Rotator cuff disease is the most common cause of shoulder symptoms, accounting for 65% to 70% of cases and leading to approximately 4.5 million physician visits per year in the United States. By 70 years of age, more than 50% of the general population has a full- or partial-thickness rotator cuff tear, regardless of the presence of symptoms. Rotator cuff tears are a significant cause of morbidity and a financial burden to the health care system. These tears require continuous reassessment of diagnostic workup and management.

Both accurate diagnosis and determination of the extent of the rotator cuff tear are critical for treatment management decisions and preoperative planning. Clinical examination alone provides limited information, making supplementary imaging an essential component of decision making. The choice of nonoperative modalities and surgical management, including the specific procedure performed, depends on the size of the lesion, retraction status, and pathologic changes determined by imaging. Previous studies recognized that muscle atrophy and fatty infiltration are important factors in the evaluation of rotator cuff tears. Ultrasound and magnetic resonance imaging (MRI) are both capable of diagnosing full-thickness rotator cuff tears. However, it is unknown which imaging modality is more accurate and precise in evaluating the characteristics of full-thickness rotator cuff tears in a surgical population. This study reviewed 114 patients who underwent arthroscopic repair of a full-thickness rotator cuff tear over a 1-year period. Of these patients, 61 had both preoperative MRI and ultrasound for review. Three musculoskeletal radiologists evaluated each ultrasound and MRI in a randomized and blinded fashion on 2 separate occasions. Tear size, retraction status, muscle atrophy, and fatty infiltration were analyzed and compared between the 2 modalities. Ultrasound measurements were statistically smaller in both tear size ($P = .001$) and retraction status ($P = .001$) compared with MRI. The 2 image modalities showed comparable intraobserver reliability in assessment of tear size and retraction status. However, MRI showed greater interobserver reliability in assessment of tear size, retraction status, and atrophy. Independent observers are more likely to agree on measurements of the characteristics of rotator cuff tears when using MRI compared with ultrasound. As tear size increases, the 2 image modalities show greater differences in measurement of tear size and retraction status. Additionally, compared with MRI, ultrasound shows consistently low reliability in detecting subtle, but clinically important, degeneration of the soft tissue envelope. Although it is inexpensive and convenient, ultrasound may be best used to identify a tear, and MRI is superior for use in surgical planning for larger tears.

The authors are from the Department of Orthopaedic Surgery, Henry Ford Hospital, Detroit, Michigan.

The authors have no relevant financial relationships to disclose.

Correspondence should be addressed to: Kelechi R. Okoroha, MD, Department of Orthopaedic Surgery, Henry Ford Hospital, 2799 W Grand Blvd (CFP-6), Detroit, MI 48202 (Krokoroha@gmail.com).

Received: April 20, 2016; Accepted: August 30, 2016.

doi: 10.3928/01477447-20161013-04
irreversible and often correlate with poor outcomes after rotator cuff tear repair.5-15

There are 2 primary imaging modalities that are used to visualize the rotator cuff: ultrasonography and magnetic resonance imaging (MRI). Historically, MRI was the method of choice, given its proven accuracy and high sensitivity in diagnosing rotator cuff pathology.16-19 The use of MRI is appealing because it is less dependent on operator factors, it provides images that are inherently easier for orthopedic surgeons to review, and it is better for the evaluation of morphologic changes in structures such as the glenoid labrum, joint capsule, articular cartilage, and surrounding muscles and bone that may contribute to symptoms. Ultrasound is among the most operator-dependent imaging methods, and its reliability and accuracy correlate directly with user experience.20,21 Because of improvements over the past 10 years in transducer strength, soft tissue penetration, and user experience, ultrasound has become a more reliable diagnostic tool for suspected rotator cuff tear. Further, unlike MRI, ultrasound is inexpensive, readily available, and well tolerated, and it offers real-time results and dynamic visualization.22 The growing interest in ultrasound vs MRI for the workup of rotator cuff tear led to a number of studies and meta-analyses comparing the 2 modalities, and these studies showed comparable results in diagnostic accuracy, sensitivity, and specificity.21,23-34 There is no general consensus on which is the preferred test, and no previous studies directly compared intra- and interobserver agreement between MRI and ultrasound in determining specific tear size and pathologic characteristics ofrotator cuff tears. In addition, no study has looked at a surgical population that included only full-thickness tears.

The current study compared the characterization of rotator cuff tears by ultrasound and MRI in terms of size, muscle atrophy, and fatty infiltration in surgical patients. In addition, the study evaluated the intra- and interobserver reliability of ultrasound and MRI when making these measurements.

**Materials and Methods**

**Study Design and Subjects**

The authors performed a retrospective review of 114 consecutive patients who underwent arthroscopic repair of a full-thickness rotator cuff tear by a single surgeon at the study institution between February 2010 and February 2011. The study included 61 patients who had both MRI and ultrasound within 16 weeks of each other before surgery. Both MRI and ultrasound studies were reviewed individually twice by 2 senior-level musculoskeletal radiologists, each with more than 20 years of experience, and twice by a musculoskeletal fellow. The study was reviewed and approved by the institutional review board.

**Magnetic Resonance Imaging**

Examinations were performed with a 1.5 Tesla magnet MRI as a baseline. The protocol for each scanner was consistent. The protocols used were coronal proton density (field of view, 16-17 cm; matrix, 512x384; repetition time/echo time range, 1500-3000/20-30); sagittal proton density (field of view, 16-17 cm; matrix range, 256-512x224-384; repetition time/echo time range, 1500-3000/20-30); axial T2 fat suppressed (field of view, 16-19 cm; matrix range, 256-304x224-235; repetition time/echo time range, 2685-2839/60-100); and coronal T2 fat suppressed (field of view, 16-18 cm; matrix range, 256x192-224; repetition time/echo time range, 2800-3643/60-75).

Sizes of the full-thickness rotator cuff tears were measured in the greatest anteroposterior dimension (width) and in length/degree of retraction. The occupational ratio of each tear was calculated to determine muscle atrophy. This calculation was based on sagittal proton density images, according to the method of Thomazeau et al.35 The cross-sectional surface area of the supraspinatus was divided by the cross-sectional area of the supraspinatus fossa on the Y-view, which is formed by the coracoid process, distal clavicle, and scapular spine. Fatty infiltration of the rotator cuff muscles was graded based on the Goutallier classification with sagittal proton density images.36 According to this classification system, grade 0=no fat, grade 1=trace fatty streaks, grade 2=less than 50% fat, grade 3=50% fat, and grade 4=more than 50% fat.36,37

**Ultrasound**

Ultrasound was performed with a Logiq E9 ultrasound machine (GE Healthcare, San Jose, California), with 9-12 MHz transducers. Shoulder ultrasound examinations were performed in a standardized fashion, with the subject in a seated position. Evaluation included the rotator cuff musculature tendons, acromioclavicular joint, long head biceps tendon, posterior labrum, and spinoglenoid notch. The supraspinatus was evaluated in both Crass and modified Crass positions. Dynamic imaging was performed for abduction/adduction of the arm in the coronal plane to evaluate for subacromial impingement and internal/external rotation of the arm at the level of the subscapularis for evaluation of subcoracoid impingement. Of the 61 patients, 52 had imaging of the supraspinatus muscle. For 9 patients, imaging of the supraspinatus muscle was not performed because these patients were examined before the institution included imaging of the muscle itself as part of the standard shoulder protocol.

Determination of the degree of muscle atrophy and evaluation of fatty infiltration were based on images of the supraspinatus within the fossa. To determine the degree of muscle atrophy, the Y-view of the MRI was re-created and the cross-sectional surface area of the supraspinatus muscle was divided by that of the fossa. Fatty infiltration was determined by looking at echogenicity and the echostructure/pennate pattern, as in previous studies by Khoury et al.38 and Goutallier et al.39 Echogenicity was graded as isoechoic, mildly hyperechoic, or markedly hyperechoic. The echostructure was graded as normal (homogeneously distrib-
uted, well-defined hyperechoic streaks corresponding to fibromuscular septa), effaced (slight loss of the pennate pattern, with blurring of margins of the hyperechoic streaks), or absent (loss of the pennate pattern, with very poor or no visibility of the streaks). In addition, the greatest anteroposterior dimension (width) and the length and degree of proximal tendon retraction of the full-thickness rotator cuff tear were measured.

**Statistical Analysis**

A biostatistician performed all statistical analyses with SPSS version 22 statistical software (IBM Corp, Armonk, New York). Analysis of inter- and intrarater reliability was performed with an intraclass correlation coefficient (ICC) for the continuous variables of length, width, and muscle atrophy. These results are shown as mean and 95% confidence interval (CI). For the ordinal variable of fatty infiltration, analysis of inter- and intrarater reliability was performed with the weighted \( \kappa \) coefficient. Similar statistics were used to assess the agreement between MRI and ultrasound methods. The \( \kappa \) statistics were interpreted based on guidelines proposed by Landis and Koch,\(^{39} \) with values less than 0 defined as poor or no agreement, values of 0 to 0.20 defined as slight agreement, values of 0.21 to 0.40 defined as fair agreement, values of 0.41 to 0.60 defined as moderate agreement, values of 0.61 to 0.80 defined as substantial agreement, and values of 0.81 to 1.00 defined as almost perfect agreement. Bland-Altman plots were used to compare the difference between MRI and ultrasound with the mean of the 2 methods. Graphs showing the distribution of percent change (100%×[MRI-ultrasound]/MRI) for continuous variables were also created to visually assess the agreement between MRI and ultrasound. Additionally, a mixed-effects model with method as the fixed effect and reviewer and patient as the random effects was used to compare actual measurements of width and length with the 2 imaging modalities. These values are shown as mean±standard error. In all analyses, \( P<.05 \) was considered statistically significant.

**RESULTS**

**Intrarater Agreement**

For length and width, intrarater agreement was almost perfect (ICC, 0.82-0.93). For muscle atrophy and fatty infiltration measures, intrarater agreement ranged from fair to almost perfect (ICC, 0.42-0.92; \( \kappa=0.51-0.97 \)), respectively (Table 1).

**Interrater Agreement**

Interrater agreement with MRI was almost perfect for width (ICC, 0.87; 95% CI, 0.81-0.92) and length (ICC, 0.89; 95% CI, 0.82-0.94). Interrater agreement for muscle atrophy was substantial, and interrater agreement for fatty infiltration was only fair (\( \kappa=0.22 \)).

Interrater agreement was lower with ultrasound, but still substantial for width (ICC, 0.70; 95% CI, 0.57-0.80) and length (ICC, 0.70; 95% CI, 0.57-0.80). For muscle atrophy and fatty infiltration, interrater agreement was moderate (Table 2).

---

**Table 1**

| Intrarater Agreement* With 95% Confidence Interval for Each Measurement Across Trials, for Each Rater and Method |
|---|---|---|---|---|
| Rater No. | Method | Width (95% CI) | Length (95% CI) | Muscle Atrophy (95% CI) | Fatty Infiltration |
| 1 | US | 0.85 (0.76-0.91) | 0.89 (0.81-0.93) | 0.52 (0.30-0.70) | 0.78 |
| 2 | US | 0.82 (0.71-0.89) | 0.83 (0.73-0.89) | 0.42 (0.16-0.62) | 0.69 |
| 3 | US | 0.85 (0.77-0.91) | 0.90 (0.84-0.94) | 0.79 (0.63-0.88) | 0.76 |
| 1 | MRI | 0.91 (0.85-0.94) | 0.91 (0.86-0.95) | 0.92 (0.87-0.95) | 0.69 |
| 2 | MRI | 0.85 (0.77-0.91) | 0.90 (0.84-0.94) | 0.79 (0.63-0.88) | 0.76 |
| 3 | MRI | 0.88 (0.80-0.92) | 0.83 (0.71-0.90) | 0.90 (0.80-0.95) | 0.51 |

Abbreviations: MRI, magnetic resonance imaging; US, ultrasound.

*For width, length, and muscle atrophy, the agreement measure is intraclass correlation. For fatty infiltration, the agreement measure is the \( \kappa \) statistic.

**Table 2**

| Interrater Agreement* With 95% Confidence Interval for Each Measurement Across Raters, for Each Method |
|---|---|---|---|---|
| Method | Width (95% CI) | Length (95% CI) | Muscle Atrophy (95% CI) | Fatty Infiltration (95% CI) |
| US | 0.70 (0.57-0.80) | 0.70 (0.57-0.80) | 0.45 (0.00-0.72) | 0.41 |
| MRI | 0.87 (0.81-0.92) | 0.89 (0.82-0.94) | 0.67 (0.20-0.85) | 0.22 |

Abbreviations: MRI, magnetic resonance imaging; US, ultrasound.

*For width, length, and muscle atrophy, the agreement measure is intraclass correlation. For fatty infiltration, the agreement measure is the \( \kappa \) statistic for more than 2 raters.
Agreement Between Magnetic Resonance Imaging and Ultrasound

The agreement between MRI and ultrasound was moderate for width (ICC, 0.55; 95% CI, 0.43-0.64) and length (ICC, 0.59; 95% CI, 0.49-0.68). However, the agreement was slight for muscle atrophy (κ=0.04) and only fair for fatty infiltration (κ=0.31). When actual values for MRI and ultrasound for width, length, and muscle atrophy were compared, ultrasound measurements were significantly lower than MRI measurements for width and length and significantly higher for muscle atrophy (P<.001 for all, Table 3).

These findings were confirmed with Bland-Altman plots, which showed that length and width were evenly distributed around zero until the mean tear size approached 20 mm. For mean values greater than 20 mm, the distribution shifted upward, reflecting a tendency for MRI values to be greater than ultrasound values at these levels (Figures 1-2).

Figure 3 and Figure 4 show the distribution of percent change of ultrasound from MRI for width and length measurements. For width, approximately 18% of observations had ultrasound values within 10% of MRI values and 34% had ultrasound values within 20% of MRI values. Approximately 26% of observations had ultrasound values with a difference of greater than 50% (positive and negative) compared with MRI. The distribution of length was similar to the distribution of width. For length, 18% of observations had ultrasound values within 10% of MRI values, 36% had ultrasound values within 20% of MRI values, and 26% had ultrasound values showing a difference of greater than 50% compared with MRI values.

**DISCUSSION**

Ultrasound is the most operator-dependent imaging study for the shoulder, and it is becoming more popular as a first-line imaging modality for evaluating rotator cuff tears. Although several studies and meta-analyses have shown comparable accuracy in diagnosing both total and partial tears, ultrasound is often considered inferior to MRI for preoperative imaging because it provides less detail on morphologic changes in the cuff musculature. Increasing collaboration between musculoskeletal radiologists and orthopedic surgeons has emphasized the importance of retraction and muscle status in predicting success in rotator cuff surgery. Previous studies evaluated the reliability of ultrasound and MRI in the diagnosis of rotator cuff tears as well as the characterization of rotator cuff tears with MRI. To the authors’ knowledge, no studies have looked at the agreement of ultrasound and MRI in characterizing full-thickness rotator cuff tears. Ultrasound had lower interobserver reliability and decreased measurement of large rotator cuff tears.

Increasing collaboration between musculoskeletal radiologists and orthopedic surgeons has emphasized the importance of retraction and muscle status in predicting success in rotator cuff surgery. Previous studies evaluated the reliability of ultrasound and MRI in the diagnosis of rotator cuff tears as well as the characterization of rotator cuff tears with MRI. To the authors’ knowledge, no studies have looked at the agreement of ultrasound and MRI in characterizing full-thickness rotator cuff tears with regard to the specificity of tear size, muscle atrophy, and fatty infiltration.

Table 3

Comparison of Ultrasound and Magnetic Resonance Imaging for Tear Size (Length and Width) and Muscle Atrophy

<table>
<thead>
<tr>
<th>Variable</th>
<th>Ultrasound Mean±Standard Error</th>
<th>Magnetic Resonance Imaging Mean±Standard Error</th>
<th>P*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length</td>
<td>20.6±0.52</td>
<td>23.4±0.75</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Width</td>
<td>20.2±0.52</td>
<td>23.0±0.78</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Muscle atrophy</td>
<td>0.76±0.01</td>
<td>0.57±0.01</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

*P value from a mixed-effects model with method as a fixed effect and reviewer and patient as random effects.

The authors believe that the inclusion of only full-thickness tears increased accuracy because partial-thickness tears are more difficult to evaluate. Ultrasound showed less reliability, with substantial agreement when assessing tear width and retraction. When the 2 image modalities were compared directly, agreement was adequate when evaluating length and width, especially with tears smaller than 20 mm. However, agreement was poor with tears greater than 20 mm, in which ultrasound measurements were smaller. Sipola et al showed that both MRI and ultrasound underestimate rotator cuff tear size compared with surgical findings.

Slabaugh et al evaluated interobserver and intraobserver reliability in classifying fatty infiltration with MRI according to the Goutallier classification. Their study showed moderate intraobserver reliability, with κ=0.56 (95% CI, 0.53-0.60), and moderate interobserver reliability, with κ=0.43 (range, 0.16-0.74). The current study showed similar results, with fair to moderate intraobserver reliability. However, when evaluating interobserver reliability, this study found only slight agreement with MRI. Ultrasound showed comparable results, with moderate intraobserver reliability and fair interobserver reliability. Additionally, when they were compared directly, agreement between the 2 imaging modalities was poor to fair.
when evaluating fatty infiltration. These findings may be explained by a lack of standardized and evidence-based protocols to characterize rotator cuff tears with ultrasound.

Spencer et al\textsuperscript{51} and Lippe et al\textsuperscript{48} evaluated interobserver reliability for muscle atrophy with MRI; they both found fair agreement when assessing atrophy with MRI ($\kappa=0.25$ and $\kappa=0.25$, respectively). The current results showed increased interobserver reliability, with substantial agreement between observers. However, with ultrasound, reliability was moderate, and agreement between the 2 imaging methods was only slight. These findings can be explained by the increased visualization of the rotator cuff musculature on MRI compared with ultrasound. The current findings suggest that muscle atrophy is better analyzed with MRI.

**Limitations**

This study had a number of important limitations. In evaluating fatty infiltration, 9 patients did not undergo ultrasound imaging of the supraspinatus muscle. These differences were the result of changing protocols at the study institution and could not be controlled for by this retrospective study. The absence of this information in such a small subset of the patient population did not have an appreciable effect on the results. Because this study was performed at a large institution with several satellite hospitals, various MRI scanners were used. Most examinations were performed with a 1.5 Tesla MRI, but some patients were examined with a more detailed 3.0 Tesla MRI. Improved image quality has been reported with 3.0 Tesla MRI vs 1.5 Tesla MRI, and this difference may affect the depiction of detail.\textsuperscript{52} However, because slice thickness was equal, the authors do not believe that the higher Tesla magnet had a profound effect on detailing size, retraction, or atrophy status in the calculations.

Another limitation of the study was the lack of comparison of ultrasound and MRI values with surgical findings. Although basic intraoperative measurements were available for all patients, the approach to measuring size and retraction status was not standardized because this was a retrospective study. Information on the size and number of suture anchors was available from operative documentation, but the authors believed that this information would not accurately predict tear size. Although MRI has been considered highly accurate, there has been little evaluation of comparison of characterization of rotator cuff tears with MRI and ultrasound.
compared with surgical findings. This is an area for future study.

**Conclusion**

Independent observers are more likely to agree on measurements of the characteristics of rotator cuff tear with MRI compared with ultrasound. As tear size increases, the 2 modalities show greater differences in measurements of tear size and retraction status. Additionally, compared with MRI, ultrasound shows consistently low reliability in detecting subtle, but clinically important, degeneration of the soft tissue envelope. Although it is inexpensive and convenient, ultrasound may be the best modality for identifying tears, and MRI is superior in surgical planning for larger tears.

**References**

35. Thomazczuk H, Rolland Y, Lucas C, Duval JM, Langlais F. Atrophy of the supraspinatus belly: assessment by MRI in 55 patients with


