Diagnostic performance of ultrasound for rotator cuff tears: a systematic review and meta-analysis

Weiyu Liang*, Huaiyu Wu*, Fajin Dong, Hongtian Tian, Jinfeng Xu
* the authors shared the first authorship

Department of Ultrasound, First Affiliated Hospital of Southern University of Science and Technology, Second Clinical College of Jinan University, Shenzhen Medical Ultrasound Engineering Center, Shenzhen People’s Hospital, Shenzhen, China

Abstract

Aim: Rotator cuff (RC) tears are considered to be the main reason for shoulder pain. Although ultrasound is a useful method to detect it, its effectiveness when diagnosing RC tears has been a heated discussion. Therefore, we aimed at evaluating RC tears’ ultrasound accuracy by conducting a systematically review and pooled comprehensive analysis. Materials and methods: Relevant articles up to May 2018 were searched from the Cochrane Library, Embase, and Pubmed databases. Either arthroscopy or magnetic resonance imaging (MRI) was considered as a reference standard. The results were estimated by pooled-sensitivity (P-SEN), pooled-specificity (P-SPE), pooled-diagnostic odds ratio (P-DOR), pooled-likelihood ratio+ (P-LR+), pooled-likelihood ratio- (P-LR-), and the area under the summary receiver operating characteristic (SROC). Result: We selected seven prospective studies in accordance with the inclusion criteria that covered 554 rotator cuff tears in 868 patients. The P-SEN, P-SPE, P-LR+, P-LR-, P-DOR, area under the SROC curve of diagnostic performance of ultrasound for RC and post-test probability were 95% (95% CI: 88 – 98), 72% (95% CI: 61 – 81), 3.41 (95% CI: 2.40 – 4.84), 0.08 (95% CI: 0.03 – 0.16), 45.31 (95% CI: 21.04 – 97.59), 89% (95% CI: 0.86 – 0.91), and 46% and 2%, respectively. Conclusion: Our meta-analysis demonstrates that ultrasound has a high efficiency for RC tears’ diagnosis. It can be a promising method in patients with suspected RC tears because of its high sensitivity, specificity and diagnostic accuracy. Keywords: ultrasound; diagnosis; meta-analysis; systematic review; rotator cuff tears.

Introduction

In the human body the shoulder can be seen as the most flexible and vulnerable joint. The primary causes of pain in shoulder and dysfunction are rotator cuff (RC) avulsions [1,2]. More than 72% of all shoulder pain can be attributed to RC pathology [3,4]. Due to the difficulty of confirming the cause of pain in the shoulder by clinical examination [5], medical imaging tends to be used to verify the diagnosis [1,6]. It is hard to achieve differential shoulder pathologies’ diagnosis because many impairments, such as instability and shoulder impingement, could have similar symptoms [7,8]. Therefore, an accurate differential diagnosis is important for clinical decision making [9]. Particularly, the accurate RC tears’ size and location measurement is important for the efficiency of pre-operative planning [10].

Various clinical examinations, such as arthrography, magnetic resonance imaging (MRI) and ultrasound, have been used to evaluate RC [11]. Both arthrography and MRI are the reference standards for diagnosing RC tears...
and the diagnostic efficiency of ultrasound may be different. There are many published studies on the use of ultrasound to diagnose RC tears, which have been divided into either complete or partial tears, and the diagnostic efficiency of ultrasound may be different in each [21]. There has been one systematic review assessing the sensitivity of ultrasound in detecting shoulder injuries [9]: however, this study separately evaluated these two types of tears. At present, no research has been undertaken on both complete and partial RC tears. Therefore, the purpose of this meta-analysis was to evaluate both types of tears in order to decide the diagnostic potential of ultrasound for RC tears.

Materials and methods

Search strategy
All relevant studies from the Cochrane Library, Embase, and Pubmed databases were retrieved. Literature was retrieved using MeSH and free words and English publications limited it. The search strategies applied by the Cochrane Library, Embase, and Pubmed are as follows: (infraspinatus AND tear) or (supraspinatus AND tear) or (tears AND minor AND tear) or (subscapularis AND tear) or (shoulder AND joint AND abnormalities) and ultrasound.

Study selection
We included all studies that used ultrasound to evaluate rotator cuff tears, had arthroscopy or MRI as the diagnostic criterion and directly or indirectly provided primordial data that could be used to calculate negative likelihood ratios (LR−), positive likelihood ratio (LR+), specificity and sensitivity. Cadaveric and animal trials, single case reports, incomplete research, meta-analyses and systematic reviews, studies using a population of fewer than 20 patients and research which did not have a comparison of reference standard were excluded. Two authors with the same level of expertise and experience selected literature meeting the inclusion criteria after reading the full text. A third author reviewed the search results and the assessment process and solved any discordances.

Data extraction and quality assessment
After collecting all qualified studies, we extracted the following content: reference standard, number of rotator cuff tears for analysis, number of patients, study design, year of publication, the surname of the first author, TP, FN, FP, TN, sensitivity and specificity.

The QUADAS-2 (Quality Assessment of Diagnostic Accuracy Studies-2) tool [22] was used to evaluate the quality of the eligible studies. It was made up of four parts: patient selection, index test, reference standard, and flow and timing. Every part was evaluated for bias risk and assigned a score of unclear, high, low. The reference standard domains, patient selection, index test were measured for the presence of attention on the applicability, and were likewise given a score of low, high, or unclear. If a research is evaluated by “low” in all parts regarding applicability or bias, it is good to comprehensively judge the “low risk of bias” or “low concern regarding applicability” for the research. If a research study is evaluated by “high” or “unclear” in one or more parts, it can be judged “at risk of bias” or “concerns regarding applicability”.

Statistical analysis
We extracted the data from the included literature and recorded it using a four-fold table. We analyzed the data using RevMan 5.3 and Stata 12.0. RevMan 5.3 was applied in order to measure the eligible studies’ methodological quality. It used stata to draw statistical graphs and pool statistical indexes such as forest plots for sensitivity, specificity, diagnostic odds ratio (DOR) with corresponding 95% confidence intervals (CI), and the area under summary receiver operating characteristic curve (SROC). Heterogeneity of the included references was verified after statistical analysis. We used the inconsistency index (I²) to estimate the heterogeneity across the included studies. If the I² was <0.5, in the included studies, no heterogeneity was found. On the contrary, it indicated the heterogeneity’s result. If heterogeneity could not be omitted, we applied a random effects model, which evaluates the combined effect’s influence containing all information (sensitivity, specificity, accuracy, positive predictive value, negative predictive value, etc) [23]. The place beneath the ROC’s curve shows the diagnostic test’s accuracy, along with the lower accurate degree defined as an area of 0.5 - 0.7, certain accuracy as an area of 0.7 - 0.9, and higher accuracy with a place >0.9 [24].

Publication bias
The Deeks’ funnel plot asymmetry test was used to evaluate potential publication bias. It was applied by regressing diagnostic log odds ratio (lnDOR) versus inverse of the square root (1/sqrt) of the effective sample size (1/ESS1/2) and measured through effective sample size.
size. A p-value <0.10 for slope coefficients demonstrated the key asymmetry.

Results

Figure 1 shows the study selection process’s flow chart. On the basis of the previously determined strategy of search, we searched an overall of 2,356 articles. Among these, 2,040 were from Pubmed, 273 were from Embase, and 43 were from the Cochrane Library. There were 145 duplicate materials that were excluded using ENDNOTE X7. There were 2,198 studies excluded according to the title and abstract, including eight reviews systematically and/or meta-analysis and 166 case reviews or/and researches. For the remaining 13 articles, we excluded six of them after reading the full text because they could not provide four-fold data directly or indirectly. In the end, the meta-analysis included seven articles. The included studies’ features are listed in Table I.

According to the methodological assessment from the QUADAS-2 checklist written by RevMan 5.3, patient selection, index test, reference standard, and flow and timing indicated good quality. The results are detailed in figure 2. The selected studies’ quality was relatively high and four of them had QUADAS full scores [11,12,21,25].

![Fig 1. Flowchart of the study selection process](image)

![Fig 2. Quality assessment using the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) tool: A) Risk of bias and applicability concerns graph: authors’ ratings for each domain are presented as percentages across included studies; B) Risk of bias and applicability concerns summary: authors’ ratings for each domain for each included study.](image)

Table I. Characteristics and diagnostic performance of included studies

<table>
<thead>
<tr>
<th>Author (year)</th>
<th>Country</th>
<th>Number of patients (men)</th>
<th>Full thickness (partial thickness rotator cuff tears)</th>
<th>Reference</th>
<th>TP</th>
<th>FP</th>
<th>FN</th>
<th>TN</th>
<th>SE (%)</th>
<th>SP (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cole (2016) [12]</td>
<td>Australia</td>
<td>238 (unclear)</td>
<td>139 (77)</td>
<td>Arthroscopy</td>
<td>197</td>
<td>3</td>
<td>19</td>
<td>19</td>
<td>91</td>
<td>86</td>
</tr>
<tr>
<td>Co (2012) [21]</td>
<td>Canada</td>
<td>37 (22)</td>
<td>12 (10)</td>
<td>MRI</td>
<td>20</td>
<td>6</td>
<td>2</td>
<td>11</td>
<td>91</td>
<td>65</td>
</tr>
<tr>
<td>Singisetti (2011) [25]</td>
<td>UK</td>
<td>82 (56)</td>
<td>55 (6)</td>
<td>Arthroscopy</td>
<td>55</td>
<td>19</td>
<td>0</td>
<td>22</td>
<td>1</td>
<td>54</td>
</tr>
<tr>
<td>Jeyam (2008) [26]</td>
<td>UK</td>
<td>64 (unclear)</td>
<td>16 (6)</td>
<td>Arthroscopy</td>
<td>21</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>95</td>
<td>86</td>
</tr>
<tr>
<td>Ziegler (2004) [27]</td>
<td>USA</td>
<td>282 (173)</td>
<td>261 (1)</td>
<td>Arthroscopy</td>
<td>261</td>
<td>1</td>
<td>14</td>
<td>6</td>
<td>95</td>
<td>86</td>
</tr>
</tbody>
</table>

SE: Sensitivity, SP: Specificity, TP: True positive rate, FP: False positive rate, FN: False negative rate, TN: True negative rate
There are three articles in which the author did not mention the time interval during the ultrasound examination and the reference standard’s implication [19,26,27].

The random effects model was used to combine research data of eligible studies because the sensitivity and the LR− with 95% CI and forest plots (fig 3a,b) displayed obvious heterogeneity, together with an I2 of 79.73% and 73.80%. Figures 3, 4a,b shows that the pooled sensitivity, specificity, LR+, LR−, DOR, area under the SROC curve of diagnostic performance of ultrasound for rotator cuff, and post-test probability were 95% (95% CI: 88 – 98), 72% (95% CI: 61 – 81), 3.41 (95% CI: 2.40 – 4.84), 0.08 (95% CI: 0.03 – 0.16), 45.31 (95% CI: 21.04 – 97.59), 89% (95% CI: 0.86 – 0.91), and 46% and 2% respec-

---

**Fig 3.** Forest plot for a) sensitivity and specificity of ultrasound for evaluating rotator cuff tears; b) a positive and negative likelihood ratio of ultrasound for evaluating rotator cuff tears; c) diagnostic odds ratio of ultrasound for evaluating rotator cuff tears.

**Fig 4.** a) Fagan’s Nomogram with pretest probability at 20%; b) Hierarchical summary receiver operating characteristic curve (HSROC) with summary point, summary estimates, 95% confidence region and 95% prediction region for all included studies of ultrasound for evaluating rotator cuff tears (n = 7); c) Deeks’ funnel plot with superimposed regression line for identifying publication bias, Log Odds Ratio versus 1/sqrt (effective sample size) (Deeks) indicates that no significant bias was found. ESS = effective sample size.
sensitivity and LR- are of a high degree of heterogeneity and all, compliant bias, language, and publication are shown. which is likely to produce differences in compliance. In finally, the studies we included are from different regions, did not contain unpublished and non-English articles. Fi-

generate the bias of patient flow. In addition, our study reference standard's implementation [19], which could the time interval between the ultrasound exam and the efficacies. For example, Elmorsy et al did not mention-

specificity, which can better reflect the diagnostic value of diagnostic experiments [30].

The rotator cuff is an important structure that maintains the stability of the shoulder joint and is important in the abduction, external rotation and shoulder’s internal rotation. It is the only tendonous tissue between the two bones of the shoulder joint, and is easily damaged due to its particular anatomical position and repetitive movement [28]. Many methods can be used to examine the shoulder joint. Radiography and CT are widely used in the diagnosis of bone damage and its extent, but the diag-
nosis of soft tissue injury is inferior to ultrasound and MRI. Although MRI is currently the most reliable, noninvasive diagnostic method for rotator cuff tears evaluation, it is an expensive imaging technique, has low reproducibility and is not suitable for use in patients with metal-mounted devices [12].

In the recent years, with the development of ultrasound technology and the emergence of high-frequency probes, ultrasound was increasingly used for musculoskeletal systems pathology. Many studies have shown that there are not obvious differences between ultrasound and MRI when detecting rotator cuff tears [19,26,29]. However, though ultrasound is inexpensive, noninvasive and convenient, its effectiveness when diagnosing RC tears is still controversial. We found that the combined sensitivity, specificity, LR- and DOR of the included studies show higher heterogeneity ($I^2 >50\%$). Even so, ultrasound had a high sensitivity and credible specificity when diagnosing the rotator cuff tear. We also found a high LR+, DOR, certain accuracy and a lower LR--. The LR is a composite indicator reflecting sensitivity and specificity, which can better reflect the diagnostic value of diagnostic experiments [30].

Although we systematically evaluated the inclusion of literature using a rigorous process, there are still deficiencies. For example, Elmorsy et al did not mention the time interval between the ultrasound exam and the reference standard’s implementation [19], which could generate the bias of patient flow. In addition, our study did not contain unpublished and non-English articles. Fin-

ally, the studies we included are from different regions, which is likely to produce differences in compliance. In all, compliant bias, language, and publication are shown.

This meta-analysis demonstrates that combined sensitivity and LR- are of a high degree of heterogeneity and we have three explanations for this. First, a portion of the included studies was selected for patients with clinically, highly suspected rotator cuff tears [19,21]. Second, some papers included in our study individually analyzed partial and complete tears [25-27], but we combined the two for analysis. Third, the frequency of the probe used and the diagnostic level of the operating physician was different in each study. These factors are likely to introduce bias and represent limitations to this study.

**Discussion**

The rotator cuff is an important structure that maintains the stability of the shoulder joint and is important in the abduction, external rotation and shoulder’s internal rotation. It is the only tendonous tissue between the two bones of the shoulder joint, and is easily damaged due to its particular anatomical position and repetitive movement [28]. Many methods can be used to examine the shoulder joint. Radiography and CT are widely used in the diagnosis of bone damage and its extent, but the diagnosis of soft tissue injury is inferior to ultrasound and MRI. Although MRI is currently the most reliable, noninvasive diagnostic method for rotator cuff tears evaluation, it is an expensive imaging technique, has low reproducibility and is not suitable for use in patients with metal-mounted devices [12].

In the recent years, with the development of ultrasound technology and the emergence of high-frequency probes, ultrasound was increasingly used for musculoskeletal systems pathology. Many studies have shown that there are not obvious differences between ultrasound and MRI when detecting rotator cuff tears [19,26,29]. However, though ultrasound is inexpensive, noninvasive and convenient, its effectiveness when diagnosing RC tears is still controversial. We found that the combined sensitivity, specificity, LR- and DOR of the included studies show higher heterogeneity ($I^2 >50\%$). Even so, ultrasound had a high sensitivity and credible specificity when diagnosing the rotator cuff tear. We also found a high LR+, DOR, certain accuracy and a lower LR--. The LR is a composite indicator reflecting sensitivity and specificity, which can better reflect the diagnostic value of diagnostic experiments [30].

Although we systematically evaluated the inclusion of literature using a rigorous process, there are still deficiencies. For example, Elmorsy et al did not mention the time interval between the ultrasound exam and the reference standard’s implementation [19], which could generate the bias of patient flow. In addition, our study did not contain unpublished and non-English articles. Finally, the studies we included are from different regions, which is likely to produce differences in compliance. In all, compliant bias, language, and publication are shown.

This meta-analysis demonstrates that combined sensitivity and LR- are of a high degree of heterogeneity and we have three explanations for this. First, a portion of the included studies was selected for patients with clinically, highly suspected rotator cuff tears [19,21]. Second, some papers included in our study individually analyzed partial and complete tears [25-27], but we combined the two for analysis. Third, the frequency of the probe used and the diagnostic level of the operating physician was different in each study. These factors are likely to introduce bias and represent limitations to this study.

**Conclusion**

In conclusion, our meta-analysis demonstrates that ultrasound has a high efficiency of diagnosis for rotator cuff tears. As for patients, it can be a promising approach with suspected rotator cuff tears due to its high diagnostic accuracy and sensitivity specificity.

**Conflicts of interest:** none

**Acknowledgements:** The project was supported by the National Natural Science Foundation of China (NO.81771841) and the Commission of Scientific and Technology of Shenzhen (No. JCYJ20170307095706 970).

**References**