REVISTA BRASILEIRA DE REUMATOLOGIA

Original article

A new musculoskeletal ultrasound scoring system (US10) of the hands and wrist joints for evaluation of early rheumatoid arthritis patients

Karine R. Luz[a], Marcelo M. Pinheiro[a], Giovanna S. Petterle[a], Marla F. dos Santos[a], Artur R.C. Fernandes[b], Jamil Natour[a],*, Rita N.V. Furtado[a]

[a] *Divisão de Reumatologia, Universidade Federal de São Paulo, São Paulo, SP, Brazil*
[b] *Departamento de Radiologia, Universidade Federal de São Paulo, São Paulo, SP, Brazil*

ABSTRACT

Objective: To propose a novel ultrasound scoring system for hand and wrist joints (US10) for evaluation of patients with early rheumatoid arthritis (RA) and to correlate the US10 with clinical, laboratory and functional variables.

Methods: Forty-eight early RA patients underwent clinical and laboratory evaluations as well as blinded ultrasound (US) examinations at baseline, three, six and 12 months. The proposed US10 system involved the assessment of the wrist, second and third metacarpophalangeal and proximal interphalangeal joints. The score consisted of inflammation parameters (synovial proliferation [SP], power Doppler [PD] and tenosynovitis [TN]) and joint damage parameters (bone erosion [BE] and cartilage damage [CD]). SP, PD, BE and CD were scored qualitatively (0–1) and semi-quantitatively (grades 0–3). Tenosynovitis was scored as presence/absence. The evaluation also involved the 28-Joint Disease Activity Score (DAS28), Health Assessment Questionnaire (HAQ) and C-reactive protein level (CRP).

Results: Mean duration of symptoms was 7.58 ± 3.59 months. Significant correlations ($p < 0.05$) were found between inflammation parameters and CRP at baseline and between the changes in these variables throughout the study. Significant correlations ($p < 0.05$) were found between DAS28 score and both PD and TN at baseline and between the changes in DAS28 score and both SP and TN throughout the follow up. Moreover, significant correlations were found between the changes in inflammation parameter scores and HAQ score throughout the follow up.

Conclusion: The proposed US10 scoring system proved to be a useful tool for monitoring inflammation and joint damage in early RA patients, demonstrating significant correlations with longitudinal changes in disease activity and functional status.

Novo sistema de escore ultrassonográfico (US10) musculoesquelético das articulações das mãos e punho para avaliação de pacientes com artrite reumatoide em fase inicial

R E S U M O

Palavras-chave:
Artrite reumatoide em fase
inicial
Mão
Ultrassom
Escore

Objetivo: Propor um novo sistema de escore ultrassonográfico das articulações da mão e punho (US10) para a avaliação de pacientes com artrite reumatoide (AR) e correlacionar o US10 com variáveis clínicas, laboratoriais e funcionais.

Métodos: Foram submetidos 48 pacientes com AR em fase inicial a avaliações clínicas e laboratoriais, bem como a exames cegos de ultrassom (US) no início do estudo e com 3, 6 e 12 meses. O sistema US10 proposto envolveu a avaliação do punho e das articulações metacarpofalângicas e interfalângicas proximais do segundo e terceiro dígitos. O escore consistiu em parâmetros inflamatórios (proliferação sinovial [PS], *Power Doppler* [PD] e tenossinovite [TN]) e parâmetros de danos articulares (erosão óssea [EO] e danos na cartilagem [DC]). PS, PD, EO e DC foram pontuados qualitativamente (0 a 1) e semiquantitativamente (graus 0 a 3). A tenossinovite foi pontuada como presença/ausência. A avaliação envolveu também o escore 28-*Joint Disease Activity* (DAS28), o *Health Assessment Questionnaire* (HAQ) e o nível de proteína C-reativa (PCR).

Resultados: A duração média dos sintomas foi de 7,58 ± 3,59 meses. Foram encontradas correlações estatisticamente significativas (p < 0,05) entre os parâmetros de inflamação e a PCR no início do estudo e entre as mudanças nessas variáveis ao longo do estudo. Foram encontradas também correlações significativas (p < 0,05) entre o escore DAS28 e a PD e TN no início do estudo e entre as mudanças no escore DAS28 e PS e TN em todo o seguimento. Além disso, foram encontradas correlações significativas entre as mudanças no escore dos parâmetros de inflamação e no escore HAQ ao longo do seguimento.

Conclusão: O sistema de escore US10 proposto provou ser uma ferramenta útil para monitorar a inflamação e o dano articular em pacientes com AR em fase inicial, demonstra correlações significativas com as alterações longitudinais na atividade da doença e no estado funcional.

Introduction

In recent years, musculoskeletal ultrasound (US) has been employed for monitoring patients with rheumatoid arthritis (RA) with regard to both disease activity and joint damage.[1–6] Disease activity is evaluated using gray-scale and Power Doppler (PD) synovitis and tenosynovitis.[4,5,7,8] However, there is no consensus on the US evaluation of synovitis (synovial proliferation and PD) in the joints, as scoring is performed with either a binary variable (presence/absence of synovitis) or a semi-quantitative variable (usually using a scale from 0 [absence of synovitis] to 3 [severe synovitis]).[9–14] US can also be used to assess joint damage and has proven to be sensitive to the detection of bone erosion, which is important to the diagnosis and evaluation of RA, especially in small joints of the hands and feet.[6,10] Indeed, US is better able to detect erosions in the metacarpophalangeal joints (MCPs) in patients with early RA than radiography.[6] Like synovitis, bone erosion can be evaluated using a semi-quantitative scoring system, with excellent interobserver agreement (ICC = 0.78 and $k = 0.68$).[10]

A number of studies have evaluated different simplified scores and report satisfactory correlations with clinical disease activity indices.[15–18] The joints of the hands (proximal interphalangeal [PIPs] and [MCPs]) and wrist are the most affected in early RA and are systematically assessed through physical examinations and for treatment decision.[19–22] However, there is insufficient evidence on which joints and synovial recesses are better for the US detection of synovitis. The prospective investigation of a simplified US scoring system for these

joints in disease-modifying antirheumatic drugs (DMARD)-naïve patients with early RA may be useful to clinical follow-up and for monitoring therapy.

Thus, the aim of the present study was to investigate a novel US scoring system for the evaluation of inflammation and joint damage in the hand and wrists joints, denominated the US10, and correlate US inflammation and joint damage parameters with clinical, laboratory, functional and radiographic findings in patients with early RA over a 12 months of follow up.

Patients and methods

Patients

A prospective cohort study was conducted involving 48 consecutive patients with early RA, with symptom duration for more than six weeks and less than one year since onset. The individuals were recruited from Rheumatology Outpatient Clinics of the Universidade Federal de São Paulo. This study received approval from the Research Ethics Committee.

The inclusion criteria were the following: diagnosis of RA based on the 1987 and 2010 classification criteria of the American College of Rheumatology (ACR) or a cutoff point of ≥8.0 on predictive validation Leiden model score.[23,24] The Leiden score consists of nine clinical and laboratory variables including age, sex, location of joint symptoms, morning stiffness, tender and swollen joint count, CRP levels and presence of RF and anti-CCP antibody. Each item has a value, and the variation of 0–14 score. A patient is classified as early RA if presents a score ≥8.[25]

The exclusion criteria were the following: previous treatment with DMARDs, use of oral glucocorticoid >10 mg/d in the previous three weeks or a parenteral glucocorticoid in the previous four weeks, serum aspartate aminotransferase or alanine aminotransferase level >3 times the upper limit of normal, bone marrow hypoplasia, overlap with any other collagen disease, suspicion of lymphoproliferative disease, positive serology for hepatitis B or C and pregnancy.

All patients underwent clinical, laboratory and US evaluations at baseline, three, six and 12 months. A tightly controlled therapeutic protocol was used for all patients by a single rheumatologist who was blinded to the US evaluation. The patients began with methotrexate (MTX) 15 mg/week, which was increased to 25 mg/week in the first three months. Subsequent steps for patients with an insufficient response (DAS28 score > 3.2 and evaluator-based global assessment of disease activity score > 4.0 [0–10 cm]) were leflunomide with MTX 15 mg/week, leflunomide and MTX 25 mg/week, adalimumab and MTX 15 mg/week and, finally, MTX 15 mg/week with a second biologic agent.

Clinical assessment

The patients were clinically evaluated during each visit. Twenty-eight joints (bilateral PIPs, MCPs, wrists, elbows, shoulders and knees) were clinically assessed for swelling and tenderness. The following instruments were also employed: Evaluator based global assessment of disease activity (0–10 cm); patient based global assessment of disease activity (0–10 cm); Brazilian version of the functional subscale of the Stanford Health Assessment Questionnaire (HAQ)[26]; and the 28-Joint Disease Activity Score (DAS28).

Laboratory evaluation

The dosage of C-reactive protein (CRP) level (mg/dliter) and erythrocyte sedimentation rate (ESR) (mm/hour) were determined at each visit. IgM rheumatoid factor and anti-cyclic citrullinated peptide (anti-CCP) antibodies were assessed at baseline.

Ultrasound assessment

The US examination was performed by a trained rheumatologist with eight years of experience in US who was blinded to all other study findings. US examinations were performed using a MyLab60 (Esaote, Biomedica – Genoa, Italy), equipped with a broadband linear probe with frequency ranging from 6 to 18 MHz.

Systematic multiplanar gray-scale US (GSUS) and PD examinations were performed on 10 joints (wrist, MCP2, MCP3, PIP2 and PIP3 in both hands) in a standardized manner based on the guidelines of the European League Against Rheumatism.[27] All joint regions were assessed using inflammation and joint damage parameters.

A. Inflammation parameters (Fig. 1):
1. Synovial proliferation (SP):
   Synovitis by GSUS – synovial proliferation, defined as an abnormal hypoechoic intra-articular tissue that is non-displaceable and poorly compressible visualized in longitudinal and transversal planes[28]; The following 20 synovial sites in 10 joints were included: wrist (dorsal carpal and ulnar carpal recess); second and third MCP (dorsal side, palmar side); second and third PIP (palmar side).
   Synovial proliferation was analyzed in all sites, as follows:
   - Semi-quantitative evaluation (SPSQ) – Grade 0 (absence), Grade 1 (small hypoechoic/anechoic line beneath joint capsule), Grade 2 (joint capsule elevated parallel to joint area) and Grade 3 (strong distension of joint capsule)[10,14];
   - Qualitative evaluation (SPQ) – binary evaluation – 0 (absent) or 1 (present, if Grade 2 or 3 semi-quantitative scores).
2. Synovial blood flow:
   Synovial blood flow was evaluated by PD in each of the intra-articular synovial sites. PD settings were standardized with a pulse repetition frequency of 750 Hz and a color-mode frequency of 12 MHz. Wall filters were set at the lowest value, while color gain was increased to the highest value, not generating PD signals under the bone cortex.
   PD intra-articular signals were evaluated and graded on a qualitative and semi-quantitative score:
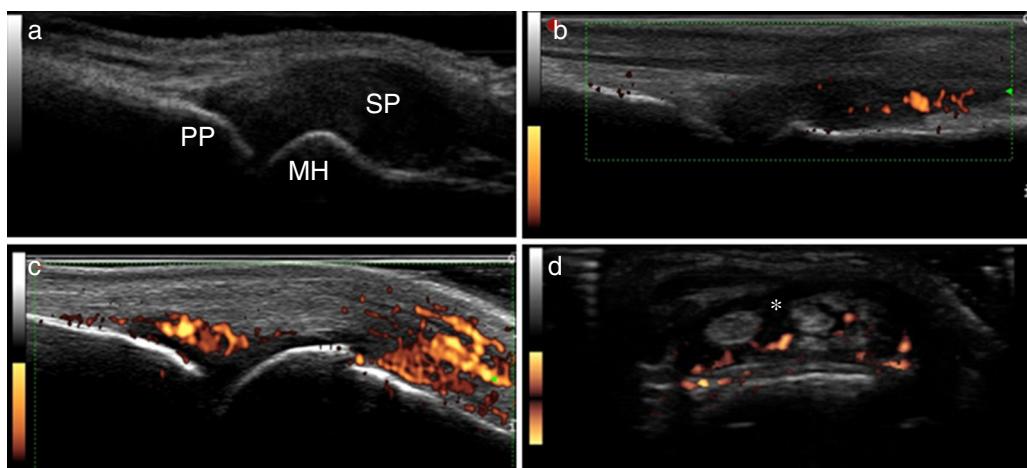
Fig. 1 – Inflammation parameters of US10. (A) Synovial proliferation grade 3; (B) synovitis by power Doppler grade 2; (C) synovitis by power Doppler grade 3; (D) tenosynovitis by power Doppler. PP, proximal phalanx; MH, metacarpal head; SP, synovial proliferation; *, tenosynovitis.

- Semi-quantitative (PDSQ) – Grade 0 (no flow in synovium), Grade 1 (single vessel signals); Grade 2 (confluent vessel signals in less than half the area of the synovium); Grade 3 (vessel signals in more than half the area of the synovium)[10];
- Qualitative (PDQ) – binary evaluation – 0 (absent) or 1 (present, if Grade 1 semi-quantitative score).

3. Tenosynovitis

Defined as a hypoechoic or anechoic thickened tissue with or without fluid within the tendon sheath[28];

The following tendons were evaluated: extensor digitorum communis; extensor carpi ulnaris; flexor digitorum communis, second and third flexor tendons.

Tenosynovitis was evaluated and graded on a GSUS and PDUS qualitative score:

- Qualitative (TNGSQ) – binary evaluation – 0 (absent) or 1 (present);

- Qualitative (TNPDQ) – binary evaluation – 0 (absent) or 1 (present).

B. Joint damage parameters (Fig. 2):

1. Bone erosion

Erosion was defined as an interruption of the bone surface on two perpendicular planes.[28]

The location of each erosion was recorded based on the bone involved, as follows:

- Dorsal quadrant of second and third metacarpal head
- Lateral quadrant of second metacarpal head
- Dorsal quadrant of second and third phalanx
- Ulnar styloid process

Bone erosions were graded on a qualitative and semi-quantitative score:

- Semi-quantitative (BESQ) – Grade 0 (regular bone surface), Grade 1 (irregular bone surface without formation of defect seen on two planes), Grade 2
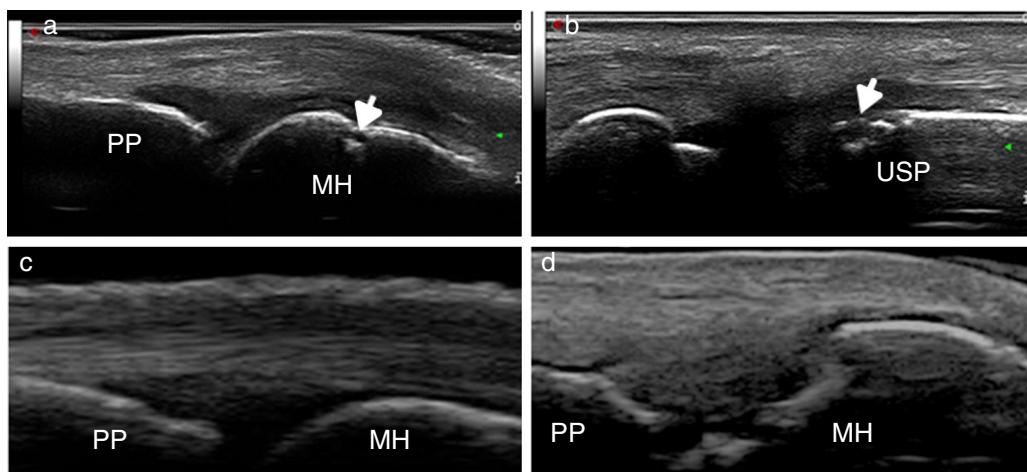

Fig. 2 – Joint damage parameters of US10. (A) Bone erosion grade 2; (B) bone erosion grade 3 (arrow); (C) cartilage damage grade 2; (D) cartilage damage grade 3. PP, proximal phalanx; MH, metacarpal head; USP, ulnar styloid process.

Table 1 – US10: Score range for each ultrasound hand parameter.

Ultrasound parameters – US10	Score range
10 Joints (second and third MCPs and PIPs and wrists)	
Inflammation parameters	
SPQ	0–16
SPSQ	0–48
PDQ	0–16
PDSQ	0–48
TNGSQ	0–10
TNPDQ	0–10
Joint damage parameters	
BEQ	0–12
BESQ	0–36
CDQ	0–4
CDSQ	0–16

US10, ultrasound score of hand joints and wrists; MCP, metacarpophalangeal joint; PIP, proximal interphalangeal jont; SPQ, qualitative synovial proliferation; SPSQ, semi-quantitative synovial proliferation; PDQ, qualitative power Doppler; PDSQ, semi-quantitative power Doppler; TNGSQ, qualitative gray-scale tenosynovitis; TNPDQ, qualitative power Doppler tenosynovitis; BEQ, qualitative bone erosion; BESQ, semi-quantitative bone erosion; CDQ, qualitative cartilage damage; CDSQ, semi-quantitative cartilage damage.

(formation of defect on bone surface seen on two planes) and Grade 3 (bone defect causing extensive bone destruction)[10];

- Qualitative (BEQ) – binary evaluation – 0 (absent) or 1 (present, if Grade 2 or 3 semi-quantitative scores).

2. Cartilage damage

US examinations were focused on the assessment of the hyaline cartilage in the dorsal view of the second and third metacarpal heads. Normal features of the hyaline cartilage at the metacarpal head included two hyperechoic sharp, regular and continuous margins delimiting a homogenous anechoic band.[29,30]

Cartilage damage was evaluated using the following semi-quantitative and qualitative scoring system:

- Semi-quantitative (CDSQ) – Grade 0 (normal hyaline cartilage); Grade 1 (loss of sharpness of superficial margin of hyaline cartilage); Grade 2 (partial thickness defect of cartilage layer); Grade 3 (full thickness defect of cartilage layer with normal subchondral bone profile); Grade 4 (complete loss of cartilage layer and subchondral bone involvement).[31,32]
- Qualitative (CDQ) – binary evaluation – 0 (absent) or 1 (present, if Grade 2 or 3 semi-quantitative scores).

The US10 scoring system analyzed 10 sonographic parameters, and each parameter was subdivided into qualitative and semi-quantitative grades. The parameters were analyzed separately, as indicated in Table 1. These 10 parameters and 10-joints constituted the US10 system, with the sum of all scores (Table 1).

Interobserver reliability

Interobserver reliability between two US operators was evaluated on recorded images from 20 randomly chosen patients for the inflammation and joint damage parameters of the joints included in the US10 score. The captured images of each US10 item of 20 patients with a total of 200 images were evaluated. The evaluation of the ultrasound images was performed by a rheumatologist, with five years of experience in musculoskeletal ultrasound.

Radiographic evaluation

Posteroanterior X-rays of the patients' hands were recorded at baseline and after 12 months. The images were analyzed at the end of the study by a radiologist who was blinded to the other evaluations. The evaluation of hand X-rays was done by a radiologist with 30 years of experience who was blind to the clinical features, ultrasound results and identification of the patient. In the radiographic evaluation it was considered only bone erosion. The presence of erosions was evaluated using the scores proposed by van der Heijde et al.[33] The correlation between narrowing of joint space on X-rays and joint cartilage damage on ultrasound was not analyzed.

Statistical analysis

Statistical analysis was performed with the aid of the SPSS program, version 17.0 (SPSS, Chicago, IL, USA). Data are expressed as mean ± standard deviation. ANOVA was performed to compare numerical variables repeated over time. Correlations between changes in the different examination modalities (clinical, laboratory and US) throughout the follow-up period were evaluated by two-tailed Spearman's correlation coefficients. Inter- and intra-reader agreement was calculated using kappa coefficients between readers. The kappa coefficients were divided as follows: $< 0.0 =$ poor, $0–0.20 =$ slight, $0.21–0.40 =$ fair, $0.4–0.60 =$ moderate, $0.6–0.80 =$ substantial and $0.81–1.0 =$ almost perfect agreement.[34] Comparisons between X-ray and US regarding bone erosion were made using the chi-square test, adjusted by the McNemar method. The statistical significance level was set to 5% ($p < 0.05$). The data were analyzed using the intention-to-treat principle.

Results

Patient characteristics

Forty-eight patients (100% women) with a mean age of 47.7 ± 11.6 years (range: 22–65) and a mean disease duration of 7.5 ± 3.5 months (range: 2–12) were examined on four occasions (baseline, 3, 6 and 12 months). Rheumatoid factor and anti-CCP were positive in 20 patients (41.7%) and 21 patients (43.8%), respectively. At baseline, 30 patients (62.5%) met the ACR criteria and 35 (72.9%) had ≥8.0 on the predictive validation Leiden model score (Table 2). After 12 months of follow up, seven patients (14.5%) continuing using MTX, 41 (85.4%) received MTX + leflunomide, 25 (52%) had switched to biological therapy and five (10.5%) had an indication to receive a second biological agent at the final evaluation (12 months).

Table 2 – Baseline demographic parameters, disease related variables in 48 RA.	
	RA patients (n = 48)
Age, years (mean ± SD)	47.7 ± 11.6
Gender (women/men)	48/0
Disease duration, months (mean ± SD)	7.5 ± 3.5
Rheumatoid factor (positive) (%)	20 (41.7%)
Anti-CCP (positive) (%)	21 (43.8%)
1987 ACR criteria (%)	30 (62.5%)
2010 ACR criteria (%)	48 (100%)
Leiden Model Score (%)	35 (72.9%)

RA, rheumatoid arthritis; ACR, American College of Rheumatology; SD, standard deviation.

Clinical, laboratory and US10 parameters

Table 3 displays the US, clinical and laboratory data. A significant reduction in mean DAS28 score was found over 12 months (6.5–3.9; $p < 0.05$). The mean HAQ score also decreased significantly (1.4–0.7; $p < 0.05$). No significant differences were found in mean CRP (14.0–6.9 mg/dliter; $p > 0.05$) or ESR (30.6–24.3 mm/h; $p > 0.05$) levels over 12 months. All US10 inflammation parameters decreased significantly over the course of the year. Mean SPQ and SPQSQ scores were respectively 12.9 and 29.1 at baseline and decreased significantly after 12 months (7.0 and 14.2, respectively) ($p < 0.05$). Mean baseline PDQ and PDSQ scores were 6.7 and 14.2, respectively, and decreased significantly and after 12 months (0.7 and 1.2, respectively) ($p < 0.05$). Mean TNGSQ and TNPDQ scores reduced respectively from 2.9 and 2.3 at baseline to 0.6 and 0.3 after 12 months ($p < 0.05$). Mean qualitative and semi-quantitative bone erosion scores increased significantly over 12 months (4.7–6.0 and 9.7–12.7, respectively) ($p < 0.05$). A significant increase in the mean qualitative cartilage damage score occurred over 12 months (0.2–1.1; $p < 0.05$). No significant difference in the mean semi-qualitative cartilage damage score occurred over 12 months (1.2–2.5; $p = 0.14$).

Baseline and longitudinal correlation between US10 parameters

At baseline, all US10 inflammation parameter scores demonstrated significant correlations ($p < 0.05$) with each other (r variation between 0.33 and 0.95). A significant correlation ($p < 0.05$) was found between changes in the SPSQ score and changes in both the PDQ ($r = 0.43$) and PDSQ ($r = 0.42$) scores and TNGSQ ($r = 0.47$) and TNPDQ ($r = 0.40$) scores over 12 months.

Baseline and longitudinal correlations between US10, clinical and laboratory parameters

All US inflammation parameters (SPQ, SPSQ, PDQ, PDSQ, TNGSQ and TNPDQ) demonstrated significant correlations with CRP levels at baseline ($r = 0.31$, $r = 0.29$, $r = 0.38$, $r = 0.39$,

$r = 0.34$ and $r = 0.33$, respectively) ($p < 0.05$). All PD (PDQ and PDSQ) and tenosynovitis (TNGSQ and TNPDQ) scores were significantly correlated ($p < 0.05$) with DAS28 ($r = 0.34$, $r = 0.34$, $r = 0.35$, $r = 0.31$, respectively). The GS and PD tenosynovitis scores correlated significantly ($p < 0.05$) with ESR ($r = 0.50$, $r = 0.40$, respectively). No correlation was observed between the change of the qualitative score of synovial proliferation (SPQ), and DAS 28 over time. However it was observed correlation between the semi-quantitative score synovial proliferation (SPSQ) and DAS28 scores. Changes in PD and tenosynovitis scores were significantly correlated with changes in HAQ over 12 months ($r = 0.32$, $r = 0.34$, $r = 0.40$, $r = 0.32$, respectively with $p < 0.05$). Changes in PD scores also correlated with changes in CRP ($r = 0.34$, $r = 0.29$, respectively). Changes in qualitative synovial proliferation and tenosynovitis scores (TNGSQ and TNPDQ) correlated with changes in ESR ($r = 0.29$, $r = 0.48$, $r = 0.49$, respectively with $p < 0.05$). There was no correlation between sonographic parameters of joint damage (cartilage and erosion) and clinical and laboratory variables. The correlations between the sonographic parameters and clinical and laboratory variables are shown in Table 4.

Inter-observer reliability regarding US10 parameters

Mean kappa values for the qualitative and semi-quantitative synovial proliferation scores were 0.49 and 0.21, respectively ($p < 0.05$). Mean kappa values for PDQ, PDSQ, TNGSQ and TNPDQ scores on stored images were 0.49, 0.56, 0.55 and 0.32, respectively ($p < 0.05$). Mean kappa values for qualitative and semi-quantitative bone erosion scores were 0.42 and 0.47, respectively ($p < 0.05$). Substantial agreement between observers was found for the qualitative and semi-quantitative cartilage damage scores ($k = 0.79$ and 0.82) ($p < 0.05$).

Comparison of conventional radiography and ultrasound for bone erosion at baseline and after 12 months

In 480 sites analyzed at baseline, radiography found erosions in 54 sites (11.3%) and US found erosions in 150 sites (31.3%). US detected 2.81-fold more bone erosions than conventional radiography in the 10 sites examined ($p < 0.001$). After 12 months, US found 171 (39.80%) bone erosions and radiography found 44 (10.20%) erosions in the 430 sites examined. The US detected 3.88-fold more bone erosions than conventional radiography in the 10 sites examined after 48 weeks of evaluation. Erosions were evaluated by the two methods described in the same places and in the same patients.

Discussion

The present study evaluated a novel US scoring system involving ten joints (2nd and 3rd MCPs and PIPs and wrists), denominated the US10. The choice of these joints was based primarily on the frequency with which they are affected in the early years of RA.[19–21] Previous studies have also proposed scoring

Table 3 – Clinical, laboratory and US10 scores over time in relation to baseline evaluation.

Time points (weeks)	Clinical parameters			
	DAS28 (n = 48)	(Mean ± SD)	HAQ (n = 48)	(Mean ± SD)
T0	6.5	(±1.3)	1.4	(±0.7)
T12	4.5	(±1.7)	0.7	(±0.5)
T24	4.7	(±1.6)	0.8	(±0.7)
T48	3.9	(±1.4)	0.7	(±0.6)
p^a	0.000		0.000	
	Laboratory parameters			
	ESR (n = 48)	(Mean ± SD)	CRP (n = 48)	(Mean ± SD)
T0	30.6	(±3.5)	14.0	(±18.0)
T12	21.9	(±2.8)	8.4	(±11.4)
T24	24.0	(±2.6)	7.8	(±8.9)
T48	24.3	(±2.3)	6.9	(±12.1)
p^a	0.058		0.325	
	US10 Parameters			
	SPQ (0–16)	(Mean ± SD)	SPSQ (0–48)	(Mean ± SD)
T0	12.9	(±2.9)	29.1	(±8.4)
T12	4.2	(±3.4)	10.3	(±8.7)
T24	8.4	(±3.7)	17.3	(±6.8)
T48	7.0	(±3.9)	14.2	(±7.9)
p^a	0.000		0.000	
	PDQ (0–16)	(Mean ± SD)	PDSQ (0–48)	(Mean ± SD)
T0	6.7	(±4.1)	14.2	(±8.9)
T12	2.7	(±2.5)	5.0	(±5.1)
T24	2.3	(±2.1)	4.2	(±3.9)
T48	0.7	(±1.3)	1.2	(±2.3)
p^a	0.000		0.000	
	TNGSQ (0–10)	(Mean ± SD)	TNPDQ (0–10)	(Mean ± SD)
T0	2.9	(±2.5)	2.3	(±2.2)
T12	1.2	(±2.0)	1.1	(±2.0)
T24	1.7	(±2.0)	1.4	(±1.9)
T48	0.6	(±2.4)	0.3	(±1.0)
p^a	0.00		0.000	
	BEQ (0–12)	(Mean ± SD)	BESQ (0–36)	(Mean ± SD)
T0	4.7	(±1.8)	9.7	(±3.8)
T12	5.1	(±2.0)	10.9	(±4.4)
T24	5.9	(±1.4)	12.1	(±2.9)
T48	6.0	(±1.9)	12.7	(±3.8)
p^a	0.003		0.010	
	CDQ (0–4)	(Mean ± SD)	CDSQ (0–12)	(Mean ± SD)
T0	0.2	(±0.5)	1.2	(±1.8)
T12	0.6	(±1.1)	1.7	(±3.1)
T24	1.0	(±1.4)	2.4	(±3.3)
T48	1.1	(±1.4)	2.5	(±3.6)
p^a	0.001		0.14	

SD, standard deviation; HAQ, health assessment questionnaire; ESR, erythrocyte sedimentation rate (mm/hour); CRP, C-reactive protein (mg/dliter); SPQ, qualitative synovial proliferation; SPSQ, semi-quantitative synovial proliferation; PDQ, qualitative power Doppler; PDSQ, semi-quantitative power Doppler; TNGS, qualitative gray-scale tenosynovitis; TNPDQ, qualitative power Doppler tenosynovitis; BEQ, qualitative bone erosion; BESQ, semi-quantitative bone erosion; CDQ, qualitative cartilage damage; CDSQ, semi-quantitative cartilage damage.

[a] ANOVA for repeated measurements.

Table 4 – Baseline and longitudinal correlations between US10 parameters and both clinical and laboratory findings.

Baseline correlation between US10 parameters and both clinical and laboratory findings*								
US10	DAS28		HAQ		CRP		ESR	
	r	p	r	p	r	p	r	p
SPQ	0.06	0.683	−0.09	0.532	0.31	0.029	0.12	0.381
SPSQ	0.06	0.687	−0.06	0.657	0.29	0.039	0.08	0.155
PDQ	0.34	0.017	0.20	0.162	0.38	0.007	0.15	0.303
PDSQ	0.34	0.015	0.22	0.126	0.39	0.006	0.15	0.291
TNGSQ	0.35	0.014	0.14	0.320	0.34	0.015	0.50	0.000
TNPDQ	0.31	0.030	0.17	0.244	0.33	0.019	0.40	0.004
BEQ	−0.15	0.280	−0.25	0.081	0.16	0.264	−0.05	0.692
BESQ	−0.12	0.387	−0.23	0.107	0.14	0.333	−0.10	0.472
CDQ	−0.14	0.344	0.02	0.878	0.08	0.558	−0.19	0.194
CDSQ	−0.19	0.190	−0.09	0.501	0.04	0.787	−0.19	0.190
Correlations regarding longitudinal changes between US10 parameters and both clinical and laboratory findings								
SPQ	0.26	0.071	0.21	0.171	0.16	0.250	0.29	0.042
SPSQ	0.34	0.018	0.26	0.079	0.14	0.334	0.27	0.060
PDQ	0.24	0.089	0.32	0.023	0.34	0.016	0.22	0.133
PDSQ	0.27	0.057	0.34	015	0.29	0.041	0.22	0.123
TNGSQ	0.54	0.000	0.40	0.004	0.21	0.136	0.48	0.000
TNPDQ	0.45	0.001	0.32	0.023	0.34	0.106	0.47	0.001
BEQ	0.02	0.879	0.04	0.772	−0.05	0.696	−0.09	0.522
BESQ	0.04	0.758	0.07	0.633	−0.03	0.796	0.02	0.879
CDQ	−0.07	0.607	−0.07	0.612	0.18	0.203	0.09	0.513
CDSQ	−0.09	0.538	−0.02	0.862	0.19	0.188	0.16	0.272

HAQ, health assessment questionnaire; ESR, erythrocyte sedimentation rate (mm/h); CRP, C-reactive protein (mg/dliter); SPQ, qualitative synovial proliferation; SPSQ, semi-quantitative synovial proliferation; PDQ, qualitative power Doppler; PDSQ, semi-quantitative power Doppler; TNGSQ, qualitative gray-scale tenosynovitis; TNPDQ, qualitative power Doppler tenosynovitis; BEQ, qualitative bone erosion; BESQ, semi-quantitative bone erosion; CDQ, qualitative cartilage damage; CDSQ, semi-quantitative cartilage damage; r, Spearman's correlation.

systems based on these joints and have demonstrated reliability and correlations with clinical and laboratory findings.[15–18]

The US10 is composed of inflammatory and joint damage parameters. Qualitative (binary) and semi-quantitative analyses were used for the evaluation of synovial proliferation, PD and joint damage. A qualitative score is simpler and faster for clinical practice. However, according to OMERACT, the use of a semi-quantitative scoring system offers greater sensitivity for the assessment of changes throughout treatment and demonstrates reproducibility in different studies.[35] In the present study, tenosynovitis was also evaluated with a qualitative score in gray scale and PD. Only a qualitative analysis was performed, since there is no semi-quantitative score yet validated for this parameter. Thus, the US10 is the first global US scoring system for the prospective analysis of changes in inflammation and joint damage in patients with early AR with no prior use of DMARDs.

Two studies have evaluated US inflammatory changes in patients with early RA and no prior use of DMARDs. The first investigated the sensitivity and predictive value of PD changes in 28 joints regarding clinical, laboratory and radiological outcomes over a one-year period.[36] However, no statistically significant change in PD was found and there was no predetermined treatment strategy applied to the patients, unlike the present study. The second study investigated synovial proliferation and PD in 44 joints in patients with early RA who achieved clinical remission after the same treatment protocol. However, no change in US inflammatory parameters was described with treatment over time in comparison to the baseline evaluation.[37] Moreover, neither of the studies cited used a global US scoring system or investigated the presence of tenosynovitis.

Previous studies have assessed US changes over time using a global scoring system on patients with established RA or other chronic inflammatory joint diseases.[15–18] A simplified 12-joint scoring system proposed by Naredo et al. (2008) was applied to patients with RA with mean disease duration of 111 months.[15] Dougados et al. (2010) investigated different US scoring systems for synovial proliferation and PD in 20, 28 and 38 joints of patients with RA with mean disease duration of 10 years[17] and found changes in the inflammatory parameters over time in relation to the baseline evaluation (as in the present study), but did not investigate tenosynovitis. Backhaus et al. (2009) employed a seven-joint ultrasound score to follow up changes in different inflammatory joint diseases (RA, psoriatic arthritis, spondyloarthritis) and found statistically significant improvements in synovitis parameters (synovial proliferation and PD) and tenosynovitis after three and six months in comparison to the baseline evaluation.[16]

The evaluation of tenosynovitis in the joints studied in the US10 system (wrist and MCPs) revealed a statistically

significant improvement over a 48-week period. A recent study[38] reports that the sum of gray-scale and PD scores of the tendons of the long flexors of the fingers and ulnar extensor of the carpus was sensitive to change, with a reduction in scores after 12 months in patients with RA using adalimumab. However, unlike the present study, there was no assessment of the tendons of the digitorum communis flexors of the fingers and the patients had long disease duration. There is evidence that patients in the early stage of the disease experience greater tenosynovitis, especially in the digitorum communis flexors.[7,8]

This study showed in addition to the correlation between tendon score and disease activity, a correlation between the improvements of tendon score with the improvement of patient's function, analyzed by HAQ. Thus, one may postulate that the tendons of the hands and wrists should be included in the follow up of patients with RA, as demonstrated by the present US findings.

The only previous US scoring system to evaluate bone erosions was the seven-joint system proposed by Backhaus et al. (2009), which revealed no changes over a six-month period (unlike in the present study).[16] However, as mentioned above, the authors evaluated patients with different inflammatory joint diseases, including established RA, rather than patients with early RA.

Over the 48-week period, the change in the semi-quantitative synovial proliferation score was moderately correlated with changes in PD and tenosynovitis and changes in these latter two variables were correlated with each other. These findings lend support to the hypothesis that a reduction occurs in the entire intra-articular and peri-articular inflammatory process following the establishment of adequate treatment and underscores the importance of the assessment of tenosynovitis as a US measure in patients with early RA.

The US10 proved to be correlated with disease activity parameters (CRP and DAS28). These correlations have also been demonstrated in other US scoring systems.[12,15–18,36] Naredo et al. (2005) found a strong correlation between CRP and the number of joints with synovial proliferation and PD signals in 60 joints in patients with established RA. The same authors found moderate correlations between synovial proliferation and both CRP ($r = 0.33$) and the DAS28 ($r = 0.43$) as well as between PD and both CRP ($r = 0.33$) and DAS28 ($r = 0.48$), in 28 joints of patients with RA with less than one year of symptoms at the onset of treatment with DMARDs.[12] These correlations are similar to those found with the US10 scoring system in the present study. The global 12-joint US scoring system validated by Naredo et al. (2008) identified moderate correlations between synovial proliferation and both PD ($r = 0.55$) and the DAS28 ($r = 0.38$).[15] Perruconi et al. (2012) also found moderate correlations between the number of joints with synovitis and both the DAS28 ($r = 0.53$) and CRP ($r = 0.51$) in a simplified six-joint scoring system administered to patients with established RA.[18] It should be stressed that the US10 scoring system was administered to a small number of joints in patients with early RA and no prior use of DMARDs and demonstrated a close relationship between the US findings and disease activity criteria, allowing better management of these patients in clinical practice.

The correlations found in the present study are in disagreement with those reported in previous cohorts with early RA. Naredo et al. (2007) found moderate cross-sectional correlations between joint count and disease activity variables (CRP and DAS28) at baseline as well as at six, nine and 12 months of follow up, but found no longitudinal correlations between changes in US evaluation and changes in the clinical and laboratory findings.

In the present study, a correlation was also found between the change in active synovitis (assessed based on PD parameters) and a changes in the HAQ score. This finding strongly supports the hypothesis of an association between an improvement in the synovial inflammatory process and improvements in both function and quality of life.

Few studies have investigated interobserver reliability in US scoring systems. In the 12-joint system proposed by Naredo et al. (2008), the authors found moderate interobserver agreement ($k = 0.5$) regarding the presence of synovial proliferation (which is similar to the agreement found in the present study) and a stronger correlation than that reported in the present study between synovial proliferation and PD ($r = 0.8$).[15] In the seven-joint system proposed by Backhaus et al. (2009), moderate agreement was found regarding the identification of synovial proliferation ($k = 0.62$), with somewhat lesser agreement ($k = 0.55$) regarding the semi-quantitative score (0–3)[16] and better agreement ($k = 0.67$) regarding the qualitative PD variable, which is similar to that of the present study. The same study was the only previous investigation to address interobserver reliability in the detection of bone erosion, for which moderate agreement was found ($k = 0.56$).[16] Regarding joint damage, the US10 is the first global scoring system to address this evaluation, with excellent agreement found for the qualitative and semi-quantitative variables ($k = 0.79$ and 0.82, respectively).

The findings of the present study indicate that the proposed US10 is a valid scoring system for the follow up of inflammation and joint damage in patients with early RA, demonstrating significant correlations with clinical and laboratory findings as well as correlations between the changes in this score and both clinical and functional measures of the disease following a specific treatment protocol.

Conflicts of interest

The authors declare no conflicts of interest.

REFERENCES

1. Manger B, Kalden JR. Joint and connective tissue ultrasonography – a rheumatologic bedside procedure? A German experience. Arthritis Rheum. 1995;38(6): 736–42.
2. Grassi W, Cervini C. Ultrasonography in rheumatology: an evolving technique. Ann Rheum Dis. 1998;57(5): 268–71.
3. Canoso JJ. Ultrasound imaging. A rheumatologist's dream. J Rheumatol. 2000;27(9):2063–4.

4. Newman JS, Adler RS, Bude RO, Rubin JM. Detection id soft tissue hyperemia: value of power Doppler sonographyc. Am J Roentgenol. 1994;163(2):385–9.

5. Backhaus M, Kamradt T, Sandrock D, Loreck D, Fritz J, Wolf KJ, et al. Arthritis of the finger joints: a comprehensive approach comparing conventional radiography, scintigraphy, ultrasound, and contrast-enhanced magnetic resonance imaging. Arthritis Rheum. 1999;42(6):1232–45.

6. Wakefield RJ, Gibbon WW, Conaghan PG, O'Connor PJ, McGonagle D, Pease C, et al. The value of sonography in the detection of bone erosion in patients with rheumatoid arthritis. A comparison with conventional radiography. Arthritis Rheum. 2000;43(12):2762–70.

7. Grassi W, Filippucci E, Farina A, Cervini C. Sonographic imaging of tendons. Arthritis Rheum. 2000;43(5): 969–76.

8. Wakefield RJ, O'Connor PJ, Conaghan PG, McGonagle D, Hensor EM, Gibbon WW, et al. Finger tendon disease in untreated early rheumatoid arthritis: a comparison of ultrasound and magnetic resonance imaging. Arthritis Rheum. 2007;57(7):1158–64.

9. Scheel AK, Hermann KA, Kahler E, Pasewaldt D, Fritz J, Hamm B, et al. A novel ultrasonographic synovitis scoring system suitable for snalysing finger joint inflammation in rheumatoid arthritis. Arthritis Rheum. 2005;52(3): 733–43.

10. Szkudlarek M, Court-Payen M, Jacobsen S, Klarlund M, Thomsen H, Ostergaard M. Interobserver agreement in ultrasonography of the finger and toe joints in rheumatoid arthritis. Arthritis Rheum. 2003;48(4):955–62.

11. Filippucci E, Farina A, Carotti M, Salaffi F, Grassi W. Grey scale and power Dopller sonographic changes induces by intra-articular steroid injection treatment. Ann Rheum Dis. 2004;63(6):740–3.

12. 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 grey scale and power Doppler ultrasonography. Ann Rheum Dis. 2005;64(3):375–81.

13. Brown AK, Quinn MA, Karim Z, Conaghan PG, Peterfy CG, Hensor E, 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(12):3761–73.

14. Luz KR, Furtado R, Mitraud SV, Porglhof J, Nunes C, Fernandes AR, et al. Interobserver reliability in ultrasound assessment of rheumatoid wrist joints. Acta Reumatol Port. 2011;36(3):245–50.

15. Naredo E, Rodríguez M, Campos C, Rodríguez-Heredia JM, Medina JA, Giner E, et al. Validity, reproducibility, and responsiveness of a twelve-joint simplified power doppler ultrasonographic assessment of joint inflammation in rheumatoid arthritis. Arthritis Rheum. 2008;59(4): 515–22.

16. Backhaus M, Ohrndorf S, Kellner H, Strunk J, Backhaus TM, Hartung W, et al. Evaluation of a novel 7-joint ultrasound score in daily rheumatologic practice: a pilot project. Arthritis Rheum. 2009;61(9):1194–201.

17. Dougados M, Jousse-Joulin S, Mistretta F, d'Agostino MA, Backhaus M, Bentin J, 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(5):828–33.

18. Perricone C, Ceccarelli F, Modesti M, Vavala C, Di Franco M, Valesini G, et al. The 6-joint ultrasonographic assessment: a valid, sensitive-to-change and feasible method for evaluating joint inflammation in RA. Rheumatology (Oxford). 2012;51(5):866–73.

19. Jacoby, Jayson RK, Cosh MIV, Onset JA. Early stages and prognosis of rheumatoid arthritis: a clinical study of 100 patients with 11 year follow-up. Br Med J. 1973;2(5858): 96.

20. Hamalainen M, Kammonen M, Lethimaki M. Epidemiology of wrist involvement in rheumatoid artritis. Rheumatol. 1992;17:1–7.

21. Visser H, le Cessie S, Vos K, Breedveld FC, Hazes JM. How to diagnose rheumatoid arthritis early: a prediction model for persistent (erosive) arthritis. Arthritis Rheum. 2002;46(2):357–65.

22. Tan AL, Tanner SF, Conaghan PG, Radjenovic A, O'Connor P, Brown AK, et al. Role of metacarpophalangeal joint anatomic factors in the distribution of synovitis and bone erosion in early rheumatoid arthritis. Arthritis Rheum. 2003;48(5):1214–22.

23. Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum. 1988;31(3):315–24.

24. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO, et al. 2010 Rheumatoid Arthritis Classification Criteria An American College of Rheumatology/European League Against Rheumatism Collaborative Initiative. Arthritis Rheum. 2010;62(9):2569–81.

25. van der Helm-van Mil AH, le Cessie S, van Dongen H, Breedveld FC, Toes RE, Huizinga TW. A prediction rule for disease outcome in patients with recent-onset undifferentiated arthritis: how to guide individual treatment decisions. Arthritis Rheum. 2007;56(2):433–40.

26. Ferraz MB, Oliveira LM, Araújo PM, Atra E, Tugwell P. Crosscultural reliability of the physical ability dimension of the health assessment questionnaire. J Rheumatol. 1990;17(6):813–7.

27. Backhaus M, Burmester GR, Gerber T, Grassi W, Machold KP, Swen WA, et al. Guidelines for musculoskeletal ultrasound in rhematology. Ann Rheum Dis. 2001;60(7):641–9.

28. Wakefield RJ, Balint PV, Szkudlarek M, Filippucci E, Backhaus M, D'Agostino MA, et al. Musculoskeletal ultrasound including definitions for ultrasonographic pathology. J Rheumatol. 2005;32(12):2485–7.

29. Grassi W, Tittarelli E, Pirani O, Avaltroni D, Cervini C. Ultrasound examination of metacarpophalangeal joints in rheumatoid arthritis. Scand J Rheumatol. 1993;22(5): 243–7.

30. Boutry N, Lardé A, Demondion X, Cortet B, Cotten H, Cotton A. Metacarpophalangeal joints at US in asymptomatic volunteers and cadaveric specimens. Radiology. 2004;232(3):716–24.

31. Möller B, Bonel H, Rotzetter M, Villiger PM, Ziswiler HR. Measuring finger joint cartilage by ultrasound as a promising alternative to conventional radiograph imaging. Arthritis Rheum. 2009;61(4):435–41.

32. Filippucci E, da Luz KR, Di Geso L, Salaffi F, Tardella M, Carotti M, et al. Interobserver reliability of ultrasonography in the assessment of cartilage damage in rheumatoid arthritis. Ann Rheum Dis. 2010;69(10):1845–8.

33. van der Heijde DM. Plain X-rays in rheumatoid arthritis: overview of scoring methods, their reliability and applicability. Baillieres Clin Rheumatol. 1996;10(3): 435–53.

34. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33(1):159–74.

35. Boers M, Brooks P, Strand CV, Tugwell P. The OMERACT filter for outcome measures in rheumatology. J Rheumatol. 1998;25(2):198–9.

36. Naredo E, Collado P, Cruz A, Palop MJ, Cabero F, Richi P, et al. Longitudinal power Doppler ultrasonographic assessment of joint inflamatory activity in early rheumatoid arthritis: predictive value in disease activity and radiologic progression. Arthritis Rheum. 2007;57(1):116–24.

37. Scirè CA, Montecucco C, Codullo V, Epis O, Todoerti M, Caporali R. Ultrasonographic evaluation of joint involvement in early rheumatoid arthritis in clinical remission: power Doppler signal predicts short-term relapse. Rheumatology (Oxford). 2009;48(9):1092–7.

38. Hammer HB, Kvien TK. Ultrasonography shows significant improvement in wrist and ankle tenosynovitis in rheumatoid arthritis patients treated with adalimumab. Scand J Rheumatol. 2011;40(3):178–82.