Office-based Versus Radiologist-based Ultrasonography for the Diagnosis of Rotator Cuff Tears

Corresponding Author:
Mr. Chetan S Modi,
Shoulder and Elbow Fellow, Toronto Western Hospital - Canada

Submitting Author:
Mr. Chetan S Modi,
Shoulder and Elbow Fellow, Toronto Western Hospital - Canada

Article ID: WMC003756
Article Type: Systematic Review
Submitted on: 06-Oct-2012, 03:20:01 AM GMT Published on: 06-Oct-2012, 05:30:52 PM GMT
Article URL: http://www.webmedcentral.com/article_view/3756
Subject Categories: ORTHOPAEDICS
Keywords: Rotator Cuff Tear, Full thickness tear, Partial thickness tear, Ultrasound, Accuracy, Office-based, Systematic review

How to cite the article: Modi CS. Office-based Versus Radiologist-based Ultrasonography for the Diagnosis of Rotator Cuff Tears. WebmedCentral ORTHOPAEDICS 2012;3(10):WMC003756

Copyright: This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC-BY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Source(s) of Funding:
No funding

Competing Interests:
No competing interests
Office-based Versus Radiologist-based Ultrasonography for the Diagnosis of Rotator Cuff Tears

Author(s): Modi CS

Abstract

Aim
The aim of this systematic review is to compare the accuracy of ultrasonography in the diagnosis of rotator cuff tears between office-based scanning by an orthopaedic surgeon and scanning by a specialist trained radiologist.

Methods
A literature search using multiple databases was performed with eligible studies being critically appraised using a validated scoring system and data pooling performed where appropriate.

Results
Seventeen studies with varying sample sizes and quality were identified. Ultrasonography was performed within the clinic setting in eight studies and by radiologists in nine studies. The best evidence within both the office-based (level 1b) and radiology department (level 2b) settings suggests that ultrasonography is a powerful tool for ruling in and ruling out full thickness (FT) tears. The best evidence for the detection of partial thickness (PT) tears (level 1b) suggests that ultrasonography is able to rule in these tears but less predictable at ruling them out with better results when performed by a radiologist (level 2b).

Conclusion
The systematic review suggests that ultrasonography is a powerful tool for the diagnosis of rotator cuff tears. The audit within the upper limb unit identified deficiencies in the detection of FT tears which requires further investigation as to the possible reasons. The evidence suggests that scanning performed within the clinic setting may improve the identification of FT tears to enable appropriate counselling of patients and treatment planning.

Introduction

High-resolution shoulder ultrasonography has many advantages over other techniques used for the diagnosis of rotator cuff pathology as it is rapid, non-invasive, relatively inexpensive and allows a dynamic examination to be performed. A disadvantage, however, is that it requires specialist training and has a significant learning curve and is therefore highly user-dependent.

The results within the literature have varied considerably with meta-analyses concluding that ultrasonography may be a useful tool for ruling in full-thickness rotator cuff tears (FTRCT) but less reliable at ruling them out with inconclusive results for the detection of partial thickness rotator cuff tears (PTRCT)\(^1\). The overall accuracy has been deemed comparable to magnetic resonance scanning\(^2\).

There has been increasing interest recently in office-based ultrasonography performed by the orthopaedic surgeon as part of the examination of the patient. This is an attractive option for patients as it requires only one visit to the clinic with an immediate diagnosis being made with imaging confirmation so that treatment can be planned.

The aim of this systematic review is to compare the accuracy of office-based with traditional radiology-based ultrasonography for the diagnosis of FTRCTs and PTRCTs.
Comparison: Arthroscopic assessment of rotator cuff (gold-standard)

Outcome: Bursitis, tendinopathy, partial-thickness RCT, full-thickness RCT

Exclusion Criteria

Non-English language publications

Narrative reviews

Database Search

The Cochrane Library of Systematic Reviews (online), Medline database and EMBASE will be searched for the evidence.

A Medline database search will be undertaken using the PubMed internet search engine to look at citations that have been published using the following MeSH (Medline/Pub Med's Subject Headings) terms:

Search date: 1950 to February 17th 2011

1. "Rotator Cuff/injuries"[Mesh]
2. "Shoulder Joint/injuries"[Mesh]
3. 1 OR 2
4. "Rotator Cuff/ultrasonography"[Mesh]
5. "Shoulder Joint/ultrasonography"[Mesh]
6. "Shoulder Pain/ultrasonography"[Mesh]
7. 4 OR 5 OR 6
8. 3 AND 7
9. Limits: Humans; English; Meta-Analysis; Review; Comparative Study; Evaluation Studies

An EMBASE search will be undertaken using MeSH headings with the following search strategy:

Search date: 1980 to 2011 Week 07

1. exp Rotator Cuff Rupture/di [Diagnosis]
2. exp Ultrasound/
3. exp diagnostic imaging/
4. exp Echography
5. 2 OR 3 OR 4
6. 1 AND 5
7. Limit 6 to (human and english language and ("diagnosis (sensitivity)" or "diagnosis (specificity)" or "diagnosis (optimized)") and english and (journal or "review"))

The studies identified from the database search will be assessed against the inclusion criteria for this review. Eligible studies will then undergo critical appraisal using a standardised validated scoring system relevant to the study design in order to identify the best evidence as in the review by Dinnes, et al. A narrative data synthesis will be performed to identify a relevant treatment effect and data pooling will be carried out where possible to generate higher study powers and levels of evidence.

Results

The Medline and EMBASE database searches revealed 123 and 272 studies respectively which were assessed against the eligibility criteria. The details of this process are shown in figures one and two and this resulted in the identification of 22 eligible studies from the Medline database and 23 studies from the EMBASE database. There were 32 studies eligible for inclusion within this systematic review after accounting for repeats from the databases. One of these studies was a Health Technology Assessment systematic review with meta-analysis of the diagnosis of shoulder pain by Dinnes, et al. This extensive review was used as a starting point for the current systematic review and included 14 of the remaining 31 studies identified from the database search. Thus 17 studies were analysed and critically appraised within this systematic review. The search within the Cochrane Library of Systematic Reviews also revealed the study by Dinnes, et al. with no other reviews identifiable. A summary of the included studies is shown in table 1 with critical appraisal shown in table 2. The studies were subdivided for further discussion into those where office-based ultrasonography was performed by the orthopaedic surgeon and those where ultrasonography was performed by a specialist-trained radiologist.

Dinnes, et al.

The study was designed to evaluate the clinical and cost-effectiveness of diagnostic imaging tests as an adjunct to clinical assessment in patients with soft tissue shoulder disorders. The evidence for a variety of imaging modalities was assessed with 38 studies identified regarding the accuracy of ultrasound. Twenty-seven studies did not differentiate the type of RCT with respect to accuracy whereas 17 studies attempted to analyse FT and PT tears as separate entities.

The sensitivities for the detection of any RCT ranged from 0.33 to 1.00 and specificities from 0.43 to 1.00. The pooled sensitivity with confidence intervals was 0.80 (0.78 to 0.83) and specificity 0.85 (0.82 to 0.87). The pooled positive likelihood ratio was 5.09 (3.44 to
The sensitivities ranged from 0.58 to 1.00 and specificities from 0.78 to 1.00 for the detection of FT tears. The pooled sensitivity was 0.87 (0.84 to 0.89) and specificity 0.96 (0.94 to 0.97). The pooled positive likelihood ratio for detecting FT tears was 13.16 (9.13 to 18.95) and negative likelihood ratio was 0.16 (0.11 to 0.24). The sensitivities for the detection of PT tears ranged from 0.25 to 0.94 with the range of specificities not reported. The pooled sensitivity was 0.67 (0.61 to 0.73) and specificity 0.94 (0.92 to 0.96). The pooled positive likelihood ratio for detecting PT tears was 8.90 (4.88 to 16.22) and negative likelihood ration was 0.36 (0.22 to 0.61).

They concluded that there was a wide variation in accuracy across the individual studies. Ultrasound may be a useful tool for ruling in FT tears as the odds of a tear being present after investigation increase by around 13-fold in a high prevalence setting such as a specialist clinic. Studies were too heterogeneous to determine whether a negative result could accurately rule out a FT tear. The ability of ultrasound to accurately rule in or out the presence of PT tears was deemed low with inconclusive evidence. Important factors that correlated with better performance of the test included a higher prevalence of the pathology and an increasing age.

**Office-based ultrasonography**

Al-Shawi, et al.¹

This study involved ultrasonography performed by one orthopaedic surgeon immediately after clinical examination using an 8-10MHz portable scanner in 364 consecutive patients. They reported a high sensitivity, specificity, PPV and NPV for the detection of PT tears with poorer results for the detection of PT tears as shown in table 1. The study scored 6 out of 14 points after critical appraisal and provided level 3b evidence. The study was limited as only 143 of the 364 patients had an arthroscopic examination and 11 underwent magnetic resonance imaging (MRI) which has previously been deemed inappropriate as a reference test.¹ The eligibility criteria for the study were poorly defined and significant bias may have occurred as the same investigator clinically examined the patients and subsequently performed the scans – this was clearly unavoidable in this study as it was set in a one-stop clinic. Disease progression also requires consideration as the study did not report the time interval between scanning and surgery.

Frei, et al.⁴

Ultrasonography was performed by two orthopaedic surgeons in a retrospectively selected sample of 20 patients with a 9-13MHz transducer in this study. They reported good results with an overall accuracy of 90% for the detection of rotator cuff tears as shown in table 1 but did not differentiate FT and PT tears. The study scored 7 points after critical appraisal and provided level 3b evidence. The study was well reported with identifiable eligibility criteria with all patients undergoing an arthroscopic examination. The main limitations were the very small numbers of patients in the study and their retrospective selection which meant that non-consecutive patients were investigated. The index test was performed by two investigators and the reference test by four investigators which may have also influenced the results.

Iannotti, et al.⁵

Ultrasonography was performed by either a physician-assistant or nurse-clinician with scans interpreted at a later date by an orthopaedic surgeon and a radiologist using a 7.5MHz transducer in 99 shoulders in 98 prospectively selected patients. They reported good sensitivity and NPV for the detection of FT tears but poor specificity and PPV. The results were poorer for the detection of PT tears as shown in table 1. The study scored 8 points and provided level 2b evidence. Although prospective and clearly designed and reported, this study has several limitations which require consideration. It is not clear whether the recruited patients were consecutive or selected resulting in possible selection bias. It is also unclear how the ultrasonographic diagnoses were reached as the methods section states that the orthopaedic surgeon and radiologist both studied recorded images at an unspecified time period after the scan was performed although only 87 scans were reviewed by the radiologist. This method of diagnosis was also inadequate as ultrasonography is a dynamic investigation with a diagnosis ideally made by the operator.¹

Moosmayer, et al.⁶

Ultrasonography was performed by a single orthopaedic surgeon on 350 consecutive patients referred with shoulder pain using a 5.5-9.4MHz transducer. The reported results showed excellent performance for the detection of FT tears but much poorer results when combined with the detection of PT tears.
tears with 13 out of 14 being missed. The study scored 9 points and provided level 1b evidence. The study was well designed with the ultrasonographer being blinded to the history and clinical examination of the patients and any other investigations they may have had and with the arthroscopist being blinded to the ultrasonography results thereby eliminating any testing bias. A limitation was that although 350 patients had the index test, only 58 subsequently had the reference test. This also significantly reduced the study power. Another limitation was that the spectrum of patients and eligibility criteria were poorly defined.

Moosmayer, et al8
Ultrasoundography was performed by a single orthopaedic surgeon on 350 consecutive patients referred with shoulder pain using a 5.5-9.4MHz transducer. The results showed good performance for the detection of FT tears regarding specificity, PPV and NPV but limited by poor sensitivity (77%) as six out of 26 FT tears were missed. The results were inferior for the detection of PT tears with six out of seven missed by ultrasonography. The study scored 8 points and provided level 2b evidence. The study was clearly reported with the same reference standard used for the patients. Ultrasonography was not used as a diagnostic criterion thereby minimising any incorporation bias (table 2). A limitation was that only 79 patients had the reference standard applied although 350 scans were performed. The sonographer was not blinded to the history, clinical examination or other test results including magnetic resonance imaging (MRI) prior to performing the scans and the spectrum of presenting patients and eligibility criteria were poorly defined.

Roberts, et al9
Ultrasoundography was performed by a single orthopaedic surgeon in the clinic setting using a 7.5MHz transducer in 24 patients with suspected rotator cuff tears. The study reported poor sensitivity (50%) for the detection of FT and PT tears and poor PPV (33%) for the detection of FT tears. The study was poorly designed and scored 5 points and provided level 4 evidence. The eligibility criteria and flow of patients through the study were clearly reported. Limitations were that it was a small study including only 24 patients with reference standards haphazardly applied to 16 patients consisting of a combination of surgery, arthrography and MRI. Although both surgery and arthrography have been established as appropriate reference tests, the true accuracy of MRI has yet to be established thereby making it an inappropriate reference test. The sonographer also examined the patients prior to scanning thus introducing bias into the study with the ultrasonography results being used as a criterion for further investigation and obtaining the final diagnosis.

Zehetgruber, et al9
This retrospective study involved ultrasonography performed by three orthopaedic surgeons using a 7.5MHz transducer one day before surgery in 332 consecutive patients with shoulder pain. The study reported excellent results for the detection of rotator cuff tears but did not differentiate between FT and PT tears as shown in table 1. The study is the largest to date and scored 11 points and provided level 2b evidence. The several good aspects included an appropriate spectrum of patients, clear and detailed reporting of the index and reference test with all patients having the reference test and the sonographers being blinded to the details of the clinical examination. Limitations of the study were that the eligibility criteria were poorly defined and ultrasound results were known to the surgeons prior to performing the reference test thereby introducing bias. The results of the ultrasound scans were also used to decide the appropriate management and to make the final diagnosis resulting in incorporation bias.

Ziegler10
This study involved ultrasonography performed by a single orthopaedic surgeon in a single-visit shoulder clinic using a portable scanner with a 7.5MHz transducer in 451 consecutive shoulders in 406 patients with shoulder pain. The study reported excellent results in the performance of ultrasonography for the detection of both FT and PT tears as shown in table 1. The study scored 7 points and provided level 2b evidence. Positive aspects of the study include an appropriate spectrum of patients being investigated with the second largest sample size within the literature (table 1). Although the setting was as a single-visit clinic, the time between scanning and surgical diagnosis is not reported and there may have been disease progression. The eligibility criteria were poorly defined with only 282 shoulders in 262 patients receiving the reference test to confirm their disease state. Although unavoidable in this setting, there was significant bias as the same orthopaedic surgeon was involved in all stages of formulating the final diagnosis.

Radiology department ultrasonography
Cullen, et al11
Ultrasonography was performed by a single
experienced fellowship-trained musculoskeletal radiologist using a 5-12MHz transducer in 68 patients with suspected rotator cuff pathology. The study reported good results for the detection of both FT and PT tears although the sensitivity for the detection of PT tears was low at 77% as shown in table 1. The study scored poorly with 5 points and provided level 2b evidence. Good aspects were that they used an appropriate reference test and the same reference test for all of patients. There were many limitations which resulted in the low score as shown in table 2, some of which include no clarity as to the spectrum of patients included in the study or their eligibility criteria, no details about how the reference test was performed and a lack of blinding of the surgeon to the ultrasound results. There also appears to be some missing data which has not been accounted for within the tables presented in the study report.

Farin, et al.\textsuperscript{12} Ultrasonography was performed dynamically by an experienced radiologist in 86 patients referred with shoulder symptoms using a 7.5MHz transducer. The analysis did not differentiate between FT and PT tears with both classed as positive findings. The results for detection of tears were good although the sensitivity was limited to 86% as shown in table 1. The study scored 7 points and provided level 2b evidence. The study was clearly reported with all patients investigated with the same appropriate reference test. The reference test was performed within one month of the index test for all patients which reduced disease progression bias. Limitations of the study include the absence of a definition of the spectrum of patients in the study and their eligibility. There was no attempt to blind the investigators in the study and ultrasound results were used to guide treatment and make the final diagnosis which increases the risk of bias. Although it appears that all patients were presented in the report, the study describes these patients were a subset but does not discuss how many patients were investigated overall.

Farin, et al.\textsuperscript{13} Ultrasonography was performed by a radiologist using four different types of equipment with a transducer specification ranging from 5-7.5MHz in 184 patients within three weeks of acute shoulder trauma. Ninety-eight of these patients required surgery and their findings were correlated with ultrasonography findings. The analysis did not differentiate between FT and PT tears with both being considered as positive findings. The study reports good results for the detection of tears although limited sensitivity of 89% and NPV of 87% as shown in table 1. The study scored poorly with 5 points and provided level 2b evidence. Good aspects of the study were that an appropriate reference test was used with the same standard for all patients. Poor aspects which reduced its quality include a lack of reporting of the eligibility criteria and the spectrum of patients. Aspects of the study such as blinding, details of the reference test and the time period between the index and reference tests were not reported. Another factor may have been the use of four different types of equipment for ultrasonography which may have affected the quality of imaging.

Fotiadou, et al.\textsuperscript{14} This study involved ultrasonography performed by an experienced radiologist using an 8-13MHz transducer in 96 consecutive patients with a clinical diagnosis of rotator cuff disorders. They reported excellent sensitivity for the detection of both FT (98%) and PT (90%) with overall accuracy of 98% and 87% respectively. They were unable to report on the NPV and the specificity as approximately 10% of patients that underwent imaging did not have surgical confirmation of their disease state. The study scored 7 points as shown in table 2 and provided level 2b evidence. Good aspects of the study included an appropriate spectrum of patients with well defined eligibility criteria and an appropriate reference test being used. Limitations were related to a lack of reporting of whether any blinding occurred and the time period between the index and reference tests. Another limitation was that only 88 of the 96 patients had confirmation of their disease status.

Goldberg, et al.\textsuperscript{15} Ultrasonography was performed by 109 different radiologists using a 7.5MHz transducer in 336 patients with a clinical diagnosis of a rotator cuff tear or impingement syndrome. The reported results were very poor as shown in table 1 with overall accuracy of only 38%. The study scored poorly with 4 points and provided level 4 evidence. Good aspects of the study included a large number of patients, all of which had an appropriate reference test. There were many limitations including the inconsistent application of different reference tests with surgical confirmation of the disease state in only 67% of patients. There is a lack of reporting regarding the spectrum of patients including the manner in which they were recruited and the time period between the index test and surgical confirmation. The ultrasound results were also used to guide further treatment and bias cannot therefore be excluded. An important factor is the large numbers of
radiologists that performed the scans as their experience may have been variable which may have had a significant impact upon the quality of the imaging and its interpretation.

Milosavljevic, et al.\textsuperscript{16} Ultrasoundography was performed by a single radiologist using a 10MHz transducer in 190 consecutive shoulders in 185 patients with suspected rotator cuff tears. The study reported excellent results for the detection of FT and PT tears although the sensitivity for the detection of PT tears was only 80\% with six false negatives and seven shoulders being incorrectly diagnosed as having FT tears. The values for PT tears were also over-estimated as true positives included whether the ultrasound scan had detected either a FT or a PT tear. The study scored 9 points and provided level 2b evidence. There were large numbers of patients with all undergoing the same appropriate reference test. The spectrum of patients and the eligibility criteria were clearly stated with all being accounted for in the report. Although the study was clearly reported, there were key areas of methodology which were not discussed in the report including whether any blinding occurred between the index and reference test, what information was available to the surgeons prior to arthroscopy and whether the ultrasonography results were used to reach the final diagnosis.

Teefey, et al.\textsuperscript{17} Ultrasoundography was performed by a specialist musculoskeletal radiologist in 124 consecutive patients referred with shoulder pain using a 7.5-9MHz transducer. The results were variable and dependent upon the method of analysis as shown in table 1. The first analysis considered PT tears and intact cuffs together which is a pragmatic approach as the surgical treatment of both PT tears and intact rotator cuffs with bursitis and tendinopathy is similar whereas repair would usually only be considered for FT tears\textsuperscript{21}. The second analysis considered both FT and PT lesions as tears. The data suggests that ultrasonography is accurate for the detection of FT tears as 45 of 46 were detected with one incorrectly diagnosed as a PT tear. There were five false positives with four diagnosed as having PT tear and one as no tear. The detection of PT tears was less accurate as only 13 out of 19 were correctly diagnosed with four incorrectly diagnosed as a FT tear and two as no tear. The study scored 8 points and provided level 2b evidence. The study was clearly reported with all patients being prospectively enrolled and accounted for and undergoing the same reference test. The spectrum of patients was appropriate although their eligibility criteria were not well defined. The radiologist was also blinded to the arthroscopy findings. Limitations include that only 71 out of the 124 patients had the reference test to confirm the diagnosis. The surgeon was not blinded to the ultrasound results and therefore incorporation bias may have occurred.

Venu, et al.\textsuperscript{18} Ultrasonography was performed by two experienced radiologists using a 5-10MHz transducer in 276 consecutive patients with clinical supraspinatus impingement syndrome. The study did not report values for sensitivity and specificity but found 100\% agreement between the ultrasound and arthroscopy findings for examining normal tendons, tendinopathy, FT tears and ruptures. The results were poorer for the detection of PT tears with agreement in only two out of seven cases. There were three false positive and two false negative studies. The study scored 7 points and provided level 2b evidence. Good aspects included an appropriate spectrum of patients within the study with an appropriate reference test which was consistently applied. The limitations of the study included poor definition of the eligibility criteria with only 41 of the 276 patients having their disease state confirmed by the reference test. The ultrasound results were also used to plan treatment and therefore introduced potential bias into the evaluation.

Yen, et al.\textsuperscript{19} This study reported ultrasonography performed by a single experienced radiologist with three different systems with transducer specification between 7.5 to 10MHz in 50 patients presenting with chronic shoulder pain. Their analysis did not differentiate FT from PT tears but reported good results for the detection of tears with an overall accuracy of 94\% as shown in table 1. The study scored 7 points and provided level 2b evidence. Good aspects of the study included all patients within the study undergoing the same appropriate reference test and within one month of the index test to reduce the risk of disease progression bias. The limitations included no definition of their eligibility criteria and the spectrum of patients within the study. The ultrasonography results were used to guide further treatment which introduces potential bias. It is unclear what information was available to the surgeons prior to surgery and whether any blinding took place.

**Discussion**
The systematic literature search identified 18 eligible studies that were reviewed in further detail. A health technology assessment systematic review with meta-analysis was performed by Dinnes, et al in 2005 which reported that there was a wide variation in accuracy across individual studies. They concluded that ultrasound was a useful tool to rule in FT tears and less useful for ruling these out with poor evidence base for the assessment of PT tears. The other 17 studies that were identified within this systematic review were not included in the study by Dinnes, et al and were critically appraised.

The mean prevalence of RCTs within the 17 studies was high at 65±23% and included a range of sample sizes between 20 to 332 patients with a total of 1966 patients. Eleven studies were prospectively designed, two retrospectively and it was unclear in four studies. When considering all of the studies together, the results varied widely with sensitivity for the detection of any RCT ranging from 0.24 to 0.98, specificity 0.61 to 0.98, PPV 0.49 to 0.94, NPV 0.34 to 1.00 and accuracy 0.38 to 0.97. In studies analysing FT and PT tears separately, the sensitivity for the detection of FT tears ranged between 0.50 to 1.00, specificity 0.80 to 1.00, PPV 0.33 to 1.00, NPV 0.89 to 1.00 and accuracy 0.85 to 0.98. The sensitivity for the detection of PT tears ranged between 0.50 to 0.97, specificity 0.67 to 0.98, PPV 0.79 to 1.00, NPV 0.59 to 0.96 and accuracy 0.76 to 0.95.

The quality of studies was assessed against assigned levels of evidence and with a checklist as used by Dinnes, et al which was modified to score studies with a maximum of 14 points. The quality of the 17 studies within this review varied between levels 1b to 4 and scored between 4 to 11 points as shown previously in table 2. One study was assigned as level 1b evidence, 12 studies as level 2b evidence, two studies as level 3b evidence and two studies as level 4 evidence.

Seven studies were judged to include an appropriate spectrum of patients for investigation with an inappropriate spectrum in five studies and insufficient detail in five studies. The eligibility criteria were described in adequate detail in only six of the 17 studies within the review. Fifteen of the 17 studies used an appropriate reference test of either arthroscopic or open surgical assessment with two studies also using MRI which is inappropriate. Partial verification occurred in most of the studies as in only six studies did all patients who underwent the index test also have the reference test to confirm their disease status. The reference standards were consistently applied in most studies except for three in which differential verification occurred as different reference tests were applied to confirm the disease status. The time delay between the index and reference test was reported in only eight of 17 studies and this was deemed a satisfactory period in only three studies. Incorporation bias, whereby the result of the index test was used to establish the final diagnosis occurred in most studies except for two where this was treated independently. Review bias occurred in most studies where either the index test or reference test result was interpreted with the knowledge of the results of the other test. Blind interpretation of both the index and reference test results occurred in only one study and this study was therefore assigned as level 1b evidence.

The 17 studies were sub-divided into those in which ultrasonography was performed within the clinic or office setting usually by an Orthopaedic surgeon and those in which ultrasound was performed by one or more radiologists. Eight studies investigated the accuracy of ultrasound for RCTs within the office-based setting whereas 9 studies involved investigations performed by radiologists. Of the eight studies investigating office-based ultrasound, two studies did not differentiate the type of tear whereas six studies attempted to differentiate between FT and PT tears. Of the nine studies investigating ultrasound performed by radiologists, four studies did not differentiate the type of tear whereas five studies attempted to differentiate between FT and PT tears.

The best evidence for the detection of any RCT in the office-based setting was level 2b evidence in the study by Zehetgruber, et al with a high study quality score of 11 points. This study reported excellent performance for ultrasonography with a sensitivity of 0.98, specificity 0.93, PPV 0.97, NPV 0.95 and overall accuracy of 0.97. The six studies within the office-based setting which attempted to differentiate between FT and PT tears ranged from levels 1b to 4 and scored between 5 and 9 points. The pooled sensitivity for the detection of FT tears was 0.93 and specificity 0.93. The pooled sensitivity for the detection of PT tears was 0.86 and specificity 0.95. The best evidence was the study by Moosmayer, et al which was assigned level 1b and scored 9 points showed good performance for the detection of FT tears with sensitivity of 1.00, specificity 0.97, PPV 0.96, NPV 1.00 and overall accuracy 0.98. The results for the detection of PT tears were poorer with a sensitivity of 0.66, specificity 0.95, PPV 0.96, NPV 0.59 and accuracy 0.76. This study was well designed but had a small number of patients (n=58). Two studies which
were assigned level 2b and scored 8 points with larger numbers of patients both showed good results for the detection of FT tears and poorer performance for the detection of PT tears in keeping with the level 1b study described above. The only study to show excellent results for the detection of both FT and PT tears within the office setting was that by Ziegler although the weight of this evidence was limited by a lower quality score of 7 points compared to other studies.

The four studies with scans performed by radiologists for the detection of any RCT ranged from levels 2b to 4 and 4 to 7 points. The pooled sensitivity for the detection of any tear type was 0.49 (0.90 if excluding outlier study) and specificity 0.77 (0.95 if excluding outlier study). The best evidence for the detection of any RCT with the radiologists performing the scans was level 2b evidence which was provided by two studies both of which scored 7 points. The sensitivity ranged from 0.86 to 0.95, specificity 0.90 to 0.98, PPV 0.96 to 0.97, NPV 0.82 to 0.93 with both studies having the same overall accuracy of 0.94. The prevalence of tears in the study by Farin, et al. was 35% whereas this was significantly greater in the study by Yen, et al. at 80% which may account for the different results although overall accuracy remained unchanged. The study by Goldberg, et al. reported poor results with a sensitivity of 0.24 and specificity of 0.61. Although there were many other limitations in this study, over 100 radiologists performed the scans and it was therefore impossible to ensure that all of the operators were experienced. This study illustrates that an important factor when considering the usefulness of ultrasonography for diagnosing RCTs is knowledge of the quality of the operator.

The five studies with scans performed by radiologists which differentiated between FT and PT tears were all assigned level 2b evidence and ranged from 5 to 9 points. The pooled sensitivity for the detection of FT tears was 0.98 and specificity 0.91. The pooled sensitivity for the detection of PT tears was 0.90 and specificity 0.96. The best evidence for the detection of FT and PT tears with scans performed by radiologists was level 2b evidence in the study by Milosevijevic, et al. with a quality score of 9 points. The sensitivity for the detection of FT tears was 1.00, specificity 0.91, PPV 0.91, NPV 1.00 and accuracy 0.95. The sensitivity for the detection of PT tears was 0.80, specificity 0.98, PPV 0.86, NPV 0.96 and accuracy 0.95. The second best study which scored 8 points by Teefey, et al. showed similar results with an accuracy of 0.94 for both FT and PT tears although the specificity and NPV for PT tears were lower at 0.67. This review of the literature regarding the accuracy of ultrasonography in the detection of RCTs has shown there to be a wide variation in its usefulness for detecting RCTs in general and in differentiating between FT and PT tears. The distinction between FT and PT tears is important as current opinion regarding the management options for these does differ as discussed by Rees, et al. Most FT tears are surgically repaired depending upon the nature of the tear, the age, symptomology and activity level of the patient. PT tears usually undergo either surgical debridement with subacromial decompression depending upon the level of pain experienced by the patient or completion and cuff repair. These different managements vary significantly with regards to the surgical time and specialist equipment required, and more importantly the rehabilitation period and potential time required off work by the patient. The studies that assessed the accuracy of ultrasonography and differentiated between FT and PT tears are therefore more useful to the clinician to decide whether it is a useful investigation.

The critical appraisal of the included studies has demonstrated the best evidence available for ultrasonography within the office-based setting with scanning usually performed by an Orthopaedic surgeon and within the traditional setting with radiologists performing the scans. The best evidence for both settings showed that ultrasonography is a powerful tool for the detection of FT tears with a sensitivity of 1.00 and specificity of 0.91 to 0.97. When comparing office-based and radiologist-performed ultrasonography, the specificity (0.97 vs. 0.91), PPV (0.96 vs. 0.91) and overall accuracy (0.98 vs. 0.95) were slightly better in the former. This may have been due to an advantage for the operator within the office-based setting of having clinically examined the patient prior to performing the scan, which may have improved accuracy. The best evidence within both settings for the detection of PT tears showed a high specificity (0.95 to 0.98) suggesting that it is a good test for ruling in PT tears. When comparing the test within the different settings, the sensitivity (0.80 vs. 0.66), NPV (0.96 vs. 0.59) and overall accuracy (0.96 vs. 0.76) were higher when ultrasonography was performed by a radiologist. This difference may have been due to the less clearly defined ultrasonographic criteria for the diagnosis of PT tears and with the difficulty in clinically differentiating between PT tears, rotator cuff tendinosis and bursitis when examining the patient.

Conclusion
The best evidence in both the office-based (level 1b) and radiology department (level 2b) settings suggests that ultrasonography is a powerful tool for ruling in and ruling out FT tears with results being slightly better when performed in the office-based setting. This may be because the Orthopaedic surgeon had the advantage of having performed a specialist clinical assessment of the patients prior to ultrasonography. The best evidence for the detection of PT tears (level 1b) suggests that ultrasonography is able to rule in these tears but less predictable at ruling them out with better results when performed by a radiologist. This is possibly due to the difficulty in clinically diagnosing these tears with less clearly defined ultrasonographic criteria.

References

20. Yen CH, ChiuJH, Chou YH, Hsu CC, Wu JJ, Ma...


Illustrations

Illustration 1

Figure 1 - Results from Medline database search

Illustration 2

Figure 2 - Results from EMBASE database search
Illustration 3

Table 1 - Summary of included studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>Weight</th>
<th>Persuence</th>
<th>Beneficiency</th>
<th>Proportion of patients with benefit</th>
<th>Proportion of patients with no benefit</th>
<th>Proportion of patients with harm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atwood 2008</td>
<td>R</td>
<td>1.15</td>
<td>0.96</td>
<td>0.85</td>
<td>0.95</td>
<td>0.95</td>
<td>0.95</td>
</tr>
<tr>
<td>Cutili 2007</td>
<td>R</td>
<td>0.9</td>
<td>0.85</td>
<td>0.75</td>
<td>0.65</td>
<td>0.55</td>
<td>0.45</td>
</tr>
<tr>
<td>Fau 1996</td>
<td>R</td>
<td>0.8</td>
<td>0.75</td>
<td>0.65</td>
<td>0.55</td>
<td>0.45</td>
<td>0.35</td>
</tr>
<tr>
<td>Fau 1995</td>
<td>R</td>
<td>0.7</td>
<td>0.65</td>
<td>0.55</td>
<td>0.45</td>
<td>0.35</td>
<td>0.25</td>
</tr>
<tr>
<td>Pindole 2008</td>
<td>R</td>
<td>0.6</td>
<td>0.55</td>
<td>0.45</td>
<td>0.35</td>
<td>0.25</td>
<td>0.15</td>
</tr>
<tr>
<td>Fui 2008</td>
<td>R</td>
<td>0.5</td>
<td>0.45</td>
<td>0.35</td>
<td>0.25</td>
<td>0.15</td>
<td>0.05</td>
</tr>
<tr>
<td>Golberg 2012</td>
<td>R</td>
<td>0.4</td>
<td>0.35</td>
<td>0.25</td>
<td>0.15</td>
<td>0.05</td>
<td>0.00</td>
</tr>
<tr>
<td>Aiki 2012</td>
<td>R</td>
<td>0.3</td>
<td>0.25</td>
<td>0.15</td>
<td>0.05</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Misaki 2012</td>
<td>R</td>
<td>0.2</td>
<td>0.15</td>
<td>0.05</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Moscato 2012</td>
<td>R</td>
<td>0.1</td>
<td>0.05</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Oberto 2016</td>
<td>R</td>
<td>0.0</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Tamba 2016</td>
<td>R</td>
<td>0.1</td>
<td>0.05</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Vus 2012</td>
<td>R</td>
<td>0.2</td>
<td>0.15</td>
<td>0.05</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Vus 2010</td>
<td>R</td>
<td>0.3</td>
<td>0.25</td>
<td>0.15</td>
<td>0.05</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Spina 2012</td>
<td>R</td>
<td>0.4</td>
<td>0.35</td>
<td>0.25</td>
<td>0.15</td>
<td>0.05</td>
<td>0.00</td>
</tr>
<tr>
<td>Deng 2014</td>
<td>R</td>
<td>0.5</td>
<td>0.45</td>
<td>0.35</td>
<td>0.25</td>
<td>0.15</td>
<td>0.05</td>
</tr>
</tbody>
</table>

Illustration 4

Table 2 - Critical appraisal of included studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Proportion of patients with benefit</th>
<th>Proportion of patients with no benefit</th>
<th>Proportion of patients with harm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atwood 2008</td>
<td>0.95</td>
<td>0.95</td>
<td>0.95</td>
</tr>
<tr>
<td>Cutili 2007</td>
<td>0.85</td>
<td>0.75</td>
<td>0.65</td>
</tr>
<tr>
<td>Fau 1996</td>
<td>0.75</td>
<td>0.65</td>
<td>0.55</td>
</tr>
<tr>
<td>Fau 1995</td>
<td>0.65</td>
<td>0.55</td>
<td>0.45</td>
</tr>
<tr>
<td>Pindole 2008</td>
<td>0.55</td>
<td>0.45</td>
<td>0.35</td>
</tr>
<tr>
<td>Fui 2008</td>
<td>0.45</td>
<td>0.35</td>
<td>0.25</td>
</tr>
<tr>
<td>Golberg 2012</td>
<td>0.35</td>
<td>0.25</td>
<td>0.15</td>
</tr>
<tr>
<td>Aiki 2012</td>
<td>0.25</td>
<td>0.15</td>
<td>0.05</td>
</tr>
<tr>
<td>Misaki 2012</td>
<td>0.15</td>
<td>0.05</td>
<td>0.00</td>
</tr>
<tr>
<td>Moscato 2012</td>
<td>0.05</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Oberto 2016</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Tamba 2016</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Vus 2012</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Vus 2010</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Spina 2012</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Deng 2014</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>
Disclaimer

This article has been downloaded from WebmedCentral. With our unique author driven post publication peer review, contents posted on this web portal do not undergo any prepublication peer or editorial review. It is completely the responsibility of the authors to ensure not only scientific and ethical standards of the manuscript but also its grammatical accuracy. Authors must ensure that they obtain all the necessary permissions before submitting any information that requires obtaining a consent or approval from a third party. Authors should also ensure not to submit any information which they do not have the copyright of or of which they have transferred the copyrights to a third party.

Contents on WebmedCentral are purely for biomedical researchers and scientists. They are not meant to cater to the needs of an individual patient. The web portal or any content(s) therein is neither designed to support, nor replace, the relationship that exists between a patient/site visitor and his/her physician. Your use of the WebmedCentral site and its contents is entirely at your own risk. We do not take any responsibility for any harm that you may suffer or inflict on a third person by following the contents of this website.