

Absence of correlation between DAS28-ESR, DAS28-CRP, SDAI, CDAI and ultrasound scoring in patients with rheumatoid arthritis in remission

Objective: Several composite scores (28-joints Disease Activity Score (DAS28), Simplified/Clinical Disease Activity Index (SDAI/CDAI)) may be used as target of therapeutic strategies in rheumatoid arthritis (RA) management. Although ultrasound (US) is a useful imaging technique in RA management, there are no validated US remission criteria. We aimed at analyzing correlations between US and clinical scores in RA patients with clinical remission.

Methods: This French multicenter cross-sectional study, in 11 Rheumatology departments, consecutively included patients from August 2015 to April 2016. Inclusions criteria were: RA diagnosis meeting ACR-EULAR 2010 criteria, <15 years of progression, DAS28-ESR<2.6 for at least 3 months, stable treatment for at least 6 months, including steroids if necessary. Correlations between US scores (Naredo12 B-mode, PD-mode, combined B+PD mode, PDUS scores) and clinical remission composite scores (DAS28-ESR (Erythrocyte Sedimentation Rate), DAS28-CRP (C Reactive Protein), SDAI, CDAI) were analyzed.

Results: 225 patients were analyzed. Intra/inter-observer reliabilities were good to excellent. Correlations between clinical and US Naredo12 scores were weak to moderate, with correlation coefficients ranging from 0.086 [-0.05; 0.21] to 0.323 [0.20; 0.44]. Best correlations were observed between PD-mode and SDAI or CDAI. US Naredo12 PD-mode score of 0/36 predicts a SDAI remission with 78% sensitivity and 49% specificity (area under curve 0.65). No association was found between US Naredo12 scores and type of treatment, remission duration or disease duration.

Conclusions: There is no strong correlation between US and clinical scores in remission RA patients in routine care. US input in the definition of RA remission remains uncertain.

Highlights:

- Ultrasound may have a role in the definition of remission, since there are some discrepancies between outcome measures of disease activity in rheumatoid arthritis patients in remission.
- Moderate correlations between clinical and ultrasound scores were shown in remission rheumatoid arthritis patients.
- Ultrasound Naredo12 PowerDoppler-mode score of 0/36 predicts a Simplified Disease Activity Index (SDAI) remission with 78% sensitivity and 49% specificity (area under curve 0.65).
- Correlation between clinical and ultrasound scores is independent of rheumatoid arthritis duration, remission duration, and type of treatment.

Keywords: rheumatoid arthritis • remission • ultrasonography

Introduction

Rheumatoid Arthritis (RA) is the most common inflammatory arthritis [1], leading to joint damage, degradation of quality of life, disability, and decrease of life expectancy [2]. Hence, RA dramatically increases the burden of medical health care systems. Clinical remission has become a realistic goal in RA management [3]. Several composite scores (28-joints Disease Activity Score (DAS28), Simplified Disease Activity Index (SDAI), Clinical Disease

Activity Index (CDAI)) are available as a target of therapeutic strategy [4]. These several outcome measures might be used with different limitations. Early diagnosis, prompt treatment initiation, and early achievement of remission provide a better chance to optimize clinical and radiographic outcomes [3].

Over the last few years, ultrasound (US) has become essential in monitoring RA patients [5-8], with a strong validity and added value over clinical and radiographic assessment. Semi-

**Mélanie Trabelsi^{1*},
Marion Cugnet²,
Athar Baillet¹, Marine Clay¹,
Stephanie Rouanet³,
Elise Contant², Muriel Piperno⁴,
Florence Merle-Vincent⁴,
Deborah Gensburger⁵,
Claire Vidon⁶,
Hubert Marotte⁷,
Béatrice Pallot-Prades⁷,
Delphine Denarié⁷,
Adamah Amouzougan⁷,
Anne Cherasse⁸,
Laurent Grimault⁸,
Clémentine Fortunet⁹,
Marie Bossert¹⁰,
Xavier Guillot¹¹,
Isabelle Brault¹¹, Benoit
Augé¹¹, Elodie Constant¹²,
Philippe Gaudin¹ &
Paul Ornetti²**

¹Department of Rheumatology, Univ Grenoble Alpes, GREPI-UGA EA7408, France

²Department of Rheumatology, CHU Dijon, INSERM 1093, University of Burgundy, France

³StathEthic, Levallois-Perret, France

⁴Department of Rheumatology, CHU Lyon Sud, France

⁵Department of Rheumatology, CHU Lyon Edouard Herriot, France

⁶Department of Rheumatology, CH Aix-les-Bains, France

⁷Department of Rheumatology, CHU Saint-Etienne, France

⁸Department of Rheumatology, CH Macon, France

⁹Department of Rheumatology, CH Chalon-sur-Saone, France

¹⁰Department of Rheumatology, CH Belfort, France

¹¹Department of Rheumatology, CH Besançon, France

¹²Department of Rheumatology, CH Valence, France

*Author for correspondence:
melaniefrancois9@gmail.com

quantitatively grey-scale assessment (B-mode), and PD-mode (Power-Doppler) predict joint damage progression [9]. In patients with RA remission, subclinical synovitis, which has been related to silent structural damage progression [10], can be detected with US, especially in deep joint. As discrepancies remain between clinical outcome measures in RA patients in remission [4], US may have a role in the definition of remission. For instance, US scores such as the US Naredo 12 joints score, have been evaluated with substantial validity, reliability, and predictive value of PD-mode on joint damage in RA patients with clinical remission or in biological disease-modifying anti-rheumatic drugs (bDMARD) tapering [11]. But the relevant variability of PD-mode signal makes it very difficult to elaborate a reliable definition of PD-mode remission [12].

Huge endeavors have been gathered to validate US in rheumatic diseases. Outcome Measures in Rheumatology (OMERACT) US group has also improved standardization of US scanning technique and definitions of abnormalities, in order to overcome the intrinsic operator-dependent nature of this imaging modality [13]. However, there are currently no validated US remission criteria, because of uncertainty on the signification of persistent synovial hypertrophy in B-mode or PD-mode in RA patients in clinical remission [14]. Studies assessing the correlation between US data and clinical evaluation have suggested either strong [15] or poor correlations [16], showing contradictory result. However, few of these studies concerned RA patients in remission. It was thus necessary to evaluate correlations between 4 clinical scores (DAS28-ESR (Erythrocyte Sedimentation Rate), DAS28-CRP (C Reactive Protein), SDAI, CDAI) and US Naredo 12 joints scores (B-mode, PD-mode, combined B+PD mode, or PDUS) in RA patients in remission. We also characterized predictive factors for US remission.

Methods

Study population

This French multicenter cross-sectional study took place in 11 Rheumatology departments (Grenoble, Dijon, Lyon HEH, Lyon Sud, Aix-les-Bains, Saint-Etienne, Macon, Chalon-sur-Saône, Belfort, Besançon, Valence), from August 2015 to April 2016. Inclusions criteria were:

- Diagnosis of RA meeting ACR-EULAR 2010 criteria,
- Less than 15 years of progression

- DAS28-ESR < 2.6 for at least 3 months,
- With a stable treatment (conventional synthetic DMARD (csDMARD), or bDMARD, or association of both) for at least 6 months, including steroids if necessary (equivalent prednisone \leq 0.1mg/kg/day) without steroid pulse or intra-articular steroid injection for 3 months.

Clinical data

Clinical data listed were : age, gender, weight, height, Body Mass Index (BMI), duration of RA, duration of remission, presence of AntiCitrullined Proteins Antibodies (ACPA), presence of Rheumatoid Factors (RF), presence of erosions, number of tender joints, number of swollen joints, patient global health visual analogic scale (mm), physician global health visual analogic scale (mm), ESR (mm/h), CRP (mg/L), type of treatment (csDMARD, bDMARD, oral steroids, nonsteroidal anti-inflammatory drugs NSAID). Clinical scores were calculated, with a remission defined as: DAS28 < 2.6, CDAI \leq 2.8, SDAI \leq 3.3, and collected before US scoring, which was fulfilled on the same day. Oral and written information about the study design were given to patients and they had to sign an informed consent before being included. The study design was based on ADEL (Association Des Epidémiologistes de Langue Française) recommendations, conducted in accordance with the Helsinki Declaration and approved by the CNIL (Comission Nationale de l'Informatique et des Libertés) with authorization number 915012. Clinical evaluation was blinded to US data.

US scoring

Standardized US machines (Esaote MyLab60, Esaote MyLab70 XVG, Hitachi Aloka Hi Vision Avius, GE Logiq-e, and Logiq-e LogiqS7) were used with linear probe (12-18MHz in superficial joints, 5-12MHz in deep joints). Because of its validity, reproducibility and responsiveness to change, the Naredo score [15] was used to assess US RA activity in B and PD-mode. An experienced ultrasonographer, blinded to clinical data, examined 12 joints (elbows, wrists, 2nd and 3rd metacarpo-phalangeals, knees, ankles), with a semi-quantitatively scale going from 0 to 3 (where 0=absence, 1=mild, 2=moderate, and 3=marked). US Naredo 12 scores in B-mode (0-36), PD-mode (Power-Doppler (0-36)), combined B+PD-mode (total of B-mode and PD-mode (0-72)) and PDUS

(highest value between B-mode or PD-mode (0-36)) were calculated. Doppler setting was individually adjusted for each patient, according to the ultrasonographer. US evaluation was blinded to clinical data.

Statistics

US scoring reliability

We analyzed inter-observer reliability in 17 experienced rheumatologists who took part of this study, by Cohen Kappa coefficient (k) and weighted k , which were considered weak for $k \leq 0.4$, moderate for $0.4 < k \leq 0.6$, good for $0.6 < k \leq 0.8$ and excellent for $k > 0.8$. 49 B-mode and 51 PD-mode imaging (wrists and hands) from a data base were assessed. A total of 20 B-mode and 19 PD-mode US scan images were similar in order to determine intra-observer reliability of the 17 rheumatologists.

Primary outcome measure

Correlation between clinical and US Naredo12 scores coefficients (95% confidence interval) were determined by Spearman's test. Correlation coefficients were interpreted as weak (< 0.3), moderate ($\geq 0.3, < 0.5$) or strong (> 0.7) [17].

The Youden Index was used as a summary measure of the receiver operating characteristic (ROC) curve in order to determine the optimal threshold of US PD-mode to predict SDAI remission.

Secondary outcome measures

The impact of disease duration, duration of remission or treatments (csDMARD and/or bDMARD) on US Naredo12 scores was also assessed by Kruskal-Wallis's test. Quantitative data were expressed as mean (standard deviation) or median (interquartile range) and qualitative data as number and percentages. Missing data had not been handled. All tests were two-sided with α risk at 5%. Statistical analyses were performed using SAS version 9.4 (SAS Institute Inc. Cary, NC, USA).

Results

A total of 238 patients were consecutively included and 225 patients were analyzed. 13 patients could not be analyzed because they were actually not in clinical remission on examination.

68.4% of patients were women. Mean \pm Standard Deviation (SD) age, RA duration and duration of remission were respectively 58.6 ± 12.4 years, 6 ± 3.7 years and 20.8 ± 19.4 months. Mean

\pm SD DAS28-ESR, DAS28-CRP, CDAI and SDAI were respectively 1.7 ± 0.5 , 1.7 ± 0.5 ; 2.6 ± 2.3 , 2.9 ± 2.4 .

A total of 196 (87.1%) patients were treated with csDMARD (182 patients (92.9%) were taking Methotrexate) and 118 (52.4%) patients were treated with bDMARD (89 patients (75.4%) were taking bDMARD with csDMARD). A total of 3.4% of patients were treated with Infliximab, 17.8% Etanercept, 8.5% Adalimumab, 4.2% Golimumab, 5.9% Certolizumab, 30.5% Tocilizumab, 19.5% Abatacept, 9.3% Rituximab, and 0.8% Anakinra. Steroids were used for 9.8% of patients and NSAID were used for 4.5% of patients (Table 1).

US scoring reliability

B-mode intra-observer reliability (mean \pm SD) was good to excellent ($k=0.77 \pm 0.16$ and weighted $k=0.87 \pm 0.10$); PD-mode intra-observer reliability was excellent ($k=0.97 \pm 0.06$ and weighted $k=0.99 \pm 0.03$). B-mode inter-observer reliability was modest with k ranging from 0.24 ± 0.12 to 0.56 ± 0.15 ; PD-mode inter-observer reliability was excellent with k ranging from 0.60 ± 0.06 to 0.82 ± 0.10 . Results were better with weighted k ranging from 0.49 ± 0.12 to 0.74 ± 0.11 in B-mode, and weighted k ranging from 0.91 ± 0.05 to 0.78 ± 0.06 in PD-mode.

Because of low reliability (particularly in B-mode intra/inter-observer reliability) 5 rheumatologists performed an improvement session before the study onset, in order to optimize reliability of US analysis as already shown by D'Agostino et al. [18].

Primary outcome

Very weak to moderate correlations between clinical scores (DAS28-ESR/CRP, CDAI, SDAI) and US Naredo12 scores were found, with Spearman's correlation coefficient ranging from $0.086 [-0.05; 0.21]$ to $0.323 [0.20; 0.44]$ (Table 2).

Best correlations were observed between US Naredo12 PD-mode and SDAI (Spearman's correlation coefficient= $0.315 [0.19; 0.43]$) or CDAI (Spearman's correlation coefficient= $0.323 [0.20; 0.44]$).

US Naredo12 PD-mode score (threshold 0/36) predicted SDAI remission with a sensitivity of 78% [95% CI: 70.4; 84.1], a specificity of 49% [95% CI: 38.1; 59.5], and an area under the curve of $0.65 [95\% \text{ CI: } 0.58; 0.71]$ (Figure 1).

Table 1. Baseline demographics.	
	n = 225 patients
Gender - (n, %)	Men : 71 (31.6%) Women : 154 (68.4%)
Age (years) - mean (SD)	58.6 (12.4)
BMI (kg/m ²) - mean (SD)	25.1 (4.8)
RA duration (years) - mean (SD)	6 (3.7)
Duration of remission (months) - mean (SD)	20.8 (19.4)
ACPA positive - (n, %)	154 (71.3%)
RF positive - (n, %)	149 (70.6%)
Erosive RA - (n, %)	112 (52.8%)
DAS28 ESR - mean (SD)	1.7 (0.5)
DAS28 CRP - mean (SD)	1.7 (0.5)
Tender joints	0.3 (0.7)
Swollen joints	0.2 (0.6)
Patient global health (mm, 0-100)	12.9 (13.7)
Physician global health (mm, 0-100)	8.4 (9)
ESR (mm/h)	9.2 (6.7)
CRP (mg/L)	2.8 (3.3)
CDAI - mean (SD)	2.6 (2.3)
CDAI - remission ≤2.8 (n, %)	127 (56.4%)
SDAI - mean (MD)	2.9 (2.4)
SDAI - remission ≤3.3 (n, %)	141 (63.8%)
csDMARD - (n, %)	196 (87.1%)
Methotrexate - (n, %DMARD)	182 (92.9%)
bDMARD - (n, %)	118 (52.4%)
Infliximab	4 (3.4%)
Etanercept	21 (17.8%)
Adalimumab	10 (8.5%)
Golimumab	5 (4.2%)
Certolizumab	7 (5.9%)
Tocilizumab	36 (30.5%)
Abatacept	23 (19.5%)
Rituximab	11 (9.3%)
Anakinra	1 (0.8%)
Monotherapy	29 (24.6%)
Association csDMARD + bDMARD	89 (75.4%)
Oral steroids - (n, %)	22 (9.8%)
Steroid dose (mg/day PREDNISONE equivalent) - mean (SD)	3.4 (1.5)
NSAID - (n,%)	10 (4.5%)

BMI: Body Mass Index; RA: Rheumatoid Arthritis; ACPA: Anti-Citrullinated Protein Antibodies; RF: Rheumatoid Factor; DAS28: Disease Activity Score; ESR: Erythrocyte Sedimentation Rate; CRP: C-Reactive Protein; CDAI: Clinical Disease Activity Index; SDAI: Simplified Disease Activity Index; csDMARD: conventional synthetic Disease-Modifying Anti Rheumatic Drugs; bDMARD: biological Disease-Modifying Anti Rheumatic Drugs; NSAID: Non Steroidal Anti Inflammatory Drug; SD: Standard Deviation

Table 2. Correlation between clinical and US Naredo12 scores (n=225).				
	DAS28-ESR	DAS28-CRP	CDAI	SDAI
(B+D)-Mode	0.113 [-0.02; 0.24]	0.281 [0.15; 0.40]	0.251 [0.12; 0.37]	0.273 [0.15; 0.39]
B-Mode	0.086 [-0.05; 0.21]	0.253 [0.12; 0.37]	0.202 [0.07; 0.32]	0.228 [0.1; 0.35]
PD-Mode	0.192 [0.06; 0.32]	0.229 [0.10; 0.35]	0.323 [0.20; 0.44]	0.315 [0.19; 0.43]
PDUS	0.089 [-0.04; 0.22]	0.257 [0.13; 0.38]	0.214 [0.09; 0.34]	0.238 [0.11; 0.36]

Data are expressed as Spearman correlation coefficient [95% CI]
DAS28: 28-joints Disease Activity Score; ESR: Erythrocyte Sedimentation Rate; CRP: C-Reactive Protein; CDAI: Clinical Disease Activity Index; SDAI: Simplified Disease Activity Index

Secondary outcomes

No association was found between US Naredo12 scores and treatments (csDMARD, bDMARD or association). In hindsight, we focused on evaluating correlations between US Naredo12 PD-mode and CDAI or SDAI (which showed best correlations as mentioned above) in each group of treatment. Best correlations were observed in the bDMARD group (Spearman's

correlation coefficient=0.438 [0.085; 0.693] when using CDAI and Spearman's correlation coefficient=0.461 [0.114; 0.708] when using SDAI). Correlations coefficient were not compared to each other but it seems that type of treatment does not impact on the association between US and clinical evaluation (Tables 3 and 4). Similarly, no association was found between US Naredo12 scores and duration of remission, nor duration of the disease (Table 5).

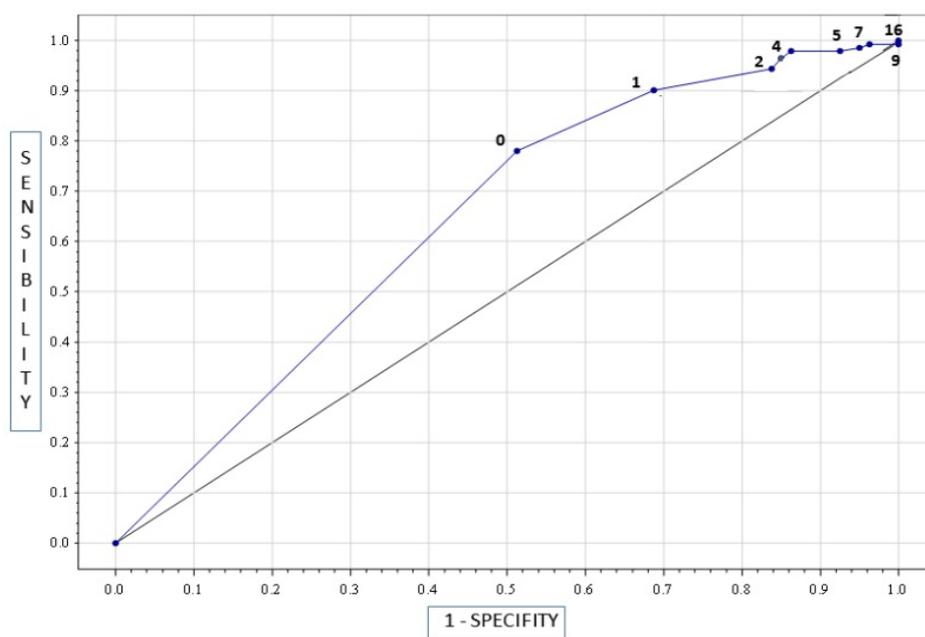


Figure 1. ROC curve of SDAI remission prediction according to US Naredo12 PD-mode.

Table 3. Association between US Naredo12 scores and type of treatment (n = 225).

	csDMARD N=107	bDMARD monotherapy N=29	bDMARD+csDMARD N=89	p
(B+D)-Mode	2.0 [1.0; 5.0]	3.0 [2.0; 5.0]	3.0 [2.0; 5.0]	0.107
B-Mode	2.0 [1.0; 4.0]	3.0 [2.0; 4.0]	3.0 [1.0; 5.0]	0.312
PD-Mode	0.0 [0.0; 0.0]	0.0 [0.0; 1.0]	0.0 [0.0; 1.0]	0.065
PDUS	2.0 [1.0; 4.0]	3.0 [2.0; 4.0]	3.0 [1.0; 5.0]	0.203

Data are expressed as Median [Q1; Q3] - Kruskal-Wallis test
csDMARD: conventional synthetic Disease-Modifying Anti Rheumatic Drugs; bDMARD: biological Disease-Modifying Anti Rheumatic Drugs

Table 4. Association between clinical scores (CDAI and SDAI), and US Naredo12 PD-mode in each treatment group (csDMARD, bDMARD, association) (n = 225).

	csDMARD N=107	bDMARD monotherapy N=29	bDMARD+csDMARD N=89
PD-Mode/CDAI	0.271 [0.086; 0.439]	0.438 [0.085; 0.693]	0.350 [0.153; 0.520]
PD-Mode/SDAI	0.259 [0.072; 0.429]	0.461 [0.114; 0.708]	0.330 [0.127; 0.507]

Data are expressed as Spearman correlation coefficient [95% CI]
bDMARD: biological Disease-Modifying Anti Rheumatic Drugs; csDMARD: conventional synthetic Disease-Modifying Anti Rheumatic Drugs; CDAI: Clinical Disease Activity Index; SDAI: Simplified Disease Activity Index

Table 5. Association between US Naredo12 scores, RA duration and duration of clinical remission (n = 225).

	(B+D)-Mode	PD-Mode	PDUS
Remission duration			
≤ 6 months (n=38)	3.0 [1.0; 5.0]	0.0 [0.0; 1.0]	3.0 [1.0; 5.0]
[6; 12] months (n=53)	3.0 [2.0; 6.0]	0.0 [0.0; 2.0]	3.0 [1.0; 5.0]
>12 months (n=134)	3.0 [1.0; 5.0]	0.0 [0.0; 1.0]	2.0 [1.0; 4.0]
p	0.201	0.124	0.28
RA duration			
≤5 years (n=112)	2.0 [1.0; 5.0]	0.0 [0.0; 1.0]	2.0 [1.0; 4.0]
[5; 10] years (n=91)	3.0 [1.0; 5.0]	0.0 [0.0; 1.0]	3.0 [1.0; 5.0]
>10 years (n=22)	3.5 [2.0; 5.0]	0.5 [0.0; 1.0]	3.0 [1.0; 5.0]
p	0.214	0.146	0.308
Data are expressed as Median [Q1; Q3] – Kruskal-Wallis test			

Discussion

This multicenter cross-sectional study shows an absence of strong correlations between clinical and US scores, in remission RA patients in daily practice, whatever the clinical scores used. US input in remission definition remains modest and we cannot support US as an essential tool in RA remission management. Even though DAS28 is the most widely used tool in everyday practice, CDAI or SDAI are more stringent remission criteria [19] and display better correlations with US Naredo12 PD-mode, enabling to evaluate the residual inflammation.

Discordant studies assessing the correlation between US data and clinical evaluation have suggested either strong correlations (correlation coefficient >0.7 for Kawashiri et al. [20] and >0.5 for Naredo et al. [15]) or poor correlations [16,21]. In our study, clinical evaluation was blinded to US data, and performed on the same day. Secondly, we defined a standardized method of data collection (time, temperature of the room, strict definition of activity scores, US machines) and good intra/inter-observer reliability. Other studies suggested higher US reliability in RA patients [22], but with fewer experts and less joints assessed [16]. Finally, we only included RA patients with relatively short disease duration (i.e. less than 15 years). Here, the mean of RA duration was actually much shorter (6 ± 3.7 years). This allows hurdling degenerative changes on US evaluation, even if sometimes degenerative changes might occur before 6 years of disease progression. In this study US evaluation was performed with a pragmatic and quick method, by rheumatologists in routine care in order to strengthen the external validity without affecting the inter-observer reliability, as already suggested by literature [23]. Recent studies bring

more information about the uncertain role of US in RA management. The TASER study [24] suggested a modest place for US-driven treat to target strategy, which was not associated with significantly better clinical or imaging outcomes than a DAS28-driven strategy. Nevertheless, there was a more frequent DAS44 remission rate in patients with US assessment (66%) compared with patients without US assessment (44%). In the same direction, the ARCTIC study [25], which is an open study, showed no superiority of the US tight control strategy over the clinical tight control strategy, concerning sustained remission and radiographic damage. US assessment may lead to overtreatment in RA patients (e.g. steroids injections, bDMARD, csDMARD changing) which imply side effects. But on the other hand, US findings may also lead to a less aggressive treatment in patients with an evident imaging remission.

Some limits of the internal validity of this study should be reported. Different US machines were used and we know US Naredo12 scores are partially machine-dependent [26]. However, US machines were standardized within all centers. Intra/inter-observer reliability was performed on recorded ultrasonography scan images of hands and wrists. While not assessing independent acquisition of ultrasonography scan image, we might overestimate intra/inter-observer reliabilities as shown by Cheung et al. [22]. Knees and ankles (included in US Naredo12 scoring) were not included in our reliability tests. This is not likely a limitation, as few patients had US ankle synovitis, and many knee synovitis may be due to osteoarthritis. Although US Naredo12 scores are well defined [15], these outcomes remain partly subjective assessments [26], whereas quantitative scores seem little more objective but are more time-consuming

[27]. Last, RA activity was evaluated on 28 joints excluding ankles in clinical scores, whereas only 12 joints excluding shoulders and proximal interphalangeal joints in US Naredo12 scores, which might decrease the correlation. Some studies showed excellent correlation between comprehensive US joints assessment and reduced US joints assessment [16], allowing to assess fewer joints in a reduced time with nearly the same validity in the rheumatologist daily practice. In a systematic review of literature, the minimal number of joints to be included in a global US score is still controversial [28]. New US scores are promising in monitoring RA patients such as the new seven-joint US score (US7) [29], US Global Synovitis Score (US-GLOSS) [30], or 44-joints score by Scirè et al. [31]. The US7 score is the first US sum score system which combines soft tissues (synovitis and tenosynovitis) and destructive lesions (erosions) in a composite scoring system [32]. But opposite to US7 score which assesses only the most clinically affected side, the US Naredo12 score we used assesses bilateral joints. These scores have never been compared to each other, explaining why there is currently no consensus about a valid, reliable and feasible US score assessing and monitoring RA patients.

Some issues regarding the external validity of this study should be raised. Subclinical synovitis may persist upon clinical remission and lead to silent structural damage progression and disease flares. They are mainly detected in US PD-mode which is confirmed by our study [10]. Recently, some studies found that US remission was related to low radiographic structural damage and longer duration of clinical remission [14]. In our study, remission duration barely impacts the correlation between US and clinical evaluation in patients with remission. This could be explained by a lack of information about US activity during the preceding period of evaluation. We cannot exclude recent US flares.

In addition, US has proved its validity for monitoring response to anti-TNF- α therapy in RA [33]. Marks et al. [34] undertook a challenging study on tapering anti-TNF- α therapy. The authors concluded that the combined clinical-US strategy may be useful to optimize the selection of patients for anti-TNF- α dose reduction. Our data suggest that type of treatment does not impact on the correlation between US and clinical evaluation. However, this analysis was beyond the scope of our study [35].

Conclusion

This cross sectional study shows an absence of strong correlations between clinical and US scores, in 225 remission RA patients in routine care. Here, US input in remission definition remains modest. Use of US as a remission criterion in RA patients is still uncertain, as recently shown by Caporali and Smolen. Other studies are needed to evaluate the potential added value of this combined clinical-US assessment of RA remission in routine care and in monitoring RA patients under treatment. Determining a US cut off of remission definition may be a strategic goal in the future.

Disclosure of interest

None of the authors have conflicts of interest to declare.

Acknowledgments

The authors thank the 11 Rheumatology departments and all the investigators who recruited and followed the patients. We also wish to thank all the patients who contributed to this study.

Funding

The authors received an unrestricted grant from ROCHE CHUGAI.

Ethical approval information

Patient consent was obtained, the study was conducted according to the Helsinki Declaration, and was approved by the CNIL (Comission Nationale de l'Informatique et des Libertés): authorization number 915012.

References

1. Ødegård S, Landewé R, van der Heijde D *et al.* Association of early radiographic damage with impaired physical function in rheumatoid arthritis: A ten-year, longitudinal observational study in 238 patients. *Arthritis. Rheum.* 54(1), 68–75 (2006).
2. Minaur NJ, Jacoby RK, Cosh JA *et al.* Outcome after 40 years with rheumatoid arthritis: a prospective study of function, disease activity, and mortality. *J. Rheumatol.* 69, 3–8 (2004).
3. Knevel R, Schoels M, Huizinga TWJ *et al.* Current evidence for a strategic approach to the management of rheumatoid arthritis with disease-modifying antirheumatic drugs: a systematic literature review informing the EULAR recommendations for the management of rheumatoid arthritis. *Ann. Rheum. Dis.* 69(6), 987–994 (2010).
4. Felson DT, Smolen JS, Wells G *et al.* American College of Rheumatology/European League against Rheumatism provisional definition of remission in rheumatoid arthritis for clinical trials. *Ann. Rheum. Dis.* 70(3), 404–413 (2011).
5. D'Agostino MA, Wakefield RJ, Berner-Hammer

- H *et al.* 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 (2016).
6. Zufferey P, Rebell C, Benaim C *et al.* Ultrasound can be useful to predict an evolution towards rheumatoid arthritis in patients with inflammatory polyarthralgia without anticitrullinated antibodies. *Joint. Bone. Spine.* 84(3), 299–303 (2017).
 7. Alcade M, D'Agostino MA, Bruyn GA *et al.* OMERACT Ultrasound Task Force. A systematic literature review of US definitions, scoring systems and validity according to the OMERACT filter for tendon lesion in RA and other inflammatory joint diseases. *Rheumatology.* 51(7), 1246–1260 (2012).
 8. Han J, Geng Y, Deng X *et al.* 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 (2016).
 9. Ohrndorf S, Backhaus M. Pro musculoskeletal ultrasonography in rheumatoid arthritis. *Clin. Exp. Rheumatol.* 33(4), S50–S53 (2015).
 10. Nguyen H, Ruyssen-Witrand A, Gandjbakhch F *et al.* 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. *Rheumatol. Oxf. Engl.* 53(11), 2110–2118 (2014).
 11. Naredo E, Valor L, De la Torre I *et al.* Predictive value of Doppler ultrasound-detected synovitis in relation to failed tapering of biologic therapy in patients with rheumatoid arthritis. *Rheumatology.* 54(8), 1408–1414 (2015).
 12. Mandl P, Naredo E, Wakefield RJ *et al.* 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 (2011).
 13. Bruyn GA, Naredo E, Iagnocco A *et al.* The OMERACT Ultrasound Working Group 10 Years On: Update at OMERACT 12. *J. Rheumatol.* 42(11), 2172–2176 (2015).
 14. Gärtner M, Alasti F, Supp G *et al.* Persistence of subclinical sonographic joint activity in rheumatoid arthritis in sustained clinical remission. *Ann. Rheum. Dis.* 74(11), 2050–2053 (2015).
 15. Naredo E, Rodríguez M, Campos C *et al.* 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 (2008).
 16. Naredo E, Bonilla G, Gamero F *et al.* Assessment of inflammatory activity in rheumatoid arthritis: a comparative study of clinical evaluation with grey scale and power Doppler ultrasonography. *Ann. Rheum. Dis.* 64(3), 375–381 (2005).
 17. Sheskin DJ. Handbook of Parametric and Nonparametric Statistical Procedures: Third Edition. USA: Chapman & Hall/CRC Press. (2003).
 18. D'Agostino MA, Maillefert JF, Said-Nahal R *et al.* Detection of small joint synovitis by ultrasonography: the learning curve of rheumatologists. *Ann. Rheum. Dis.* 63(10), 1284–1287 (2004).
 19. Aletaha D, Smolen JS. The Simplified Disease Activity Index (SDAI) and Clinical Disease Activity Index (CDAI) to monitor patients in standard clinical care. *Best. Pract. Res. Clin. Rheumatol.* 21(4), 663–675 (2007).
 20. Kawashiri S, Kawakami A, Iwamoto N *et al.* The power Doppler ultrasonography score from 24 synovial sites or 6 simplified synovial sites, including the metacarpophalangeal joints, reflects the clinical disease activity and level of serum biomarkers in patients with rheumatoid arthritis. *Rheumatol. Oxf. Engl.* 50(5), 962–965 (2011).
 21. Geng Y, Han J, Deng X *et al.* Presence of power Doppler synovitis in rheumatoid arthritis patients with synthetic and/or biological disease-modifying anti-rheumatic drug-induced clinical remission: experience from a Chinese cohort. *Clin. Rheumatol.* 33(8), 1061–1066 (2014).
 22. Cheung PP, Dougados M, Gossec L. Reliability of ultrasonography to detect synovitis in rheumatoid arthritis: A systematic literature review of 35 studies. *Arthritis. Care. Res.* 62(3), 323–334 (2010).
 23. Rosenberg C, Etchepare F, Fautrel B *et al.* Ultrasound diagnosis of synovitis in rheumatoid arthritis: one year of experience is sufficient to interpret static images. *Rev. Rum.* 76, 39–42 (2009).
 24. Dale J, Stirling A, Zhang R *et al.* Targeting ultrasound remission in early rheumatoid arthritis: the results of the TaSER study, a randomised clinical trial. *Ann. Rheum. Dis.* 75(6), 1043–1050 (2016).
 25. Haavardsholm EA, Aga AB, Olsen IC *et al.* Ultrasound in management of rheumatoid arthritis: ARCTIC randomised controlled strategy trial. *BMJ.* 354, i4205 (2016).

26. Ten Cate DF, Luime JJ, van der Ven M *et al.* Very different performance of the power Doppler modalities of several ultrasound machines ascertained by a microvessel flow phantom. *Arthritis. Res. Ther.* 15(5), R162 (2013).
27. Schmidt WA, Schönau V, Reiche BE *et al.* Grading of ultrasound Doppler signals in synovitis: does it need an update? *Rheumatol. Oxf. Engl.* 54(10), 1897–1903 (2015).
28. Terslev L, Ellegaard K, Christensen R *et al.* Head-to-head comparison of quantitative and semi-quantitative ultrasound scoring systems for rheumatoid arthritis: reliability, agreement and construct validity. *Rheumatology.* 51(11), 2034–2038 (2012).
29. Backhaus TM, Ohrndorf S, Kellner H *et al.* 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 (2013).
30. Iagnocco A, Naredo E, Wakefield R *et al.* Responsiveness in rheumatoid arthritis. a report from the OMERACT 11 ultrasound workshop. *J. Rheumatol.* 41(2), 379–382 (2014).
31. Scirè CA, Montecucco C, Codullo V *et al.* Ultrasonographic evaluation of joint involvement in early rheumatoid arthritis in clinical remission: power Doppler signal predicts short-term relapse. *Rheumatol Oxf. Engl.* 48(9), 1092–1097 (2009).
32. Ohrndorf S, Backhaus M. Advances in sonographic scoring of rheumatoid arthritis. *Ann. Rheum. Dis.* 72 (Suppl 2), ii69–75 (2013).
33. Iagnocco A, Finucci A, Ceccarelli F *et al.* Power Doppler ultrasound monitoring of response to anti-tumour necrosis factor alpha treatment in patients with rheumatoid arthritis. *Rheumatology.* 54(10), 1890–1896 (2015).
34. Marks JL, Holroyd CR, Dimitrov BD *et al.* Does combined clinical and ultrasound assessment allow selection of individuals with rheumatoid arthritis for sustained reduction of anti-tumor necrosis factor therapy? *Arthritis. Care. Res.* 67(6), 746–753 (2015).
35. Caporali R, Smolen JS. Back to the future: forget ultrasound and focus on clinical assessment in rheumatoid arthritis management. *Ann. Rheum. Dis.* 77(1), 18–20 (2018).