulders of 183 consecutive patients with acute or long-standing shoulder pain or disability.

Patients were referred to the radiology department of a tertiary teaching hospital by orthopedic surgeons and general practitioners. Our study population comprised 89 men and 94 women whose ages ranged from 14.1 to 87.2 years, with a mean age of 49.9 years. Informed consent was obtained from all patients before the ultrasound examinations.

Two radiologists, one with more than 15 years of experience and one general radiologist with no shoulder ultrasound experience but with more than 15 years of experience in abdominal ultrasound, performed the examinations during the same patient visit. At the start of the study, the general radiologist acquired knowledge of shoulder anatomy and received instructions from the experienced radiologist on the use of a standard ultrasound scanning technique [4] and about the diagnostic ultrasound criteria for partial-thickness [8, 9] and full-thickness rotator cuff tears [9]. Both radiologists were blinded to each other’s images and were unaware of each other’s interpretations.

**Ultrasound**
Sonographic examinations were performed with an APLIO (Toshiba Medical Systems) using a 7.5–14 MHz broadband linear-array transducer (PLF-805ST, Toshiba Medical Systems). Both radiologists used a similar scanning protocol [4]. The deltoid muscle, subacromial-subdeltoid bursa, long head of the biceps tendon, and entire rotator cuff were examined, with special emphasis on the integrity of the subscapularis tendon (SSC), supraspinatus tendon (SSP), infraspinatus tendon (ISP), and teres minor tendon (TM). The tendons were scanned along their long and short axes. Both radiologists used the same ultrasound imaging criteria, which were derived from the literature, for partial-thickness [8, 9] and full-thickness rotator cuff tears [9]. The rotator cuff was evaluated for primary abnormalities (i.e., ultrasound criteria) of full-thickness rotator cuff tears, such as tendon nonvisualization, full-thickness rotator cuff discontinuity (i.e., focal hypoechoic zone or mixed hyper- and hypoechoic zone [defect or cleft] extending through the entire substance of the rotator cuff), and focal thinning or substance loss of the tendon with visible margins of the tear.

The ultrasound criteria used for the diagnosis of partial-thickness rotator cuff tears were focal partial-thickness rotator cuff discontinuity (i.e., a focal hypoechoic zone or mixed hyper- and hypoechoic zone) involving the articular or bursal side or located within the tendon, and focal thinning (flat-

**TABLE 1: Cross-Tabulations of Ultrasound-Based Judgments of Two Radiologists of the Presence of Rotator Cuff Tears in 200 Shoulders**

<table>
<thead>
<tr>
<th>General Radiologist</th>
<th>Experienced Radiologist</th>
<th>RC (Any Tendon)</th>
<th>Total</th>
<th>Agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>N</td>
<td>40</td>
<td>0</td>
<td>Chance corrected % agreement (κ): 91.0</td>
</tr>
<tr>
<td>P</td>
<td>P</td>
<td>3</td>
<td>2</td>
<td>Overall % agreement any type of tear: 93.0</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>0</td>
<td>49</td>
<td>200</td>
</tr>
<tr>
<td>SSC</td>
<td>N</td>
<td>109</td>
<td>0</td>
<td>Overall % agreement: 92.5</td>
</tr>
<tr>
<td>P</td>
<td>P</td>
<td>3</td>
<td>0</td>
<td>Overall % agreement any type of tear: 93.0</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>1</td>
<td>14</td>
<td>200</td>
</tr>
<tr>
<td>SSP</td>
<td>N</td>
<td>47</td>
<td>0</td>
<td>Overall % agreement: 91.5</td>
</tr>
<tr>
<td>P</td>
<td>P</td>
<td>6</td>
<td>2</td>
<td>Overall % agreement any type of tear: 93.5</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>0</td>
<td>48</td>
<td>200</td>
</tr>
<tr>
<td>ISP/TM</td>
<td>N</td>
<td>188</td>
<td>0</td>
<td>Overall % agreement: 99.0</td>
</tr>
<tr>
<td>P</td>
<td>P</td>
<td>0</td>
<td>1</td>
<td>Overall % agreement any type of tear: 99.5</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>0</td>
<td>5</td>
<td>200</td>
</tr>
</tbody>
</table>

tening) of the rotator cuff or loss of convexity of the outer border (bursal surface) of the rotator cuff.

The ultrasound images were also evaluated for secondary findings for rotator cuff tears, including the cartilage interface sign (a thin, markedly hyperechoic line at the surface of the hypoechoic, hyaline articular cartilage of the humeral head), cortical pitting or irregularity of the greater tuberosity, and fluid located intraarticularly or in the subacromial-subsutdeltoideus bursa.

**Examination Analysis**

After the ultrasound examinations, each radiologist filled out a data sheet. Each rotator cuff constituent was graded as normal or torn, and a torn rotator cuff tendon was graded as partial-thickness or full-thickness.

**Statistical Analysis**

To evaluate observer agreement, a kappa value of agreement (normal, partial-thickness, or full-thickness rotator cuff tear) was calculated for the entire rotator cuff and for the cuff constituents (SSC, SSP, ISP together with TM) separately. The agreement was calculated for 200 shoulders and for four consecutive series of 50 shoulders each. The kappa-coefficient was interpreted as follows: 0.00–0.20, poor agreement; 0.21–0.40, fair agreement; 0.41–0.60, moderate agreement; 0.61–0.80, good agreement; and 0.81–1.00, excellent agreement.

Sensitivity, specificity, accuracy, and positive and negative predictive values for diagnosing partial-thickness and full-thickness rotator cuff tears by sonography were calculated with surgical findings as reference. To evaluate observer agreement, a kappa coefficient was calculated for 200 shoulders and for the following examination groups.

**Interobserver Agreement**

Table 1 compares the general radiologist’s interpretive results for the sonographic detection of partial- and full-thickness rotator cuff tears in 200 consecutively examined shoulders against the findings by the experienced radiologist for the entire rotator cuff and for the three rotator cuff constituents (SSC, SSP, ISP/TM) separately.

Taking into account an agreement in 200 ultrasound examinations with the experienced radiologist, the overall kappa for the detection of partial- and full-thickness tears was 0.95 [standard error, 0.03] (98%) and for partial-thickness tears 0.79 [0.05] (90%) (Table 2).

**Full-Thickness Rotator Cuff Tears**

The overall agreement between the general and experienced radiologists for the detection of full-thickness tears during the first 50 examinations was good (κ = 0.81 [0.11]) and excellent (κ = 0.96 [0.05]) – 1.00 [0.00] thereafter (Table 2).

After having performed 100 ultrasound examinations, the general radiologist showed excellent agreement for the detection of full-thickness tears.

**Partial-Thickness Rotator Cuff Tears**

The overall agreement between the general and experienced radiologist for the detection of partial-thickness tears was moderate (κ = 0.60 [0.11]) during the first 50 examinations (Table 2) and varied from moderate (κ = 0.68 [0.12]) to excellent (κ = 0.92 [0.08]) in the following examination groups.

**Sensitivity, Specificity, and Accuracy Related to Surgery**

Surgical findings for 71 shoulders were available as a reference standard (Table 3). Table 4 lists the sensitivities, specificities, and accuracies for the ultrasound detection of partial-thickness and full-thickness tears by the experienced and general radiologist, respectively, with surgical findings as reference. To be able to calculate these diagnostic parameters, the finding of a partial-thickness tear at surgery or ultrasound was arbitrarily ignored (i.e., considered with “no tear” as one group) when evaluating the diagnosis of full-thickness torn and vice versa. The experienced radiologist showed a high degree of sensitivity, specificity, and accuracy (94%, 94%, and 94%, respectively) for the detection of a full-thickness rotator cuff tear and acceptable but lower values for the detection of partial-thickness tears (100%, 32%, and 57%).

The general radiologist showed sensitivity, specificity, and accuracy for the detection of full- and partial-thickness rotator cuff tears of 89%, 91%, and 90% and 84%, 35%, and 53%, respectively.

**Discussion**

Ultrasound has become an important imaging technique in the evaluation of suspected rotator cuff tears. Many studies have shown the good diagnostic performance of ultrasound in the detection of partial- and full-thickness tears [1, 2, 10]. In the current study, sensitivity, specificity, and accuracy could be determined in 71 patients, with surgery as the standard of reference. For the detection of full-thickness tears, both radiologists performed within the range of prior studies [1, 2, 10]. For the detection of partial-thickness tears, both radiologists showed good sensitivities, yet relatively low specificities and accuracies (Table 4) due to a high number of false-positive findings. The sensitivities for both full- and partial-thickness tears are higher than in the literature, but the specificities for partial tears are quite low for both examiners. For example, in the recent meta-analysis by de Jesus et al. [10], the specificities for partial-thickness tears on ultrasound were in the 90s, with sensitivities only in the 60s. This can be attributed to a difference in reading styles and overcalling partial tears, because accuracies of only 57% and 53% respectively are quite low. In this study, however, a considerable number of false-positive findings may be true-positive findings given surgical limitations on detecting partial-thickness rotator cuff tears in an articular or intratendinous location [11, 12] (Fig. 1).

Many of the 71 patients underwent acromioplasty or bursectomy, or both procedures, in the management of subacromial impingement syndrome of the shoulder. Because surgical findings were analyzed in this study retrospectively, the orthopedic surgeons often only inspected the bursal side and not the articular side of the rotator cuff during open surgery or arthroscopy. To rule out full-thickness rotator cuff tears, they frequently inject a blue-colored liquid into the glenohumeral joint. By doing this they may miss partial-thickness rotator cuff tears located articularly and intratendinously. In this study, this is probably the most important reason the specificities of both readers are low, the more so because both readers show a relatively high agreement (κ = 0.79 [90%]) for the ultrasound detection of partial-thickness tears.

**TABLE 2: Kappa Values and Percentage Agreement Between Two Radiologists for the Detection of Rotator Cuff Tears in Four Consecutive Groups of 50 Shoulders Each**

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>Kappa Values (Agreement) in 200 Ultrasound Examinations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1–50</td>
</tr>
<tr>
<td>Partial-thickness tears</td>
<td>0.60 (80)</td>
</tr>
<tr>
<td>Full-thickness tears</td>
<td>0.81 (94)</td>
</tr>
</tbody>
</table>

Note—Percentage agreement in parentheses.
Learning Curve
To our knowledge, there are no previous reports that document the agreement with increasing experience (i.e., the learning curve) for the detection of rotator cuff tears with ultrasound. In the current study, the general radiologist, with no experience in musculoskeletal ultrasound, was able to detect full-thickness tears at a level comparable to an experienced musculoskeletal radiologist after brief instruction. After 50 examinations the general radiologist was able to detect partial-thickness tears at a level almost equal to that of the experienced radiologist. The probable reason for this longer learning curve is that differentiation between hypoechogenic tendinosis and partial-thickness tears can be difficult and requires more experience [3]. To assess this learning curve we examined a large number (200) of shoulders in a relatively short time period. The results were consistent and improved with time.

Interobserver Agreement
In the literature, shoulder ultrasound has frequently been considered an operator-dependent imaging technique. Yet, only three studies evaluated the interobserver agreement of ultrasound in the detection of rotator cuff tears [5–7]. These studies show that the interobserver agreement in the sonographic detection and characterization of rotator cuff tears is excellent for experienced operators/examiners (κ = 0.82–1.0) [5, 6]. In the case of an inexperienced reader, the agreement with an experienced reader is poor (κ = 0.18–0.22) [6]. However, if the inexperienced reader, as in our study, is a general radiologist with abdominal ultrasound experience, the agreement with the experienced radiologist appears to be good to excellent for the detection of full-thickness tears (κ = 0.90) and moderate for partial-thickness tears (κ = 0.63) and intratendinous tears (κ = 0.57) [7]. The number of examined patients in these other studies was relatively small, varying from 24 [6] to 35 [5] to 65 patients [7], and in two [5, 7] of the three studies the a priori chance for having a rotator cuff tear was relatively high because only patients with a high clinical suspicion for having a rotator cuff tear were included. Because the level of agreement is related to the prevalence of disease, we performed a study with a lower a priori chance for having a rotator cuff tear was relatively high because only patients with a high clinical suspicion for having a rotator cuff tear were included. Because the level of agreement is related to the prevalence of disease, we performed a study with a lower a priori chance for having a rotator cuff tear. Nevertheless, the interobserver agreement for the detection of partial- and full-thickness tears was comparably high. The two observers agreed on the classification of the cuff status in 93% of patients. The observers disagreed on only 18 of the 200 shoulders. There were no cases in which one observer diagnosed a full-thickness tear and the other diagnosed a normal rotator cuff. There were 14 cases in which the difference in categorization of the rotator cuff status was between a normal cuff or tendinosis and a partial-thickness rotator cuff tear (Fig. 2). In only four discrepant cases was the difference in categorization between a partial-thickness tear and a full-thickness rotator cuff tear (Fig. 3). The distinction between tendinosis and partial-thickness tear may not be clinically relevant given the size of the abnormality. Therefore, even when there was a disagreement, the discrepancy was relatively limited.

Interobserver agreement studies with MRI between experienced and inexperienced readers showed (Table 2) slightly worse results for the detection of full-thickness tears (κ = 0.67–

### TABLE 3: Cross-Tabulations of the Ultrasound Findings of Both Radiologists Related to Surgical Findings in 71 Patients

<table>
<thead>
<tr>
<th>Reference</th>
<th>Finding</th>
<th>Experienced Radiologist</th>
<th>General Radiologist</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surgery</td>
<td>FTT</td>
<td>Sensitivity</td>
<td>Specificity</td>
</tr>
<tr>
<td>PTT</td>
<td>100</td>
<td>32</td>
<td>57</td>
</tr>
<tr>
<td>Overall</td>
<td>65</td>
<td>50</td>
<td>0.50 (0.07)</td>
</tr>
</tbody>
</table>

Note—SE = standard error, PPV = positive predictive value, NPV = negative predictive value, N = Normal, P = partial-thickness rotator cuff tear, F = full-thickness rotator cuff tear.

---

### TABLE 4: Sensitivity, Specificity, Accuracy, and Positive and Negative Predictive Values for the Detection of Partial- and Full-Thickness Rotator Cuff Tears of the Experienced and General Radiologists With Surgical Findings as Standard of Reference

<table>
<thead>
<tr>
<th>Reference</th>
<th>Finding</th>
<th>Experienced Radiologist</th>
<th>General Radiologist</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surgery</td>
<td>FTT</td>
<td>Sensitivity</td>
<td>Specificity</td>
</tr>
<tr>
<td>PTT</td>
<td>84</td>
<td>35</td>
<td>53</td>
</tr>
<tr>
<td>Overall</td>
<td>59</td>
<td>0.41 (0.08)</td>
<td></td>
</tr>
</tbody>
</table>

Note—PPV = positive predictive value, NPV = negative predictive value, PTT = partial-thickness rotator cuff tear, FTT = full-thickness rotator cuff tear, Overall = overall evaluation: normal, PTT, and FTT. Values are percentages. Kappa values = measurement of agreement (κ) with the standard error in parentheses.
0.84) relative to ultrasound and significantly worse interobserver agreement results for the detection of partial-thickness tears \( \kappa = 0.13–0.44 \) [13, 14]. In these MRI studies, results of a single MRI technique were interpreted. If the same patient were to be scanned with different scanning parameters (field strength, sequences, scan direction, slice thickness, and so on) the agreement might be even less. To the best of our knowledge, the operator dependency in obtaining scan planes in shoulder MRI has not been evaluated until now. In our opinion, in concordance with ultrasound, the performance and interpretation of shoulder MRI should also be considered an operator-dependent imaging technique.

The strength of this study is that, as advocated by Middleton et al [5], two independent examiners performed two completely separate examinations and interpreted the results on the basis of real-time examinations and their own set of stored images.

This study does have some limitations. First, it lacks an independent reference standard in all patients. However, the focus of this study is not to reconfirm accuracy but to evaluate interobserver agreement, which in our opinion justifies the inclusion of patients who lacked surgical confirmation. Second, although in a number of patients surgery was performed because of the ultrasound results, a considerable number of patients were operated on regardless of the ultrasound findings, because of their complaints related to the subacromial impingement syndrome. In this group, especially, the cuff was not always inspected from both the bursal and the articular sides, which leads to an imperfect standard of reference bias. Third, the number of observers is
Ultrasound of Rotator Cuff Tears

56-year-old man with shoulder pain and disability after trauma. A, Long-axis sonogram of supraspinatus tendon (SSP) shows hypoechoic area in insertion (arrows), which was interpreted by inexperienced radiologist as tendinosis with partial-thickness rotator cuff tear. Experienced radiologist read this as full-thickness tear of SSP.

B, Coronal T1-weighted, fast spin-echo image with fat saturation of MR arthrogram (1.5 T) (TR/TE, 700/16; field of view: 160–160 mm, slice: 3 mm, flip angle: 90°, matrix: 512 × 512) confirmed surgical findings of full-thickness rotator cuff tear (arrows) by showing leakage of contrast media from glenohumeral joint through SSP to subacromial-subdeltoid bursa (arrowheads). H = Humeral head.

References


