Clinical and epidemiological research

EXTENDED REPORT

Reliability of a consensus-based ultrasound score for tenosynovitis in rheumatoid arthritis

Esperanza Naredo,[1] Maria Antonietta D'Agostino,[2] Richard J Wakefield,[3] Ingrid Möller,[4] Peter V Balint,[5] Emilio Filippucci,[6] Annamaria Iagnocco,[7] Zunaid Karim,[8] Lene Terslev,[9] David A Bong,[10] Jesús Garrido,[11] David Martínez-Hernández,[12] George A W Bruyn,[13] on behalf of the OMERACT Ultrasound Task Force*

Handling editor Tore K Kvien

For numbered affiliations see end of article.

Correspondence to
Dr Esperanza Naredo, Department of Rheumatology, General Universitario Gregorio Marañón, Doctor Alvarez Sierra 4, 4° A, Madrid 28033, Spain;
enaredo@ser.es

* Outcome Measures in Rheumatology in Clinical Trials Ultrasound Task Force members: Sibel Aydin, Marina Backhaus, Artur Batcha, Paz Collado, Cristina Estrach, Frederique Gandjbakhch, Marwin Gutierrez, Hilda B. Hammer, Kei Ikeda, Frederick Joshua, Sandrine Jousse-Joulin, David Kane, Helen I. Keen, Juhani M. Koski, Peter Mandl, Levent Ozcakar, Carlos Pineda, Nanno Swen, Wolfgang A. Schmidt, Philip G Conaghan.

ABSTRACT

Objective To produce consensus-based scoring systems for ultrasound (US) tenosynovitis and to assess the intraobserver and interobserver reliability of these scoring systems in rheumatoid arthritis (RA).

Methods We undertook a Delphi process on US-defined tenosynovitis and US scoring system of tenosynovitis in RA among 35 rheumatologists, experts in musculoskeletal US (MSUS), from 16 countries. Then, we assessed the intraobserver and interobserver reliability of US in scoring tenosynovitis on B-mode and with a power Doppler (PD) technique. Ten patients with RA with symptoms in the hands or feet were recruited. Ten rheumatologists expert in MSUS blindly, independently and consecutively scored for tenosynovitis in B-mode and PD mode three wrist extensor compartments, two finger flexor tendons and two ankle tendons of each patient in two rounds in a blinded fashion. Intraobserver reliability was assessed by Cohen's κ. Interobserver reliability was assessed by Light's κ. Weighted κ coefficients with absolute weighting were computed for B-mode and PD signal.

Results Four-grade semiquantitative scoring systems were agreed upon for scoring tenosynovitis in B-mode and for scoring pathological peritendinous Doppler signal within the synovial sheath. The intraobserver reliability for tenosynovitis scoring on B-mode and PD mode was good (κ value 0.72 for B-mode; κ value 0.78 for PD mode). Interobserver reliability assessment showed good κ values for PD tenosynovitis scoring (first round, 0.64; second round, 0.65) and moderate κ values for B-mode tenosynovitis scoring (first round, 0.47; second round, 0.45).

Conclusions US appears to be a reproducible tool for evaluating and monitoring tenosynovitis in RA.

INTRODUCTION

Rheumatoid arthritis (RA) is a chronic inflammatory disease characterised by intra-articular and periarticular synovial inflammation (ie, synovial proliferation and angiogenesis).[1–3] Intra-articular synovitis can damage the cartilage, bones, capsule and ligaments,[1] and tenosynovitis can produce tendon adhesion and rupture with consequent severe joint function impairment.[4]

Accurate assessment of inflammation in RA is essential in rheumatological practice to reach therapeutic decisions and to evaluate the response to treatment. Within the past decade, technological improvements in ultrasound (US) image resolution of musculoskeletal structures have led to an increasingly important role for this imaging modality in the evaluation and monitoring of patients with RA and other inflammatory arthritides, based mainly on its greater ability compared with clinical examination to detect synovitis and tenosynovitis.[5–7] In addition, colour Doppler and power Doppler (PD) techniques can detect synovial blood flow, which is an indirect sign of inflammatory activity.[8–11] US is a routinely available, non-invasive, relatively inexpensive bedside technique, that can be repeated as many times as required at the time of consultation as it is characterised by high patient acceptability.

Despite the increasing implementation of US in clinical management of patients with RA, this imaging modality is still regarded as too dependent on the examiner to be incorporated into clinical trials. This is mainly because its accuracy depends on both acquisition and interpretation of US images. Since 2004 the Outcome Measures in Rheumatology in Clinical Trials (OMERACT) Ultrasound Task Force, an international collaborative group of musculoskeletal US (MSUS) experts, have examined the metric qualities of MSUS in RA and other inflammatory arthritides, according to criteria specified by the OMERACT filter.[12] Since then, the group effort has focused on assessing the reliability of MSUS for detecting and scoring inflammatory findings in RA. In 2005, the above group proposed preliminary definitions for inflammatory diseases,[13] including bone erosion, synovial fluid, synovial hypertrophy, enthesopathy and tenosynovitis. Over the past 6 years the group has developed a standardised scoring system for synovitis in RA which combines B-mode and PD on a 0–3 scale; this has demonstrated intraobserver and interobserver reliability and is applicable to all joints and consistent between US machines.[14] Now the group work is focusing, among other activities, on the metric properties of MSUS for evaluating tendon inflammation and tendon damage in RA,[14] planning first to generate a reliable scoring system for tenosynovitis components.

In a previous study[15] we tested the intraobserver and interobserver reliability of US for detecting B-mode tenosynovitis and tenosynovial PD signal according to the preliminary OMERACT definition.[3] Intraob-server reliability was moderate to good.

Interobserver reliability showed high substantial agreement but only fair κ results, partially owing to the low prevalence of tenosynovitis in the studied population.

The purposes of this study were the following: (1) to reach consensus on elementary lesions and definition of tenosynovitis in RA; (2) to generate agreed scoring systems for tenosynovitis on B-mode US and with PD US in RA; (3) to test the intraobserver and interobserver reliability of the developed scoring systems in patients with RA among rheumatologists expert in MSUS.

METHODS

Study design

This study comprised two sections: (1) consensus on the US definition of tenosynovitis and the US scoring system of tenosynovitis; (2) patient-based exercise to assess the reliability of US in scoring tenosynovitis.

Consensus process

This part of the study consisted of three phases: (1) a Delphi consensus process on US-defined tenosynovitis and US scoring system of tenosynovitis among experts in MSUS; (2) collection of US images of tendons representative of the tenosynovitis scores, agreed in the previous phase by the experts in MSUS, from patients with RA seen in their daily practice; (3) consensus on the assigned scores of the collected images of tendons that were shown during a meeting of experts before the reliability exercise on patients with RA.

Delphi process

We undertook a two-round Delphi consensus process through two consecutive written questionnaires sent by email to 35 rheumatologists, experts in MSUS, from 16 countries (ie, Australia, Denmark, Finland, France, Germany, Hungary, Ireland, Italy, Japan, Mexico, Netherlands, Norway, Spain, Turkey, UK and USA). They were selected because of their declared interest in participating in the OMERACT US task force on tenosynovitis.

The first questionnaire included 30 statements divided into three sections on the following topics: (1) US-defined normal tendons and related anatomical structures; (2) US-defined elementary lesions of tenosynovitis on B-mode and Doppler mode and definition of tenosynovitis and (3) US scoring systems for tenosynovitis on B-mode and Doppler mode.

The participants were asked to rate their level of agreement or disagreement with each statement according to a 1–5 Likert scale (1=strongly disagree; 5=strongly agree). Space for additional free comments was also included at the end of each statement. The participants were asked to respond within 1 month; after 2 weeks email reminders were sent to non-responders.

The second questionnaire included 14 statements divided into the above three sections. The second questionnaire and the results from the first questionnaire were sent by email to the respondents of the first questionnaire. The content of the second questionnaire consisted of several statements not previously agreed and some new statements generated from the comments supplied in the first questionnaire. Again, the participants were asked to rate their level of agreement or disagreement for each statement according to a 1–5 Likert scale (1=strongly disagree; 5=strongly agree). They were asked to respond within 1 month and after 2 weeks email reminders were also sent to the non-responders. The results from the

second questionnaire were sent to the respondents of both questionnaires.

Group agreement was considered if ≥75% of responders scored an item as either 4 or 5.

Collection of US images representative of the agreed scoring system for tenosynovitis

The respondents to both questionnaires were asked to collect US images of tendons in patients with RA that represented the tenosynovitis scores agreed in the Delphi process from their daily practice within 2 months. Each expert was asked to collect at least one US image in both transverse and longitudinal planes, representative of each B-mode and each Doppler grade of tenosynovitis. The images were sent by email to the investigator who coordinated the study (EN).

Consensus meeting

A meeting of the experts who participated in the reliability exercise was held the day before the actual exercise. During this meeting the above collected images were shown and the assigned scores were either agreed immediately or after discussion by the group. The final consensus on the tenosynovitis scoring system reached at this meeting was used in the reliability exercise on patients with RA the next day.

US reliability assessment

The second part of the study consisted of a reliability exercise on patients with RA carried out over 2 days in Madrid, Spain. The exercise lasted for 16 h divided into four sessions, a 4 h morning session and a 4 h afternoon session each day. This exercise included intraobserver and interobserver reliability assessment of US in scoring tenosynovitis on B-mode and with the PD technique.

Patients

Ten patients with RA according to the American College of Rheumatology 1987 criteria[16] with moderate or severe disease activity (ie, 28-joint count Disease Activity Score (DAS) 28>3.2) and symptoms in their hands or feet were recruited for the US reliability assessment (five patients for each day of the reliability exercise) from the outpatient rheumatology clinic (Hospital Universitario Severo Ochoa). The following data were recorded for each patient at study entry: demographics, RA characteristics, RA treatment and DAS28.

Each patient was randomly assigned to a scanner where they remained during both the morning and afternoon sessions. The more symptomatic hand and foot of each patient were selected for the US investigation.

The study was conducted in accordance with the Declaration of Helsinki and was approved by the local ethics committee of Hospital Universitario Severo Ochoa. Written informed consent was obtained from all patients before the study.

Ultrasonographers

The investigators comprised 10 rheumatologists with more than 10 years of experience in MSUS who had participated in the full consensus process.

Tendon assessed

In accordance with our previous study,[15] we selected the following hand/wrist and foot/ankle tendons with synovial sheath: wrist extensor compartment 2 (ie, extensor carpi radialis brevis and longus), 4 (ie, extensor digitorum communis and extensor indices propius) and 6 (ie, extensor carpi ulnaris),

finger flexor digitorum superficialis and profundus tendons 3 and 4 at the metacarpophalangeal level, tibialis posterior tendon and peroneal tendons (ie, peroneus longus and brevis). Flexor digitorum superficialis and profundus at the wrist were not selected because of the frequent variability in level of differentiation into distinct tendon slips and consequent anisotropy. Flexor pollicis longus, flexor carpi radialis and extensor compartment 1 were excluded owing to the proximity of the radial artery which can produce Doppler artefacts. Finger flexors 1, 2 and 5 were excluded owing to the almost constant presence of sesamoid bones that makes US evaluation of tenosynovitis difficult. The metacarpophalangeal level was selected for evaluating finger flexor tendons to avoid confusing pathological with normal distal tenosynovial vascularisation.

US examination

The US investigation was carried out using five commercially available real-time scanners (ie, two Mylab 70 X Vision, two Mylab 60 and one Mylab class C; Esaote, Genoa, Italy) equipped with multifrequency linear transducers (6–18 or 4–13 MHz). The B-mode and PD settings of each type of US machine were optimised for maximal image resolution and sensitivity to detect flow, respectively, in superficial anatomical areas by an application specialist before the reliability exercise. The ultrasonographers were not allowed to change these settings during the reliability exercise except for the position of the foci according to the depth of the scanned structure.

The 10 ultrasonographers blindly, independently and consecutively performed a longitudinal and transverse B-mode and PD US examination of the synovial sheath covered area of the selected tendons at the established locations in one hand and one foot of each patient in two rounds (ie, morning and afternoon) in a blinded fashion. The scanning technique had been previously standardised.[17] The selected tendons were scored for tenosynovitis on B-mode and PD mode according to the scoring systems agreed at the consensus meeting. During the morning and afternoon sessions the ultrasonographers were assigned to the US machines in a different order. They were unaware of the clinical details. Each ultrasonographer was given a maximum of 15 min to scan each patient and fill in a standardised report sheet with the US findings. Each examiner was informed of the selected anatomical region (right/left). An application specialist from the US company was near each machine to solve technical adjustment problems. A statistician (JG) was present to collect the filled score sheets after each US examination.

Statistical analysis

Statistical analysis was performed using SPSS V.15.0 (SPSS, Chicago, Illinois, USA). Simple summary statistics were calculated from the responses to the Delphi questionnaires. The results from the Delphi process were presented as the percentage of responders who scored a statement as either 4 or 5. Quantitative variables (ie, patient characteristics, prevalence of detected US abnormalities) were presented as the mean and range or as percentages.

Intraobserver reliability was assessed by Cohen's κ. Interobserver reliability was assessed by Light's κ (mean κ for all pairs of observations). Weighted κ coefficients with absolute weighting were computed for B-mode and PD signals.

κ Values of 0–0.20 were considered poor, >0.20–0.40 fair, >0.40–0.60 moderate, >0.60–0.80 good and >0.80–1 excellent.[18]

RESULTS

Delphi process

The response rate was 80% (28/35) from the first questionnaire and 100% (28/28) from the second questionnaire. There was group agreement after the two rounds about the following statements from the three sections:

1. US-defined normal tendons and related anatomical structures.
    1.1. Definition of normal tendon structure. Hyperechoic (relative to subdermal fat) fibrillar pattern (ie, hyperechoic parallel lines in longitudinal planes and hyperechoic dots in transverse planes) (agreement 100%).
    1.2. Definition of normal tendon synovial sheath. A thin regular hypoechoic (relative to tendon fibres) halo surrounding/thin regular hypoechoic lines above and below the tendon structure in transverse/longitudinal plane respectively at anatomical sites where synovial sheaths are known to exist and which can be distinguished from pulleys and retinaculae (agreement 85.7%).
    1.3. Definition of normal retinaculae (wrist and ankle level) and pulleys (finger flexor level). Focal hypoechoic (relative to tendon fibres) thickening of the peritendinous tendon sheath with fibrillar pattern in the area located perpendicular to the probe, at its expected normal anatomical location (agreement 88.9%).
2. US-defined elementary lesions of tenosynovitis on B-mode and Doppler mode and definition of tenosynovitis.
    2.1. Tenosynovitis can be defined on B-mode as abnormal anechoic and/or hypoechoic (relative to tendon fibres) tendon sheath widening which can be related to both the presence of tenosynovial abnormal fluid and/or hypertrophy (agreement 96.4%).
    2.2. Definition of tendon sheath effusion can be as follows: presence of abnormal anechoic or hypoechoic (relative to tendon fibres) material within the synovial sheath, either localised (eg, in the synovial sheath cul-de-sacs) or surrounding the tendon that is displaceable and seen in two perpendicular planes (agreement 89.3%).
    2.3. Definition of tenosynovial hypertrophy can be as follows: presence of abnormal hypoechoic (relative to tendon fibres) tissue within the synovial sheath that is not displaceable and poorly compressible and seen in two perpendicular planes (agreement 89.3%).
    2.4. Tenosynovitis can be characterised on Doppler mode by the presence of peritendinous Doppler signal within the synovial sheath, seen in two perpendicular planes, excluding normal feeding vessels (ie, vessels at the mesotenon or vinculae or vessels entering the synovial sheath from surrounding tissues) only if the tendon shows peritendinous synovial sheath widening on B-mode (agreement 78.6%).
3. US scoring system for tenosynovitis on B-mode and Doppler mode.
    3.1. The grade of tenosynovitis should be assessed in both longitudinal and transverse planes (agreement 82.1%).
    3.2. A four-grade semiquantitative scoring system (ie, grade 0, normal; grade 1, minimal; grade 2, moderate; grade 3, severe) can be used to score tenosynovitis on B-mode (agreement 85.7%).
    3.3. A four-grade semiquantitative scoring system (ie, grade 0, no Doppler signal; grade 1, minimal; grade 2, moderate; grade 3, severe) can be used to score pathological peritendinous Doppler signal within the synovial sheath (agreement 96.3%).

There was no group agreement about including an abnormal intratendinous Doppler signal in the elementary lesions of tenosynovitis because it could correspond to intratendinous tenosynovial angiogenesis (ie, invasive tenosynovium), vasodilatation of intratendinous feeding vessels or hypervascularisation in areas of tendon repair. However, 60.7% of the participants agreed that an abnormal intratendinous Doppler signal can be considered as elementary lesions of tenosynovitis if the tendon also shows an abnormal peritendinous Doppler signal within the synovial sheath. Nor was there group agreement on how to score an intratendinous Doppler signal, although 71.4% of the participants agreed that an abnormal peritendinous and intratendinous Doppler signal could be scored together on a four-grade semiquantitative scoring system. There was no group agreement about those B-mode and Doppler scoring systems based on the measurement of tenosynovial thickness on B-mode and the percentage of tenosynovial widening showing a Doppler signal.

Collection of US images of tendons and consensus meeting

Nineteen of 28 (68%) experts collected and sent the requested US images. During the meeting it was noted that a pathological intratendinous Doppler signal was taken into consideration in scoring tenosynovitis on Doppler mode by most experts. Based on this and the Doppler scores assigned by the experts, a scoring system for tenosynovitis on Doppler mode was proposed and agreed as follows: grade 0, no signal; grade 1, peritendinous focal signal within the widened synovial sheath (ie, signals in only one area of the widened sheath), seen in two perpendicular planes, excluding normal feeding vessels; grade 2, peritendinous multifocal signal within the widened synovial sheath (ie, signals in more than one area of the widened sheath), seen in two perpendicular planes, excluding normal feeding vessels; grade 3, peritendinous diffuse signal within the widened synovial sheath (ie, signals filling most of the widened sheath), seen in two perpendicular planes, excluding normal feeding vessels. If in addition to an abnormal peritendinous (ie, intra-sheath) signal there was an abnormal intratendinous signal seen in two perpendicular planes (ie, excluding intratendinous small isolated signals that can correspond to normal feeding vessels detectable by US), then grades 1 and 2 would be increased by one point.

Illustrative images of the B-mode and PD scores of tenosynovitis are shown in figures 1A–D and 2A–D, respectively.

US reliability assessment
Patient characteristics

Patients comprised eight women and two men with mean age 58.1 (35–69) years and mean disease duration 10.1 (3–18) years. Rheumatoid factor and anti-cyclic citrullinated protein antibodies were positive in eight patients. Six patients had radiological erosions. Mean (range) DAS28 was 4.5 (3.3–5.5). All patients were receiving synthetic disease-modifying antirheumatic drugs (DMARDs), alone (five patients) or in combination with biological DMARDs (three patients).
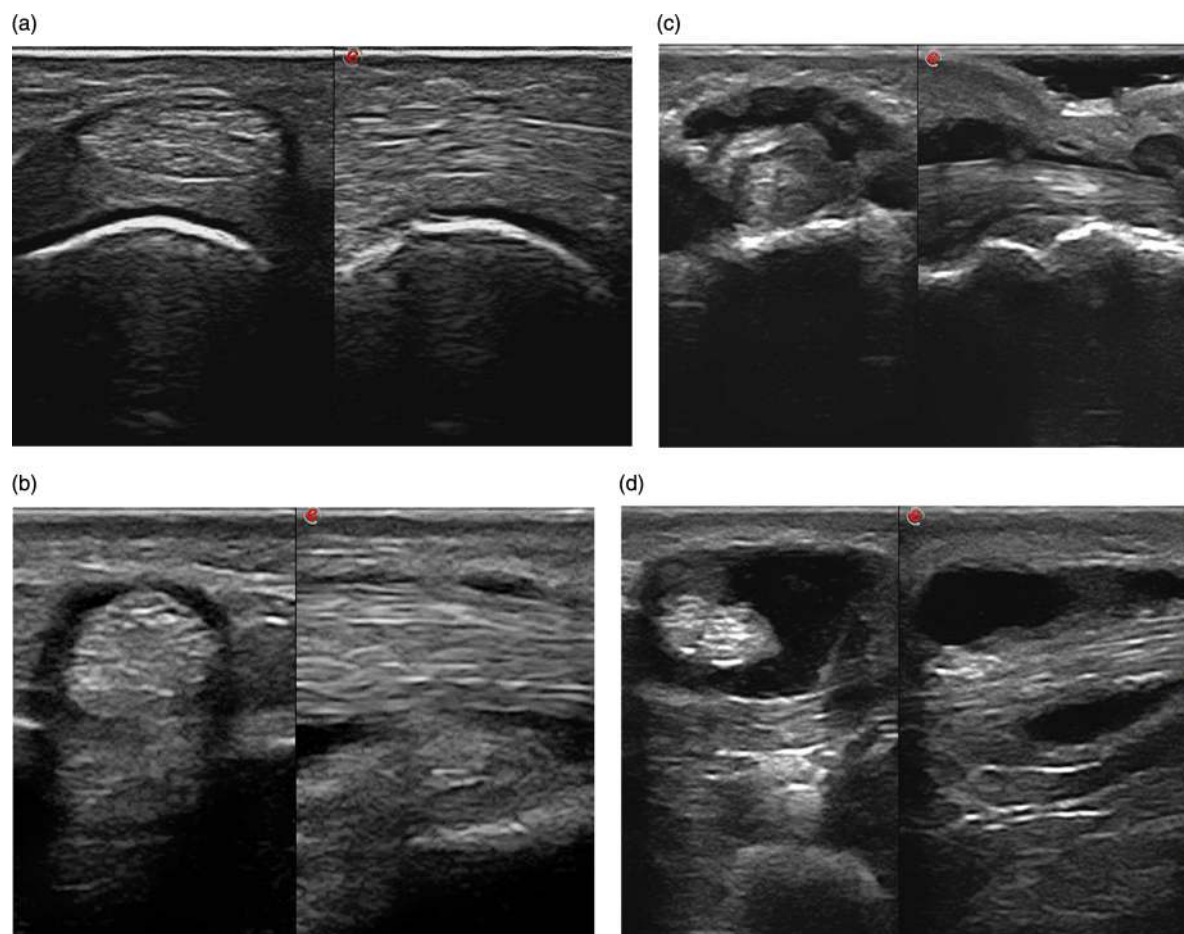
Figure 1 Ultrasound score of tenosynovitis in B-mode in transverse (left) and longitudinal (right) scan. (A) Grade 0 (finger flexors); (B) grade 1 (finger flexor); (C) grade 2 (finger flexor); (D) grade 3 (finger flexor).
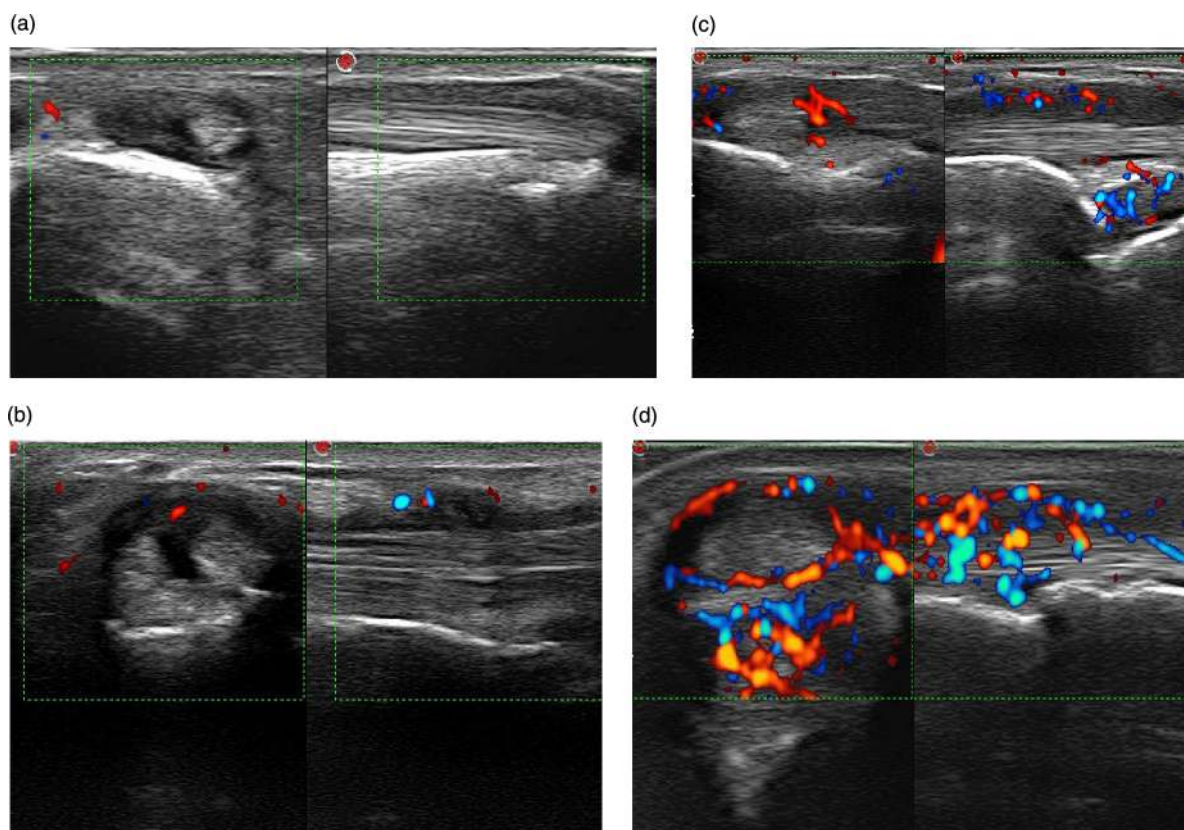
Figure 2 Ultrasound score of tenosynovitis in power Doppler mode in transverse (left) and longitudinal (right) scan. (A) Grade 0 (extensor carpi ulnaris); (B) grade 1 (extensor carpi radialis longus and brevis); (C) grade 2 (extensor carpi ulnaris); (D) grade 3 (extensor carpi ulnaris).

Prevalence of US abnormalities

Considering the two rounds, the mean prevalence of US-detected tenosynovitis on B-mode was 43% of tendons. The distribution of the assigned scores was as follows: grade 1 in 29% of tendons, grade 2 in 10% and grade 3 in 4%. In 45% of B-mode tenosynovitis an abnormal PD signal was detected and the distribution of scores was 15% of tendons for grade 1, 18% for grade 2 and 12% for grade 3.

Intraobserver reliability

The κ values and CI for the intraobserver concordance are shown in table 1. Both B-mode and Doppler scores showed good intraobserver reliability.

Interobserver reliability

Table 2 displays the κ values and CI for the interobserver concordance in the two US rounds. For the PD score the interobserver reliability was good, although for B-mode tenosynovitis it was only moderate.

DISCUSSION

The pronounced clinical and prognostic relevance of the involvement of tendons in RA makes their early and accurate assessment of utmost importance for therapeutic decisions that can prevent irreversible structural damage. The high image resolution and Doppler sensitivity offered by US technology within the past decade make this imaging modality a potentially powerful tool for evaluating superficial tendons, particularly those of the hands and feet which are target anatomical areas for inflammation and damage in RA.

Although there are no published studies on the concurrent validity of US versus a reference method such as surgical findings or histology for assessing tenosynovitis, a high specificity of US as compared with MRI in detection of hand and foot tenosynovitis has been proved in some studies.[19 20] US-detected tenosynovitis evaluated by both B-mode and PD US has demonstrated sensitivity to change in patients with RA who had begun treatment with a biological agent.[21] Tenosynovitis (ie, extensor carpi ulnaris) has shown predictive value in relation to radiological and MRI progression of bone damage in RA.[22] The reproducibility between two ultrasonographers of US acquisition and interpretation of inflammatory tendon lesions has been reported in some single-centre studies.[20 23 24] Despite the promising results above, the paucity of reported data on the metric properties of ultrasound assessment of

Table 1 Intraobserver reliability results

US abnormality	Weighted κ coefficients (mean)	95% CI
B-mode tenosynovitis	0.72	0.66 to 0.78
PD signal	0.78	0.72 to 0.84

PD, power Doppler; US, ultrasound.

Table 2 Interobserver concordance in the two US rounds

US abnormality	Weighted κ coefficients	95% CI
B-mode tenosynovitis, round 1	0.47	0.43 to 0.50
B-mode tenosynovitis, round 2	0.45	0.41 to 0.48
PD signal, round 1	0.64	0.61 to 0.67
PD signal, round 2	0.65	0.62 to 0.68

PD, power Doppler; US, ultrasound.

tendons in RA and other chronic arthritides suggests that these properties require attention.[25] This has led the MSUS OMERACT group to start research into the metric properties of US in assessment of tenosynovitis and tendon damage in RA.[14] This project has started with standardisation of the tendon scanning technique, consensus on elementary lesions, definition and scoring system of tenosynovitis and assessment of US reliability in detecting and scoring tenosynovitis in RA.

To the best of our knowledge, this is the first study that has assessed the multi-examiner reproducibility of US in scoring tenosynovitis after a consensus process among international experts. In addition, we have proposed a new scoring system for the Doppler component of tenosynovitis. This scoring system is based on the extension of Doppler signals within the widened synovial sheath excluding those in characteristic locations of feeding blood supply. In addition, confluent intratendinous Doppler signals which are not detectable by US in normal tendons were also included in the scoring system.

Some previous studies on inflammatory tenosynovitis have used semiquantitative US scoring for B-mode[21] [26–34] and/or Doppler mode.[21] [27–32] [34] [35] These scores were purely subjective or based on the measurement of tenosynovial thickness on B-mode and percentage of tenosynovial widening showing PD or colour Doppler flow. Of particular note is the great heterogeneity of the morphology of US tenosynovitis in B-mode images due to anatomical details of each tendon/group of tendons, and also in RA depending on the integrity of retinaculae and pulleys.[36] Consequently, the distribution of tenosynovial effusion and tenosynovial proliferation and its pathological vascularisation is highly variable throughout the area of the tendon covered by the synovial sheath. This probably led the majority of the expert panel to agree on subjective scores for tenosynovitis in both B-mode and Doppler made owing to the difficulty in standardising quantitative scoring systems based on measures of tenosynovial diameters or number of Doppler signals in relation to the tenosynovial area. Despite the qualitative nature of our scoring systems, we obtained good intraobserver reliability for both B-mode and Doppler tenosynovitis scoring, and acceptable (ie, B-mode tenosynovitis) to good (ie, Doppler tenosynovitis) multi-examiner reliability. The wide experience in MSUS of the ultrasonographers, together with a standardised scanning technique, possibly contributed to the good results. The interobserver reliability for B-mode tenosynovitis showed poorer results than Doppler reliability, probably because of the extreme subjectivity of the scoring system used. However, we can speculate that the Doppler component of tenosynovitis is probably more important than the B-mode component in assessing inflammatory activity and predicting damage, as has been shown for intra-articular synovitis.[31] [37–39]

Some limitations in our study should be noted. We tested the US reliability in a small group of patients with active RA. Unfortunately, this type of international meeting for real-time scanning of patients is not feasible beyond a few days. In addition, we did not use MRI or any other comparator for our US findings. However, this study was not a validation study but a reliability study as a first step in implementing US for the assessment and monitoring of tenosynovitis in clinical practice and trials.

In conclusion, our results seem to be sufficiently promising to support the reproducibility of US in scoring tenosynovitis in RA, especially for Doppler assessment of tendon inflammatory activity. Further studies should confirm our results in other RA populations and evaluate other metric properties of US in tenosynovitis assessment in RA and other inflammatory joint diseases.

Author affiliations
[1]Department of Rheumatology, Hospital General Universitario Gregorio Marañón, Complutense University, and Hospital Universitario Severo Ochoa, Madrid, Spain
[2]Department of Rheumatology, Université Paris Ouest-Versailles-Saint Quentin en Yvelines, Hôpital Ambroise Paré, APHP, Boulogne-Billancourt, France
[3]Department of Rheumatology, Section of Musculoskeletal Disease, University of Leeds, Leeds, UK
[4]Department of Rheumatology, Instituto Poal de Reumatologia, Barcelona, Spain
[5]3rd Rheumatology Department, National Institute of Rheumatology and Physiotherapy, Budapest, Hungary
[6]Department of Rheumatology, Clinica Reumatologica, Universitá Politecnica delle Marche, Jesi, Ancona, Italy
[7]Department of Rheumatology, Sapienza Universitá di Roma, Rome, Italy
[8]Department of Rheumatology, Mid Yorkshire NHS trust, Wakefield, UK
[9]Department of Rheumatology, Copenhagen University Hospital at Glostrup, Copenhagen, Denmark
[10]Department of Rheumatology, Instituto Poal, Barcelona, Spain
[11]Department of Social Psychology and Methodology, Autonoma University, Madrid, Spain
[12]Department of Preventive Medicine and Public Health, Complutense University, Madrid, Spain
[13]Department of Rheumatology, MC Groep, Lelystad, The Netherlands

REFERENCES

1. **Pap T,** Distler O. Linking angiogenesis to bone destruction in arthritis. *Arthritis Rheum* 2005;**52**:1346–8.
2. **Boutry N,** Larde A, Lapegue F, *et al.* Magnetic resonance imaging appearance of the hands and feet in patients with early rheumatoid arthritis. *J Rheumatol* 2003;**30**:671–9.
3. **Eshedı E,** Feist CE, Althoff , *et al.* Tenosynovitis of the flexor tendons of the hand detected by MRI: an early indicator of rheumatoid arthritis. *Rheumatology* 2009;**48**:887–91.
4. **Jain A,** Nanchahal J, Troeberg L, *et al.* Production of cytokines, vascular endothelial growth factor, matrix metalloproteinases and tissue inhibitor of metalloproteinases by tenosynovium demonstrates its potential for tendon destruction in rheumatoid arthritis. *Arthritis Rheum* 2001;**44**:1754–60.
5. **Wakefield RJ,** Green MJ, Marzo-Ortega H, *et al.* Should oligoarthritis be reclassified? Ultrasound reveals a high prevalence of subclinical disease. *Ann Rheum Dis* 2004;**63**:382–5.
6. **Grassi W,** Filippucci E, Farina A, *et al.* Sonographic imaging of tendons. *Arthritis Rheum* 2000;**43**:969–76.
7. **Filippucci E,** Gabba A, Di Geso L, *et al.* Hand tendon Involvement in rheumatoid arthritis: an ultrasound study. *Semin Arthritis Rheum* 2012;**41**:752–60.
8. **Naredo E,** Bonilla G, Gamero F, *et al.* Assessment of inflammatory activity in rheumatoid arthritis: a comparative study of clinical evaluation with grey-scale and power Doppler ultrasonography. *Ann Rheum Dis* 2005;**64**:375–81.
9. **Walther M,** Harms H, Krenn V, *et al.* Correlation of power Doppler sonography with vascularity of the synovial tissue of the knee joint in patients with osteoarthritis and rheumatoid arthritis. *Arthritis Rheum* 2001;**44**:331–8.
10. **Szkudlarek M,** Court-Payen M, Strandberg C, *et al.* Power Doppler ultrasonography for assessment of synovitis in the metacarpophalangeal joints of patients with rheumatoid arthritis: a comparison with dynamic magnetic resonance imaging. *Arthritis Rheum* 2001;**44**:2018–23.
11. **Terslev L,** Torp-Pedersen S, Savnik A, *et al.* Doppler ultrasound and magnetic resonance imaging of synovial inflammation of the hand in rheumatoid arthritis. A comparative study. *Arthritis Rheum* 2003;**48**:2434–41.
12. **Boers M,** Brooks P, Strand CV, *et al.* The OMERACT filter for outcome measures in rheumatology. *J Rheumatology* 1998;**25**:198–9.
13. **Wakefield R,** Balint PV, Szkudlarek M, *et al.* Musculoskeletal ultrasound including definitions for ultrasonographic pathology. *J Rheumatol* 2005;**32**:2485–7.

14. **Naredo E,** Wakefield RJ, Iagnocco A, et al. The OMERACT ultrasound task force. Summary of advances and priorities. J Rheumatol 2011;**38**:2063–7.

15. **Bruyn GW,** Möller I, Garrido J, et al. Reliabiliity testing of tendon disease using two different scanning methods in patients with rheumatoid arthritis. First step towards an ultrasonography scoring index. Rheumatology (Oxford) 2012;**51**:1655–61.

16. **Arnett FC,** Edworthy SM, Bloch DA, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988;**31**:315–24.

17. **Backhaus M,** Burmester GR, Gerber T, et al., the Working Group for Musculoskeletal Ultrasound in the EULAR Standing Committee on International Clinical Studies Including Therapeutic Trials. Guidelines for musculoskeletal ultrasound in rheumatology. Ann Rheum Dis 2001;**60**:641–9.

18. **Landis JR,** Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;**33**:159–74.

19. **Wakefield RJ,** O'Connor PJ, Conaghan PG, et al. Finger tendon disease in untreated early rheumatoid arthritis: a comparison of ultrasound and magnetic resonance imaging. Arthritis Rheum 2007;**57**:1158–64.

20. **Wakefield RJ,** Freeston JE, O'Connor P, et al. The optimal assessment of the rheumatoid arthritis hindfoot: a comparative study of clinical examination, ultrasound and high field MRI. Ann Rheum Dis 2008;**67**:1678–82.

21. **Hammer HB,** Kvien TK. Ultrasonography shows significant improvement in wrist and ankle tenosynovitis in rheumatoid arthritis patients treated with adalimumab. Scand J Rheumatol 2011;**40**:178–82.

22. **Lillegraven S,** Bøyesen P, Hammer HB, et al. Tenosynovitis of the extensor carpi ulnaris tendon predicts erosive progression in early rheumatoid arthritis. Ann Rheum Dis 2011;**70**:2049–50.

23. **Wiell C,** Szkudlarek M, Hasselquist M, et al. Ultrasonography, magnetic resonance imaging, radiography and clinical assessment of inflammatory and destructive changes in fingers and toes of patients with psoriatic arthritis. Arthritis Res Ther 2007;**9**:R119. doi:10.1186/ar2327

24. **Micu MC,** Serra S, Fodor D, et al. Inter-observer reliability of ultrasound detection of tendon abnormalities at the wrist and ankle in patients with rheumatoid arthritis. Rheumatology (Oxford) 2011;**50**:1120–4.

25. **Alcalde M,** D'Agostino MA, Bruyn GAW, et al. A systematic literature review of ultrasound definitions, scoring systems and validity according to the OMERACT filter for tendon lesion in rheumatoid arthritis and other inflammatory joint disease. Rheumatology (Oxford) 2012;**51**:1246–60.

26. **Lehtinen A,** Paimela L, Kreula J, et al. Painful ankle region in rheumatoid arthritis. Analysis of soft-tissue changes with ultrasonography and MR imaging. Acta Radiol 1996;**37**:572–7.

27. **Hoving JL,** Buchbinder R, Hall S, et al. A comparison of magnetic resonance imaging, sonography and radiography of the hand in patients with early rheumatoid arthritis. J Rheumatol 2004;**31**:663–75.

28. **Milosavljevic J,** Lindqvist U, Elvin A. Ultrasound and power Doppler evaluation of the hand and wrist in patients with psoriatic arthritis. Acta Radiol 2005;**46**:374–85.

29. **Iagnocco A,** Perella C, Naredo E, et al. Etanercept in the treatment of Rheumatoid Arthritis: clinical follow-up over 1 year by ultrasonography. Clin Rheumatol 2008;**27**:491–6.

30. **Iagnocco A,** Filippucci E, Perella C, et al. Clinical and ultrasonographic monitoring of response to adalimumab treatment in rheumatoid arthritis. J Rheumatol 2008;**35**:35–40.

31. **Naredo E,** Möller I, Cruz A, et al. Power doppler ultrasound monitoring of response to anti-tumor necrosis factor therapy in patients with rheumatoid arthritis. Arthritis Rheum 2008;**58**:2248–56.

32. **Naredo E,** Rodriguez M, Campos C, et al. Validity, reproducibility and responsiveness of a twelve-joint simplified power Doppler ultrasonographic assessment of joint inflammation in Rheumatoid Arthritis. Arthritis Rheum 2008;**59**:215–22.

33. **Haavardsholm EA,** Ostergaard M, Hammer HB, et al. Monitoring anti-TNFα treatment in Rheumatoid Arthritis: responsiveness of magnetic resonance imaging and ultrasonography of the dominant wrist joint compared with conventional measures of disease activity and structural damage. Ann Rheum Dis 2009;**68**:1572–79.

34. **Klauser A,** Franz M, Arora R, et al. Detection of vascularity in wrist tenosynovitis: power Doppler-Ultrasound compared to contrast-enhanced gray-scale ultrasound. Arthritis Res Ther 2010;**12**:R209.

35. **Backhaus M,** Ohrndorf S, Kellner H, et al. Evaluation of a novel 7-joint ultrasound score in daily rheumatologic practice: a pilot project. Arthritis Rheum 2009;**61**:1194–201.

36. **Benjamin M,** Kaiser E, Milz S. Structure-function relationships in tendons: a review. J Anat 2008;**212**:211–28.

37. **Naredo E,** Collado P, Cruz A, et al. Longitudinal power Doppler ultrasonographic assessment of joint inflammatory activity in early rheumatoid arthritis: predictive value in disease activity and radiologic progression. Arthritis Rheum 2007;**57**:116–24.

38. **Brown AK,** Conaghan PG, Karim Z, et al. An explanation for the apparent dissociation between clinical remission and continued structural deterioration in rheumatoid arthritis. Arthritis Rheum 2008;**58**:2958–67.

39. **Døhn UM,** Ejbjerg B, Boonen A, et al. No overall progression and occasional repair of erosions despite persistent infl ammation in adalimumabtreated rheumatoid arthritis patients: results from a longitudinal comparative MRI, ultrasonography, CT and radiography study. Ann Rheum Dis 2011;**70**:252–8.

Reliability of a consensus-based ultrasound score for tenosynovitis in rheumatoid arthritis

Esperanza Naredo, Maria Antonietta D'Agostino, Richard J Wakefield, Ingrid Möller, Peter V Balint, Emilio Filippucci, Annamaria Iagnocco, Zunaid Karim, Lene Terslev, David A Bong, Jesús Garrido, David Martínez-Hernández, George A W Bruyn and on behalf of the OMERACT Ultrasound Task Force*

Updated information and services can be found at:

http://ard.bmj.com/content/72/8/1328

These include:

References	This article cites 39 articles, 18 of which you can access for free at: **http://ard.bmj.com/content/72/8/1328**#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 Open access (562) Musculoskeletal syndromes (4931) Clinical diagnostic tests (1279) Connective tissue disease (4236) Degenerative joint disease (4622) Immunology (including allergy) (5120) Radiology (1112) Radiology (diagnostics) (750) Rheumatoid arthritis (3245)

Notes