EXTENDED REPORT

Ultrasound definition of tendon damage in patients with rheumatoid arthritis. Results of a OMERACT consensus-based ultrasound score focussing on the diagnostic reliability

George A W Bruyn,1 Petra Hanova,2 Annamaria Iagnocco,3 Maria-Antonietta d’Agostino,4 Ingrid Möller,5 Lene Terslev,6 Marina Backhaus,7 Peter V Balint,8 Emilio Filippucci,9 Paul Baudoin,1 Richard van Vught,10 Carlos Pineda,11 Richard Wakefield,12 Jesus Garrido,13 Ondrej Pecha,14 Esperanza Naredo,15 on behalf of the OMERACT Ultrasound Task Force

ABSTRACT

Objective To develop the first ultrasound scoring system of tendon damage in rheumatoid arthritis (RA) and assess its intraobserver and interobserver reliability.

Methods We conducted a Delphi study on ultrasound-defined tendon damage and ultrasound scoring system of tendon damage in RA among 35 international rheumatologists with experience in musculoskeletal ultrasound. Twelve patients with RA were included and assessed twice by 12 rheumatologists-sonographers. Ultrasound examination for tendon damage in B mode of five wrist extensor compartments (extensor carpi radialis brevis and longus; extensor pollicis longus; extensor digitorum communis; extensor digiti minimi; extensor carpi ulnaris) and one ankle tendon (tibialis posterior) was performed blindly, independently and bilaterally in each patient. Intraobserver and interobserver reliability were calculated by \( \kappa \) coefficients.

Results A three-grade semiquantitative scoring system was agreed for scoring tendon damage in B mode. The mean intraobserver reliability for tendon damage scoring was excellent (\( \kappa \) value 0.91). The mean interobserver reliability assessment showed good \( \kappa \) values (\( \kappa \) value 0.75). The most reliable were the extensor digitii minimi, the extensor carpi ulnaris, and the tibialis posterior tendons. An ultrasound reference image atlas of tenosynovitis and tendon damage was also developed.

Conclusions Ultrasound is a reproducible tool for evaluating tendon damage in RA. This study strongly supports a new reliable ultrasound scoring system for tendon damage.

INTRODUCTION

Tenosynovitis is one of the key features of the clinical pattern in patients with rheumatoid arthritis (RA).

PATIENTS AND METHODS

A two-step study

The study was carried out in two steps. The first step consisted of a Delphi exercise, aiming to find agreement on ultrasound definitions of normal tendons, periarticular structures, tenosynovitis and tendon damage in RA; furthermore, the Delphi
exercise was done to reach consensus on the ultrasound grading of tenosynovitis and tendon damage in RA patients. Details on the methodology of the first step have previously been reported by Naredo et al.\textsuperscript{13}

**ultrasound reliability assessment**

The first step of the study was followed by a two-day patient-reliability exercise, which took place in Amsterdam, The Netherlands. Each day was divided in a morning and an afternoon session. The afternoon session was a repetition of the morning session in order to assess the intraobserver reliability.

**Patients**

Twelve patients with RA according to the American College of Rheumatology 1987 criteria\textsuperscript{14} representing all degrees of disease activity (severe, moderate, low and remission as defined by DAS28) were recruited from the outpatient rheumatology clinic (MC Groep hospitals). Demographic and clinical data were recorded for all patients.

The 12 patients were equally divided over 2 days. Both wrists and ankles were studied for the ultrasound investigation. All patients were assessed twice, that is, during the morning and again in the afternoon. The local ethics committee approved the study and all patients gave written consent according to the Declaration of Helsinki.

**Ultrasonographers**

Twelve rheumatologists with extensive experience in ultrasound, that is, more than 10 years, participated in the present study.

**Tendons**

At the wrist, the following extensor tendons enclosed in a synovial sheath were selected: the second extensor compartment, that is, the extensor carpi radialis brevis and longus; the third, that is, the extensor pollicis longus (EPL); the fourth, that is, the extensor digitorum communis (EDC); the fifth, that is, the extensor digiti minimi (EDM); and the sixth, the extensor carpi ulnaris (ECU). At the ankle, the tibialis posterior tendon was included. Since flexor tendons at the wrist may show a high level of anisotropy making ultrasound evaluation of tendon damage difficult, they were not included in the ultrasound evaluation.\textsuperscript{12,13}

**Ultrasonography**

Bilateral ultrasound investigation was performed with six Esaote ultrasound scanners (one Mylab 70 XVision and five Mylab Class C; Esaote, Genoa, Italy) by means of linear array transducers (6–18 MHz or 4–13 MHz). The B mode settings of each ultrasound machine were optimised and fixed. Dynamic investigation by flexion and extension of particular fingers was allowed to improve differentiation of tendon pathologies.

The 12 ultrasonographers independently, consecutively and blinded to the clinical data performed the ultrasound examination of the selected tendons and assessed tendon damage in B mode according to the agreed scoring system. The extensor tendons of the wrist were scanned from the level of Lister’s tubercle downwards to the level of the extensor retinaculum; the tibialis posterior tendon was scanned from a level proximal to the medial malleolus to slightly distal of it.\textsuperscript{12,13} Maximal scanning time was 15 min per patient. The scanning time included the time to fill out the scoring sheet.

**Atlas**

All members of the OMERACT US task group collected images which were used to develop an US reference image atlas of tenosynovitis and tendon damage.

**Statistical analysis**

Statistical analysis was performed with the software package SPSS, version 17.0. Normally distributed continuous data were summarised with means and SDs or 95% CIs; non-normally distributed data were summarised with median and range.

Intra- and interobserver agreement was assessed by $\kappa$ coefficients. Cohen’s $\kappa$ coefficient was calculated for intraobserver agreement, whereas Light’s $\kappa$ was calculated for interobserver agreement.\textsuperscript{15,16} The comparison of the $\kappa$s between first and second occasion was conducted using the Root Mean Square Difference index, and by the product-moment correlation coefficient. Basic statistics and interobserver reliability represented by the intraclass correlation coefficient (ICC) with 95% CI were determined for each tendon compartment separately.

ICC and $\kappa$ values are comparable; $\kappa$ values were interpreted as follows: 0–0.20 poor, 0.20–0.40 fair, 0.40–0.60 moderate, 0.60–0.80 good and 0.80–1 excellent agreement.

**RESULTS**

**Delphi process**

The results of the Delphi exercise regarding ultrasound definitions of normal tendons, anatomically related structures and tenosynovitis have been previously reported.\textsuperscript{13} Regarding the statements on tendon damage, there was group agreement on

![Figure 1](http://ard.bmj.com/)

**Figure 1** (A) Transverse scan in B mode of a normal extensor carpi ulnaris tendon, residing in its groove on the distal ulna. Tendon damage grade 0. Dimension unit indicates 10 mm. (B) Longitudinal scan in B mode of a normal extensor carpi ulnaris tendon. Tendon damage score grade 0. Dimension unit indicates 10 mm.
the definition of tendon damage and the scoring system after two rounds.

In particular, group agreement was achieved on the following items: tendon damage can be defined on B-mode as internal and/or peripheral focal tendon defect (ie, absence of fibres) in the region enclosed by tendon sheath, seen in two perpendicular planes; the grade of tendon damage should be assessed in both longitudinal and transverse planes; and, a four-grade semiquantitative scoring system (ie, grade 0, normal; grade 1, minimal; grade 2, moderate; grade 3, severe) can be used to score tendon damage on B mode.

Review of sent ultrasound images of tendons and consensus finding on scoring system

Out of 28 consulted experts, 19 (68%) sent a set of ultrasound images covering all grades of tendon damage to the organisers of the study (GA WB and EN). All the participants of the reliability exercise reviewed these images in a consensus meeting on the evening prior to the exercise. During this review process, it was noted that a four-grade semiquantitative scoring system did not work for most experts. Based on their opinion, the following scoring system for tendon damage in B mode was concurred: grade 0, normal tendon; grade 1, partial tendon damage; grade 2, moderate tendon damage; grade 3, severe tendon damage.

**Figure 2** (A) Longitudinal scan of extensor carpi ulnaris tendon. Tendon damage score grade 1. Asterisks indicate an area of synovial proliferation within the tendon sheath. Arrow point to partial rupture. (B) Transverse scan of extensor carpi ulnaris tendon. Tendon damage score grade 1. Asterisks indicate tenosynovitis, arrows point to partial rupture.

**Figure 3** (A) Transverse sonogram showing a stump (arrow) of the completely ruptured extensor digiti minimi tendon. The stump is surrounded by fluid and synovial proliferation. (B) Longitudinal sonogram of the extensor digiti minimi tendon, showing the site of complete rupture (***) and distension of the tendon sheath due to fluid and synovial proliferation (arrow).

**Table 1** Demographical, disease-related characteristics and ultrasound grading

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>n (%)</th>
<th>Median, range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of patients</td>
<td>12 (100)</td>
<td>–</td>
</tr>
<tr>
<td>Gender (F/M)</td>
<td>7/5 (58/42)</td>
<td>–</td>
</tr>
<tr>
<td>Age (years), median and range</td>
<td>66.5, 36–74</td>
<td>76, 12–104</td>
</tr>
<tr>
<td>Disease duration (months)</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Rheumatoid factor (RF)+</td>
<td>8 (67)</td>
<td>–</td>
</tr>
<tr>
<td>Anti CCP+</td>
<td>5 (42)</td>
<td>–</td>
</tr>
<tr>
<td>RF+ anti CCP+</td>
<td>3 (25)</td>
<td>–</td>
</tr>
<tr>
<td>RF– anti CCP–</td>
<td>2 (17)</td>
<td>–</td>
</tr>
<tr>
<td>DAS 28</td>
<td>–</td>
<td>3.79, 2.41–5.6</td>
</tr>
<tr>
<td>Remission (DAS 28 &lt; 2.6)</td>
<td>3 (25)</td>
<td>–</td>
</tr>
<tr>
<td>Low disease activity (DAS 28 &lt; 3.2)</td>
<td>4 (33)</td>
<td>–</td>
</tr>
<tr>
<td>Moderate disease activity DAS 28 &lt; 5.2)</td>
<td>3 (25)</td>
<td>–</td>
</tr>
<tr>
<td>High disease activity (DAS 28 &gt; 5.2)</td>
<td>2 (17)</td>
<td>–</td>
</tr>
<tr>
<td>Current use of anti-TNFa agents</td>
<td>7 (58)</td>
<td>–</td>
</tr>
<tr>
<td>Erosive disease</td>
<td>8 (67)</td>
<td>–</td>
</tr>
<tr>
<td>Cumulative prednisone dose (g)</td>
<td>19.1, 0–30.5</td>
<td>25, 4–44</td>
</tr>
<tr>
<td>Total vdHSS</td>
<td>25, 4–44</td>
<td>–</td>
</tr>
<tr>
<td>Grade 0 tendon damage</td>
<td>2652 (76.7)</td>
<td>–</td>
</tr>
<tr>
<td>Grade 1 tendon damage</td>
<td>727 (21)</td>
<td>–</td>
</tr>
<tr>
<td>Grade 2 tendon damage</td>
<td>77 (2.3)</td>
<td>–</td>
</tr>
</tbody>
</table>

CCP, anti-cyclic citrullinated peptide.
Intraobserver and interobserver agreement

Table 3 shows the κ coefficient estimates of interobserver agreement calculated for pairs of investigators at first and second (italic) occasion. Intraobserver agreement is shown as bold numbers on the diagonal line. Means are calculated below the table.

In table 4, the interobserver reliability of the tested scoring system within particular compartments is estimated.

Atlas

GS and PD US images of tendon lesions were collected into a US reference image atlas. The reference images include US images of tenosynovitis and tendon damage of various grades affecting tendons frequently involved in RA. Multiple examples covering semiquantitative grades of tenosynovitis (0–3) and tendon damage (0–2) are shown in the online supplementary material. In addition to the typical images, the atlas comprises a series of challenging Doppler images. With the guidance of the reference images displayed in this atlas, US scans of tendon abnormalities in RA can easily be scored for various grades of tenosynovitis and tendon damage both in clinical practice and in research trials (see online supplementary material).

**DISCUSSION**

To our knowledge, the present exercise is the first multiobserver study that assesses the reproducibility of ultrasound in scoring tendon damage in patients with RA. The results show a high intraobserver and interobserver reliability among experienced rheumatology ultrasonographers. The findings may be relevant for both daily clinical practice and trials. As yet, no ultrasound studies have assessed the grading of tendon damage in RA; a reliable imaging scoring system may be used to identify and follow-up tendons at risk of rupture. Furthermore, a ultrasound scoring system for tendon disease may serve as an imaging biomarker for clinical drug trials.

Only a few studies have looked at the ultrasound assessment of tendon damage in RA, observing a wide variability. Filippucci and colleagues found partial tendon tears in 12% and complete tears in 3% in a cross-sectional analysis of 90 patients with RA. Micu and colleagues found tendon damage in over 50% seen in two orthogonal planes, and grade 2, complete tendon rupture seen in two orthogonal planes. Greyscale examples of tendon damage are shown in figures 1–3.

**Patient characteristics**

The demographics and disease-related characteristics of the patients with RA are summarised in table 1.

**Prevalence of ultrasound abnormalities**

Overall, 3456 tendon compartments were assessed by ultrasound in B mode (144 per investigator). Of these, 804 ultrasound investigations showed either a grade 1 or a grade 2 tendon lesion (23%). The prevalence of lesions per tendon compartment is shown in table 2.
The presence of involvement of the ECU tendons have been reported to be frequently due to various difficulties in interpretation of tendon damage, especially in transverse view of the EDC tendon. The EPL tendon can be difficult to follow in its course while crossing other compartments in RA patients, with the best scores in the EDM, ECU and tibialis posterior tendons.

Some differences in reliability of scoring of tendon damage within particular compartments were found. This is probably due to various difficulties of investigations of particular tendons. The most reliable tendons were the ECU and tibialis posterior tendon—both non-splitting and relatively thick and straight-running tendons. One other very reliable tendon was the EDM. The ECU proved also to be the most often damaged tendon of all tendons investigated in the study. All the above mentioned tendons have been reported to be frequently involved in RA.

The least reliable tendons were EDC and EPL, probably due to splitting in some finger extensor tendons which may cause difficulties in interpretation of tendon damage, especially in transverse view of the EDC tendon. The EPL tendon can be difficult to follow in its course while crossing other tendons. There are limitations inherent in our study. First, only 12 patients were assessed. However, similar numbers of patients have been assessed in other multitester reliability studies for feasibility reasons. The difference in reliability noted between the right and left compartments three and four is related to the small prevalence of positive findings on the left side compared with the right, rather than true differences among observers in scoring lesions. Second, the lack of a gold standard, for example, MRI or surgery prevented to determine the true prevalence of tendon damage lesions. However, a concurrent validity study showed comparative accuracy in diagnosing tendon damage between ultrasound and MRI. Additionally, this was not a validity but a reliability study. Finally, the rheumatologists involved were all expert in ultrasonography. Thus, it is not taken for granted that these reliability results can be extrapolated to a population of less experienced rheumatologists. It is reassuring, however, that the broad US7 experience in Germany has revealed a good correlation between experts and less experienced rheumatologists.

A strength of this study is the inclusion of the acquisition phase of ultrasound images. Another strength is that the reliability was assessed separately for particular tendon compartments.

In conclusion, the present study suggests that rheumatologists-ultrasongraphers can have a high reliability in their performance of ultrasound assessment of 12 target tendon compartments in RA patients, with the best scores in the EDM, ECU and tibialis posterior tendons.

Table 4 Descriptives and interobserver reliability (ICC) of the tested scoring system within particular tendon compartments

<table>
<thead>
<tr>
<th>Tendon compartment</th>
<th>Mean</th>
<th>SD</th>
<th>ICC</th>
<th>95% CI</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>2R</td>
<td>0.094</td>
<td>0.283</td>
<td>0.978</td>
<td>(0.954 to 0.993)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>2L</td>
<td>0.118</td>
<td>0.301</td>
<td>0.898</td>
<td>(0.786 to 0.965)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>3R</td>
<td>0.222</td>
<td>0.551</td>
<td>0.977</td>
<td>(0.953 to 0.992)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>3L</td>
<td>0.035</td>
<td>0.174</td>
<td>0.425</td>
<td>(−0.208 to 0.805)</td>
<td>0.073</td>
</tr>
<tr>
<td>4R</td>
<td>0.163</td>
<td>0.413</td>
<td>0.903</td>
<td>(0.797 to 0.967)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>4L</td>
<td>0.101</td>
<td>0.072</td>
<td>0.428</td>
<td>(−0.202 to 0.806)</td>
<td>0.071</td>
</tr>
<tr>
<td>5R</td>
<td>0.094</td>
<td>0.289</td>
<td>0.961</td>
<td>(0.919 to 0.987)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>5L</td>
<td>0.163</td>
<td>0.395</td>
<td>0.966</td>
<td>(0.929 to 0.988)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>6R</td>
<td>0.493</td>
<td>0.491</td>
<td>0.973</td>
<td>(0.943 to 0.991)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>6L</td>
<td>0.795</td>
<td>0.688</td>
<td>0.991</td>
<td>(0.982 to 0.997)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Tib R</td>
<td>0.368</td>
<td>0.477</td>
<td>0.969</td>
<td>(0.935 to 0.989)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Tib L</td>
<td>0.507</td>
<td>0.502</td>
<td>0.986</td>
<td>(0.971 to 0.995)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Note. Scores were averaged over both occasions; Number of observers was 12; 2R denotes right second extensor compartment, 3L means third left extensor compartment, etc. Tib R means right tibialis posterior tendon. Grading scores (0, 1, 2) were dichotomised to 0 and 1. ICC, intraclass correlation coefficient.
REFERENCES


Ultrasound definition of tendon damage in patients with rheumatoid arthritis. Results of a OMERACT consensus-based ultrasound score focussing on the diagnostic reliability

George A W Bruyn, Petra Hanova, Annamaria Iagnocco, Maria-Antonietta d'Agostino, Ingrid Möller, Lene Terslev, Marina Backhaus, Peter V Balint, Emilio Filippucci, Paul Baudoin, Richard van Vugt, Carlos Pineda, Richard Wakefield, Jesus Garrido, Ondrej Pecha and Esperanza Naredo

Ann Rheum Dis 2014 73: 1929-1934 originally published online August 12, 2013
doi: 10.1136/annrheumdis-2013-203596

Updated information and services can be found at:
http://ard.bmj.com/content/73/11/1929

These include:

Supplementary Material
Supplementary material can be found at:
http://ard.bmj.com/content/suppl/2014/09/16/annrheumdis-2013-203596.DC1.html

References
This article cites 23 articles, 12 of which you can access for free at:
http://ard.bmj.com/content/73/11/1929#BIBL

Email alerting service
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Topic Collections
Articles on similar topics can be found in the following collections

- Clinical diagnostic tests (1279)
- Radiology (1112)
- Radiology (diagnostics) (750)
- Connective tissue disease (4236)
- Degenerative joint disease (4622)
- Immunology (including allergy) (5120)
- Musculoskeletal syndromes (4931)
- Rheumatoid arthritis (3245)

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/