EXTENDED REPORT

Interobserver reliability of rheumatologists performing musculoskeletal ultrasonography: results from a EULAR ''Train the trainers'' course

A K Scheel, W A Schmidt, K-G A Hermann, G A Bruyn, M A D'Agostino, W Grassi, A Iagnocco, J M Koski, K P Machold, E Naredo, H Sattler, N Swen, M Szkudlarek, R J Wakefield, H R Ziswiler, D Pasewaldt, C Werner, M Backhaus

See end of article for authors' affiliations
....

Correspondence to:
Dr A K Scheel, Department of Medicine, Nephrology and Rheumatology, Robert-Koch-Strasse 40, D-37075 Göttingen, Germany; ascheel@gwdg.de

Objective: To evaluate the interobserver reliability among 14 experts in musculoskeletal ultrasonography (US) and to determine the overall agreement about the US results compared with magnetic resonance imaging (MRI), which served as the imaging ''gold standard''.

Methods: The clinically dominant joint regions (shoulder, knee, ankle/toe, wrist/finger) of four patients with inflammatory rheumatic diseases were ultrasonographically examined by 14 experts. US results were compared with MRI. Overall agreements, sensitivities, specificities, and interobserver reliabilities were assessed.

Results: Taking an agreement in US examination of 10 out of 14 experts into account, the overall κ for all examined joints was 0.76. Calculations for each joint region showed high κ values for the knee (1), moderate values for the shoulder (0.76) and hand/finger (0.59), and low agreement for ankle/toe joints (0.28). κ Values for bone lesions, bursitis, and tendon tears were high (κ = 1). Relatively good agreement for most US findings, compared with MRI, was found for the shoulder (overall agreement 81%, sensitivity 76%, specificity 89%) and knee joint (overall agreement 88%, sensitivity 91%, specificity 88%). Sensitivities were lower for wrist/finger (overall agreement 73%, sensitivity 66%, specificity 88%) and ankle/toe joints (overall agreement 82%, sensitivity 61%, specificity 92%).

Conclusion: Interobserver reliabilities, sensitivities, and specificities in comparison with MRI were moderate to good. Further standardisation of US scanning techniques and definitions of different pathological US lesions are necessary to increase the interobserver agreement in musculoskeletal US.

R ecent technological advances have made musculo-skeletal ultrasonography (US) a promising tool for the assessment of patients with rheumatic diseases. US has strengths in visualising soft tissue inflammatory processes and bone erosions in different joints.[1–8] Synovitis can be detected in joints, bursae, and tendon sheaths.[9–11] In addition, multiple studies have shown that Doppler US allows the visualisation of vessels in arthritic joints.[5 11–15]

Because US is relatively quick to perform and an easily accessible bedside procedure with low running costs,[16] it is highly relevant to consider the validity of musculoskeletal US measures. Despite increasing evidence of the potential applications of US in the evaluation of arthritic joints, data on its accuracy and reproducibility remain limited.[17] The current validity measures for US in rheumatoid arthritis (RA) have recently been summarised, concluding that further evaluation of its discriminant validity and reproducibility is needed.[18] It has been stated, but underinvestigated, that musculoskeletal US is highly operator dependent. Inter- and intraobserver variations have been tested in only a few studies, however, while interobserver agreement has only been assessed between two observers.[6 17 19] Most studies investigating interobserver variation for musculoskeletal US, radiography, and magnetic resonance imaging (MRI) have assessed the interpretation of images only and not image acquisition.[4 6]

Since 1998, EULAR US courses have been organised. The trainers of these courses are rheumatologists from many European countries. They are considered experts in the field, yet have varied educational backgrounds. We therefore decided to perform a ''Train the trainers'' study, which was held before the 8th EULAR US course conducted in association with the 2004 annual EULAR congress in Berlin (Germany), to contribute to validity in musculoskeletal US. All 14 trainer rheumatologists who comprised the faculty of this US course participated in the study to examine three main issues: firstly, to collect data on standardisation of musculoskeletal US by evaluation of a detailed questionnaire; secondly, to evaluate the interobserver reliability between the sonographers; and thirdly, to determine the overall agreement of US results compared with MRI, which served as the imaging ''gold standard''.

PATIENTS AND METHODS

The 14 trainers represented 10 different countries (1 Austria, 1 Denmark, 1 Finland, 1 France, 3 Germany, 1 UK, 2 Italy, 2 The Netherlands, 1 Spain, 1 Switzerland). Their mean age was 44 years (range 33–56). The study was divided into two parts. Firstly, a detailed questionnaire on general information about the examiners, the equipment used, and the most commonly examined joints was mailed to all trainers before the meeting and evaluated anonymously. Additionally, questions about standardisation (transducer orientation, documentation, positioning, and adherence of standard scans according to EULAR recommendations) were asked.

....

Secondly, each trainer performed practical US examinations of predefined joints of four patients who had previously been assessed by MRI.

Practical scanning

All US examinations were performed anonymously and separately by all the experts. Each sonographer was given a maximum of 10 minutes for US examination for each joint region. To ensure standardised documentation, each participant was given a report sheet that listed possible pathological findings against which a yes or no tick box was used, indicating the presence or absence of each particular finding.

Patients

Four patients were recruited from the outpatient clinics of the Medical Centre for Rheumatology, Berlin-Buch (patients 1 and 2) and the Department of Rheumatology and Clinical Immunology, Charité-University-Hospital, Berlin (patients 3 and 4) 1 week before the study. Drug treatment was kept constant from the day of recruitment. Based on the questionnaire, the most often examined joint regions were selected: shoulder, knee, ankle/toe, and wrist/finger. The predefined joint region of the following patients was ultrasonographically examined by each expert:

Patient 1: Woman (aged 38 years) with erosive RA (disease duration 11 years) underwent examination of the right shoulder. She received prednisolone 7.5 mg/day, methotrexate 15 mg/week, and infliximab 300 mg every 4 weeks; C reactive protein (CRP) and erythrocyte sedimentation rate (ESR) were normal.

Patient 2: Man (69 years) with remitting seronegative symmetrical synovitis with pitting oedema (RS3PE) syndrome (disease duration 3 months) underwent examination of the right wrist and finger joints. He received rofecoxib (25 mg/day), CRP was slightly raised (7 mg/l, normal<5), and ESR was normal.

Patient 3: Man (aged 59 years) with gout and swelling and pain for 4 weeks underwent examination of the right knee. He received celecoxib 100 mg and colchicine 0.5 mg twice daily, allopurinol 300 mg once daily; CRP and ESR were normal, uric acid was raised (600 μmol/l, normal 180–410).

Patient 4: Man (aged 28 years) with reactive arthritis of the left ankle and toe joints with pain and swelling of both knees and the left ankle. He received rofecoxib (25 mg/day) and antibiotic treatment (doxycycline 200 mg once daily). The patient was HLA-B27 positive; ESR (61/84 mm/1st h) and CRP (60 mg/l) were greatly raised.

Ultrasonography

We employed a linear probe for all investigations (LA 523, 13–4 MHz; length of the probe, 45 mm; Esaote Technos MPX; Esaote SpA, Genova, Italy). Scanner settings were

Table 1 Joints mainly examined by expert ultrasonographers

Joint	1st	2nd	3rd
Shoulder	9	0	1
Elbow	0	0	0
Wrist	0	3	5
Finger joints	2	1	1
Hip	0	1	2
Knee	2	5	4
Ankle	1	3	1
Toe joints	0	1	0
Total	14	14	14

1st, 2nd, 3rd: most frequently, second and third frequently examined joint.

Table 2 US examination of individual joints transducers

Joint	Array	Frequency*	Range (MHz)
Shoulder	14 linear	7.5–10	(4–10)
Elbow	14 linear	7.5–10	(4–13)–(10–14)
Wrist	14 linear	10	(4–13)–(8–16)
Finger joints	14 linear	10	(5–10)–(10–22)
Hip	10 linear	5–10	(5)–(5–12)
	4 curved		
Knee	14 linear	7.5–10	(4–13)–(10–14)
Ankle	14 linear	7.5–10	(4–13)–(8–16)

*All participants use frequencies within this range.

Table 3 Scanning of individual joints: standard scans as demanded by EULAR*

Joint	Standard scans*	Mean scans performed No (%)
Shoulder	9	6.4 (71)
Elbow	8	6.6 (83)
Wrist	7	7 (100)
Finger	6	5.3 (88)
Hip	3	2.5 (83)
Knee	10	8.2 (82)
Ankle	8	7.4 (93)

uniform for all measurements: frequency setting, 12.5 MHz for wrist/finger and ankle/toe, 10 MHz for shoulder and knee investigations; B mode gain, 100%; one focus point position in the region of measurement. An introduction to the US device was given to the observers before US examinations. Two application specialists from Esaote were present to help in case of problems with machine adjustments during the investigation.

Four joint regions were examined by each ultrasonographer:

Shoulder joint: examination of the biceps tendon, rotator cuff (subscapularis, supraspinatus, and infraspinatus tendons), glenohumeral joint cavity, humeral bone surface, and subacromial-subdeltoid bursa, was required.

Wrist/finger joints: evaluation was performed of the right wrist and metacarpophalangeal (MCP) II joint as well as tenosynovitis of the extensor carpi ulnaris, flexor, and extensor tendons II.

Knee joint: the suprapatellar recess, infrapatellar bursae, popliteal cysts, and patella ligament were evaluated.

Ankle joint: US was performed for the tibiotalar and talonavicular joints as well as the Achilles tendon, plantar fascia and the extensor, flexor, and peroneus tendons.

Two sonographers (MB, WAS), who were unaware of the other results, performed the US examination exactly as described above 4 days before the "practical examination", and their findings were included in the final evaluation. One hour before the "practical examination" performed by the 12 other experts, patients were re-examined by MB and WAS to ensure that the pathological findings were still present.

Magnetic resonance imaging

MRI of the above mentioned joint regions, which served as the imaging "gold standard", was performed in our four patients by a musculoskeletal radiologist (K-GAH). MRI was performed with a 1.5 T whole body magnet (MAGNETOM Sonata Maestro Class, Siemens AG Medical Solutions, Erlangen, Germany) 4 days before the course. Standard imaging protocols were applied for all joints:

Shoulder: the protocol comprised T_1 weighted fast spin echo (T1/FSE) sequences (slice thickness (SL) 4 mm) in axial and oblique coronal views, a short τ inversion recovery (STIR)

Table 4 Practical examinations

Anatomical structure	Pathology	MRI*	Ultrasound findings	
			No	%
A Shoulder joint (right)				
Long biceps tendon	Tenosynovitis	Yes	13/14	93
	Tear	No	13/14	93
	Dislocation	No	14/14	100
Subscapularis tendon	Tear	No	12/14	86
Supraspinatus tendon	Tear	Yes	12/14	86
Infraspinatus tendon	Tear	No	11/14	79
Humeral head	Erosion	Yes	14/14	100
Shoulder	Glenohumeral joint, synovitis/effusion	Yes	7/14	50
	Subdeltoid bursitis	Yes	14/14	100
Acromioclavicular joint	Synovitis/effusion	Yes	4/14	29
Overall agreement			11.4/14	81
Sensitivity/specificity (%)	76.2/89.3			
B Wrist/finger joints (right)				
Wrist	Synovitis/effusion	Yes	12/14	86
	Dorsal tenosynovitis	Yes	11/14	79
	Palmar tenosynovitis	Yes	7/14	50
	Extensor carpi ulnaris, tenosynovitis	Yes	10/14	71
MCP II	Synovitis/effusion	Yes	7/14	50
	Erosion	No	13/14	93
	Osteophyte	No	12/14	86
Flexor tendon II	Tenosynovitis	Yes	8/14	57
Extensor tendon II	Tenosynovitis	No	12/14	86
Overall agreement			10.2/14	73
Sensitivity/specificity (%)	65.5/88.1			
C Knee joint (right)				
Suprapatellar recessus	Effusion	Yes	14/14	100
	Synovial hypertrophy	Yes	11/14	79
Bone surface at joint space	Osteophyte medial	No	11/14	79
	Osteophyte lateral	No	10/14	71
	Erosion medial	No	11/14	79
	Erosion lateral	No	11/14	79
Patella tendon	Tendinitis	No	14/14	100
	Enthesitis	No	13/14	93
Anterior bursae	Prepatellar bursitis	No	14/14	100
	Superficial infrapatellar	No	14/14	100
Posterior bursae	Baker's cyst	Yes	13/14	93
Overall agreement			12.4/14	88
Sensitivity/specificity (%)	90.5/87.5			
D Ankle and foot joints (right)				
Ankle (tibiotalar)	Synovitis/effusion	Yes	8/14	57
Talonavicular joint	Synovitis/effusion	Yes	8/14	57
	Erosion	No	12/14	86
	Osteophyte	No	12/14	86
Extensor tendons	Tenosynovitis	Yes	5/14	36
Flexor tendons	Tenosynovitis	Yes	12/14	86
Peroneus tendons	Tenosynovitis	Yes	10/14	71
Achilles tendon	Tendinitis	No	14/14	100
	Paratendonitis	No	14/14	100
	Retrocalcaneal bursitis	No	13/14	93
Posterior aspect of calcaneus	Erosion/irregularity	No	12/14	86
	Calcaneal spur	No	11/14	79
Plantar aspect of calcaneus	Erosion/irregularity	No	13/14	93
	Calcaneal spur	No	14/14	100
	Fasciitis	No	14/14	100
Overall agreement			11.5/14	82
Sensitivity/specificity (%)	61.4/92.1			

Overall total agreement (45 sonographic findings) 81.6% (specificity 89.7%, sensitivity 71.4%).
*Serving as gold standard.

sequence in oblique coronal view, and T_1 weighted FSE sequences with fat saturation after application of gadolinium diethylenetriamine pentaacetic acid (T_1/FSE-Gd).

Wrist/finger joints: coronal and axial T_1/FSE (SL 3 mm), a coronal STIR, and coronal and axial T_1/FSE-Gd sequences were used.

Knee: coronal T_1/FSE (SL 4 mm), coronal STIR, sagittal proton density/T_2 weighted sequence, and coronal, sagittal, and axial T_1/FSE-Gd sequences were used.

Ankle joint: Sagittal and coronal T_1/FSE (SL 4 mm), sagittal STIR, and coronal, sagittal, and axial T_1/FSE-Gd sequences were used.
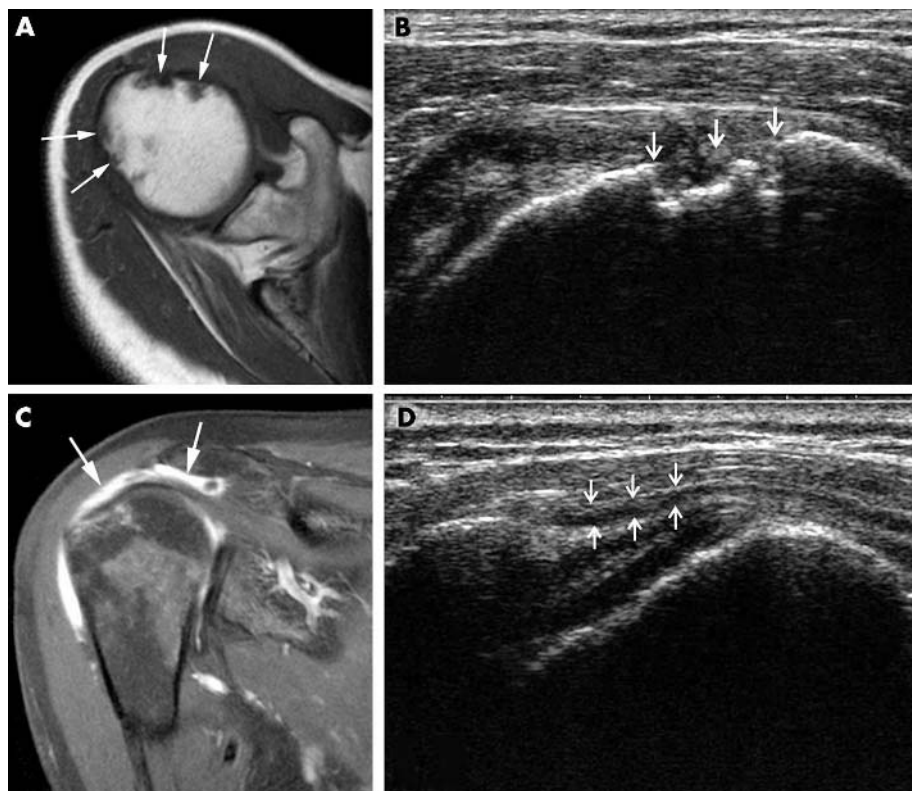
Figure 1 Shoulder joint. (A and B) Humeral head erosions. (A) In MRI, multiple erosions can be seen from the anterior and posterior sides of the humeral head as bone defects with sharp margins (arrows). (B) Distinct bone defects below the bone surface (erosions, arrows) can also be detected by US. This image is taken from the anterior side with maximum inner rotation (transverse scan). (C and D) Glenohumeral joint synovitis. (C) In MRI, contrast enhancement clearly depicts a subdeltoid/subacromial bursitis (arrows) and synovitis within the joint. (D) The US image shows a lateral longitudinal scan of the shoulder joint. Subdeltoid bursitis can be visualised as an anechoic area below the deltoid muscle (arrows).

The musculoskeletal radiologist carried out the evaluations without knowing the diagnoses and clinical data using the same standardised report form as for US examinations.

Figure 2 Finger joint (MCP II). (A) The MR image shows the MCP joints II–V in transverse section. Focusing on MCP joint II shows slight contrast enhancement from the dorsal and palmar aspects, representing synovitis (arrowheads). Also, tenosynovitis is seen at the flexor tendons (arrow). (B) The US longitudinal image from the palmar side displays an anechoic to hypoechoic area at the region of the diaphysis reflecting synovitis (arrows). Also, there is tenosynovitis along the flexor tendon (upper arrows).

Statistical analysis

Interobserver agreement was estimated using a modified κ index for majority agreement.[20] The majority was defined as 10 out of 14, which corresponds to 71% agreement among the raters. Overall agreement (defined as the percentage of observed exact agreements) as well as sensitivity and specificity were calculated using the statistical software package SAS 8.02 (SAS Institute Inc, Cary, NC, USA).
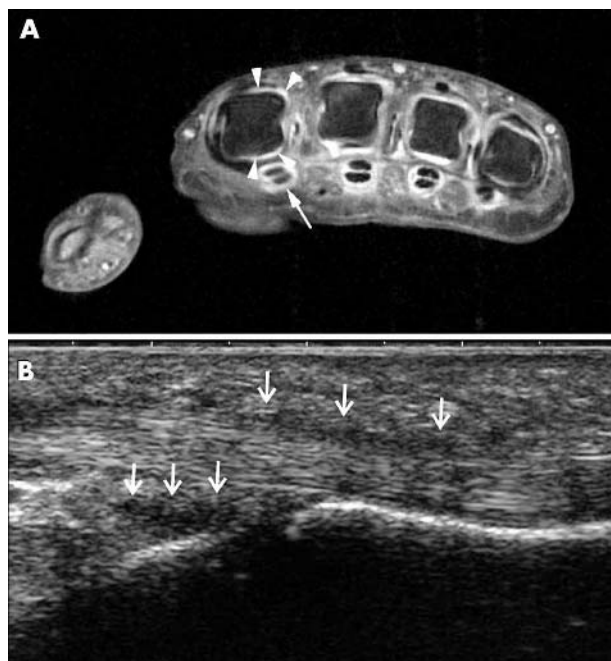
RESULTS
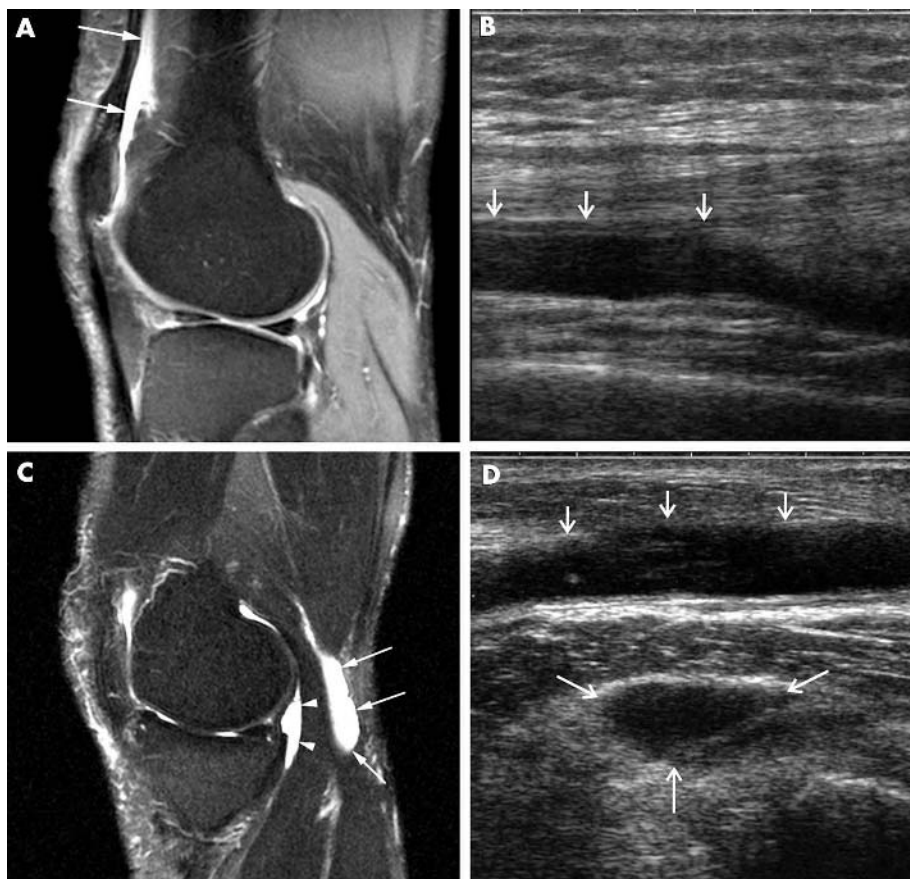### Results of questionnaire

The expertise of the sonographers was documented by a total of 89 original articles as first author in the field of musculoskeletal US. All participants frequently perform musculoskeletal US (9/14, >20 examinations a week, range 10 to >40). Table 1 lists the joints most often examined. The transducer orientation for longitudinal scans was applied by all sonographers according to the EULAR recommendations[21] (left of screen, proximal or cranial of the patient). On the contrary, for transverse scans 6/14 experts assigned their left side to the left of the screen as opposed to the EULAR recommendations (left of screen, medial of patient, 5/14). Doppler US is frequently used in musculoskeletal US (13/14). Documentation of normal and pathological findings is performed by all participants. Table 2 lists the transducers mainly used by the experts for musculoskeletal US. The sonographers perform most of the EULAR standard scans.[21] However, 29% of the demanded scans were not performed for the shoulder joint (table 3). The less frequently performed scans demanded a special position during dynamic examination—for example, the anterior transverse and longitudinal scans in maximal inner rotation of the shoulder were only performed by 5/14 sonographers, each.

Results of practical examinations

Tables 4A–D display the normal and pathological findings of the joint structures detected by MRI and US.

Figure 3 Knee joint. (A) MRI shows some contrast agent enhancement in the suprapatellar recess, reflecting inflammatory effusion (two arrows). (B) US also clearly depicts the effusion in the suprapatellar recess (arrows). (C) In MRI, a popliteal cyst is visualised in the sagittal view with a deep part (arrowheads) and a superficial part (arrows). (D) Both parts can also clearly be detected by US as anechoic areas (arrows).

Interobserver agreement

Taking an agreement in US examination of 10 out of 14 experts as a point of reference, the overall κ for all examined joints was 0.76. Calculations for each joint region showed good κ values for the knee (1) and shoulder (0.76) joints, moderate agreement for the hand/finger joints (0.59), and low agreement in ankle/toe joints (0.28). κ Values for bone lesions, bursitis, and tendon tears (evaluation includes all joint regions) were excellent (κ = 1, for all). There was also a moderate κ value for the detection of tenosynovitis (0.49), but κ values were low for the detection of synovitis/effusion, mainly because small amounts (for example, acromioclavicular joint) were missed. The overall agreements between the 14 experts were 81% for the shoulder, 73% for the wrist/fingers, 88% for the knee joint, and 82% for the ankle/toe joints.

Overall agreement of US findings compared with MRI

In US examination, we found a good agreement for most findings for the shoulder joint (for example, humeral head erosions, 100%; figs 1A and B) with an overall agreement of 81% when compared with MRI findings (table 4A). However, the detection of synovitis/effusion in the shoulder was moderate (50%; figs1C and D) and poor for the acromio-clavicular joints (29%). US examinations of the shoulder joint showed a sensitivity of 76% and specificity of 89%.

For the wrist and finger joints, we found a lower overall agreement of US findings compared with MRI (73%; table 4B). In particular, palmar tenosynovitis and synovitis/effusion in the MCP II joint showed a low agreement with MRI findings (50% each; figs 2A and B). However, agreement for dorsal synovitis findings (for example, tenosynovitis 79%) was clearly better, possibly because evaluation of joints is—in many countries—routinely performed only from the dorsal

sides and palmar inflammation may be missed. The overall sensitivity for US of the wrist/finger joints was rather low (66%), with a higher specificity (88%).

For the knee joint the overall agreement of US findings with MRI was 88% (table 4C). Effusion in the suprapatellar recess was seen by all participants (figs 3A and B), and almost all (13/14) detected the popliteal cyst (figs 3C and D). The overall sensitivity for US of the knee joint was 91%, the overall specificity 88%.

Results for US examination of the ankle and foot joints were similar to the results for the wrist and finger joints. Although the overall agreement was 82%, we found a rather low overall sensitivity of 61%, with a high specificity (92%) (table 4D). There was low concordance in the detection of synovitis/effusion in both tibiotalar (figs 4A and B) and talonavicular joints (57%, each) as well as the extensor tendons (36%), whereas agreement was better in the detection of flexor (86%) and peroneus (71%) tenosynovitis.

The overall total agreement of US findings as compared with MRI for all the joints examined (45 sonographic findings) was 82% (sensitivity 71%, specificity 90%).

DISCUSSION

There is increasing evidence that musculoskeletal US has an important role in the management of patients with arthritis.[16] However, operator dependence remains one of the perceived major limitations to its widespread use.[5 17 18 22 23] Currently, available information about reproducibility, in particular for a large number of observers, is limited. This study has demonstrated moderate to good correlations between 14 independent observers.

The main result of an open blinded questionnaire sent to all experts was that most standard scans as published by the
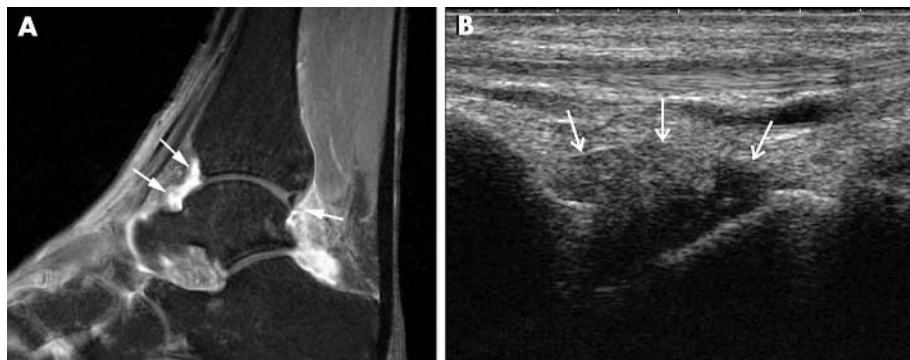
Figure 4 Ankle/toe joints. (A) MRI of the ankle shows contrast enhancement in the tibiotalar joint from anterior and posterior aspects (arrows). (B) The longitudinal US image is an example of the anterior side of the tibiotalar joint. The anechoic area displays effusion (anechoic) and synovitis (hypoechoic; arrows).

EULAR working group for musculoskeletal US[21] were performed by the sonographers: however, for the shoulder joint standardisation was less (71%) than for the other joints (ranging from 82% for the knee to 100% for the wrist joints). Possibly, the reason for this is that some additional EULAR scans that demand special positions for dynamic examination are not performed by all. These scans, however, are helpful for detecting subtle amounts of effusion that can only be visualised by moving the limbs and/or transducer.

For comparing musculoskeletal US results, interobserver agreement has so far only been calculated between two observers.[6 17 19] The more observers who participate in interobserver computations, the lower is the probability of simultaneous agreement among all observers, resulting in κ values which are too low.[20] In our majority agreement[20] there is already a contribution to agreement if at least 10 out of 14 judgments of a joint are the same, resulting in an overall κ of 0.76. Our overall κ value is higher than found when two observers were compared (range 0.48–0.68),[6 17] which means the results are relatively considering that we compared 14 observers. This difference may be due to several reasons. Firstly, all participants in our study were experienced sonographers; this means that agreement should be tested in further studies with less experienced sonographers. Secondly, the sonographers may have paid more attention than usual, so that any possible lesion was reported.

The OMERACT MRI in RA working group studied interreader agreement for a simple scoring system in wrist and MCP joints of patients with RA among five different centres.[24 25] They found mean unweighted κ values of 0.62, suggesting that the basic interpretation of MRI changes is relatively consistent among readers from different countries, but that further training and standardisation are necessary to achieve better intergroup reproducibility.[24] This is currently underway as part of the OMERACT process.[26]

In a follow up study, the OMERACT working group obtained an improvement in interreader agreement as reflected by acceptable intraclass correlation coefficients (range 0.6–0.91).[25] This measure, however, is not applicable to our data owing to the limited number of cases. Nevertheless, because MRI allows for detailed documentation of joint examinations, these studies compare the reading of images taken at one MR examination, whereas in US the reliability of producing and reading images is considered. κ Values were slightly higher between ultrasonographers than among MRI readers, thus disproving the generally held opinion that US is highly observer dependent. Calculations of US examination results for each joint region showed good κ values for the knee (1) and shoulder (0.76) joints, acceptable agreement for the hand/finger (0.59), and low agreement in ankle/toe joints (0.28). However, the κ values of the ankle/toe joints are not fully applicable because there was an asymmetric distribution of positive and negative findings,

and the overall agreement between observers should be taken into account (84%).

Overall, we found a moderate to good agreement between the expert ultrasonographers and MRI, with a high concordance for the main findings for both bone surface and soft tissue abnormalities.

For the shoulder, the overall agreement of US findings compared with MRI was 81%. We found a relatively good agreement with most detected pathologies. More discrete findings, such as minimal effusion in the acromioclavicular and glenohumeral joints, were detected to a far lesser extent (29% v 50%), which was also reported in a recent study.[27] However, inflammation within the joint cavity could only be seen in full inner rotation, again supporting the need for a full dynamic US examination.

For the wrist and finger joints, we found a high overall specificity (88%) and a moderate sensitivity (66%) owing to low sensitivities in the detection of palmar tenosynovitis and MCP II joint synovitis (50% each). However, because finger joint synovitis was mainly present at the palmar side of the finger in our patients, these findings might have been missed when evaluation was solely performed from the dorsal side. Similar observations have been reported recently in 42 patients with RA and finger joint inflammation.[28]

For the knee, US resulted in a high overall sensitivity (91%) and specificity (88%). In particular, US was very sensitive in the detection of suprapatellar effusion (100%) and popliteal cysts (93%). However, there might have been some over-interpretation with US in the detection of bone lesions because no bone defects were detected by MRI (specificity range 71–79%).

The sensitivity for the ankle and toe joints was rather low (61%), with a high specificity (92%), most probably owing to the inability in dynamic and both plantar and dorsal examinations to detect even subtle pathologies, which can only be seen in special positions and during movement.

Although guidelines have been published, scanning techniques vary to a certain extent in the European countries and between the experts. Ten of the 14 sonographers were not familiar with the equipment and the scanner settings, and the level of experience with the US device was different for each sonographer. In addition, the scanner settings were not variable, the sonographers were unaware of the clinical diagnosis, and a symptom driven clinical examination of the affected joint region was not performed. These aspects have a relevant influence on the information that can be obtained in 10 minutes' scanning of complex anatomical areas, and reliability might have been better if longer training on the US devices had been given.

In conclusion, our results show that musculoskeletal US is a reliable technique with moderate to good interobserver reliability in an expert setting between a large number of observers. Interpretations of the US images by sonographers

differ considerably for some joints. In addition, the study underpins the need for dynamic examination for complete detection of subtle pathological findings. Training and standardisation of musculoskeletal US are necessary to achieve higher interobserver reproducibility. For our next step we aim at performing interobserver testing on semi-quantitative and quantitative grading of pathological structures. As a result of this study, the participants decided to accelerate efforts to standardise the musculoskeletal US investigation techniques of both EULAR and OMERACT.

ACKNOWLEDGEMENTS

. . . .

Authors' affiliations

A K Scheel, D Pasewaldt, Department of Nephrology and Rheumatology, Georg-August-University Göttingen, Germany

W A Schmidt, Medical Centre for Rheumatology Berlin-Buch, Berlin, Germany

K-G A Hermann, Department of Radiology, Charité University Hospital, Berlin, Germany

G A Bruyn, Department of Rheumatology, Medisch Centrum Leeuwarden, Leeuwarden, The Netherlands

M A D'Agostino, Department of Rheumatology, Ambroise Paré Hospital, UVSQ University, Boulogne Billancourt, France

W Grassi, Department of Rheumatology, Universita degli Studi di Ancona, Ospedale A. Murri, Jesi (Ancona), Italy

A Iagnocco, Department of Rheumatology, University La Sapienza, Rome, Italy

J M Koski, Department of Rheumatology, Mikkeli Central Hospital, Mikkeli, Finland

K P Machold, Department of Rheumatology, Vienna General Hospital, University of Vienna, Austria

E Naredo, Department of Rheumatology, the Research Unit, and the Epidemiology Unit, Severo Ochoa Hospital, Madrid, Spain

H Sattler, Department of Rheumatology, Parkklinik, Bad Durkheim, Germany

N Swen, Department of Rheumatology, Medisch Centrum Alkmaar, Alkmaar, The Netherlands

M Szkudlarek, Department of Rheumatology, University of Copenhagen, Hvidovre Hospital, Denmark

R J Wakefield, Academic Department of Musculoskeletal Medicine, Leeds General Infirmary, UK

H R Ziswiler, Department of Rheumatology and Clinical Immunology, Inselspital, Bern, Switzerland

C Werner, Department of Medical Statistics, Georg-August-University Göttingen, Göttingen, Germany

M Backhaus, Department of Rheumatology and Clinical Immunology, Charité University Hospital, Berlin, Germany

REFERENCES

1 **Sattler H**. [Value of arthrosonography of the shoulder in rheumatologic diagnosis. Examination technique, findings and their interpretation.] *Z Rheumatol*, 1993;**52**:90–6.
2 **Koski JM**. Ultrasonography of the subtalar and midtarsal joints. *J Rheumatol* 1993;**20**:1753–5.
3 **Grassi W**, Filippucci E, Farina A, Salaffi F, Cervini C. Ultrasonography in the evaluation of bone erosions. *Ann Rheum Dis* 2001;**60**:98–103.
4 **Backhaus M**, Kamradt T, Sandrock D, Loreck D, Fritz J, Wolf KJ, et al. Arthritis of the finger joints: a comprehensive approach comparing conventional radiography, scintigraphy, ultrasound, and contrast-enhanced magnetic resonance imaging. *Arthritis Rheum* 1999;**42**:1232–45.
5 **Szkudlarek M**, Court-Payen M, Strandberg C, Klarlund M, Klausen T, Ostergaard M. 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.
6 **Wakefield RJ**, Gibbon WW, Conaghan PG, O'Connor P, McGonagle D, Pease C, et al. The value of sonography in the detection of bone erosions in patients with rheumatoid arthritis: a comparison with conventional radiography. *Arthritis Rheum* 2000;**43**:2762–70.
7 **Naredo E**, Iagnocco A, Valesini G, Uson J, Beneyto P, Crespo M. Ultrasonographic study of painful shoulder. *Ann Rheum Dis* 2003;**62**:1026–7.
8 **Iagnocco A**, Coari G, Palombi G, Valesini G. Sonography in the study of metatarsalgia. *J Rheumatol* 2001;**28**:1338–40.
9 **Koski JM**. Detection of plantar tenosynovitis of the forefoot by ultrasound in patients with early arthritis. *Scand J Rheumatol* 1995;**24**:312–13.
10 **Grassi W**, Filippucci E, Farina A, Cervini C. Sonographic imaging of tendons. *Arthritis Rheum* 2000;**43**:969–76.
11 **D'Agostino MA**, Said-Nahal R, Hacquard-Bouder C, Brasseur JL, Dougados M, Breban M. Assessment of peripheral enthesitis in the spondylarthropathies by ultrasonography combined with power Doppler: a cross-sectional study. *Arthritis Rheum* 2003;**48**:523–33.
12 **Terslev L**, Torp-Pedersen S, Savnik A, von der Recke P, Qvistgaard E, Danneskiold-Samsoe B, 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.
13 **Schmidt WA**, Volker L, Zacher J, Schlafke M, Ruhnke M, Gromnica-Ihle E. Colour Doppler ultrasonography to detect pannus in knee joint synovitis. *Clin Exp Rheumatol* 2000;**18**:439–44.
14 **Carotti M**, Salaffi F, Manganelli P, Salera D, Simonetti B, Grassi W. Power Doppler sonography in the assessment of synovial tissue of the knee joint in rheumatoid arthritis: a preliminary experience. *Ann Rheum Dis* 2002;**61**:877–82.
15 **Wakefield RJ**, Brown AK, O'Connor PJ, Emery P. Power Doppler sonography: improving disease activity assessment in inflammatory musculoskeletal disease. *Arthritis Rheum* 2003;**48**:285–8.
16 **Manger B**, Kalden JR. Joint and connective tissue ultrasonography—a rheumatologic bedside procedure? A German experience. *Arthritis Rheum* 1995;**38**:736–42.
17 **Szkudlarek M**, Court-Payen M, Jacobsen S, Klarlund M, Thomsen HS, Ostergaard M. Interobserver agreement in ultrasonography of the finger and toe joints in rheumatoid arthritis. *Arthritis Rheum* 2003;**48**:955–62.
18 **Ostergaard M**, Wiell C. Ultrasonography in rheumatoid arthritis: a very promising method still needing more validation. *Curr Opin Rheumatol* 2004;**16**:223–30.
19 **Ribbens C**, Andre B, Marcelis S, Kaye O, Mathy L, Bonnet V, et al. Rheumatoid hand joint synovitis: gray-scale and power Doppler US quantifications following anti-tumor necrosis factor-alpha treatment: pilot study. *Radiology* 2003;**229**:562–9.
20 **Craig R**. Generalization of Scott's index of intercoder agreement. *Public Opinion Quarterly* 1981;**45**:260–4.
21 **Backhaus M**, Burmester GR, Gerber T, Grassi W, Machold KP, Swen WA, et al. Guidelines for musculoskeletal ultrasound in rheumatology. *Ann Rheum Dis* 2001;**60**:641–9.
22 **Balint PV**, Sturrock RD. Intraobserver repeatability and interobserver reproducibility in musculoskeletal ultrasound imaging measurements. *Clin Exp Rheumatol* 2001;**19**:89–92.
23 **Scheel AK**, Backhaus M. Ultrasonographic assessment of finger and toe joint inflammation in rheumatoid arthritis: comment on the article by Szkudlarek et al. *Arthritis Rheum* 2004;**50**:1008; author reply 1008–9.
24 **Ostergaard M**, Klarlund M, Lassere M, Conaghan P, Peterfy C, McQueen F, et al. Interreader agreement in the assessment of magnetic resonance images of rheumatoid arthritis wrist and finger joints—an international multicenter study. *J Rheumatol* 2001;**28**:1143–50.
25 **Lassere M**, McQueen F, Ostergaard M, Conaghan P, Shnier R, Peterfy C, et al. OMERACT Rheumatoid Arthritis Magnetic Resonance Imaging Studies. Exercise 3: an international multicenter reliability study using the RA-MRI Score, *J Rheumatol* 2003;**30**:1366–75.
26 **Ostergaard M**, Peterfy C, Conaghan P, McQueen F, Bird P, Ejbjerg B, et al. OMERACT Rheumatoid Arthritis Magnetic Resonance Imaging Studies. Core set of MRI acquisitions, joint pathology definitions, and the OMERACT RA-MRI scoring system. *J Rheumatol* 2003;**30**:1385–6.
27 **Hermann KG**, Backhaus M, Schneider U, Labs K, Loreck D, Zuhlsdorf S, et al. Rheumatoid arthritis of the shoulder joint: comparison of conventional radiography, ultrasound, and dynamic contrast-enhanced magnetic resonance imaging. *Arthritis Rheum* 2003;**48**:3338–49.
28 **Scheel AK**, Hermann, KG, Kahler E, Pasewaldt D, Brunner E, Müller GA, et al. A novel ultrasonographic synovitis score suitable for analyzing rheumatoid arthritis finger joint inflammation. *Ann Rheum Dis* 2004;**63**:poster SAT, 0295;(suppl).