Rheumatology
INTERNATIONAL

VALIDATION STUDIES

Measuring inflammation in rheumatoid arthritis with a new clinical and ultrasound index: development and initial validation

Juan José de Agustín[1] · Alba Erra[2] · Andrés Ponce[3] · Carmen Moragues[4,5] · Cesar Díaz-Torné[5] · Delia Reina[6] · Estefanía Moreno[7] · Julio Ramírez[8] · Lourdes Mateo[9] · Manel Pujol Busquets[10] · Patricia Moya[5] · Pilar Santo-Panero[11] · Sergi Ros-Expósito[12] · Javier Narváez[13] · Raimon Sanmartí[8] · Mª Jesús García de Yébenes[14] · Loreto Carmona[14]

Abstract

With the aim to develop and validate a clinical + ultrasound (US) inflammation score in rheumatoid arthritis (RA) for use in clinical practice, a mixed-method study was conducted. The theoretical development of the index was achieved with qualitative methodology (discussion group and Delphi survey). Subsequently, a cross-sectional study was carried out to analyse issues related to scoring and validation of the new index. RA patients underwent clinical [28 swollen and tender joints count, patient and physician global assessment (PhGA), erythrocyte sedimentation rate (ESR), C-reactive protein (CRP)], and US assessments [synovitis or tenosynovitis by grey-scale (GS) and power Doppler (PD) of 42 structures]. An index was created based on statistical models and expert interaction. Construct validity was tested by correlation with DAS28, SDAI, CDAI, and PhGA. Reliability was evaluated in a subgroup of patients with the intraclass correlation coefficient (ICC). US assessment, CRP, and swollen joints were the items that passed the prioritization phase (Delphi study). For the cross-sectional study, 281 patients were randomly divided into design ($n = 141$) and validation samples ($n = 140$). The combination of US sites chosen (7 bilaterally) detected the maximum proportion of synovitis and PD present. Three scoring methods were tested: semiquantitative (0–3 GS + 0–3 PD), dichotomous (0/1 GS + 0/1 PD), and qualitative (0/1 based on algorithm). All showed strong correlation with activity measures ($\rho \geq 0.60$), and reliability (ICC 0.89–0.93). The index with best parameters of validity, feasibility, and reliability was the qualitative. The final index chosen was the sum of swollen joint count, US qualitative score, and CRP. The UltraSound Activity score is a valid and reliable measure of inflammation in RA equal to the sum of 28 SJC, a simplified (0/1) US assessment of 11 structures and CRP. It is necessary further investigation to demonstrate additional value over existing indices.

Keywords Rheumatoid arthritis · Validation studies · Ultrasound · Composite score · Index

Introduction

The prognosis of patients with rheumatoid arthritis (RA) has improved markedly due to a combination of factors, such as the development of biological agents and new small molecules, a better use of conventional therapies—probably in relation to the high price of the newer drugs—and the use

✉ Juan José de Agustín
    jjagor@hotmail.com

Extended author information available on the last page of the article

of strategies for treatment adjustment according to disease activity. These latter strategies imply the use of quantitative measures—so-called clinical indices—to avoid basing decisions solely on physician impression.

In general, quantitative measures for RA are composite indices based on a core set of measurements, including number of tender and swollen joints, physician assessment of disease activity, patient evaluation of pain, activity, and physical function, plus laboratory measurements, basically acute phase reactants, either the erythrocyte sedimentation rate (ESR), or C-reactive protein (CRP) [1, 2].

The composite indices developed for RA in the last decades are based on this core set; nevertheless, none of them includes all dimensions of disease activity. The most

commonly used indices are: the Disease Activity score (DAS) [3, 4], which combines, in a continuous measure, the Ritchie's index (number of tender joints out of 44), a global assessment of disease activity by the patient through an Visual Analogue scale (VAS) from 0 to 10, and ESR; simplified versions of the DAS, using the counts of 28 joints instead of 44 [5] and using CRP instead of ESR in the formula [6, 7]; the Simplified Disease Activity index (SDAI), which is the arithmetic sum of the variables included in DAS28 plus the physician's assessment of disease activity [8]; and a version without phase reactants, the Clinical Disease Activity index (CDAI) [9, 10].

The DAS28, used as part of the endpoint in almost all RA clinical trials, has some limitations, the main being its complexity for calculation in clinical practice. In addition, in the formula, the tender joint count is weighted twice as much as the swollen joint count when the latter is a more specific characteristic for RA than the former; the same for ESR, which is strongly weighted and may induce changes in the index even when within the normal range [11.] Some researchers argue that some parameters included in the DAS28 are very subjective, as they may be affected by the psychological state, not solely by inflammation, not accurately reflecting the patient's clinical status, but his or her perception [12]. All in all, and despite being an excellent measure for clinical trials, DAS28 may result in wrong measures of disease activity at the individual patient level. As a result, several groups are working on adapting this index to clinical practice. In addition, it is more evident the necessity to implement new methods to identify better the real inflammatory state of these patients. Remission and low activity state are probably the most important targets for these new tools.

High-frequency ultrasound (US) can be used to assess objectively inflammation in RA [13]. Adding US to a clinical measure should, in principle, improve the reliability and validity of the measurement of disease activity, and several US indices have been developed, using various combinations of joints and scoring systems [14–16]. All of them use information based solely on US. The correlation of these indices with the DAS28 is good, but none of the US-based indices has been extensively used in clinical practice because of different reasons, such as the large number of joints to be evaluated, the use of a semiquantitative scale for synovitis, and the long time required to perform the evaluation [17].

An index based on essential clinical measures plus a US measure, focused on simplicity, with appropriate validation, would allow a better classification of patients at different levels of disease activity than a clinical only or US only index. On the other hand, in cases with low activity or remission, a combined index including ultrasound measure could provide a more objective measure detecting subclinical synovitis to help the rheumatologist being more effective in using medications. The main objective of this study was to develop and validate a mixed clinical-US index to reflect disease inflammatory activity in RA for use in clinical practice avoiding confounding variables.

Methods

Study design

This study was carried out with mixed methods, qualitative and quantitative.

For the theoretical development of the index, discussion group and Delphi techniques were used. The purpose of the discussion group—composed of ten rheumatologists from the US group of the Catalan Society of Rheumatology plus two renowned RA experts—was to elicit items or potential elements to be included in a disease activity index and to define them as clearly as possible, including measurement variants. After the meeting, a Delphi survey (19 investigators) was carried out to prioritize items among all elicited ones to include in the index. For this task, all items were anonymously graded as to "perceived degree of objectivity", "capacity to reflect RA activity and its changes", and "feasibility in clinical practice" in 1–5 scales. The grades were then averaged into a global ranking score from 1 to 5. Only those items which scores were ≥ 4 were forwarded to a second quantitative phase.

For the construction and validation of the index, a cross-sectional multicentre study was conducted.

Centres in Catalonia with at least one rheumatologist with high experience in US and availability of US machines at the rheumatology offices were invited to participate. All ultrasonographers had demonstrated experience performing musculoskeletal US. All of them had passed the advance level of Spanish ultrasound school of our Rheumatology Society and had more than 5 years of experience performing US exams.

Patients

Patients were selected from the participating centres through a consecutive sampling among those with enough inflammatory activity to evaluate all the items needed for the index. All centres included the same number of patients during the recruitment period (2 months). Neither treatment nor comorbidities were considered as exclusion criteria. Patients who had required joint surgery in the past were excluded.

The inclusion criteria were:

1   Diagnosis of RA according to EULAR (*European League Against Rheumatism*)/ACR (*American College of Rheumatology*) 2010 criteria [18];
2   Any degree of inflammatory activity with any treatment;

All patients signed the inform consent prior to inclusion in the study.

Variables

The gold standard was disease activity (present or absent), defined by consensus of the panel of rheumatologists as based on all available information in each case.

In addition, the following clinical variables were collected: swollen and tender joints from a total of 28 [shoulders, elbows, wrists, metacarpophalangeal (MCP) 1–5, interphalangeal (IP) 1–5, and knees], patient and physician global assessment (VAS), ESR, CRP, rheumatoid factor (RF), and anticyclic citrullinated peptide antibodies (ACPA); age; disease duration; and previous or current treatment.

We used ultrasound machines of General electric (GE) and Esaote (E) brand: E Mylab Six, Mylab seven and Mylab twice, and GE S8 and E9. All machines were equipped with a multifrequency linear transducer and frequency used was the maximum possible in each machine (10–15 MHz) and in each joint to obtain the best quality image as possible. In each ultrasound machine, the same settings were used for all patients.

PD settings were: medium dynamic range, medium persistence, medium frame rate, low wall filter, and 0.5–0.8 Hz pulse repetition frequency. In each machine, these parameters were adjusted to obtain the maximum sensitivity to identify Doppler signal at the lowest possible value for each joint. The US assessment included: synovitis or tenosynovitis by grey scale (GS) and PD from a total of 42 anatomical structures [bilateral shoulder, bilateral elbow, bilateral wrist, bilateral wrist flexor tendons (all grouped), bilateral wrist extensor tendons first–sixth compartment (all grouped), bilateral MCP 2–5, bilateral IP 2–5, bilateral finger flexor tendons (all grouped), bilateral knee, bilateral posterior tibial tendon, bilateral peroneal tendons (long and brevis grouped), bilateral tibio-talar joints, bilateral subtalar joints, and bilateral metatarsophalangeal (MTP) second and third]. The standard US method was used (OMERACT definitions) [19].

Each structure was scanned in longitudinal and transverse view following EULAR ultrasound recommendations.

Blinding between clinical and US assessments was maintained by having patient's data collected by two independent rheumatologists in each centre.

Statistical analysis

After the descriptive analysis, the study sample was divided into two random sub-samples, one for the construction of the index and the other for its validation.

For the construction of the index, a procedure in different and successive steps was used: (1) selection of US locations;

(2) selection of US scoring method; and (3) creation of the index.

For the selection of the most suitable US sites to include in the index, we used combinations based on: (a) frequency of US abnormalities, typical of RA, in GS and PD; (b) locations used in the indices published by Naredo et al. formed by 12 structures [15], and in the APPRAISE study [19]; and (c) feasibility and time spent on US assessment in the opinion of the researchers. We then tested the capacity of the different combinations to detect > 90% of the structures with synovitis and PD signal (sensitivity).

Once the locations to be included were selected, the scoring method needed to be defined. For this, we used three approaches: (a) a semiquantitative scale (0–3) used for GS and PD; (b) a dichotomous scale also for GS and PD (0 and 1 if GS > 2 or PD ≥ 1); and (c) a qualitative scale (0 and 1 total, not by GS and PD) based on a decision tree proposed by the participating researchers (Fig. 1).

The composite index was calculated as the arithmetic sum of three constituent sub-scales (US assessment, swollen joints count, and CRP) selected from the previous steps. Three indices were created according to the different scoring modalities of the US evaluation.

After the creation of the index, a validation study was carried out in the subsample intended for this purpose. The dimensions of validity analysed were construct validity and reliability.

Construct validity was tested by a correlation analysis between the new index and different external measures of activity. Convergent validity was based on the
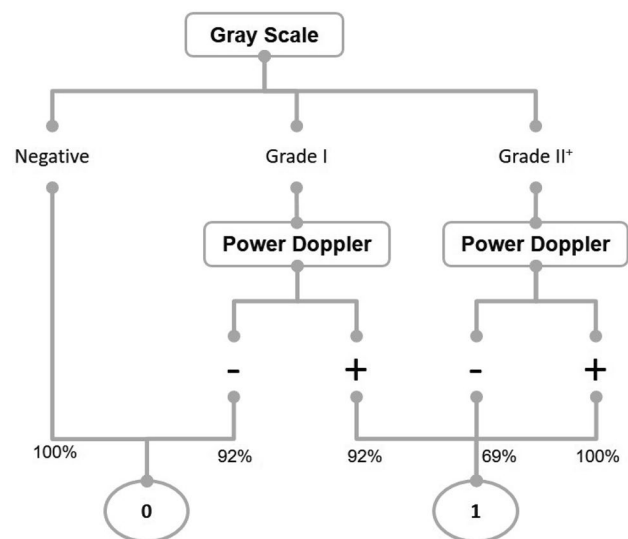

Fig. 1 Decision algorithm for the qualitative ultrasound score of inflammation. All researchers were asked to define the value (0 or 1) of each branch of the algorithm as to whether there was inflammation or not. The percentages (%) reflect the agreement with the final value of the branch

correlation between the new index/ices and DAS28, SDAI, and CDAI, calculated according to the following formulas: $DAS28 = (0.56 \times \sqrt{\text{tender joints count}}) + (0.28 \times \sqrt{\text{swollen joint count}}) + (0.70 \times \ln(\text{ESR})) + (0.014 \times \text{VAS physician})$; $SDAI = \text{swollen joint count} + \text{tender joints count} + \text{VAS (patient)} + \text{VAS (physician)} + \text{CRP (mg/dl)}$; and $CDAI = \text{tender joints count} + \text{swollen joint count} + \text{VAS (patient)} + \text{VAS (physician)}$.

Correlation between the new composite US indices and the calculated external activity measures [(DAS28, CDAI, and SDAI plus the physician's overall assessment (PGA)] was analysed with the Spearman correlation coefficient (ρ).

The reliability of the index was evaluated by a test–retest analysis in a subgroup of patients re-evaluated in a week time (5 per centre) with the intraclass correlation coefficient (ICC).

Results

Delphi study

Responses with values ≥ 4 over 5 were considered a priority to include in the assessment. The only parameters that met this requirement were: bilateral US evaluation of tendons and joints, CRP, and swollen joint count.

Descriptive analysis

A total of 13 hospitals participated in the study. The sample consisted of 281 RA patients, mainly women in their fifties, with a mean of 11 years from diagnosis, and various levels of disease activity, being 80% receiving treatment with DMARD and 46% with biological (Table 1).

After the descriptive study, a random sample split was performed in two sub-samples, one for the index design ($n = 141$) and the other for the validation study ($n = 140$).

Selection of ultrasound locations

Two summary variables were constructed for GS and PD based on the frequency of abnormalities at each US location. The presence of synovitis and PD signal was defined by values greater than zero in any location. Among the 141 patients included, 130 (92%) had synovitis at some location, and in 89 (63%), a PD signal was detected. The most frequently affected joint was the wrist and the sites that presented fewer abnormalities were the wrist flexors and peroneal tendons. Table 2 shows the sensitivity of the different combinations of sites to detect synovitis and PD.

Due to the similarity of the results on the capacity of these combinations to detect synovitis and PD, it was decided to use the last combination for the construction of the index,

Table 1 Description of the complete sample (creation plus validation)

Variable	Value
Age, m ± SD	59.7 ± 13.1
Female sex, n (%)	217 (77)
Rheumatoid factor positive, n (%)	216 (77)
Disease duration (years), m ± SD	11.6 ± 10.5
ACPA+, n (%)	216 (79)
ESR (mm/h), m ± SD	24.3 ± 23.1
CRP (mg/L), m ± SD	10.0 ± 26.5
Tender joint count, m ± SD	3.2 ± 3.9
Swollen joint count, m ± SD	2.3 ± 2.7
Global physician assessment, m ± SD	3.2 ± 2.2
Global patient assessment, m ± SD	4.0 ± 2.5
Morning stiffness (min), m ± SD	23.0 ± 41.7
Previous DMARDs, n (%)	191 (70)
Number of DMARDs, m ± SD	1.8 ± 1.2
Current DMARDs, n (%)	224 (80)
Previous biologics, n (%)	90 (32)
Number of previous biologics, m ± SD	1.7 ± 1.2
Current biologics, n (%)	117 (45)

m mean, *SD* standard deviation, *DMARDs* disease modifying antirheumatic drugs, *ESR* erythrocyte sedimentation rate, *CRP* C-reactive protein, *ACPA* anticyclic citrullinated peptide antibodies

as it was considered more feasible and faster to perform. This combination contains 14 structures (7 bilateral): wrist (including flexors and extensors of the wrist), MCP (2 and 3), knee, tibio-talar joint (plus posterior tibial tendon and peroneal tendon), and MTP (2 and 3).

Methods of ultrasound scoring

We then calculated the scores by the three previously defined scoring scales: (1) semiquantitative: from 0 to 3 (each location is based on 4 evaluations (right and left side, GS and PD); therefore, the score ranges from 0 to 12); (2) Dichotomous: 0/1 (0 and 1 in GS count as 0, and 2 and 3 as 1; in PD, any value ≥ 1 counts as 1; the total score per location ranging 0–4); and (3) Qualitative: 0/1 based on the algorithm (Fig. 1) (range per location: 0–2). In the qualitative score, if one area has a 1, for instance a tendon, the rest of the area needs no assessment, as it will be a 1 in any case.

Final indices to validate

Three composite indices were created based on the sum of US assessment, swollen joint counts, and CRP value (mg/dL) with each of the three scales (index 1 semiquantitative, index 2 dichotomous, and index 3 qualitative).

Table 2 Sensitivity of different locations for the detection of synovitis or power Doppler signal

Possible combinations	Synovitis: n (%) (Total > 0; $n = 130$)	PD signal: n (%) (Total > 0; $n = 89$)
Criteria: frequency of US synovitis		
Wrist	86 (66.1)	72 (80.9)
Wrist, 2 MCP	101 (77.7)	80 (89.9)
Wrist, 2 MCP, 3 MCP	105 (80.8)	83 (93.2)
Wrist, 2 MCP, 3 MCP, Knee	119 (91.5)	87 (97.7)
Wrist, 2 MCP, 3 MCP, Knee, 2 MTP	123 (94.6)	87 (97.7)
Wrist, 2 MCP, 3 MCP, Knee, 2 MTP, 3 MTP	123 (94.6)	87 (97.7)
Wrist, 2 MCP, 3 MCP, Knee, 2 MTP, tibial posterior	123 (94.6)	87 (97.7)
Wrist, 2 MCP, 3 MCP, Knee, 2 MTP, shoulder	126 (96.9)	87 (97.7)
Wrist, 2 MCP, 3 MCP, Knee, 2 MTP, shoulder, elbow	126 (96.9)	87 (97.7)
Wrist, 2 MCP, 3 MCP, Knee, 2 MTP, shoulder, carpal extensor tendons	126 (96.9)	87 (97.7)
Criteria: published reduced indexes		
Wrist, 2 MCP, 3 MCP, Knee, ankle, elbow [15]	119 (91.5)	87 (97.7)
Wrist, 2 MCP, 3 MCP, 2 IFP, knee, 2 MTP, 3 MTP, shoulder, elbow [19]	126 (96.9)	87 (97.7)
Criteria: feasibility		
Wrist, 2 MCP, 3 MCP, Knee, 2 MTP, tibio-talar, elbow	123 (94.6)	87 (97.7)
Wrist, wrist flexor tendons, wrist extensor tendons, MCP 2–3, knee, tibial posterior tendons, peroneal tendons, tibio-talar joint, MTP 2–3	125 (96.1)	87 (97.7)
Wrist extended*, MCP 2–3, knee, tibio-talar extended†, MTP 2–3	125 (96.1)	87 (97.7)

All combinations are bilateral and include the assessment of GS + PD

MTP metatarsophalangeal, *MCP* metacarpophalangeal, *IP* interphalangeal 1–5, *PD* power Doppler

*Wrist + flexor + extensor tendons, acting as a single region

†Tibio-talar joint + tibial posterior tendons, peroneal tendons, as a single region

Validation study

In the validation subsample, 126 patients had valid information for the construction of the three composite indices. The mean ± SD of the three indices in this subsample were 11.9 ± 9.7 for the semiquantitative; 3.3 ± 3.1 for the dichotomous; and 4.6 ± 3.8 for the qualitative. The respective ranges were: 0–43; 0–13.7; and 0–16.4.

Correlations between the three US indices were high. Correlation with external measures of activity was higher for the dichotomous and qualitative indexes than for the semiquantitative. The highest correlations were obtained with the physician's overall assessment (PGA) (0.702 and 0.771), followed by DAS-28 (0.694 and 0.678), SDAI (0.661 and 0.666), and finally CDAI (0.652 and 0.658) (Table 3).

For the reliability analysis, a sample of 44 patients with two US examinations separated by a week was used. The ICC values obtained were high for all three indices, ranging from 0.89 to 0.93 (Table 4). No modification in treatment was done between both US examinations.

Table 3 Correlation between composite ultrasound indices and external activity measures

	Index 1	Index 2	Index 3	DAS28	SDAI	CDAI	VGM
Index 1	1.000						
Index 2	0.779	1.000					
Index 3	0.901	0.941	1.000				
DAS-28	0.627	0.694	0.678	1.000			
SDAI	0.604	0.661	0.666	0.796	1.000		
CDAI	0.599	0.652	0.658	0.791	0.991	1.000	
PGA	0.648	0.702	0.771	0.716	0.898	0.897	1.000

Cells contain the ρ value between the variables in column and row

Index 1: semiquantitative scale, index 2: dichotomous scale, index 3: qualitative scale

PGA physician global assessment of disease activity (gold standard)

Table 4 Interobserver reliability of the three indices created

Composite index	ICC (95% CI)
Semiquantitative score (index 1)	0.93 (0.87–0.96)
Dichotomous score (index 2)	0.78 (0.64–0.88)
Qualitative score (index 3)	0.89 (0.80–0.94)

ICC intraclass correlation coefficient, *CI* confidence interval

Given the most adequate validity, reliability, and feasibility, the index selected was index 3, from now on, USAS (UltraSound Activity score).

Discussion

We have created a new tool, USAS, combining clinical, US, and biological information to measure and monitor inflammatory activity in RA patients in clinical practice (See Supplementary material for an example of scoring sheet). Experts agreed upon the face validity of the components finally included in the proposed tool, as a reflection of inflammatory state, rather than other aspects of RA activity. Further development of the index stressed on the feasibility of the tool without losing validity. Finally, the validation showed a good behaviour against widely used indices of disease activity, and high reproducibility.

The utility of clinical parameters in decision making in RA is out of doubt [20]. However, evidence shows over and underestimation of the inflammatory activity in individual patients [19, 21–23]. US have demonstrated greater sensitivity than clinical exam for detecting synovitis, and PD is an excellent outcome marker for flare and structural damage and sensitive to change in patients under different treatment strategies [15, 19, 24, 25]. US is not systematically used in the clinic to monitor disease activity for different reasons, mainly related to feasibility and scoring complexity. We created the USAS with the aim to be as simple as possible, so that rheumatologists, with any level of expertise on US, could use it in the clinical practice. For this reason, we included a large number of centres with US equipment of varying quality, and rheumatologists with sufficient qualification to perform an US but without a specified level of expertise. Such heterogeneity, instead of becoming a problem, enhanced the possibility to design an index useful in a real-life context. Other authors, as Naredo et al. have used similar strategies for their US studies [15].

The originality of USAS relies in that it combines exam, laboratory and US parameters. Most studies on US scores did not integrate other measures, and show poor concordance with clinical parameters, probably because they measure distinct aspects of the patient's clinical situation [14, 15, 26–32]. The DASECO is an US-based DAS28 index that uses GS and PD measurements as an alternative to tender and swollen joints by physical exam plus the rest of parameters in the formula of the DAS28 [33]. The correlation between the DAS28 and the DASECO is good. However, we should bear in mind subtle differences between the DASECO and the original DAS28 evaluations: (1) in DASECO, PD is performed in the MTP joints, whereas DAS28 does not include foot joints; (2) in the DAS28 the joint count is binary (yes/no for tenderness and inflammation), while the US score is semiquantitative; and (3) DASECO differentiates between tenosynovitis and synovitis, this latter more representative of RA, whereas this differentiation is not possible in DAS28. Our group did not aim to create a variant of the DAS28—same as when we enter CRP instead of ESR in the formula—but to generate a reduced measure—to enhance feasibility—with enough face validity of inflammation and feasibility. We could have used a weighted index or formula, as in the DASECO, but decided to use a sum instead, as in the CDAI. These decisions improve uptake in the clinic, and we proved that are valid, at least to the point we have validated the USAS. On the other hand, the time to perform the USAS index was less than 30 min, including joint count and ultrasound examination. In our opinion, this time is feasible if we intend to carry out an exhaustive evaluation of the patient's inflammatory state.

Compared to other US scores, USAS uses a qualitative scoring based on an easy algorithm, while others, except for the DASECO [33], use semiquantitative scores. Semiquantitative scoring may not be the most appropriate in clinical practice, because is time consuming and there is significant variability in the assessment of different degrees of activity in GS and PD. In addition, PD is probably the best outcome parameter of inflammation in musculoskeletal US [34–36]. Based on this, the algorithm used in the USAS weighs PD heavily, what increases the face validity even more.

Regarding the structure selected for US assessment, most of the relevant US scores include anatomical locations similar to ours [14, 15, 29, 30, 37]. We decided to add tendons to complete the best possible evaluation, in line with previous studies showing the need to evaluate joints and tendons jointly in RA patients [38–40]. Whether merging tendons with joints may have an impact on sensitivity to change must be tested.

Recently, another group has created a combined clinical and ultrasound score, the US-CLARA [41]. This index is clearly different to our, presented here. They use different parameters, as self-administered tender joint counts, that can be considered subjective and very influenced by other pathologies as osteoarthritis, fibromyalgia, or pain. Our score is specifically created to avoid this problem, and the main advantage is the utility to identify inflammatory activity in patients in clinical remission or help clinicians to better identify the extension of inflammatory activity. Both are

created to control RA patients, but our score is specifically thought to study real inflammatory state in clinical practice, avoiding subjective clinical variables.

A final recall to the objectives of this new development. Despite needing further validation, we created the USAS with the aim to reflect inflammation and to be used in clinical practice. The fact that USAS reflects inflammation may need further assessment—independently of the hurdles of finding an adequate gold standard—but it may be the reason why the correlation with DAS28, SDAI, and CDAI is not perfect, but it is with PhGA. We believe that classical activity measures are measuring something more than pure inflammation. On the other hand, we are not suggesting that USAS should be used in all patients in clinical practice, but mostly in those in which there might be a discordance between the patient and doctor assessment, namely those in which inflammation might not be as clear as in others. Further validation to confirm this and other hypotheses is under way. Once the validity and reliability of the index have been demonstrated, a prospective study shall be carried out to assess its responsiveness, or ability to detect a change in the construct of interest (activity). For methodological reason we decided to perform the project in two different phases. First, the creation and internal validation, presented here. In second place, external validation in a prospective study to analyse the responsiveness of the index and identify levels of inflammatory activity using USAS score. In summary, USAS combines clinical, laboratory, and a simplified US assessment in a single score with good metric properties, easy to perform. Although further validation in other setting is needed, the USAS is able to help the knowledge of inflammatory process and outcome as well as facilitate the evaluation of RA patients in whom inflammation may be unclear.

Compliance with ethical standards

References

1. Buchbinder R, Bombardier C, Yeung M, Tugwell P (1995) Which outcome measures should be used in rheumatoid arthritis clinical trials? Clinical and quality-of-life measures' responsiveness to treatment in a randomized controlled trial. Arthritis Rheum 38(11):1568–1580 **(PubMed PMID: 7488277)**

2. Kirwan JR (1993) Outcome measures in rheumatoid arthritis clinical trials: assessing improvement. J Rheumatol 20(3):543–545 **(PubMed PMID: 8478868)**

3. van der Heijde DM, van't Hof MA, van Riel PL, Theunisse LA, Lubberts EW, van Leeuwen MA et al (1990) Judging disease activity in clinical practice in rheumatoid arthritis: first step in the development of a disease activity score. Ann Rheum Dis 49(11):916–920 **(PubMed PMID: 2256738. Epub 1990/11/01. Eng)**

4. van der Heijde DM, van't Hof M, van Riel PL, van de Putte LB (1993) Development of a disease activity score based on judgment in clinical practice by rheumatologists. J Rheumatol 20(3):579–581 **(PubMed PMID: 8478878. Epub 1993/03/01. Eng)**

5. Prevoo ML, van't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LB, van Riel PL (1995) Modified disease activity scores that include twenty-eight-joint counts development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. Arthritis Rheum 38(1):44–48 **(PubMed PMID: 7818570. Epub 1995/01/01. Eng)**

6. Inoue E, Yamanaka H, Hara M, Tomatsu T, Kamatani N (2007) Comparison of Disease Activity Score (DAS)28-erythrocyte sedimentation rate and DAS28-C-reactive protein threshold values. Ann Rheum Dis 66(3):407–409 **(PubMed PMID: 16926186. Epub 2006/08/24. Eng)**

7. Matsui T, Kuga Y, Kaneko A, Nishino J, Eto Y, Chiba N et al (2007) Disease Activity Score 28 (DAS28) using C-reactive protein underestimates disease activity and overestimates EULAR response criteria compared with DAS28 using erythrocyte sedimentation rate in a large observational cohort of rheumatoid arthritis patients in Japan. Ann Rheum Dis 66(9):1221–1226 **(PubMed PMID: 17369281. Epub 2007/03/21. Eng)**

8. Smolen JS, Breedveld FC, Schiff MH, Kalden JR, Emery P, Eberl G et al (2003) A simplified disease activity index for rheumatoid arthritis for use in clinical practice. Rheumatology (Oxford) 42(2):244–257 **(PubMed PMID: 12595618. Epub 2003/02/22. Eng)**

9. Aletaha D, Nell VP, Stamm T, Uffmann M, Pflugbeil S, Machold K et al (2005) Acute phase reactants add little to composite disease activity indices for rheumatoid arthritis: validation of a clinical activity score. Arthritis Res Ther 7(4):R796–R806 **(PubMed PMID: 15987481. Epub 2005/07/01. Eng)**

10. Aletaha D, Smolen J (2005) The Simplified Disease Activity Index (SDAI) and the Clinical Disease Activity Index (CDAI): a review of their usefulness and validity in rheumatoid arthritis. Clin Exp Rheumatol 23(5 Suppl 39):S100–S108 **(PubMed PMID: 16273793. Epub 2005/11/09. Eng)**

11. Bakker MF, Jacobs JW, Verstappen SM, Bijlsma JW (2007) Tight control in the treatment of rheumatoid arthritis: efficacy and feasibility. Ann Rheum Dis 66(Suppl 3: iii):56–60 **(PubMed PMID: 17934098. Pubmed Central PMCID: 2095293. Epub 2007/11/21. Eng)**

12. Taylor WJ, Harrison AA, Highton J, Chapman P, Stamp L, Dockerty J et al (2008) Disease Activity Score 28-ESR bears a similar relationship to treatment decisions across different rheumatologists, but misclassification is too frequent to replace physician judgement. Rheumatology (Oxford) 47(4):514–518 **(PubMed PMID: 18321947. Epub 2008/03/07. Eng)**

13. Szkudlarek M, Court-Payen M, Strandberg C, Klarlund M, Klausen T, Ostergaard M (2001) Power Doppler ultrasonography for assessment of synovitis in the metacarpophalangeal joints of patients with rheumatoid arthritis: a comparison with dynamic magnetic resonance imaging. Arthritis Rheum 44(9):2018–2023 **(PubMed PMID: 11592362. Epub 2001/10/11. Eng)**

14. Backhaus M, Ohrndorf S, Kellner H, Strunk J, Backhaus TM, Hartung W et al (2009) Evaluation of a novel 7-joint ultrasound score in daily rheumatologic practice: a pilot project. Arthritis Rheum 61(9):1194–1201 **(PubMed PMID: 19714611. Epub 2009/08/29. Eng)**

15. Naredo E, Rodriguez M, Campos C, Rodriguez-Heredia JM, Medina JA, Giner E et al (2008) Validity, reproducibility, and responsiveness of a twelve-joint simplified power doppler ultrasonographic assessment of joint inflammation in rheumatoid arthritis. Arthritis Rheum 59(4):515–522 **(PubMed PMID: 18383408. Epub 2008/04/03. Eng)**

16. Hameed B, Pilcher J, Heron C, Kiely PD (2008) The relation between composite ultrasound measures and the DAS28 score, its components and acute phase markers in adult RA. Rheumatology (Oxford) 47(4):476–480 **(PubMed PMID: 18281367. Epub 2008/02/19. Eng)**

17. Mandl P, Naredo E, Wakefield RJ, Conaghan PG, D'Agostino MA (2011) A systematic literature review analysis of ultrasound joint count and scoring systems to assess synovitis in rheumatoid arthritis according to the OMERACT filter. J Rheumatol 38(9):2055–2062 **(PubMed PMID: 21885517. Epub 2011/09/03. Eng)**

18. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO 3rd et al (2010) Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheum 62(9):2569–2581 **(PubMed PMID: 20872595. Epub 2010/09/28. Eng)**

19. D'Agostino MA, Wakefield RJ, Berner-Hammer H, Vittecoq O, Filippou G, Balint P et al (2016) Value of ultrasonography as a marker of early response to abatacept in patients with rheumatoid arthritis and an inadequate response to methotrexate: results from the APPRAISE study. Ann Rheum Dis 75(10):1763–1769 **(PubMed PMID: 26590174. Pubmed Central PMCID: 5036216. Epub 2015/11/22. Eng)**

20. van Riel PL, van Gestel AM (2000) Clinical outcome measures in rheumatoid arthritis. Ann Rheum Dis 59(Suppl 1):i28–i31 **(PubMed PMID: 11053082. Pubmed Central PMCID: PMC1766622)**

21. Zammurrad S, Munir W, Farooqi A (2013) Disease activity score in rheumatoid arthritis with or without secondary fibromyalgia. J Coll Physicians Surg Pak 23(6):413–417 **(PubMed PMID: 23763802)**

22. Leeb BF, Haindl PM, Maktari A, Nothnagl T, Rintelen B (2007) Disease activity score-28 values differ considerably depending on patient's pain perception and sex. J Rheumatol 34(12):2382–2387 **(PubMed PMID: 17985407)**

23. Balsa A, Carmona L, Gonzalez-Alvaro I, Belmonte MA, Tena X, Sanmarti R et al (2004) Value of Disease Activity Score 28 (DAS28) and DAS28-3 compared to American College of Rheumatology-defined remission in rheumatoid arthritis. J Rheumatol 31(1):40–46 **(PubMed PMID: 14705217)**

24. Naredo E, Collado P, Cruz A, Palop MJ, Cabero F, Richi P et al (2007) Longitudinal power Doppler ultrasonographic assessment of joint inflammatory activity in early rheumatoid arthritis: predictive value in disease activity and radiologic progression. Arthritis Rheum 57(1):116–124 **(PubMed PMID: 17266071)**

25. Backhaus TM, Ohrndorf S, Kellner H, Strunk J, Hartung W, Sattler H et al (2013) The US7 score is sensitive to change in a large cohort of patients with rheumatoid arthritis over 12 months of therapy. Ann Rheum Dis 72(7):1163–1169 **(PubMed PMID: 22956596. Pubmed Central PMCID: PMC3686255)**

26. Ohrndorf S, Backhaus M (2013) Advances in sonographic scoring of rheumatoid arthritis. Ann Rheum Dis 72(Suppl 2):ii69–ii75 **(PubMed PMID: 23253922)**

27. Yoshimi R, Ihata A, Kunishita Y, Kishimoto D, Kamiyama R, Minegishi K et al (2015) A novel 8-joint ultrasound score is useful in daily practice for rheumatoid arthritis. Mod Rheumatol 25(3):379–385 **(PubMed PMID: 25401228)**

28. Leng XF, Zhu Y, Wang GP, Jin J, Xian L, Zhang YH (2016) Accuracy of ultrasound for the diagnosis of cervical lymph node metastasis in esophageal cancer: a systematic review and meta-analysis. J Thorac Dis 8(8):2146–2157 **(PubMed PMID: 27621871. Pubmed Central PMCID: PMC4999759)**

29. Dougados M, Jousse-Joulin S, Mistretta F, d'Agostino MA, Backhaus M, Bentin J et al (2010) 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 69(5):828–833 **(PubMed PMID: 19740905)**

30. Hartung W, Kellner H, Strunk J, Sattler H, Schmidt WA, Ehrenstein B et al (2012) Development and evaluation of a novel ultrasound score for large joints in rheumatoid arthritis: one year of experience in daily clinical practice. Arthritis Care Res 64(5):675–682 **(PubMed PMID: 22183834)**

31. Boesen M, Ellegaard K, Boesen L, Cimmino MA, Jensen PS, Terslev L et al (2012) Ultrasound Doppler score correlates with OMERACT RAMRIS bone marrow oedema and synovitis score in the wrist joint of patients with rheumatoid arthritis. Ultraschall Med 33(7):E166–E172 **(PubMed PMID: 21259184)**

32. de Miguel E, Pecondon-Espanol A, Castano-Sanchez M, Corrales A, Gutierrez-Polo R, Rodriguez-Gomez M et al (2017) A reduced 12-joint ultrasound examination predicts lack of X-ray progression better than clinical remission criteria in patients with rheumatoid arthritis. Rheumatol Int 37(8):1347–1356 **(PubMed PMID: 28389854)**

33. Damjanov N, Radunovic G, Prodanovic S, Vukovic V, Milic V, Simic Pasalic K et al (2012) Construct validity and reliability of ultrasound disease activity score in assessing joint inflammation in RA: comparison with DAS-28. Rheumatology (Oxford) 51(1):120–128 **(PubMed PMID: 22072084. Epub 2011/11/11. Eng)**

34. Han J, Geng Y, Deng X, Zhang Z (2016) Subclinical synovitis assessed by ultrasound predicts flare and progressive bone erosion in rheumatoid arthritis patients with clinical remission: a

systematic review and metaanalysis. J Rheumatol 43(11):2010–2018 (PubMed PMID: 27803342)

35. Nguyen H, Ruyssen-Witrand A, Gandjbakhch F, Constantin A, Foltz V, Cantagrel A (2014) Prevalence of ultrasound-detected residual synovitis and risk of relapse and structural progression in rheumatoid arthritis patients in clinical remission: a systematic review and meta-analysis. Rheumatology (Oxford) 53(11):2110–2118 (PubMed PMID: 24929634)

36. Raffeiner B, Grisan E, Botsios C, Stramare R, Rizzo G, Bernardi L et al (2017) Grade and location of power Doppler are predictive of damage progression in rheumatoid arthritis patients in clinical remission by anti-tumour necrosis factor α. Rheumatology (Oxford) 56(8):1320–1325 (PubMed PMID: 28431141)

37. Naredo E, Gamero F, Bonilla G, Uson J, Carmona L, Laffon A (2005) Ultrasonographic assessment of inflammatory activity in rheumatoid arthritis: comparison of extended versus reduced joint evaluation. Clin Exp Rheumatol 23(6):881–884 (PubMed PMID: 16396709)

38. Alcalde M, D'Agostino MA, Bruyn GA, Moller I, Iagnocco A, Wakefield RJ et al (2012) A systematic literature review of US definitions, scoring systems and validity according to the OMER-ACT filter for tendon lesion in RA and other inflammatory joint diseases. Rheumatology (Oxford) 51(7):1246–1260 (PubMed PMID: 22378717)

39. Aga AB, Berner Hammer H, Christoffer Olsen I, Uhlig T, Kvien TK, van der Heijde D et al (2016) Development of a feasible and responsive ultrasound inflammation score for rheumatoid arthritis through a data-driven approach. RMD Open 2(2):e000325 (PubMed PMID: 28074154. Pubmed Central PMCID: PMC5174791)

40. Hammer HB, Kvien TK (2011) Ultrasonography shows significant improvement in wrist and ankle tenosynovitis in rheumatoid arthritis patients treated with adalimumab. Scand J Rheumatol 40(3):178–182 (PubMed PMID: 21091275)

41. Salaffi F, Di Carlo M, Iannone F, Fedele AL, Epis OM, Pellerito R et al (2018) The UltraSound-CLinical ARthritis Activity (US-CLARA) index: properties of a new composite disease activity index for rheumatoid arthritis. Semin Arthritis Rheum 47(5):619–629 (PubMed PMID: 29102157)

Affiliations

Juan José de Agustín[1] · Alba Erra[2] · Andrés Ponce[3] · Carmen Moragues[4,5] · Cesar Díaz-Torné[5] · Delia Reina[6] · Estefanía Moreno[7] · Julio Ramírez[8] · Lourdes Mateo[9] · Manel Pujol Busquets[10] · Patricia Moya[5] · Pilar Santo-Panero[11] · Sergi Ros-Expósito[12] · Javier Narváez[13] · Raimon Sanmartí[8] · Mª Jesús García de Yébenes[14] · Loreto Carmona[14]

Alba Erra
aerra@hsrafael.com

Andrés Ponce
aponce@fhag.es

Carmen Moragues
cmoraguesp@gmail.com

Cesar Díaz-Torné
cesardiaztorne@gmail.com

Delia Reina
deliareinasanz@gmail.com

Estefanía Moreno
esmoru@yahoo.es

Julio Ramírez
julramga@gmail.com

Lourdes Mateo
lmateo.germanstrias@gencat.cat

Manel Pujol Busquets
manelpuj@gmail.com

Patricia Moya
PMoyaA@santpau.cat

Pilar Santo-Panero
pili_santo@yahoo.es

Sergi Ros-Expósito
28243sre@comb.cat

Javier Narváez
fjnarvaez@bellvitgehospital.cat

Raimon Sanmartí
SANMARTI@clinic.cat

Mª Jesús García de Yébenes
mjgarciadeyebenes@inmusc.eu

Loreto Carmona
loreto.carmona@inmusc.eu

1   Servicio de Reumatología, Hospital Universitari Vall d´Hebron, Passeig de la Vall d'Hebron, 119-129, 08035 Barcelona, Spain

2   Hospital Sant Rafael, Barcelona, Spain

3   Hospital de Granollers, Granollers, Spain

4   Hospital Plató, Barcelona, Spain

5   Hospital Sant Pau, Barcelona, Spain

6   Hospital Moisés Broggi, Sant Joan Despí, Spain

7   Hospital Comarcal de l'Alt Penedès, Vilafranca, Spain

8   Hospital Clinic, Barcelona, Spain

9   Hospital Germans Trias i Pujol, Badalona, Spain

10 Hospital Universitari Mútua Terrassa, Terrassa, Spain

11 Parc Sanitari Sant Joan de Déu, Sant Boi de Llobregat, Spain

12 Hospital de Viladecans, Viladecans, Spain

13 Hospital Universitari de Bellvitge, L'Hospitalet de Llobregat, Spain

14 Instituto de Salud Musculoesquelética (InMusc), Madrid, Spain