Original article

The 6-joint ultrasonographic assessment: a valid, sensitive-to-change and feasible method for evaluating joint inflammation in RA

Carlo Perricone[1], Fulvia Ceccarelli[1], Mariagrazia Modesti[1], Caterina Vavala[1], Manuela Di Franco[1], Guido Valesini[1] and Annamaria Iagnocco[1]

Abstract

Objective. Musculoskeletal US can be useful in monitoring RA. It can be time-consuming and there is no consensus in defining the joints to evaluate. We assessed the validity, sensitivity to change and feasibility of a reduced 6-joint US score in patients with RA starting therapy with an anti-TNF agent.

Methods. A group of consecutive RA patients starting etanercept were investigated. The patients underwent clinical evaluation, laboratory tests and US assessment at baseline and 3 months. A semi-quantitative score (0–3) was used to evaluate synovial effusion (SE), synovial proliferation (SP) and power Doppler (PD) signal in 12 joints. A process of data reduction, based on the frequency of synovial site involvement by US-SE, US-SP and US-PD signal, was conducted to investigate the validity of a 6-joint US assessment.

Results. Forty-five RA patients were evaluated. A significant decrease in all clinical, serological and 12-joint US parameters was found at follow-up. A significant correlation between changes in the DAS-28 and changes in the US scores in the 12-joint assessment was observed at follow-up ($P < 0.001$). A reduced 6-joint US score was obtained, including wrist, second MCP and knee joints of both sides, detecting US-SE in 97.78% of patients, US-SP in 100% of patients and positive US-PD in 100% of patients. The 6-joint US score showed a highly significant correlation with changes in DAS-28 ($P < 0.001$). The 6-joint evaluation was quick and easy to do.

Conclusion. A 6-joint US assessment may be a valid, sensitive-to-change and feasible method for evaluating joint inflammation in RA.

Key words: imaging, ultrasound, rheumatoid arthritis, anti-TNF.

CLINICAL SCIENCE

Introduction

The assessment of joint inflammation is essential in diagnosis and in monitoring response to therapies in patients affected by inflammatory arthropathies, such as RA. For this purpose, use of musculoskeletal US, with application of the power Doppler (PD) method, has been increasing over the past decade, thanks to its high sensitivity for detecting synovitis [1, 2]. Several studies have

demonstrated the capability of musculoskeletal US in monitoring response to different biological drugs and in analysing different joints and synovial recesses [3–9]. However, there is no evidence regarding which joints and synovial recesses should be evaluated to assess disease activity and response to biologic therapy in RA patients. Nonetheless, a comprehensive evaluation including multiple recesses of all accessible peripheral joints may be time consuming in daily practice and clinical trials.

Several studies have evaluated different simplified scores, showing good correlation with clinical disease activity indices [3–5]. Recently Naredo et al. [6] published a longitudinal study demonstrating the validity, reliability and sensitivity to change of a 12-joint simplified musculoskeletal US assessment compared with a comprehensive 44-joint US evaluation of joint inflammation in RA patients

starting biological treatment. However, the mean time required to perform this 12-joint US examination was 22 min, probably still too long to transform US into the stethoscope for the rheumatologist. Thus the aim of the present study was to investigate the validity, sensitivity to change and feasibility of a 6-joint US assessment in the evaluation of joint inflammation as compared with a 12-joint US examination in patients with established RA who started treatment with etanercept.

Methods

Patients

In this prospective study we included consecutive patients affected with RA, diagnosed according to the 1987 ACR criteria [10]. The patients were recruited in the Rheumatology Unit of the Dipartimento di Medicina Interna e Specialità Mediche, Sapienza Università di Roma. All the enrolled patients started therapy with etanercept, an anti-TNF agent, administered according to the Italian consensus on the use of biologic drugs for the treatment of RA because of inefficacy or intolerance to conventional DMARDs. The patients underwent clinical, laboratory and ultrasonographic evaluation at baseline before starting biological treatment and after 3 months. The study was conducted in accordance with the Declaration of Helsinki and was approved by the local ethics committee (Comitato Etico Sapienza Università di Roma). Informed consent was obtained from all patients before entry into the study.

Clinical and laboratory assessment

A single rheumatologist, who was blinded to the US findings, performed the clinical evaluation. Data, including demographics, date of diagnosis, comorbidities, past and present treatments, date of the beginning of therapy with etanercept and concomitant medications, were recorded on a standardized computerized form. RF (Behring, Germany; normal values <40 IU/ml) and anti-cyclic citrullinated protein/peptides antibodies (ACPAs; normal values <25 IU/ml) (Axis Shield, Dundee, Scotland) were detected by ELISA following the manufacturer's instructions. For each patient, ESR (mm/h) with the Westergen method and CRP (mg/dl) were also evaluated. A count of tender and swollen joints was performed, and a visual analogue scale (0–100 mm) for the patient's assessment of disease activity was administered. Disease activity was evaluated by calculating the 28-joint DAS and response to therapy was assessed according to EULAR guidelines [11].

Ultrasonographic assessment

Each patient underwent a musculoskeletal US assessment with application of PD. The ultrasonographic evaluation was performed by a single rheumatologist sonographer, experienced in musculoskeletal US, who was blinded to the clinical and laboratory findings. A systematic multiplanar grey-scale and PD examination of 12 joints (elbow, wrist, second MCP, third MCP, knee

and ankle of both sides) was performed using a MyLab 70 XVisionGold (Esaote, Firenze, Italy) machine equipped with a multifrequency linear array transducer (6–18 MHz). B-mode frequency ranged from 12 to 18 MHz (12 MHz for elbow, knee and ankle assessment, 15 MHz for wrist and 18 MHz for second and third MCP); PD pulse repetition frequency was 750 Hz; Doppler frequency was 6.7–11.1 MHz; low wall filters were used. At the beginning of each scanning session at different sites, focus was positioned at the level of the region of interest. Colour gain was adjusted just below the degree that caused the appearance of noise artefacts. The colour box was positioned at the level of the assessed site, enlarging the box to the upper part of the image.

The US assessment included 24 synovial sites in 12 joints: elbow (anterior and posterior recesses), wrist (dorsal carpal recesses), second and third MCP (dorsal side, palmar side), knee (suprapatellar recess, lateral parapatellar recess) and ankle (anterior tibiotalar recess, medial tibiotalar recess, lateral tibiotalar recess). These joints and synovial sites were chosen from the simplified 12-joint score previously described by Naredo et al. [6].

We then considered each joint as a unique structure, and we assessed the presence of synovial effusion (SE) and synovial proliferation (SP) by B-mode US and PD within the SP in each joint. According to the OMERACT definitions [12], SE and SP were defined as follows: SE as an abnormal hypoechoic or anechoic IA material that is displaceable and compressible, but does not exhibit PD signal; SP as an abnormal hypoechoic IA tissue that is non-displaceable and poorly compressible and may exhibit PD signal.

US-detected elementary lesions (US-SE, US-SP and US-PD) were scored according to a semi-quantitative scale (0 = absent, 1 = mild, 2 = moderate and 3 = severe). The higher score obtained for each of the US elementary lesions (US-SE, US-SP and US-PD) at each synovial site was then considered for the scoring of each joint as a unique structure.

Fig. 1 shows representative images of the four different degrees of US-SE in the anterior recess of the knee. Thus each of the 12 joints had a US score resulting from the sum of US-SE, US-SP and US-PD scores ranging from 0 to 9. Finally, from the sum of the scores at all joint sites, we obtained a US-total 12-joint score (ranging from 0 to 108).

Statistical analysis

Statistical analysis was performed using SPSS statistical software, version 13.0 (SPSS, Chicago, IL, USA). Quantitative variables (DAS-28, US parameters) were given as the mean (s.d.) and range. Comparisons between groups were performed using contingency tables and Pearson's χ^2. Corrections were made where necessary for the sample size (Fisher's exact test). The comparisons between parametric variables were performed with the Wilcoxon's test. One-way analysis of variance was applied to evaluate the comparisons between multiple

Fɪɢ. 1 Representative images of the four different degrees of US-SE in the anterior recess of the knee. Clockwise from upper left: (**a**) normal knee, grade = 0; (**b**) mild effusion, grade = 1; (**c**) moderate effusion, grade = 2; and (**d**) severe effusion, grade = 3. p: patella; t: tendon; f: femur; *: synovial effusion.
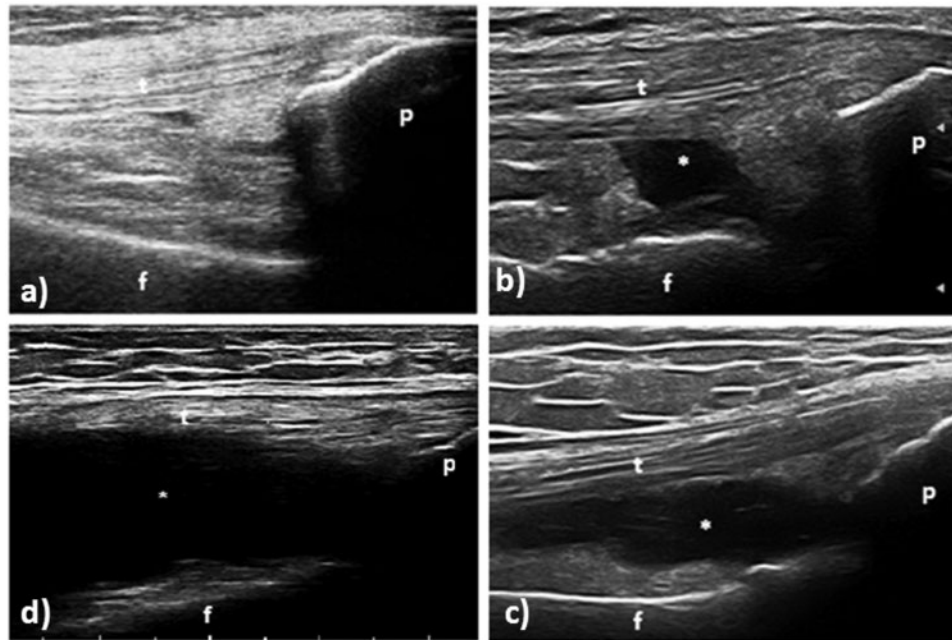
groups. Pearson's and Spearman's tests were used to perform the correlation analysis.

As a first step, we undertook a process of data reduction based on the frequency of joint involvement within the 12 joints by US-SE, US-SP and US-PD signal at baseline. A reduced US assessment was selected from different joint combinations. The reduced model was chosen considering that the joints selected should have allowed us to detect >97% of the joints involved by US-SE, US-SP and US-PD signal. The final identified model was named reduced US assessment. US-detected elementary lesions (US-SE, US-SP and US-PD) were scored for the reduced US assessment by the same method used for scoring the 12 joints, thus obtaining a reduced US count.

Afterwards, the sonographer was asked to record the time taken to perform the 12-joint US assessment and the reduced US assessment in two consecutive patients with RA. Content validity was evaluated by correlating the 12-joint US count (including US-SE, US-SP and US-PD scores) with the total reduced US count (obtained with the reduced US assessment) using Pearson's rank correlation coefficient. Construct validity was evaluated by testing the association between the 12-joint US count, the reduced US count and the EULAR response, and by correlating the 12-joint US count, the reduced US assessment count and the disease activity index (DAS-28) using Pearson's correlation coefficient. Sensitivity to change of the US variables was tested by comparing the mean change in reduced US assessment from baseline to 3 months; in addition, we evaluated the correlation between the changes in the reduced US assessment and

the variations in DAS-28 from baseline to 3 months. The feasibility of the reduced US assessment was estimated by comparing the time spent on the 12-joint US examination and the reduced US assessment by the independent-samples *t*-test. *P*-values <0.05 were considered statistically significant.

Results

Forty-five Caucasian patients (8 males and 37 females) were included in the study. The main demographic, clinical and laboratory parameters of the enrolled patients at baseline and after 3 months of etanercept treatment are reported in Table 1.

Clinical and laboratory features

The disease activity was moderate to severe in all patients at baseline [mean DAS-28 (s.d.) 4.5 (1.2)]. After 3 months of therapy with etanercept, a good to moderate response, according to EULAR criteria was registered in 14 of the 45 patients (31.1%; good response in 7 of the 45, and moderate in 7 of the 45). A significant decrease in clinical and laboratory parameters was found at 3 months follow-up. DAS-28 was reduced to 3.6 (1.3) ($P = 0.0016$), ESR decreased from 27.1 (20.3) to 23.5 (21.2) mm/h ($P < 0.001$) and CRP decreased from 10.5 (18.7) to 8.6 (10.6) mg/dl ($P < 0.001$).

12-joint US assessment

US-SE, US-SP and US-PD values assessed at 12 joints at baseline and 3 months are reported in Table 2. All the

TABLE 1 Clinical features of the 45 RA patients studied at baseline and after 3 months of etanercept therapy

Feature	Baseline	3 months	P-values
Age, mean (S.D.), months	638.5 (173.4)		
Disease duration, mean (S.D.), months	120.7 (98.7)		
Male/female	8/37		
Smoke, n (%)	7 (15.5)		
Concomitant MTX, n (%)	10 (22.2)		
Anti-TNF monotherapy, n (%)	10 (22.2)		
Prednisone (PDN) equivalent, mean (S.D.), mg/day	7.1 (6.2)		
ESR, mean (S.D.), mm/h	27.1 (20.3)	23.5 (21.2)	<0.001
CRP, mean (S.D.), mg/dl	10.5 (18.7)	8.6 (10.6)	<0.001
RF, mean (S.D.), U/l	240.1 (264.6) (range 9.4–600)	192.1 (240.0)	NS
ACPA, mean (S.D.), U/l	70.7 (46) (range 9–121)	64.8 (43.8)	NS
DAS-28, mean (S.D.)	4.5 (1.2)	3.6 (1.3)	0.0016
EULAR response,[a] n (%)	—	14 (31.1)	

[a]Good or moderate response.

scores were significantly reduced after 3 months of anti-TNF therapy ($P < 0.001$ for all parameters).

Six-joint US assessment

We obtained a reduced 6-joint US model after adoption of different models, as shown in Table 3. At first, we screened the single joints from the 12-joint assessment and found that US-SE, US-SP and US-PD were mostly present at the wrist, knee and second MCP joints. Among these joints, the one with the highest prevalence of abnormalities was the wrist. Thus we assumed this joint to be the reference joint and then each of the remaining joints was added to the count. We found that adding the data from the knee assessment allowed us to obtain a very high sensitivity, although lower than the requested 97% for each the three parameters. Thus we added the US data on each of the remaining joints to the count and found that the highest sensitivity was reached when the US data from the second MCP joint were added. We thus obtained a 6-joint US assessment that was able to detect 97.7% of patients with 12-joint US-SE, 100% of patients with 12-joint US-SP and 100% of those with 12-joint PD.

Correlation between the 12-joint and the 6-joint US assessment

The 12-joint US count decreased from 20.87 (16.86) at baseline to 13.67 (10.62) ($P < 0.001$). The 12-joint US variables significantly correlated with the 6-joint US variables both at baseline and at the 3-month follow-up. All the variables (12-joint US-SE, US-SP and US-PD) showed a positive correlation with the respective 6-joint variables ($P < 0.0001$ for all comparisons) at baseline. The 12-joint US count at baseline showed a positive correlation with the 6-joint US total count ($P = 1.07 \times 10^{-22}$, $R = 0.943$, Fig. 2). Also, the 12-joint US count at 3 months positively

TABLE 2 Baseline and 3 months values of US-SE, US-SP and US-PD for the 12- and the 6-joint US assessment

	Baseline, mean (S.D.)	After 3 months, mean (S.D.)	P-values
US-SE 12 joints range (0–36)	8.71 (6.23)	5.84 (4.05)	<0.001
US-SP 12 joints (0–36)	8.27 (6.14)	5.47 (4.06)	<0.001
US-PD 12 joints (0–36)	4.36 (5.44)	2.67 (3.38)	<0.001
US 12 joints count (0–108)	20.87 (16.86)	13.67 (10.62)	<0.001
US-SE 6 joints (0–18)	5.71 (3.92)	4.00 (2.59)	<0.001
US-SP 6 joints (0–18)	5.80 (4.21)	4.07 (2.82)	<0.001
US-PD 6 joints (0–18)	3.62 (4.35)	2.16 (2.69)	<0.001
US 6 joints count (0–54)	14.80 (12.1)	10.00 (7.5)	0.0037

correlated with the 6-joint US count at the same time point ($P = 2.7 \times 10^{-19}$, $R = 0.918$).
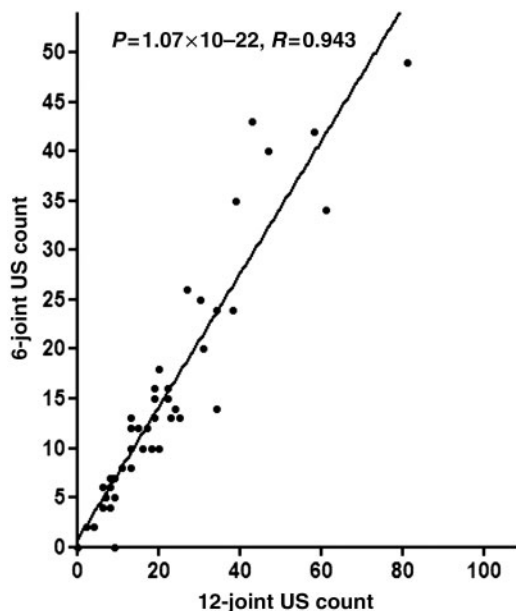
Correlation between the 12-joint US assessment, the 6-joint US assessment and the clinical features

The 12-joint US count at baseline positively correlated with CRP at baseline ($P = 0.005$, $R = 0.461$). Changes in the 12-joint US count showed a positive correlation with changes in DAS-28 ($P = 0.047$, $R = 0.338$). When considering the changes in the US-SE, US-SP and US-PD in 12 joints, only US-SP correlated with changes in DAS-28 ($P < 0.05$, $R = 0.339$).

The 6-joint US count decreased from 14.80 (12.1) at baseline to 10.00 (7.5) ($P = 0.0037$, Table 2) and showed

TABLE 3 Sensitivity for detecting US-SE, US-SP and US-PD of the combination of joints selected for the reduced US assessment

	US 12 joints		
	US-SE (*n* = 45)	US-SP (*n* = 45)	US-PD (*n* = 45)
Wrist, *n* (%)	43/45 (95.5)	42/45 (93.3)	42/45 (93.3)
Second MCP, *n* (%)	30/45 (66.7)	27/45 (60)	20/45 (44.4)
Third MCP, *n* (%)	18/45 (17.8)	16/45 (35.5)	19/45 (42.2)
Elbow, *n* (%)	15/45 (33.3)	9/45 (20)	13/45 (28.9)
Knee, *n* (%)	33/45 (73.3)	32/45 (71.1)	25/45 (55.5)
Ankle, *n* (%)	22/45 (48.9)	20/45 (44.4)	22/45 (48.9)
Wrist + knee, *n* (%)	44/45 (97.7)	43/45 (95.5)	44/45 (97.7)
Wrist + knee + second MCP, *n* (%)	44/45 (97.7)	45/45 (100)	45/45 (100)

FIG. 2 Correlation between the 12-joint US count and the 6-joint US count at baseline ($P = 1.07 \times 10^{-22}$, $R = 0.943$).

positive correlation with DAS-28 at baseline ($P = 0.002$, $R = 0.535$, Fig. 3), as well as with CRP at the same time point ($P = 0.001$, $R = 0.519$).

Feasibility

There was a significant difference between the mean time spent on the 12-joint US examination [23.4 (2.6) min] and the mean time spent on the 6-joint US examination [14.1 (3) min, $P < 0.001$].

FIG. 3 Correlation between the 6-joint US count and the DAS-28 at baseline ($P = 0.002$, $R = 0.535$).
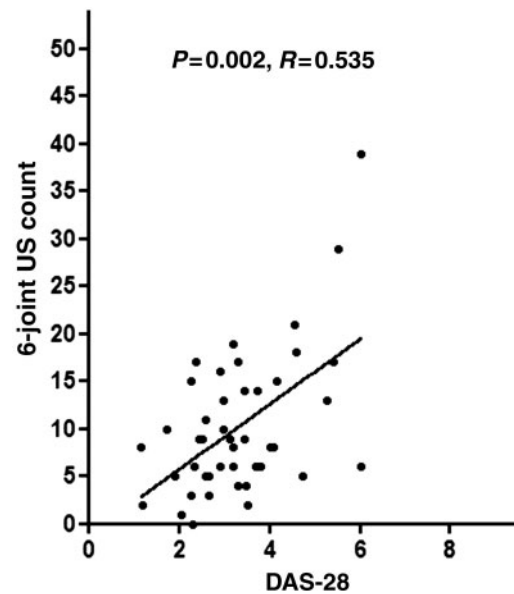
Discussion

Our results provide evidence that a 6-joint US assessment of joint inflammation may be a valid, sensitive-to-change and feasible method for monitoring the response to etanercept therapy in patients affected with RA. In the last decade, several studies have demonstrated that US assessment can be useful in the management of RA and in monitoring the course of the disease at all stages [1]. The application of US is helpful in the evaluation of RA patients and represents a complementary tool to classic methods used to detect inflammatory status, such as clinical evaluation and radiography, particularly when MCP, PIP and MTP joints are considered [13–15].

Short-term follow-up studies have widely demonstrated the correlation between disease activity and degree of inflammation of synovial tissue, as shown by grey-scale and US-PD evaluation [16–18]. US can be used in the evaluation of response to biological drugs, such as TNF antagonists. It was demonstrated that the effective administration of TNF blocking agents results in a significant decrease in vascularization assessed with US. For instance, Naredo *et al.* [7] found a significant parallel improvement in DAS-28 and US parameters in RA patients undergoing therapy with a TNF blocking agent. The authors suggested that US evaluation with application of PD could be a valid method for monitoring response to anti-TNF therapy in RA patients.

However, there is no evidence regarding which joints and synovial recesses are appropriate for studying and monitoring RA patients by means of US. Nonetheless, a remarkable variability in the synovial sites assessed can be observed in published studies. An appropriate choice

of the joints to be assessed is fundamental. It is necessary to identify those joints that can be considered sensitive to represent the patient global inflammatory status. However, a comprensive evaluation of all accessible joints is time consuming. On the other hand, assessing only the swollen or tender joints can lead to a lack of information due to the subjectivity of the examination and the possibility of the presence of subclinical inflammation. For these reasons, US assessment for clinical practice must result from a compromise between an extensive and an informative report. Nonetheless, standardization of the US technique will allow comparison between groups of RA patients and, analogous to DAS-28, may result in important clinical trials. Several authors have faced the problem with different approaches and results.

Two studies adopted the joints present in several disease activity indices for US assessement. Hammer et al. [9] suggested a 78-joint US assessment. They evaluated 20 RA patients starting adalimumab as the first biological agent and found an association between US scores and clinical and laboratory parameters [9]. US detected higher numbers of inflamed joints when compared with clinical assessment. However, the average time spent for each US examination of all 78 joints was about 70 min, thus it is not compatible with daily clinical practice [9].

Dougados et al. [8] conducted a multicentre study on RA patients requiring TNF blocker. US evaluation was performed in the joints included in the DAS-28, plus the MTP joints. Three different US scoring systems using a range of joint counts [20] were analysed using either a binary (yes/no) or a semi-quantitative score (0–3). The MTP joints evaluation was added because of their frequent involvement in the early phases of the disease. The authors found that US evaluation of synovitis could represent an outcome measure at least as good as, and possibly more accurate than, physical examination. The time spent by investigators in collecting the US data ranged from 10 to 25 min, depending on the number of joints evaluated, thus it was satisfactory for patient acceptance [8]. However, this short time (relative to the number of joints assessed) was the main concern with this study. Indeed, most of the studies present in the literature, and our study as well, reported a longer time needed for US examination.

Several authors arbitrarily selected the joints to be evaluated with US. This selection was generally performed on the basis of the most frequent involvement in RA found in the clinical data. In our experience, we have previously evaluated the response to treatment with etanercept and adalimumab by choosing the second and fifth MCP, the third PIP, the wrist and the knee joints [3, 4]. The US score applied in those studies showed a significant decrease after a long-term follow-up (24 months in patients treated with adalimumab and 12 months in those treated with etanercept) and a significant correlation with disease activity (DAS-28) [3, 4]. In 2009, Backhaus et al. [5] used a novel US score using the wrist, the second and third MCP, the second and third PIP, and the second and fifth MTP joints of the clinically dominant side of RA patients, the so-called US7 score. In this study, a significant

correlation between changes in the US parameters for synovitis and the DAS-28 was registered. This US7 score may represent a valuable tool for US examination of inflamed joint activity in rheumatological diseases, especially in RA. Concentration on a small number of active joint regions reduced examination time ($\sim$10–20 min), suggesting the possibility of integrating the US7 score in daily rheumatological practice [5]. However, possible limitations could have been that only one side of the body was assessed, thus possibly excluding a number of active patients with RA.

Naredo and colleagues [6] performed one of the more interesting studies. The authors chose the joints to be evaluated by a process of data reduction starting from the 44 joints included in the DAS-44 index. Afterwards, they evaluated the frequency of involvement of synovial sites by both synovitis and PD signal and then obtained a simplified assessment evaluating 12 joints. The results showed that this 12-joint US score correlated with the non-simplified 44-joint US score. This simplified US assessment thus showed content and construct validity. Considering the feasibility, there was a significant difference between the mean time spent on the 44-joint US examination (83.6 min) and the mean time spent on the 12-joint US examination (22 min) [6].

Starting from this study, we thought that 22 min could still represent a long time expense that cannot be afforded in daily clinical practice as well as in clinical trials. Thus, applying the same process of data reduction used by Naredo et al. [6], we aimed to investigate the validity, responsiveness and feasibility of a 6-joint US score in assessing joint inflammation as compared with the already described 12-joint US evaluation.

Our 6-joint US assessment detected 97.7% of patients with 12-joint US-SE, 100% of patients with 12-joint US-SP and 100% of 12-joint PD. These percentages indicate that evaluation of the six selected joints was very sensitive. Our 6-joint US score showed a significant correlation with DAS-28, thus showing it to be at least as sensitive as the clinical data. Importantly, the score was sensitive to change after 3 months of follow-up of anti-TNF therapy, suggesting that the score can be used in the short-term monitoring of the response to anti-TNF treatment.

We evaluated patients treated with anti-TNF due to the intrinsic features of these drugs in influencing joint inflammation and the specific features of patients naïve to anti-TNF who have high disease activity indexes and amelioration of symptoms is expected in a significant percentage of patients and a relatively short period of time. Nonetheless, US can also recognize subclinical synovitis in RA patients treated with DMARDS, as shown by Brown et al. [19]. When considering the usage of US scoring systems in the follow-up of patients with RA, Backhaus et al. [5] used the German US7 in patients treated with DMARDs, anti-TNF or a combination of the two. More recently, Saleem et al. [20] used US in the assessment of remission in RA. They evaluated patients treated with either DMARDs or a combination of TNF blockers and MTX, showing that US is superior to clinical

evaluation. Indeed, the clinical criteria may underestimate the detection of low but clinically relevant levels of inflammation.

Finally, Peluso *et al*. [21] showed that US-PD can be used in the assessment of remission in patients with early RA as well as in those with long-standing RA, independently of the use of DMARDs (specifically MTX) and/or anti-TNF.

In conclusion, the application of US assessment in clinical practice should include a comprehensive evaluation of patient inflammatory status and feasibility in order to reduce the time needed for the US examination. The importance of testing the feasibility of US was included in the research agenda of the OMERACT US task force in 2009, being a fundamental aspect of the OMERACT filter [22]. In our study, we achieved a significantly shorter time of execution, suggesting that this 6-joint model could be more feasible than others previously described.

Rheumatology key messages

- A comprehensive US evaluation of inflammation in RA is time consuming.
- The 6-joint US assessment was able to detect synovitis in ~100% of patients evaluated with the 12-joint assessment.
- US assessment of wrists, second MCP and knees is valid, sensitive to change and feasible in the evaluation of RA synovitis.

Acknowledgements

References

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

2 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.

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

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

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

6 Naredo E, Rodríguez M, Campos C *et al*. Ultrasound Group of the Spanish Society of Rheumatology. Validity, reproducibility, and responsiveness of a twelve-joint simplified power doppler ultrasonographic assessment of joint inflammation in rheumatoid arthritis. Arthritis Rheum 2008;59:515–22.

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

8 Dougados M, Jousse-Joulin S, Mistretta F *et al*. Evaluation of several ultrasonography scoring systems for synovitis and comparison to clinical examination: results from a prospective multicentre study of rheumatoid arthritis. Ann Rheum Dis 2010;69:828–33.

9 Hammer HB, Sveinsson M, Kongtorp AK *et al*. A 78-joints ultrasonographic assessment is associated with clinical assessments and is highly responsive to improvement in a longitudinal study of patients with rheumatoid arthritis starting adalimumab treatment. Ann Rheum Dis 2010;69: 1349–51.

10 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.

11 van Gestel AM, Prevoo ML, van 't Hof MA *et al*. Development and validation of the European League Against Rheumatism response criteria for rheumatoid arthritis. Comparison with the preliminary American College of Rheumatology and the World Health Organization/International League Against Rheumatism Criteria. Arthritis Rheum 1996;39:34–40.

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

13 Backhaus M, Kamradt T, Sandrock D *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.

14 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.

15 Szkudlarek M, Klarlund M, Narvestad E *et al*. Ultrasonography of the metacarpophalangeal and proximal interphalangeal joints in rheumatoid arthritis: a comparison with magnetic resonance imaging, conventional radiography and clinical examination. Arthritis Res Ther 2006;8:R52.

16 Backhaus M, Burmester GR, Sandrock D *et al*. Prospective two year follow up study comparing novel and

conventional imaging procedures in patients with arthritic finger joints. Ann Rheum Dis 2002;61:895–904.

17 Scheel AK, Hermann KG, Ohrndorf S et al. Prospective 7 year follow up imaging study comparing radiography, ultrasonography, and magnetic resonance imaging in rheumatoid arthritis finger joints. Ann Rheum Dis 2006;65: 595–600.

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

19 Brown AK, Quinn MA, Karim Z et al. Presence of significant synovitis in rheumatoid arthritis patients with disease-modifying antirheumatic drug-induced clinical remission: evidence from an imaging study may explain structural progression. Arthritis Rheum 2006;54: 3761–73.

20 Saleem B, Brown AK, Keen H et al. Should imaging be a component of rheumatoid arthritis remission criteria? A comparison between traditional and modified composite remission scores and imaging assessments. Ann Rheum Dis 2011;70:792–8.

21 Peluso G, Michelutti A, Bosello S et al. Clinical and ultrasonographic remission determines different chances of relapse in early and long standing rheumatoid arthritis. Ann Rheum Dis 2011;70:172–5.

22 D'Agostino MA, Conaghan PG, Naredo E et al. The OMERACT ultrasound task force—advances and priorities. J Rheumatol 2009;36:1829–32.