EXTENDED REPORT

Interobserver reliability in musculoskeletal ultrasonography: results from a ''Teach the Teachers'' rheumatologist course

E Naredo, I Möller, C Moragues, J J de Agustín, A K Scheel, W Grassi, E de Miguel,
M Backhaus, P Balint, G A W Bruyn, M A D'Agostino, E Filippucci, A Iagnocco,
D Kane, J M Koski, L Mayordomo, W A Schmidt, W A A Swen, M Szkudlarek,
L Terslev, S Torp-Pedersen, J Uson, R J Wakefield, C Werner, EULAR Working Group
for Musculoskeletal Ultrasound

Objective: To assess the interobserver reliability of the main periarticular and intra-articular ultrasonographic pathologies and to establish the principal disagreements on scanning technique and diagnostic criteria between a group of experts in musculoskeletal ultrasonography.

Methods: The shoulder, wrist/hand, ankle/foot, or knee of 24 patients with rheumatic diseases were evaluated by 23 musculoskeletal ultrasound experts from different European countries randomly assigned to six groups. The participants did not reach consensus on scanning method or diagnostic criteria before the investigation. They were unaware of the patients' clinical and imaging data. The experts from each group undertook a blinded ultrasound examination of the four anatomical regions. The ultrasound investigation included the presence/absence of joint effusion/synovitis, bony cortex abnormalities, tenosynovitis, tendon lesions, bursitis, and power Doppler signal. Afterwards they compared the ultrasound findings and re-examined the patients together while discussing their results.

Results: Overall agreements were 91% for joint effusion/synovitis and tendon lesions, 87% for cortical abnormalities, 84% for tenosynovitis, 83.5% for bursitis, and 83% for power Doppler signal; κ values were good for the wrist/hand and knee (0.61 and 0.60) and fair for the shoulder and ankle/foot (0.50 and 0.54). The principal differences in scanning method and diagnostic criteria between experts were related to dynamic examination, definition of tendon lesions, and pathological v physiological fluid within joints, tendon sheaths, and bursae.

Conclusions: Musculoskeletal ultrasound has a moderate to good interobserver reliability. Further consensus on standardisation of scanning technique and diagnostic criteria is necessary to improve musculoskeletal ultrasonography reproducibility.

See end of article for authors' affiliations

Correspondence to:
Dr E Naredo, Calle Arturo Soria 259, 4° A, 28033 Madrid, Spain; esnaredo@eresmas.com

High resolution musculoskeletal ultrasonography effectively depicts superficial periarticular and intra-articular structures involved in rheumatic diseases.[1] [2] Ultrasonography has considerable advantages over other imaging methods, including non-invasiveness, rapidity of performance, relatively low cost, ability to scan multiple joints, repeatability, and high patient acceptability. In addition, it can be used routinely for dynamic examinations. Last but not least, rheumatologists can undertake in-office ultrasonography, avoiding referral to radiologists and saving time and money.

Recently, ultrasonography has shown better sensitivity than clinical evaluation and plain radiography for the detection of rheumatoid synovitis[3–6] and joint erosions.[6] [7] These encouraging reports have directed ultrasonography research towards the assessment of early inflammatory and structural changes and monitoring therapeutic response in patients with chronic inflammatory arthritis.

However, ultrasonography has been viewed as one of the most operator dependent imaging techniques. This partly reflects the intrinsic real time nature of ultrasonographic image acquisition. The recorded images largely display the subjective findings observed by the individual performing the examination. In addition, the intra-observer and interobserver reliability of ultrasonography has been assessed in only a

minority of papers.[4] [5] [7–13] Thus a strictly standardised scanning technique and diagnostic criteria are urgently needed in order to compare the results of ultrasonography reports, develop multicentre studies, and teach the technique uniformly.

The European League Against Rheumatism (EULAR) working group for musculoskeletal ultrasound consists of the faculty of the EULAR Sonography Courses. In 2001, the group published the guidelines for musculoskeletal ultrasonography scanning in rheumatology,[14] and in 2004 it reached a consensus on the first preliminary ultrasonography pathological definitions at OMERACT 7 (Asilomar, California, USA) and carried out the first interobserver variability study between 14 experts in musculoskeletal ultrasonography during a ''Train the Trainers'' course in Berlin.[13]

Afterwards, the group decided to organise a second longer ''Teach the Teacher'' meeting in Sitges (Barcelona, Spain) in October, 2004. This course had 23 participants expert in musculoskeletal ultrasonography and two objectives: first, to assess the interobserver reliability of ultrasonography for detecting the main rheumatic periarticular and intra-articular pathological features; and second, to compare and discuss the ultrasonographic scanning techniques, image interpretation, and diagnostic criteria between the experts by examining patients together and re-evaluating recorded video clips or

Table 1 Ultrasonographic investigation

Anatomical structure	Ultrasonographic findings
Shoulder	
Long biceps tendon	Tenosynovitis, partial tear, complete tear
Subscapularis, supraspinatus, and infraspinatus tendons	Tendinosis, partial tear, full thickness tear, calcification
Subacromial-subdeltoid and subcoracoid bursae	Bursitis
Acromio-clavicular joint	Effusion/synovitis, erosions, osteophytes
Glenuhumeral joint (axillary, anterior, and posterior recesses)	Effusion/synovitis
Humeral head (anterior, lateral, posterior, and axillary aspects)	Erosions
Supraspinatus tendon	Impingement
Wrist/hand	
Radiocarpal joint (dorsal and palmar aspects)	Effusion/synovitis, power Doppler signal
Wrist tendons (dorsal, palmar, extensor carpi ulnaris)	Tenosynovitis
First carpometacarpal joint	Effusion/synovitis, erosions, osteophytes
Second metacarpophalangeal joint	Effusion/synovitis, erosions, osteophytes, power Doppler signal
Second proximal interphalangeal joint	Effusion/synovitis, erosions, osteophytes
Second finger flexor tendons	Tenosynovitis
Ankle/foot	
Tibiotalar joint (anterior and posterior aspects)	Effusion/synovitis
Talonavicular joint	Effusion/synovitis, erosions, osteophytes
Anterior tibialis, flexor, peroneus, and posterior tibialis tendons	Tenosynovitis
Posterior tibialis tendon	Partial tear, complete tear
Achilles tendon	Tendinosis, enthesitis, paratenonitis, partial tear, complete tear
Retrocalcaneal bursa	Bursitis
Calcaneous (posterior and plantar aspects)	Erosions, spur
Plantar fascia	Fasciitis
First metatarsophalangeal joint	Effusion/synovitis, erosions, osteophytes, power Doppler signal
Fifth metatarsophalangeal joint	Effusion/synovitis, erosions, osteophytes
Knee	
Suprapatellar recess	Effusion, synovial hypertrophy
Bone surfaces	Erosions, osteophytes
Quadriceps and patellar tendons	Tendinosis, enthesitis, partial tear, complete tear
Prepatellar, superficial and deep infrapatellar, anserin bursae	Bursitis
Gastrocnemius-semimembranous bursa	Baker's cyst
Femoral articular cartilage	Lesion
Medial collateral ligament	Partial tear, complete tear

images, in order to address the main differences to be standardised in future EULAR/OMERACT exercises.

METHODS
Twenty two rheumatologists and one radiologist expert in musculoskeletal ultrasonography—members of the EULAR working group for musculoskeletal ultrasound—from nine European countries (Denmark 3; France 1; Finland 1; Germany 3; Hungary 1; Italy 3; Netherlands 2; United Kingdom 2; Spain 7) participated in the Teach the Teacher course. The meeting started on Friday afternoon and finished on Sunday morning. It took 16 hours, divided into four consecutive sessions, each for ultrasonographic examination of an anatomical region and each lasting four hours: session 1, shoulder; session 2, wrist/hand; session 3 ankle/foot; session 4, knee. The 23 experts were assigned to six groups of three members (one group) or four members (five groups) in each part of the study. The members of the six groups were then rotated for each anatomical region examined. The distribution of the experts was done randomly while avoiding, as far as possible, participants from the same country being in the same group.

Patients
Twenty four patients (eight men, 16 women, mean (SD) age 56.9 (14.2) years, range 26 to 75) were recruited from the outpatient rheumatology clinic of Instituto Poal, Hospital de Bellvitge and Hospital Vall d'Hebron, Barcelona, Spain. Six patients were selected for shoulder examination, six for wrist/hand examination, six for ankle/foot examination, and six for knee examination. Patients had been diagnosed by clinical evaluation, plain radiography, and magnetic resonance

imaging (MRI) or ultrasonography carried out within one month of the start of the study by staff members of the hospitals. Diagnoses were degenerative shoulder disorder (3) and rheumatoid arthritis (3) for the shoulder session; rheumatoid arthritis (6) for the wrist/hand session; rheumatoid arthritis (3), spondylarthropathy (2), and osteoarthritis (1) for the ankle/foot session; and rheumatoid arthritis (2) and osteoarthritis (4) for the knee session. All patients were symptomatic at the time of the study. An ultrasonographic examination was carried out in all patients within two days of the start of the study to confirm the presence of abnormalities in the anatomical region of interest. Their clinically dominant region was selected for ultrasonographic examination: right shoulder in four patients, left shoulder in two, right wrist/hand in five, left wrist/hand in one, right ankle/foot in three, left ankle/foot in three, right knee in three, and left knee in three.

Ultrasonographic examination
Ultrasonography was carried out using six commercially available ultrasound real time scanners (three Logiq 5 Pro, General Electric Medical Systems, Kyunngi, Korea; two Technos MPX, Esaote, Genoa, Italy; and one Sonoline Antares, Siemens, Mountainview, California, USA), using multifrequency linear transducer (7–14 MHz) and power Doppler function.

Each group was randomly assigned to an ultrasonography machine and a patient for assessing shoulder, wrist/hand, ankle/foot, and knee. The participants did not reach consensus on scanning method or diagnostic criteria before the investigation. They were asked to carry out their routine scanning technique and diagnose according to their usual

Table 2 Overall agreements by ultrasonographic diagnosis in each region and in all regions

Diagnosis	Shoulder	Wrist/hand	Ankle/foot	Knee	All regions
Joint effusion/synovitis	88.5%	95%	89%	91.5%	91%
Bony cortex abnormalities	84.5%	88.5%	92%	84.5%	87%
Tenosynovitis/paratenonitis	76%	88.5%	88.5%	NE	84%
Tendon lesions	88%	NE	92%	94%	91%
Bursitis/cyst	85.5%	NE	73.5%	92%	83.5%
Power Doppler signal	NE	92.5%	73.5%	NE	83%

NE, diagnosis not evaluated.

diagnostic criteria. They were blinded to patients' diagnosis and previous ultrasonography and MRI data.

The ultrasonography investigation was quite similar to that used in the Train the Trainer study in Berlin[13] and included the presence or absence of the ultrasonographic pathological findings listed in table 1.

During the first part of each session (one hour), the three or four members of each group blindly, independently, and consecutively examined the patient assigned. Each expert was given a maximum of 15 minutes for scanning the corresponding anatomical region and anonymously filling in a standardised report sheet with the ultrasonographic findings. Each examiner was informed of the selected anatomical region (right/left). An application specialist from the ultrasonography company was near each machine to solve technical adjustment problems. The results of the blinded three or four examinations for each group were used to estimate interobserver reliability.

For the following one and a half hours of each session, the three or four experts of each group compared their results. Then they re-examined their patient together while discussing the scanning method and diagnostic criteria used by each of them and recording their different results.

During the last part of each session (one and a half hours), each group was given 15 minutes for explaining and discussing with the rest of the experts the main agreements and differences in scanning technique or diagnostic criteria found, using recorded video clips or images.

Statistical analysis

The ultrasonography findings were grouped for statistical analysis according to the following ultrasonographic diagnoses: joint effusion/synovitis; bony cortex abnormalities including bone erosions and osteophytes; tenosynovitis or paratenonitis; tendon lesions including tendinosis, enthesopathy, calcification, partial and complete tear; bursitis, including Baker's cyst; power Doppler signal.

Overall agreement, defined as the percentage of exact agreement observed, was calculated for each ultrasonographic diagnosis in each region and in all regions. Interobserver reliability was calculated for each group and anatomical region using the unweighted κ; the κ value could also be calculated for those ultrasonographic diagnoses that were investigated in more than four locations in a region. Values of κ of <0.40 reflect poor agreement, between 0.40
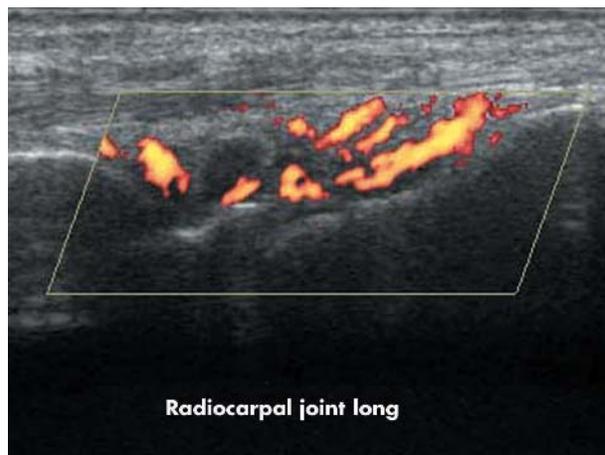

Figure 1 Longitudinal ultrasonographic image of the radiocarpal joint with synovitis and intense power Doppler signal in a patient with rheumatoid arthritis. long, longitudinal.

and 0.75 fair to good agreement, and >0.75 excellent agreement.[15]

RESULTS
Interobserver reliability
The overall agreements by ultrasonographic diagnosis in each region and in all regions are given in table 2. They ranged from 83% for power Doppler signal to 91% for joint effusion/synovitis and tendon lesions. The overall agreement for rotator cuff impingement, plantar fasciits, femoral articular cartilage lesion, medial collateral ligament partial, and complete tear were 87.5%, 96%, 86%, 87.5%, and 100%, respectively. Table 3 shows the κ values by group and anatomical region and the overall κ by region. Interobserver agreement was good for the wrist/hand and knee and fair for the shoulder and ankle/foot. The mean κ values for the detection of wrist/hand and ankle/foot effusion /synovitis were 0.73 and 0.69, respectively. There was a good agreement for the diagnosis of ankle and knee tendon lesions (κ = 0.71 and 0.72, respectively) while agreement was fair for shoulder tendon lesions (κ = 0.50). The κ value was excellent for the detection of knee bursitis and Baker's cyst (κ = 0.82), good for wrist/hand and ankle/foot cortical abnormalities (κ = 0.64

Table 3 κ Values by group and anatomical region and overall (mean) κ by region

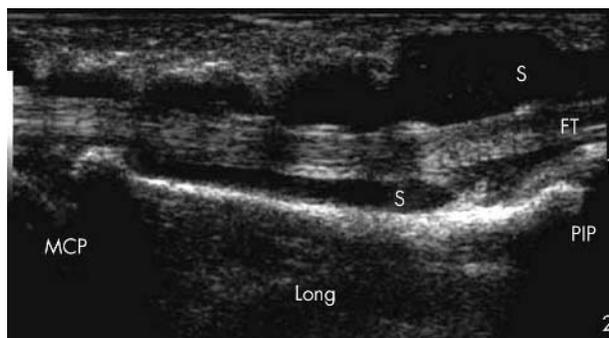| Region	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Mean κ
Shoulder	0.62	0.67	0.14	0.59	0.35	0.64	0.50
Wrist/hand	0.61	0.50	0.83	0.28	0.59	0.83	0.61
Ankle/foot	0.53	0.41	0.30	0.58	0.79	0.60	0.54
Knee	0.70	0.58	0.38	0.68	0.75	0.54	0.60

Figure 2 Second finger flexor tenosynovitis in a patient with rheumatoid arthritis. Longitudinal ultrasonographic image shows increased hypoechoic synovitis (S) within the sheath surrounding the tendon (FT). Long, longitudinal; MCP, metacarpophalangeal joint; PIP, proximal interphalangeal joint.

and 0.63, respectively), and fair for shoulder cortical abnormalities and ankle tenosynovitis ($\kappa = 0.50$ and 0.47, respectively).

Differences identified

The principal differences in scanning method and diagnostic criteria between experts were as follows:

- Although all experts carried out most of the standard scans recommended by the EULAR guidelines,[13] some used more multiplanar and dynamic image acquisition which facilitate the detection of subtle abnormalities. However, all experts agreed on scanning the various recess in each joint for detecting effusion/synovitis because there are not enough studies comparing their sensitivity.

- There was no agreement on the definition of rotator cuff tendon lesions such as tendinosis and partial and full thickness tears. These discrepancies caused different interpretations by the experts of the same pathological ultrasonographic findings.

- There was disagreement on the definition of normality/ pathology with regard to the minimum fluid within synovial recesses, tendon sheaths, and large bursae such as the subacromial-subdeltoid and retrocalcaneal bursa found in both rheumatological patients and many normal subjects. Neither the measure of normal versus pathological fluid nor the location for detecting it were standardised between the experts. This resulted in a diagnosis of mild tenosynovitis, bursitis, and joint effusion by some experts, while others considered the findings normal. Some experts argued that the presence of local clinical symptoms should be decisive for this differential diagnosis. In addition, ultrasonographic findings in the opposite side should be taken into account.

Two illustrative ultrasonography images are shown in figs 1 and 2.

DISCUSSION

Ultrasonography has been considered the most operator dependent imaging technique. The paucity of studies on its validity, reliability, and sensitivity to change has largely contributed to this and has limited the development of multicentre and longitudinal ultrasonographic studies.

European rheumatologists highly experienced in musculoskeletal ultrasonography have comprised the faculty of the nine training courses on musculoskeletal ultrasonography organised in different European countries under the auspices of the EULAR Standing Committee for Education and

Training since 1998. They have a teaching and research curriculum in this field. Many of them chair and organise ultrasonography training for rheumatologists in their countries.

For the last four years, the EULAR working group for musculoskeletal ultrasound has made an effort to standardise ultrasonographic scanning methods[14] and diagnostic criteria and to develop reliability studies.

The first official ultrasound special interest group (SIG) met at OMERACT 7 (Asilomar, California) in May 2004. The principal activities of the ultrasonography SIG have been a systematic review of published reports and a consensus on preliminary pathological definitions of synovial hypertrophy, tenosynovitis, enthesopathy, and bone erosion.

The first Train the Trainers meeting was held in Berlin before the eighth EULAR sonography course organised by M Backhaus and W A Schmidt in June 2004. Fourteen teachers from that course participated in the present study, which had two main objectives: to assess the interscanner variability between the 14 examiners and to evaluate agreement in ultrasonographic diagnosis, with MRI findings as the gold standard, in four anatomical regions (shoulder, knee, wrist/ finger, and ankle/toe) of four patients, respectively, with inflammatory rheumatic diseases.[13]

Before the study by Scheel et al,[13] ultrasonographic interobserver reliability had only been tested between two examiners.[4 5 7–12] Swen et al[8] reported a good κ value (0.63) in the detection of rotator cuff full thickness tear. Middleton et al[12] found a high agreement (92%) in the diagnosis of rotator cuff partial and full thickness tear. The κ values for ultrasonography detection of wrist synovitis, tenosynovitis, and erosions were from 0.73 to 0.89 according to Iagnocco et al.[11] In the study by Szkudlarek et al,[4] the overall agreement/κ values for the semiquantitative assessment of effusion, synovitis, power Doppler signal, and erosions in small joints of the hand and foot were 79%/0.48, 86%/0.63, 87%/0.55, and 91%/0.68, respectively. However, Filippucci et al[10] reported higher κ values for the detection of effusion/synovitis and power Doppler signal (0.86 and 0.95, respectively) in the wrist and small joints of the hand and foot. In addition, the κ value for ultrasonographic identification of metacarpophalangeal erosions was 0.76 in the study by Wakefield et al.[7] Finally, Hauzeur[9] and Karim[5] reported κ values of 0.90 and 0.71 for the detection of knee effusion and synovitis, respectively.

Although the results of the Train the Trainers interobserver study were moderate to good (overall κ for all examined joints = 0.76), we organised the Teach the Teacher course four months later in order to re-evaluate the interobserver reliability of the main periarticular and intra-articular ultrasonographic diagnoses and reveal the principal disagreements between the participants by scanning patients together in real time.

Even though we showed a high level of overall agreement, our κ values were lower than those communicated in individual studies by some rheumatologists of the group.[4 5 7 8 10 11] There may be several reasons for these differences. In previous reliability reports the two examiners worked at the same hospital and used the same machine, probably had a common ultrasonographic background, and usually reached consensus on scanning and diagnostic criteria before the study. However, in the present study as well as in the one by Scheel et al,[13] the experts—despite meeting for a few days on several occasions in the past six years—work in different hospitals and countries and many were not familiar with the ultrasonographic equipment. The latter may explain the interobserver variability for power Doppler findings among participants. In addition, the examiners were unaware of the patient's clinical data and

did not train together before the investigation—indeed, the aim of the study was to assess the interobserver reliability of the spontaneous ultrasonographic evaluation carried out by experts within the usual time spent on it in daily clinical practice. Nevertheless, our ultrasonographic interobserver reliability was similar to or better than that described in studies on MRI reliability in the detection of rotator cuff disorders[8 16] or joint synovitis, erosions, and tenosynovitis,[17] or on interobserver variability of the clinical examination of joint inflammation.[12 18] Both clinical evaluation and MRI are widely considered to be the gold standard in clinical trials.

With regard to the second objective of our study, some issues should be explored. As Scheel et al[13] reported, multiplanar and dynamic scans were not carried out by all the experts. Dynamic ultrasonography is very useful for detecting subtle musculoskeletal abnormalities such as small bone erosions, tendon tears, and minimum fluid within synovial recesses and tendon sheaths, and probably should be used for all musculoskeletal ultrasonographic studies. A more intensive training in standardisation of scanning methods is likely to improve the sensitivity and reliability of musculoskeletal ultrasonography.

Another point of interest is to identify which recesses of each joint should be scanned for detecting synovitis. As the sensitivity of ultrasonographic detection of synovitis has not yet been compared in the different joint recesses, most experts scan all of them, although it makes the examination longer. Future studies providing evidence of the more sensitive joint recesses for detecting intra-articular inflammation would be very useful to shorten scanning time.

In addition, more accurate definitions of tendinosis, tendon partial tear, and complete tear—mainly rotator cuff lesions—based on validation studies of the ultrasonographic semiology are needed to improve interobserver agreement.

Finally, it was not easy to reach consensus among experts on the subjective diagnosis of pathological mild joint effusion, tenosynovitis, or bursitis versus normality. Physiological fluid in joint recesses, synovial sheath of tendons, and large bursae, as well as hypoechoic rims in joints that correspond to normal synovial fluid or articular cartilage, or both, are commonly detected with high resolution ultrasonography machines in normal subjects.[19] Although in our study the experts used the same machine for scanning the same patient, their different ultrasonography backgrounds could have influenced the final diagnosis. Objective diagnostic criteria of pathological fluid within joints, tendon sheaths, and bursae are necessary to distinguish normality from mild pathology, independent of the ultrasonography machine used. This emphasises the relevance of the study by Schmidt et al,[19] who determined standard reference values for musculoskeletal ultrasonography in a large series of healthy adults. Nevertheless, a rheumatological ultrasonography approach correlating ultrasonographic findings with clinical symptoms is always recommended.

Some limitations of our study should be mentioned. For example, κ values could not be calculated for each ultrasonographic diagnosis in all the regions because the observers in each group changed during the study. This is inconvenient from a statistical point of view. However, the main goal of the Teach the Teacher course was to work with as large a number of different experts as possible.

Conclusion

Further meetings of the EULAR/OMERACT musculoskeletal ultrasonography group for training in standardisation of scanning method, establishing definitions, quantifying ultrasonographic pathologies, and assessing reproducibility, sensitivity to change, and intermachine variability are necessary.

These future exercises will contribute to the expanding use of musculoskeletal ultrasonography in clinical and research rheumatology to improve the evaluation of inflammatory activity and therapeutic response in patients with rheumatic diseases.

. . . .

Authors' affiliations

E Naredo, Rheumatology, Hospital Severo Ochoa, Madrid, Spain
I Möller, Department of Rheumatology, Instituto Poal, Barcelona, Spain
C Moragues, Rheumatology, Hospital de Bellvitge, Barcelona
J J de Agustín, Rheumatology, Hospital Vall d'Hebron, Barcelona
A K Scheel, Nephrology and Rheumatology, Georg-August-University, Göttingen, Germany
W Grassi, Department of Rheumatology, Universita Politecnica delle Marche, Ancona, Italy
E de Miguel, Rheumatology, Hospital La Paz, Madrid, Spain
M Backhaus, Rheumatology and Clinical Immunology, Charité University Hospital, Berlin, Germany
P Balint, Third General and Paediatric Rheumatology Department, National Institute of Rheumatology and Physiotherapy, Budapest, Hungary
G A W Bruyn, Rheumatology, Medisch Centrum Leeuwarden, Leeuwarden, Netherlands
M A D'Agostino, Rheumatology, Ambroise Paré Hospital, UVSQ University, Boulogne Billancourt, France
E Filippucci, Rheumatology, Universita Politecnica delle Marche, Ancona, Italy
A Iagnocco, Rheumatology, University La Sapienza, Rome, Italy
D Kane, Rheumatology, University of Newcastle upon Tyne, Newcastle upon Tyne, UK
J M Koski, Rheumatology, Mikkeli Central Hospital, Mikkeli, Finland
L Mayordomo, Rheumatology, Hospital de Valme, Sevilla, Spain
W A Schmidt, Medical Centre for Rheumatology Berlin-Buch, Berlin, Germany
W A A Swen, Rheumatology, Medisch Centrum Alkmaar, Alkmaar, Netherlands
M Szkudlarek, Rheumatology, University of Copenhagen, Hvidovre Hospital, Copenhagen, Denmark
L Terslev, Rheumatology, The Parker Institute, Frederiksberg Hospital, Frederiksberg, Denmark
S Torp-Pedersen, Radiology, The Parker Institute, Frederiksberg Hospital
J Uson, Rheumatology, Hospital de Mostoles, Madrid, Spain
R J Wakefield, Academic Department of Musculoskeletal Medicine, Leeds General Infirmary, Leeds, UK
C Werner, Medical Statistics, Georg-August-University Göttingen, Göttingen, Germany

REFERENCES

1 **Grassi W**, Cervini C. Ultrasonography in rheumatology: an evolving technique. *Ann Rheum Dis* 1998;**57**:268–70.
2 **Wakefield RJ**, Gibbon WW, Emey P. The current status of ultrasonography in rheumatology. *Rheumatology* 1999;**38**:195–201.
3 **Kane D**, Balint PV, Sturrock RD. Ultrasonography is superior to clinical examination in the detection and localization of knee joint effusion in rheumatoid arthritis. *J Rheumatol* 2003;**30**:966–71.
4 **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.
5 **Karim Z**, Wakefield RJ, Quinn M, Conaghan PG, Brown AK, Veale DJ, et al. Validation and reproducibility of ultrasonography in the detection of synovitis in the knee. *Arthritis Rheum* 2004;**50**:387–94.
6 **Szkudlarek M**, Narvestad E, Klarlund M, Court-Payen M, Thomsen HS, Ostergaard M. Ultrasonography of the metatarsophalangeal joints in rheumatoid arthritis. *Arthritis Rheum* 2004;**50**:2103–12.

7 **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. *Arthritis Rheum* 2001;**43**:2762–70.

8 **Swen WAA**, Jacobs JWG, Algra PR, Manoliu RA, Rijkmans J, Willems WJ, *et al.* Sonography and magnetic resonance imaging equivalent for the assessment of full-thickness rotator cuff tears. *Arthritis Rheum* 1999;**42**:2231–8.

9 **Middleton WD**, Teefey SA, Yamaguchi K. Sonography of the rotator cuff: analysis of interobserver variability. *Am J Roentgenol* 2004;**183**:1465–8.

10 **Iagnocco A**, Ossandon A, Coari G, Conti F, Priori R, Alessandri C, *et al.* Wrist joint involvement in systemic lupus erythematosus. An ultrasonographic study. *Clin Exp Rheumatol* 2004;**22**:621–4.

11 **Filippucci E**, Farina A, Carotti M, Salaffi F, Grassi W. Grey scale and power Doppler sonographic changes induced by intra-articular steroid injection treatment. *Ann Rheum Dis* 2004;**63**:740–3.

12 **Hauzeur EP**, Mathy L, De Maertelaer V. Comparison between clinical evaluation and ultrasonography in detecting hydrarthrosis of the knee. *J Rheumatol* 1999;**26**:2681–3.

13 **Scheel AK**, Schmidt WA, Hermann KG, Bruyn GA, D'Agostino MA, Grassi W, *et al.* Interobserver reliability of rheumatologist performing musculoskeletal ultrasonography: results from a EULAR ''Train the Trainer'' course. *Ann Rheum Dis* 2005;**64**:1043–9.

14 **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.

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

16 **Robertson PL**, Schweitzer ME, Mitchell DG, Schlesinger F, Epstein RE, Frieman BG, *et al.* Rotator cuff disorders: interobserver and intraobserver variation in diagnosis with MR imaging. *Radiology* 1995;**194**:831–5.

17 **Hoving JL**, Buchbinder R, Hall S, Lawler G, Coombs P, McNealy 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.

18 **Naredo E**, Bonilla G, Gamero F, Uson J, Carmona L, Laffon A. Assessment of inflammatory activity in rheumatoid arthritis: a comparative study of clinical evaluation with gray-scale and power Doppler ultrasonography. *Ann Rheum Dis* 2005;**64**:375–81.

19 **Schmidt WA**, Schmidt H, Schicke B, Gromnica-Ihle E. Standard reference values for musculoskeletal ultrasonography. *Ann Rheum Dis* 2004;**63**:988–94.