RHEUMATOLOGY

doi:10.1093/rheumatology/kes161
Advance Access publication 27 July 2012

Original article

The interobserver reliability of ultrasound in knee osteoarthritis

Annamaria Iagnocco[1], Carlo Perricone[1], Chiara Scirocco[1], Fulvia Ceccarelli[1], Mariagrazia Modesti[1], Angelica Gattamelata[1], Caterina Vavala[1], Iolanda Maria Rutigliano[1], Anca Musetescu[2] and Guido Valesini[1]

Abstract

Objective. To assess the interobserver reliability between sonographers with different levels of experience in detecting inflammatory and structural damage abnormalities in patients with knee OA.

Methods. After achieving consensus on definitions and scanning protocols, three ultrasonographers with different levels of experience in musculoskeletal US examined the knees of nine patients with OA. US examinations were conducted with independent blinded evaluations of inflammatory (joint effusion, synovial hypertrophy, power Doppler signal, Baker's cysts) and structural (osteophytes, cortical bone irregularities, femoral hyaline cartilage abnormalities, protrusion of the medial meniscus) lesions. All abnormalities were scored by applying a dichotomous scale (0–1). In addition, at each knee joint site global scores for joint inflammation, cortical bone abnormalities and cartilage damage were calculated by summing the single-lesion scores. Reliability was assessed using kappa (κ) coefficients.

Results. Seventeen knees were examined. Inflammatory abnormalities were observed with moderate to very good agreement ($\kappa = 0.55$–0.88) between the observers. From fair to very good agreement ($\kappa = 0.31$–0.82) was registered between sonographers for structural damage lesions. The overall κ was 0.716 for junior and 0.571 for beginner sonographers comparing their findings with those of senior sonographers.

Conclusion. This represents the first ultrasonographic study focusing on the analysis of interobserver reliability between sonographers with different levels of experience in demonstrating inflammatory and structural abnormalities in knee OA. Globally, even considering some variable results that were mainly obtained by the evaluation of single components of bone involvement, US offered a reliable assessment of a wide set of abnormalities in knee OA.

Key words: ultrasound, knee osteoarthritis, reliability.

Introduction

OA is a common rheumatic disease of the joints, which affects primarily older people (70% older than 65 years) [1]. Both small and large joints can be involved and the knee is a prime site of involvement. The pathological characteristics of the disease consist of predominant cartilage damage with concomitant bone abnormalities and inflammatory processes within the synovium. Over time OA causes a loss of joint function, causing disability and worsening quality of life [2–6]. Imaging techniques have a fundamental role in the assessment of OA. Conventional radiography (CR) is the traditional tool for imaging the osteoarthritic joint and has been demonstrated to be readily available, inexpensive and reliable [7–10]. However, it gives only a two-dimensional image of a tridimensional joint site, it cannot detect inflammation and soft-tissue abnormalities and it exposes patients to ionizing radiation [11].

[1]Rheumatology Unit, Dipartimento di Medicina Interna e Specialità Mediche, Sapienza Università di Roma, Rome, Italy and [2]Rheumatology Department, Emergency Clinical County Hospital Craiova, University of Medicine and Pharmacy Craiova, Craiova, Romania.

Submitted 7 March 2012; revised version accepted 8 May 2012.

Correspondence to: Annamaria Iagnocco, Dipartimento di Medicina Interna e Specialità Mediche: Reumatologia, Sapienza Università di Roma, Policlinico Umberto 1°, Viale del Policlinico 155, 00161, Roma. E-mail: annamaria.iagnocco@uniroma1.it

Musculoskeletal US is an imaging tool with an increasing role in the assessment of OA [12, 13]. It has been demonstrated to show findings related to both inflammation and structural damage [1, 2, 4–7, 14–19]. In addition, it is characterized by a wide set of advantages over other imaging modalities, being safe, easily accessible, relatively cheap, not invasive and lacking any contraindications [20]. Moreover, Doppler modalities are able to differentiate active and inactive inflammation within joints and periarticular soft tissues. However, a limitation to its widespread use consists of the common perception in the medical community that it is a highly operator-dependent technique [21–23]. This is related to various aspects that may influence the skill of the sonographer, including operator experience in the phases of image acquisition and interpretation, knowledge of anatomical details and ability to exploit all the functions of the equipment. Variable intra- and interobserver reliability of US has been reported in the literature in recent years and has been assessed mainly in inflammatory arthritis [23–25]. However, the different influencing factors have not been adequately addressed thus far, and only limited focus has been developed in OA.

The aim of the present study was to assess the interobserver reliability between sonographers with different levels of experience in detecting inflammatory and structural damage abnormalities in patients with knee OA. As secondary objectives, we aimed to assess the relationship of US parameters to clinical and radiographic parameters.

Methods

Patients

Consecutive patients with knee OA, fulfilling the American College of Rheumatology (ACR) classification criteria for knee OA [26], were recruited from the rheumatology outpatient clinic of the Dipartimento di Medicina Interna e Specialità Mediche, Sapienza Università di Roma, Rome, Italy. Inclusion criteria were age >18 years, primary knee OA and symptoms >6 months. Exclusion criteria were the presence of any other rheumatic diseases, history of trauma or knee surgery, intra-articular corticosteroid or hyaluronic acid injections in the previous 6 months. Patients underwent clinical examination and power Doppler (PD) US assessment of both knees on the same day. This study was conducted in compliance with good clinical practice. Ethics committee approval was obtained from the Ethics Committee of the Policlinico Umberto I – Sapienza Università di Roma, Rome, Italy, and patients gave written informed consent.

Clinical assessment

Demographic and clinical data were recorded using standardized forms by an expert rheumatologist who took the clinical history and performed the physical examination.

Age, sex, disease duration, BMI and therapy assumption during the last 4 weeks were recorded. Patient's pain and severity were assessed with a visual analogue scale (VAS) of 0–100 mm, and the WOMAC index was used [27]. Concerning the presence of pain and stiffness, we applied the WOMAC scale separately to the right and left knee. Physical examination included the evaluation of effusion that was scored with a four-point scale. Standard CR was performed in all patients; OA severity was classified according to the Kellgren and Lawrence (KL) radiological score [28].

US assessment

Before patient enrolment, US examination methodology was clarified among sonographers, and a consensus was obtained on scanning protocol and image interpretation. The sonographers were three separate operators, with different levels of experience in musculoskeletal US. The senior, the junior and the beginner had a history of 24 years, 5 years and 3 months, respectively, of active scanning. Before the reliability study the senior ultrasonographer had performed more than 10 000 US examinations of the knee; the junior more than 3000 and the beginner 60. Moreover, the beginner, during her training period, was supervised by an expert (the senior). The backgrounds of the junior and the beginner ultrasonographers were similar, both having been trained by the senior ultrasonographer. All the operators performed the US examinations independently on the same day in each patient and registered their findings, being unaware of the other operators' results and of clinical data.

In all cases the examination was performed using a MyLab70 XVG (Esaote Biomedica, Genoa, Italy) machine equipped with a linear multifrequency (4–13 MHz) transducer, operating at a frequency of 13 MHz; in addition, PD modality was applied (PRF 750 Hz, gain 50%, frequency 6.3 MHz). The same settings were used in all cases. At the beginning of each scanning session focus was positioned at the level of the region of interest. Colour gain was adjusted below the degree that caused the appearance of noise artefacts [29].

US scans were carried out following a protocol based on European League Against Rheumatism (EULAR) guidelines for musculoskeletal ultrasonography [30]. After having applied gel to the skin to provide an acoustic interface, US examinations were started, paying attention to not applying probe pressure on the anatomical structures under examination. During the same scanning session, US was initially performed in B-mode modality with the aim of detecting morphological changes and immediately afterwards using PD technique searching for synovial abnormal vascularization.

Basic lesions related both to inflammation and structural damage were assessed according to international definitions [31–32] as follows: joint effusion (JE), synovial hypertrophy (SH), pathological hypervascularization at PD, Baker's cyst (BC), osteophytes, cortical bone irregularities, femoral hyaline cartilage abnormalities and protrusion of the medial meniscus. All abnormalities were scored according to a dichotomous scale (0-1). Patients were examined in the supine position with the knee flexed

at 30°. US assessment of the femoral hyaline cartilage was performed with the knee flexed at 120°.

In addition, global scores for joint inflammation, cortical bone abnormalities and cartilage damage were calculated. At each knee joint site, the inflammatory score was obtained as the sum of JE, SH, PD and BC scores (total 0–8); the bone damage score was calculated as the sum of medial and lateral osteophytes and bone irregularity scores (total 0–8); the cartilage damage score was obtained as the sum of loss of anechoic structure, loss of sharpness of at least one margin, irregularity of at least one margin and thinning of the layer scores (total 0–8).

Statistical analysis

Statistical analysis was accomplished using Statistical Package for Social Sciences 13.0 (SPSS 13.0). The results of each operator were reported as absolute numbers and percentages; the mean value (±S.D.) was also calculated. The junior's and the beginner's results were compared with those obtained by the senior and interobserver reliability using standard Cohen's kappa (κ) coefficients [33]. K values for each abnormality were calculated and were considered representative of specific agreement levels. K coefficients were interpreted according to Landis and Koch (<0 absence of agreement, 0.1–0.20 slight, 0.21–0.40 fair, 0.41–0.60 moderate, 0.61–0.80 good and 0.81–1 almost perfect agreement) [34]. We used Wilcoxon's test and *t*-test to obtain the difference among median and mean values, respectively, of quantitative variables. The correlation between clinical, radiographic and ultrasonographic data was calculated with Spearman's test. Two-tailed $P \leqslant 0.05$ was considered significant.

Results

Demographic, clinical and radiographic characteristics of patients included in the study as well as the concomitant therapy are reported in Table 1. No significant differences for clinically detected features were found between the right and left knees.

A total of 17 knees from nine patients were examined. In one patient, only one knee was examined, due to previous arthroprosthesis operation. Only two patients reported knee pain at the time of the examination: one at both the medial and lateral femorotibial aspects of the right knee, the other at the lateral femorotibial portion of the right knee.

When the patients were studied by applying VAS pain, VAS severity (patient's and physician's), WOMAC pain, WOMAC stiffness and WOMAC physical function, there were no differences between right and left knees. Four patients reported morning stiffness that had a duration of no longer than 15 min. Clinically detected knee JE was present in two patients. Three patients were receiving treatment consisting of NSAIDs (two patients) or chondroprotective agents (one subject). The radiological severity by applying the KL score was similar between right and left knees.

The prevalence of US-detected abnormalities is reported in Table 2. Considering the results of the senior sonographers as the reference data, findings related to both inflammation and structural damage were detected. A high prevalence of abnormalities was registered at both right and left knees, JE being the most frequent inflammatory finding (50–66.67%), followed by SH (25–55.50%) BC (37.5–44.44%) and PD signal (0–11.11%). Structural abnormalities also showed a high prevalence, with evidence of cartilage lesions in 50%–100% of cases, bone abnormalities in 55.56–88.89% and, finally, meniscal extrusion in 37.5%–55.56%. The US findings on the total of the knees and the interobserver reliability results (senior *vs* junior and senior *vs* beginner κ values) are reported in Table 3.

Inflammatory abnormalities were observed with moderate to very good agreement (κ = 0.55–0.88) between the

Table 1 Demographic, clinical and radiographic features of patients and concomitant therapy

Demographic and clinical features (patients $n = 9$)	Mean (S.D.)	P^a
Sex	6 F (75%); 2 M (25%)	
Age, years	61.33 (7.35)	
Disease duration, months	82.00 (93.43)	
Weight, kg	71.00 (8.21)	
Height, cm	161.63 (6.41)	
BMI, kg/cm^2	27.50 (4.31)	
VAS pain (0–100 mm)		
Right	38.89 (27.13)	0.2676
Left	40.63 (26.25)	
VAS severity (patient) (0–100 mm)		
Right	41.11 (21.47)	0.2021
Left	48.75 (18.85)	
VAS severity (physician) (0–100 mm)		
Right	34.44 (20.68)	0.0579
Left	46.25 (22.00)	
WOMAC pain		
Right	0.71 (0.93)	0.2476
Left	1.03 (1.14)	
WOMAC stiffness		
Right	0.22 (0.67)	0.25
Left	0.69 (1.03)	
WOMAC physical function		
Right	0.75 (0.79)	0.1755
Left	0.99 (0.82)	
Knee joint effusion (clinically evaluated)	2 pts (22.2%)	
Concomitant therapy	Aceclofenac 1 pt Diclofenac 1 pt Galactosamineglucuronoglycan sulphate 800 mg daily 1 pt	
KL knee score (0–4)		
Right	2.22 (0.83)	1.0
Left	2.25 (0.89)	

[a]Wilcoxon's matched paired test.

TABLE 2 Sonographic findings of left and right knees obtained by the senior, junior and beginner ultrasonographers (number and percentage of positive cases for each abnormality)

US findings	Senior ultrasonographer *N* (%)	Junior ultrasonographer *N* (%)	Beginner ultrasonographer *N* (%)
Right knee ($n = 9$)			
Joint effusion	6 (66.67)	6 (66.67)	6 (66.67)
Synovial hypertrophy	5 (55.56)	3 (33.33)	6 (66.67)
Power Doppler signal	1 (11.11)	0 (0)	1 (11.11)
BC	4 (44.44)	3 (33.33)	4 (44.44)
Medial meniscal extrusion	5 (55.56)	6 (66.67)	6 (66.67)
Medial osteophytes	8 (88.89)	9 (100)	8 (88.89)
Lateral osteophytes	7 (77.78)	8 (88.89)	8 (88.89)
Bone irregularities medial	5 (55.56	8 (88.89)	9 (100)
Bone irregularities lateral	8 (88.89)	9 (100)	8 (88.89)
Cartilage: thinning of the layer	7 (77.78)	7 (77.78)	8 (88.89)
Cartilage: loss of anechoic structure	9 (100)	9 (100)	8 (88.89)
Cartilage: irregularity of the superficial margin	9 (100)	8 (88.89)	8 (88.89)
Cartilage: loss of sharpness of at least one margin	9 (100)	8 (88.89)	8 (88.89)
Left knee ($n = 8$)			
Joint effusion	4 (50)	4 (50)	5 (62.5)
Synovial hypertrophy	2 (25)	2 (25)	5 (62.5)
Power Doppler signal	0 (0)	0 (0)	0 (0)
BC	3 (37.5)	3 (37.5)	4 (50)
Medial meniscal extrusion	3 (37.5)	4 (50)	6 (75)
Medial osteophytes	7 (87.5)	8 (100)	8 (100)
Lateral osteophytes	6 (75)	8 (100)	6 (75)
Bone irregularities medial	6 (75)	7 (87.5)	8 (100)
Bone irregularities lateral	6 (75)	7 (87.5)	6 (75)
Cartilage: thinning of the layer	4 (50)	7 (87.5)	6 (75)
Cartilage: loss of anechoic structure	6 (75)	7 (87.5)	7 (87.5)
Cartilage: irregularity of the superficial margin	6 (75)	5 (62.5)	7 (87.5)
Cartilage: loss of sharpness of at least one margin	6 (75)	6 (75)	7 (87.5)

observers. Due to the very low prevalence of PD signal (only one observation by the senior investigator), the data are not conclusive for this finding and may generate an unbalanced marginal total on calculation (paradox of the low κ and good agreement). Fair to very good agreement ($\kappa = 0.31$–0.82) was registered between sonographers for structural damage lesions. Due to the constant findings (100%) of medial osteophytosis and medial bone irregularities, data were considered inconclusive concerning the agreement between sonographers.

Overall κ was 0.716 for the junior and 0.571 for the beginner sonographer comparing their findings with those of the senior sonographer. Concerning the inflammatory score, an excellent agreement ($\kappa = 0.81$) was registered between the junior's and senior's findings, and a substantial agreement ($\kappa = 0.699$) between the beginner's and senior's findings. Concerning the bone structural damage score, the junior's and the beginner's findings showed, respectively, moderate ($\kappa = 0.479$) and fair ($\kappa = 0.247$) agreement when compared with the senior's findings. Regarding the cartilage damage score, the junior's findings showed good agreement ($\kappa = 0.634$) when compared with the results of the senior, while the beginner's findings demonstrated moderate agreement ($\kappa = 0.495$). Representative sonographic images of US findings evaluated by the ultrasonographers are shown in Figure 1.

Clinical correlations

WOMAC pain positively correlated with patient weight ($r = 0.813$, $P = 0.014$) as well as with the BMI ($r = 0.740$, $P = 0.036$). VAS severity assessed by the physician correlated with WOMAC (WOMAC pain, WOMAC stiffness and WOMAC physical function: $r = 0.705$, $P = 0.002$; $r = 0.656$, $P = 0.004$; $r = 0.689$, $P = 0.002$, respectively).

Correlation between clinical features and sonographic findings

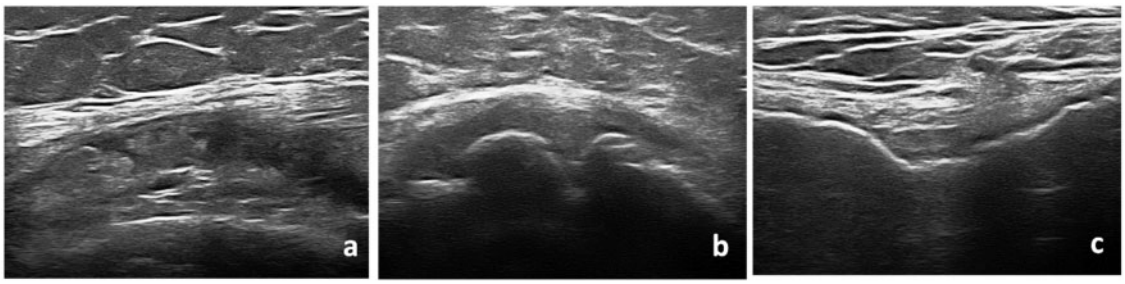
VAS pain positively correlated with the overall cartilage damage assessed by the senior ($r = 0.528$, $P = 0.028$), and by the junior investigators ($r = 0.639$, $P = 0.008$). The patient VAS severity correlated with the global inflammatory findings assessed by the senior sonographer ($r = 0.557$, $P = 0.025$) and the physician VAS severity correlated with global cartilage damage ($r = 0.557$, $P = 0.025$). The pain VAS correlated with the KL score ($r = 0.609$, $P = 0.012$).

TABLE 3 Global sonographic findings for the 17 joints

US findings ($n = 17$)	Senior ultrasonographer *N* (%)	Junior ultrasonographer *N* (%)	Beginner ultrasonographer *N* (%)	Senior *vs* junior		Senior *vs* beginner	
				κ	Agreement	κ	Agreement
Joint effusion	10 (58.82)	10 (58.82)	11 (64.71)	0.757	Good	0.628	Good
Synovial hypertrophy	7 (41.18)	5 (29.41)	11 (64.71)	0.746	Good	0.553	Moderate
Power Doppler signal	1 (5.88)	0 (0)	1 (5.88)	—[a]		—[a]	—[a]
BC	7 (41.18)	6 (35.29)	8 (47.06)	0.876	Very good	0.881	Very good
Medial meniscal extrusion	8 (47.06)	10 (58.82)	12 (70.59)	0.767	Good	0.311	Fair
Medial osteophytosis	15 (88.24)	17 (100)	16 (94.12)	—[b]	—[b]	−0.85	Poor
Lateral osteophytosis	13 (76.47)	16 (94.12)	14 (82.35)	0.338	Fair	0.821	Very good
Medial bone irregularities	11 (64.71)	15 (88.24)	17 (100)	0.89	Very good	—[b]	—[b]
Lateral bone irregularities	14 (82.35)	16 (94.12)	14 (82.35)	0.452	Moderate	−0.214	Poor
Cartilage							
Thinning of the layer	11 (64.71)	14 (82.35)	14 (82.35)	0.564	Moderate	0.564	Moderate
Loss of anechoic structure	15 (88.24)	16 (94.12)	15 (88.24)	0.638	Good	0.433	Moderate
Irregularity of the superficial margin	15 (88.24)	13 (76.47)	15 (88.24)	0.605	Good	0.433	Moderate
Loss of margin sharpness	15 (88.24)	14 (82.35)	15 (88.24)	0.767	Good	0.433	Moderate

Number and percentage of the single abnormalities detected by the senior, junior and beginner operators. Inter-observer reliability considering the senior's findings as the reference data, κ values and agreement. [a]The low prevalence of the lesions created unbalanced marginal totals on calculation (paradox of the low κ and high agreement). [b]Agreement cannot be calculated as one of the findings is a constant.

FIG. 1 Knee US in a patient with OA. Evidence of synovial hypertrophy and joint effusion (**a**), osteophytes over the medial aspect of the femorotibial joint (**b**) and irregularities of the margins, loss of anechoic structure and thinning of the cartilage layer (**c**).



Correlation between radiographic and sonographic findings

The KL score correlated with global structural damage assessed by the senior ultrasonographer ($r = 0.523$, $P = 0.038$) and by the junior sonographer ($r = 0.625$, $P = 0.01$). Finally, the KL score correlated with the cartilage damage assessed by the senior operator ($r = 0.525$, $P = 0.037$).

Discussion

As far as we know, this represents the first ultrasonographic study focusing on the analysis of interobserver reliability between sonographers with different levels of experience in demonstrating inflammatory and structural abnormalities in knee OA. Globally, even considering variability in the results that were mainly obtained by the evaluation of single components of bone involvement, US offered a reliable assessment of a wide set of abnormalities in knee OA.

In particular, satisfactory results were obtained in the detection of inflammatory findings both between the senior and junior and between the senior and beginner ultrasonographers. Considering the general perception that sonography is a highly operator-dependent technique that also requires a long learning curve and great skill,

these findings appear significantly encouraging and represent a relevant step in support of widespread application of US.

However, analysis of structural abnormalities showed variable results. These findings were evident particularly for the assessment of bone lesions, which demonstrated fair to very good agreement between sonographers. More satisfactory results were obtained for the detection of cartilage abnormalities, which showed moderate to good levels of agreement both between the senior and junior and between the senior and beginner operators. These variable results are probably due to persistent difficulties in the depiction and interpretation of bone lesions by US, even after the obtained consensus for scanning technique and image interpretation.

Today, US has acquired increasing importance and widespread use in the assessment of rheumatic diseases, particularly in the evaluation of joint involvement in OA patients [35]. However, its widespread use is still limited by the general belief that it is an operator-dependent technique. This conviction often makes its use in routine clinical practice as well as in clinical trials difficult. The belief that US needs a long learning curve and a long period of training accounts for the hesitant approach to this imaging tool by a number of rheumatologists. However, the numerous advantages of US over other imaging modalities (i.e. its lower costs, its safety and its widespread availability) make its use advisable for the assessment of various lesions in knee OA. In addition, our findings support the widespread use of US in clinical practice, showing that this imaging tool is a reliable modality that requires neither particular skill nor a long apprenticeship. Moreover, US offers the opportunity to assess certain individual aspects of OA pathology that contribute to the global knee joint involvement.

Our results are in agreement with those obtained by previous studies in inflammatory arthritis that investigated the reliability among differently experienced sonographers [36, 37] in analysing different aspects of a wide set of abnormalities in tenosynovitis and arthritis and demonstrated, similarly to our study, variable agreement that ranged from moderate to good.

Concerning correlations between clinical abnormalities, our findings highlighted that weight and BMI correlate with pain (as measured by WOMAC). In addition, the correlations between the US signs of inflammatory/structural lesions and the clinical/radiographic indices of disease may be interpreted as a greater adherence of US experience to the clinical findings.

In conclusion, the present study demonstrates that US is a reliable tool for assessing different abnormalities in patients with knee OA, even when the exam is performed by operators with limited experience. The variable results obtained by the sonographic analysis of single components of bone involvement may be improved by the application of strict definitions and the standardization of US scanning techniques. In addition, these results represent a starting point for further sonographic studies on inflammatory and structural damage lesions in knee OA.

Rheumatology key messages

- US is a reliable tool for detecting inflammatory and structural abnormalities in knee OA.
- US offers a reliable assessment of a wide set of abnormalities in knee OA.
- Knee US in OA can be reliably performed even by operators with limited experience.

Acknowledgements

A.I. made substantial contributions to conception and design, acquisition of data, analysis and interpretation of data and writing the manuscript; C.P. made substantial contributions to analysis and interpretation of data and drafting the manuscript; C.S. made substantial contributions to interpretation of data and drafting the manuscript; F.C. made substantial contributions to acquisition of data and drafting the manuscript; M.M. made substantial contributions to interpretation of data and drafting the manuscript; A.G. made substantial contributions to interpretation of data and drafting the manuscript; C.V. made substantial contributions to interpretation of data and drafting the manuscript; I.R. made substantial contributions to interpretation of data and drafting the manuscript; A.M. made substantial contributions to acquisition of data and drafting the manuscript; G.V. made substantial contributions to conception and design, interpretation of data and drafting the manuscript. All authors read and approved the final manuscript.

Disclosure statement: The authors have declared no conflicts of interest.

References

1 Felson DT. The epidemiology of osteoarthritis: prevalence and risk factors. In: Kuettner KE, Goldberg VM, eds. Osteoarthritic disorders. Rosemont, IL: American Academy of Orthopedic Surgeons 1995:13–24.

2 Krasnokutsky S, Samuels J, Abramson SB. Osteoarthritis in 2007. Bull NYU Hosp Jt Dis 2007;65:222–8.

3 Zhang W, Doherty M, Peat G *et al*. EULAR evidence-based recommendations for the diagnosis of knee osteoarthritis. Ann Rheum Dis 2010;69:483–9.

4 Samuels J, Krasnokutsky S, Abramson SB. Osteoarthritis: a tale of three tissues. Bull NYU Hosp Jt Dis 2008;66: 244–50.

5 Altman RD. Early management of osteoarthritis. Am J Manag Care 2010;16:S41–7.

6 Abramson SB, Attur M. Developments in the scientific understanding of osteoarthritis. Arthritis Res Ther 2009;11: 227.

7 Bliddal H, Boesen M, Christensen R *et al*. Imaging as a follow-up tool in clinical trials and clinical practice. Best Pract Res Clin Rheumatol 2008;22:1109–26.

8 Burstein D, Hunter DJ. ''Why aren't we there yet?'' Reexamining standard paradigms in imaging of OA. Summary of the 2nd annual workshop on imaging based

measures of osteoarthritis. Osteoarthris Cartil 2009;17:571–8.

9 Kubassova O, Boesen M, Peloschek P *et al*. Quantifying disease activity and damage by imaging in rheumatoid arthritis and osteoarthritis. Ann NY Acad Sci 2009;1154:207–38.

10 Hunter DJ, Conaghan PG. Imaging outcomes and their role in determining outcomes in osteoarthritis and rheumatoid arthritis. Curr Opin Rheumatol 2006;18:157–62.

11 Teichtahl AJ, Wluka AE, Davies-Tuck ML *et al*. Imaging of knee osteoarthritis. Best Pract Res Clin Rheum 2008;22:1061–74.

12 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.

13 Filippucci E, Iagnocco A, Meenagh G *et al*. Ultrasound imaging for the rheumatologist VII. Ultrasound imaging in rheumatoid arthritis. Clin Exp Rheumatol 2007;25:5–10.

14 Moller I, Bong D, Naredo E *et al*. Ultrasound in the study and monitoring of osteoarthritis. Osteoarthr Cartil 2008;16:S4–7.

15 D'Agostino MA, Conaghan P, Le Bars M *et al*. EULAR report on the use of ultrasonography in painful knee osteoarthris. Part 1: prevalence of inflammation in osteoarthritis. Ann Rheum Dis 2005;64:1703–9.

16 Wakefield RJ, Gibbon WW, Conaghan PG *et al*. The value of sonography in the detection of bone erosions in patients with rheumatoid arthritis: a comparison to conventional radiography. Arthritis Rheum 2000;43:2762–70.

17 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.

18 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.

19 Szkudlarek M, Narvestad E, Klarlund M *et al*. Ultrasonography of the metatarsophalangeal joints in rheumatoid arthritis: comparison with magnetic resonance imaging, conventional radiography and clinical examination. Arthritis Rheum 2004;50:2103–12.

20 Filippucci E, Iagnocco A, Meenagh G *et al*. Ultrasound imaging for the rheumatologist. Clin Exp Rheumatol 2006;24:1–5.

21 Karim Z, Wakefield RJ, Qiunn M *et al*. Validation and reproducibility of ultrasound in the detection of synovitis in the knee. Arthritis Rheum 2004;50:387–94.

22 Scheel AK, Schmidt WA, Hermann KG *et al*. Interobserver reliability of rheumatologists performing musculoskeletal ultrasonography: results from a EULAR 'Train the trainers' course. Ann Rheum Dis 2008;67:1678–82.

23 Naredo E, Moller I, Moragues C *et al*. Interobserver reliability in musculoskeletal ultrasonography: results from a 'Teach the Teachers rheumatologist course'. Ann Rheum Dis 2006;65:14–9.

24 Scheel AK, Schmidt WA, Hermann KG *et al*. Interobserver reliability of rheumatologists performing musculoskeletal ultrasonography: results from a EULAR 'Train the trainers' course. Ann Rheum Dis 2005;64:1043–9.

25 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.

26 Altman R, Asch E, Bloch D *et al*. Development of criteria for the classification and reporting of osteoarthritis: classification of osteoarthritis of the knee. Arthritis Rheum 1986;29:1039–49.

27 Bellamy N, Buchanan WW, Goldsmith CH *et al*. Validation of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in subjects with osteoarthritis of the hip or knee. J Rheumatol 1988;15:1833–40.

28 Kellgren J, Lawrence J. Radiological assessment of osteo-arthrosis. Ann Rheum Dis 1957;16:494.

29 Iagnocco A, Epis O, Delle Sedie A *et al*. Ultrasound imaging for the rheumatologist. XVII. Role of colour Doppler and power Doppler. Clin Exp Rheumatol 2008;26:759–62.

30 Backhaus M, Burmester GR, Gerber T *et al*. Guidelines for musculoskeletal ultrasound in rheumatology. Ann Rheum Dis 2001;60:641–9.

31 Wakefield RJ, Balint PV, Szkudlarek M *et al*. Muskuloskeletal ultrasound including definitions for ultrasonographic pathology. J Rheumatol 2005;32:2485–7.

32 Iagnocco A, Meenagh G, Riente L *et al*. Ultrasound imaging for the rheumatologist XXIX. Sonographic assessment of the knee in patients with osteoarthritis. Clin Exp Rheumatol 2010;28:643–6.

33 Cohen J. A coefficient of agreement for nominal scale. Educ Psychol Measure 1960;20:37–4.

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

35 Iagnocco A. Imaging the joint in osteoarthritis: a place for ultrasound? Best Pract Res Clin Rheumatol 2010;24:27–38.

36 Micu MC, Serra S, Fodor D *et al*. Interobserver reliability of ultrasound detection of tendon abnormalities at the wrist and ankle in patients with rheumatoid arthritis. Rheumatology 2011;33:1060–6.

37 Ohrndorf S, Naumann L, Grundey J *et al*. Is musculoskeletal ultrasonography an operator-dependent method or a fast and reliably teachable diagnostic tool? Interreader agreements of three ultrasonographers with different training levels. Int J Rheumatol 2010;2010:164518.