Musculoskeletal ultrasound to complement clinical evaluation and drive treatment of rheumatoid arthritis patients in remission does not prevent worsening of patient-reported outcomes: the ULTRAPRO randomised controlled study

C. Sifuentes-Cantú[1], D. Butrón-Hernández[2], I. Contreras-Yáñez[1], V. Pascual-Ramos[1]

[1]Department of Immunology and Rheumatology; [2]Department of Radiology, Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, Mexico City, Mexico.

Abstract

Objective

The primary objective was to determine the impact of sharing musculoskeletal ultrasound (MUS) results with rheumatologists on worsening patient-reported outcomes (PROs) at 6 months of follow-up in rheumatoid arthritis (RA) patients with clinical remission. Secondary objectives were to describe MUS findings and to compare the proportion of patients with flares, according to the DAS28-ESR, following the intervention.

Methods

Ninety-four consecutive outpatients with clinical remission had PROs and a treatment proposal recorded at study entry. MUS was then performed by trained specialists who were blinded to clinical assessments. Forty-seven patients were randomised (1:1) to either the intervention group (MUS data shared with the primary rheumatologist) or the control group (data not shared); changes in the treatment proposal were recorded. PROs worsening and the proportion of patients with flares were compared between both groups at 6±2 months of follow-up. The study received IRB approval. Appropriate statistics were used.

Results

At baseline, patients from the intervention and control groups had similar characteristics; 43 and 41 patients, respectively, completed the 6-month follow-up period. PROs worsening at 6 months of follow-up were similar between groups, as were the DAS28-ESR and the proportion of patients who flared. In general, MUS findings were in accordance with the clinical remission status, although power Doppler synovitis was detected in up to 37% of the patients. RA-related treatment was increased in all the patients from the intervention group with discordant findings between clinical and MUS assessments.

Conclusion

The addition of MUS to clinical evaluation of RA outpatients in remission did not prevent worsening PROs at 6 months.

Key words

musculoskeletal ultrasound, rheumatoid arthritis, patient-reported outcomes

César Sifuentes-Cantu, MD, MSc
David Butrón-Hernández, MD
Irazú Contreras-Yánez, SW, MSc
Virginia Pascual-Ramos, MD, Bioethics MSc

Please address correspondence to:
Virginia Pascual-Ramos,
Department of Immunology and Rheumatology,
Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán,
Vasco de Quiroga 15, colonia sección XVI,
Belisario Domínguez,
Tlalpan 14080, México City, México.
E-mail: virtichu@gmail.com

Received on September 6, 2019; accepted in revised form on January 28, 2020.

Competing interests: none declared.

Introduction

Disease activity (DA) is a central aspect in the evaluation of rheumatoid arthritis (RA) patients and can be assessed through different tools, such as joint counts, acute reactant phase determinations and validated indices such as the Disease Activity Scores on 28 joints (DAS28) (1).

Current guidelines recommend the treat-to-target strategy (T2T), which sets the goal of remission or low DA, because either target enhances clinical and quality-of-life outcomes (2). Rates of remission have varied upon the criteria used, although the majority of the rheumatologists agree that composite measures better define remission (2, 3). Nonetheless, these indices have limitations; patients with painful comorbidities can exhibit higher scores that do not reflect active synovitis; also, a considerable number of patients who achieved the T2T goal continue to show structural damage and functional impairment, which has been explained by "subclinical disease activity" (4, 5). Based on these considerations, it seems necessary to identify tools to define a true remission status.

Musculoskeletal ultrasound (MUS) can be easily incorporated into practice, it is more sensitive and reliable than clinical examination to improve the accuracy of detecting the level of DA, and it is better correlated with joint damage than the medical assessment is (6, 7). Nonetheless, in real clinical settings, MUS findings modified treatment in about 20% to 37% of the clinical scenarios (8-10). Two randomised clinical trials demonstrated contradictory results in terms of reducing radiographic progression and in the percentage of patients who achieved clinical remission at the end of their follow-up. The TaSER study, a MUS-driven T2T strategy, led to more intensive treatment but was not associated with significantly better clinical or radiographic outcomes than a DAS28-driven strategy (11); meanwhile, in the ARCTIC trial, authors found a very subtle difference between groups in terms of radiographic progression after 24 months of follow-up that favoured the MUS-driven therapy (12).

Currently, there is an increasing focus on patient-centered care and patient-reported outcomes (PROs) emerge as both time- and cost-efficient tools for monitoring chronic diseases (13). PROs add unique information regarding treatment efficacy and quality-of-life outcomes from the patient perspective (14-15) and are predictors of disease progression and mortality (16). The Food and Drug Administration (FDA) and the Outcome Measures in Rheumatology (OMERACT) group recommend PROs measures to define the level of DA and to assess treatment response (17-21).

To date, different studies have assessed the impact of a MUS-driven treatment strategy on physician and radiographic outcomes, but PROs have never been the primary aim evaluated. The primary objective of the study was to determine the impact of sharing MUS results with rheumatologists on PROs, in terms of PROs worsening, in RA patients in clinical remission. Secondary objectives were to describe MUS findings and to compare the proportion of patients with flares, following the intervention. We report herein the results at 6 months of follow-up.

Materials and methods

Study design

The study was a single-centre, blinded, randomised controlled study. Patients and physicians who performed MUS were blinded to the intervention maneuver, which was defined as sharing the MUS findings with the treating rheumatologist.

Participants

The ULTRAPRO study was conducted within the Immunology and Rheumatology Department of the Instituto Nacional De Ciencias Médicas y Nutrición Salvador Zubirán (INCMyN-SZ), a national referral centre for rheumatic diseases in Mexico City, between May 2017 and June 2018.

Patients were eligible if they were aged over 18 years, currently attending the outpatient clinic, had the clinical diagnosis of RA and were in remission. Patients met the remission definition if they were considered in remission according to their primary rheumatologist *and* if RA-related treatment remained unchanged at study entry. Patients were excluded if they had overlapping syndrome.

Prior to patients' enrolment, the study was presented to all the rheumatologists assigned to outpatient clinic (11 senior rheumatologists and 10 trainees in rheumatology) and they agreed to participate.

Interventions

Briefly, the patients included were identified after they received their scheduled rheumatic assessment at the outpatient clinic. Their remission status was established by their primary rheumatologist who additionally indicated a treatment proposal which was to maintain the RA-related treatment. Once the patient was enrolled, outcome measures were also obtained (see paragraph below), and the first treatment proposal was recorded on standardised formats. Then, MUS was performed to all the patients included. After that, patients were randomised 1:1, to either the intervention group (MUS information was shared with the primary rheumatologist) or to the control group (MUS information was not shared). Finally, each primary rheumatologist was able to modify his/her first treatment proposal and treatment modifications were recorded when applicable. At the follow-up visit (6±2 months), PROs were obtained, in addition to the DAS28-ESR, which was used to define flares.

Outcome measures

The primary objective was defined based on worsening of the following PROs: pain score using a visual analogue scale (pain-VAS), patient global assessment of disease using a VAS (PtGA-VAS), disability according to the Health Assessment Questionnaire (HAQ), quality of life based on the Short Form-36 Questionnaire (both components, physical and mental) (SF-36), and disease activity measured with the Rheumatoid Arthritis Disease Activity Index (RADAI), (19-23).

PROs worsening were defined based on published minimal clinically important improvement estimates (19-21, 24); the same values were used to define worsening. Accordingly, for pain-VAS, PtGA-VAS, HAQ and RADAI, worsening was defined as an increase in the score (of at least 20.4 mm, 18.4 mm, 0.375 and 1, respectively), and for the SF-36, a decrease in the score (of at least 7.1 for the physical component and 3.1 for the mental component). In addition, PROs worsening was also assessed as the number (%) of patients who had baseline PROs within normal range but scored 6 months follow-up PROS out of normal range. PROs within the normal range, were defined according to the following cut-offs: pain-VAS ≤30 mm (22), PtGA-VAS ≤20 mm (1), HAQ score ≤0.5 (23), SF-36 physical component score ≥79 and mental component score ≥77 (20), and RADAI score ≤2 (1, 19).

Finally, the proportion of patients with a flare was derived from the DAS28-ESR score; a flare was defined when the DAS28-ESR increased to a score >3.2 and there was an increase in the EULAR DA category (1).

Musculoskeletal ultrasound

MUS was performed by a trained rheumatologist and a radiologist, both blinded to the primary rheumatologist evaluation; the description of the findings was a consensus between these two physicians. The joints were assessed according to the German Ultrasound Score, GUS-7 (25, 26). Synovitis was defined as previously published, and the definition used was included in the most recent EULAR/OMERACT ultrasound definition and quantification system for RA synovitis (27).

A Logiq e GE machine was used with a 10–16 MHz linear probe (SP10-16RS) and the following standardised power Doppler (PD) settings: high frequency (10 MHz, machine preset), pulse repetition frequency (PRF) of 0.8 kHz, low wall filter (machine preset) and gain adjusted to just below the level at which Doppler artifacts appeared beneath the bone.

For each patient, total grey-scale (GS) and PD scores were obtained as the sum of individual joint scores; also, the number of identified erosions/patient was recorded (24). Finally, DA according to MUS was defined if ≥ grade-1 PD activity was detected in at least one joint/area examined; however, in the wrist joint, ≥ grade-2 PD activity was required as suggested (25, 28).

Sample size

The patients included were in clinical remission, and their PROs were expected to be close to normality; accordingly, a sample size of 37 pairs of evaluations was estimated, assuming a difference between the intervention group and the control group of 25% (of the patients with unfavourable PROs), with a 95% two-sided confidence level and 80% power.

Randomisation

We conducted a balanced stratified, randomised study with the use of a computer-generated schedule (www.randomisation.com). The randomisation list was generated by a computer and used to prepare the groups; the list included a consecutive number and the assigned group. A priori, an independent collaborator generated the random allocation sequence; this collaborator always protected the randomisation list. Each time a patient was included, a different investigator consecutively assigned a number to the patient; this number was announced by telephone to the collaborator who guarded the list, who in turn indicated the assigned intervention group.

Blinding

All the patients and the physicians who performed MUS assessments were blinded to the intervention maneuver. To maintain the blind, at baseline, measures of PROs and MUS assessments were performed in all the instances before the intervention assignment was made.

Statistical methods

We performed a descriptive statistical analysis, presenting frequencies for categorical variables and measures of position and dispersion for numerical variables. The Mann-Whitney U-test was used to compare continuous variables. Fisher's exact test was used to compare proportions. Additionally, a sensitivity analysis was performed to explore whether the patients included in the study with remission according to their rheumatologist had remission-low DA based on the DAS28-ESR cut-off (n=74 patients [78%]). The statistical analysis was performed using SPSS IBM V.21.

Ethics

The study was approved by the IRB (reference no.: IRE-1613-15-16-1) and registered in ClinicalTrials.Gov (NCT03228342). Written informed consent was obtained from all the patients.

Results

Participant's flow, recruitment and baseline data (Fig. 1; Table I).

Between May 2017 and June 2018, 110 consecutive RA outpatients in clinical remission were invited to participate; 94 patients were included (10 patients denied and 6 patients did not meet inclusion criteria). In addition to disease diagnosis according to their primary rheumatologist, all the patients met ≥4 ACR 1987 classification criteria (29). For the current report, the follow-up period concluded six months after the last patient was included.

The patients were primarily middle-aged females, had disease-specific autoantibodies and long-standing disease; as expected, the median DAS28-ESR was 2.4, and the majority of the patients (54%) were in remission according to DAS28-ESR cut-off; these characteristics translated to most of the patients having PROs within the normal range. Findings from MUS examination were in accordance with disease characteristics and DA status, but 35 patients (37%) were considered to have MUS activity. Finally, the majority of the patients were on disease-modifying anti-rheumatic drugs (DMARDs), meanwhile few were receiving (low oral doses of) corticosteroids; median DMARDs/patient was 1.

The randomisation process assigned 47 patients to the control group and 47 patients to the intervention group; baseline socio-demographic and disease characteristics, DAS28-ESR, patients with MUS-defined disease activity, PROs and treatment were similar between both groups (Table I). Forty-one patients in the control group (87%) and 43 patients in the intervention group (91%) completed the 6-month follow-up assessments. The baseline data from the 10 patients lost to follow-up did not differ from that of those who completed the 6-month follow-up (data not shown).

Table I. Baseline socio-demographic and disease characteristics, MUS activity and PROs in the population, and comparison of these between the intervention and the control group.

	Population, n=94	Intervention group, n=47	Control group, n=47	*p*-value
Socio-demographic and disease characteristics				
Age, years[1]	49 (39-60)	49.2 (40-60)	49 (37-60)	0.60
Females[2]	87 (92)	45 (95.7)	42 (89.4)	0.23
Patients with RF or ACPA[2]	85 (90)	43 (91.5)	42 (89.4)	0.72
Disease duration, years[1]	11 (5-15)	11 (5-15)	10.5 (5-16.5)	0.84
DAS28-ESR[1]	2.4 (1.8-2.9)	2.5 (1.8-2.9)	2.3 (1.6-3)	0.75
Patients with DAS28-ESR <2.6[2]	51 (54)	25 (53)	26 (55)	0.92
Patients with MUS-disease activity[2]	35 (37)	20 (42)	15 (32)	0.39
PROs				
Pain-VAS (0-100 mm)[1]	13 (1-28)	13 (3-25)	13 (1-35)	0.86
Pain-VAS within NR[2]	74 (79)	40 (85)	34 (72)	0.13
PtGA-VAS (0-100 mm)[1]	13 (1-32)	14 (1-30)	10 (1-34)	0.95
PtGA-VAS within NR[2]	58 (62)	30 (63)	28 (59)	0.67
HAQ score (0-3)[1]	0.4 (0-0.8)	0.5 (0.1-0.8)	0.3 (0-0.8)	0.70
HAQ score within NR[2]	59 (63)	31 (66)	28 (59.6)	0.52
RADAI score (0-10)[1]	1.4 (0.3-2.1)	1.6 (0.4-2.1)	1.0 (0.2-2.1)	0.29
RADAI score within NR[2]	63 (67)	30 (64)	33 (70)	0.66
SF-36, physical component score[1]	70 (54-86)	69 (56-85)	71 (52-86)	0.69
SF-36 physical component score within NR[2]	33 (35)	16 (34)	17 (36)	0.82
SF-36, mental component score[1]	62 (52-77)	61 (53-74)	64 (51-77)	0.74
SF-36 mental component score within NR[2]	24 (25)	10 (21)	14 (29)	0.47
RA-related treatment				
Patients on DMARDs[2]	87 (92)	44 (94)	43 (92)	0.5
Number of DMARDs/patient[1,3]	1 (1-2)	1 (1-2)	1 (1-2)	0.31
Patients on corticosteroids[2]	18 (19)	7 (15)	11 (23)	0.23
Corticosteroids dose[1,3]	5 (5-7.5)	5 (5-7.5)	5 (5-7.5)	0.72

[1]Median (IQR); [2]number (%) of patients; [3]restricted to the patients on DMARDs/corticosteroids.
RF: rheumatoid factor; ACPA: antibodies to cyclic citrullinated peptides; DAS28-ESR: Disease activity score (28 joints evaluated) erythrocyte sedimentation rate; MUS: musculoskeletal ultrasound. VAS: visual analogue scale; NR: normal range; PtGA: Patient global assessment of disease; HAQ: Health Assessment Questionnaire; RADAI: Rheumatoid Arthritis Disease Activity Index; SF-36: Short form 36; RA: rheumatoid arthritis; DMARDs: disease-modifying anti-rheumatic drugs.

Table II. Comparison of 6 months-PROs between patients from the intervention group and patients from the control group.

	Intervention group, n=41	Control group, n=43	*p*-value
Pain-VAS (0-100 mm)[1]	11 (2-31.5)	22 (2-45)	0.30
Patients with pain-VAS worsening[2,*]	6 (14.6)	10 (23.3)	0.31
PtGA-VAS (0-100 mm)[1]	10 (0.5-35)	17 (1-46)	0.41
Patients with PtGA-VAS worsening[2,*]	9 (22)	11 (25.6)	0.69
HAQ score (0-3)[1]	0.25 (0-0.6)	0.5 (0-1)	0.30
Patients with HAQ worsening	4 (9.8)	8 (18.6)	0.24
RADAI score (0-10)[1]	0.8 (0.4-1.9)	1.6 (0.04-3.3)	0.38
Patients with RADAI worsening[2,*]	4 (9.8)	11 (25.6)	0.058
SF-36 PC score[1]	73.1 (59.1-85.1)	66.3 (48.8-86.9)	0.22
Patients with SF-36 PC score worsening[2,*]	10 (24.4)	8 (18.6)	0.51
SF-36 MC score[1]	68 (59-75)	63.7 (52-74)	0.34
Patients with SF-36 MC score worsening[2,*]	12 (29.3)	17 (39.5)	0.32

[1]Median (IQR); [2,*]number (%) of patients, according to minimal clinically important worsening; VAS: visual analogue scale; PtGA: Patient global assessment of disease actvity; HAQ: Health Assessment Questionnaire; RADAI: Rheumatoid Arthritis Disease Activity Index; SF-36=Short form 36; PC: physical component; MC: mental component.

Primary aim: worsening PROs

Table II summarises comparison of 6 months-PROs and of the proportion of patients who deteriorated their PROs, between the intervention group and the control group. Similar results were obtained when scores and percentages were compared, although a tendency

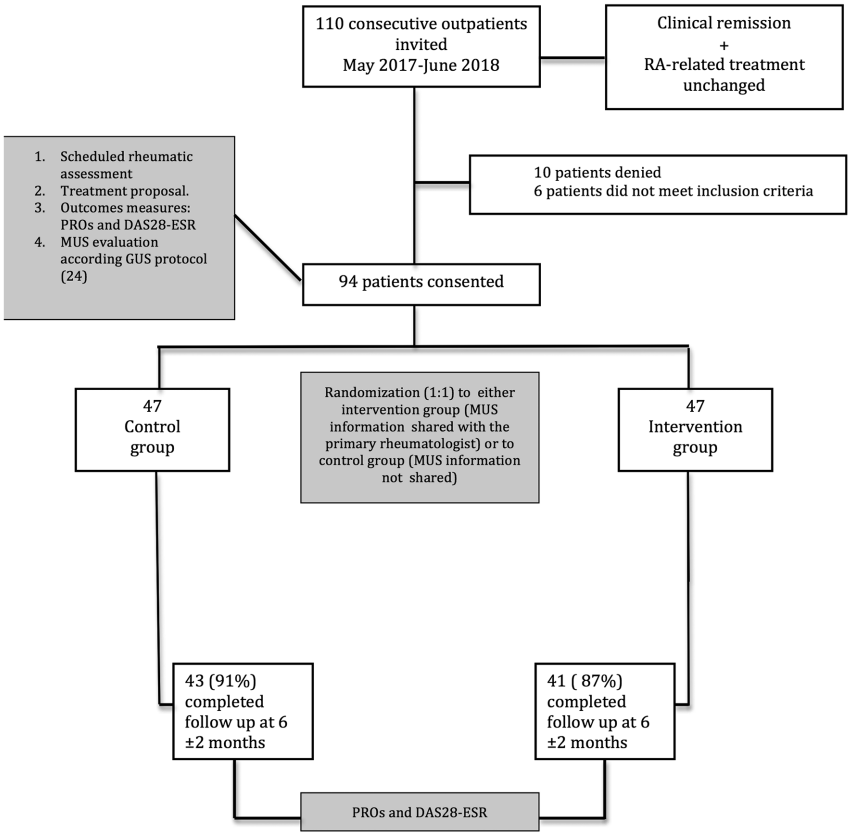


Fig. 1. Study description.

Table III. Comparison of 6-month follow-up PROs, between patients from the intervention group and patients from the control group, in the restricted population of patients with baseline PROs within normal range.

	Intervention group	Control group	*p*-value
Baseline pain-VAS within NR (n=67, 36 in the intervention group)			
Pain-VAS out of NR	8 (22)	9 (29)	0.52
Baseline PtGA-VAS within NR (n=54, 29 in the intervention group)			
PtGA-VAS out of NR	10 (34)	3 (12)	0.07
Baseline HAQ score within NR (n=83, 41 in the intervention group)			
HAQ score out of NR	4 (13.8)	4 (14.8)	0.91
Baseline RADAI score within NR (n=58, 28 in the intervention group)			
RADAI score out of NR	2 (7.1)	8 (26)	0.08
Baseline SF-36 physical component score within NR (n=31, 15 in the intervention group)			
SF-36 physical component score out of NR	5 33)	2 (12.5)	0.16
Baseline SF-36 mental component score within NR (n=22, 10 in the intervention group)			
SF-36 physical component score out of NR	5 (50)	5 (41.7)	1.0

Data are presented as number (%) of patients; VAS: visual analogue scale; NR: normal range; PtGA: Patient global assessment of disease activity; HAQ: Health Assessment Questionnaire; RADAI: Rheumatoid Arthritis Disease Activity Index; SF-36: Short form 36.

was seen regarding the number of patients with RADAI worsening favouring the intervention group (9.8% in the intervention group *vs*. 25.6% in the control group, p=0.058).

The proportion of patients with 6 months-PROs out of normal range was also compared between both groups, in the restricted population of patients with baseline PROS within normal range; results are summarised in Table III; in general, percentage were similar between groups.

Secondary aims:
MUS findings and flares
• *MUS findings*

Table IV summarises (baseline) MUS findings in the whole population and compares ultrasound characteristics between the intervention and the control groups. In general, MUS findings were similar in both groups and were in accordance with the clinical remission status, although PD synovitis was detected in up to 37% of the patients, as were erosions, in up to 60% of the patients.

• *Flares*

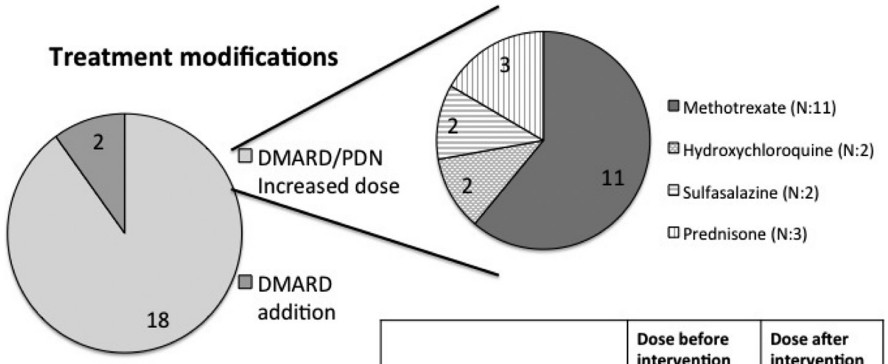
Six months DAS28-ESR were similar in the patients randomised to the intervention and the control groups (median, [IQR] DAS28-ESR of 1.98, [1.2-2.7] *vs*. 2.5, [1.6-3.2], p=0.11), as were the proportion of patients who flared (10 [12.5%] *vs*. 5 [24.4%], p=0.16).

Treatment

As expected, none of the patients allocated to the control group had treatment modifications. Meanwhile, in the intervention group, there was change in the first treatment proposal in 20 patients (42%), and all of them belonged to the category of patients with clinical remission but MUS activity. The RA-related treatment was intensified in the 20 patients referred: in 2 patients, leflunomide was added meanwhile in the 18 patients left, the dose of the background DMARD or the (oral) prednisone was increased, as summarised in Figure 2.

Discussion

The ULTRAPRO study was developed in the outpatient clinic of rheumatic diseases from a tertiary care-level centre. We focused on patients in remission status, as both, sustained remission in the context of recent-onset disease, and low DA in the context of established disease had been proposed



	Dose before intervention	Dose after intervention
Methotrexate (mg/week)	13.6±9.9	20±3.7
Hydroxychloroquine (mg/day)	100±141	300±141
Sulfasalazine (gr/day)	0.7 ±1.0	1.75±1.0
Prednisone (mg/day)	6.6±1.4	11.6±3.8

Fig. 2. Treatment modifications in the 20 patients with clinical remission but MUS activity, assigned to the intervention group.

Table IV. Baseline MUS findings of the population, and comparison of these between the intervention and the control group.

	Population n=94	Intervention, group, n=47	Control group, n=47	*p*-value
Patients with GS[1]	94 (100)	47 (100)	47 (100)	
Patients with GS and PD[1] (=patients with MUS activity)	35 (37)	20 (42)	15 (31.9)	0.28
Patients with tenosynovitis[1]	31 (33)	16 (34)	15 (32)	0.81
Patients with PD tenosyniovitis[1]	17 (18)	9 (19)	8 (17)	0.78
Patients with erosions[1]	60 (64)	32 (68)	28 (59)	0.40
Total GS score[2] (per patient), (n=94)	8 (6-9)	8 (6-9)	8 (6-10)	0.98
Total PD score[2] (per patient), (n=35)	2 (2-4)	3 (2-4)	2 (2-4)	0.26
Total GS + PD score[2] (per patient), (n=94)	8 (6-12)	8 (6-13)	8 (6-11)	0.64
Erosions[2] (per patient), (n=60)	2 (1-3)	2 (1-3)	2 (1-3)	0.74

Data presented as median (IQR) but [1]N° (%) of patients. [2]Restricted to the subpopulation of patients with the outcome, [2]median (IQR). GS: grey-scale; PD: power Doppler; MUS: musculoskeletal ultrasound.

as the treat-to-target paradigm (30-32), and patients who achieve the status of remission are recommended to undergo MUS to determine the possible need for treatment modifications (33). In ULTRAPRO, patients were defined as in remission according to their rheumatologist criteria which reflects current daily practice, where indices that evaluate DA are exceptionally calculated (1). Outcomes were evaluated at the 6-months follow-up, which also reflects the standard of care for patients with long-term follow-up and clinical remission (2). The majority of the patients were treated with methotrexate as monotherapy or combined with additional DMARDs and (low doses of oral) corticosteroids. Accordingly, we consider our results to be of practical relevance because they reflect the daily condition of the patients. Importantly, in ULTRAPRO, the ultrasonographer differed from the clinician, and all the patients had a MUS performed, although MUS data were shared with the primary rheumatologist only in patients randomised to the intervention group; this approach does not represent clinical practice but was considered mandatory to avoid bias (34).

Our primary objective included worsening of relevant PROs; worsening was selected as patients included were in clinical remission and their baseline PROs were expected to be near normal range. As far as we know, this is the first study that evaluates the impact of MUS to complement clinical assessments and drive treatment in terms of PROs, although flares (a physician-reported outcome) was selected as a secondary objective. Recently, a group of MUS experts reviewed the available literature and discussed the best approach for developing pragmatic suggestions for the use of MUS in the daily management of RA (33); the authors also identified areas with a paucity of evidence and produced a research agenda, where the impact of MUS-defined remission on PROs was highlighted (33). The importance of PROs was emphasised on the 2016 update of the EULAR recommendations for the management of RA patients (31). In these discussions, patients suggested that the list of recommendations should end with an item about PROs to convey their importance in treatment and to facilitate the shared-decision making process (31, 35).

We first found that incorporating MUS information into the clinical assessment of RA patients in remission did not prevent 6 months follow-up PROs worsening. The results are in accordance with 2 randomised clinical studies, the TaSER study (11) and the ARTIC study (12). Both studies were designed to test the hypothesis that incorporating MUS disease activity assessment into a T2T strategy would produce superior clinical and imaging outcomes compared with a strategy driven by a composite DA score. In both studies, PROs were considered as secondary outcomes. Both trials concluded that the systematic use of MUS to inform treatment decisions does not add to clinical management of RA in terms of clinical or radiographic outcomes (36). Nonetheless, in the TaSER study, patients from the intervention group had a greater change in the EuroQoL 5D-3L index between the baseline and the 18-month assessments than patients from the control group, and a similar tendency was found in the HAQ score (p=0.06). Meanwhile, in our study there was a tendency in the percentage of patients with RADAI worsening at 6 months, favouring the intervention group. The consistency in the (primarily negative) results observed should be emphasised, although important differences in the patients and studies characteristics

were present. First, ARTIC and TaSER studies included early RA patients naive to DMARDs or who were receiving DMARDs initiated within the 6 months prior to the baseline evaluation; meanwhile, our patients had longstanding disease and long-term treatment with DMARDs. Second, patients in ULTRAPRO were in remission according to the their rheumatologist criteria; meanwhile, patients from the ARTIC and the TaSER studies had (at least) moderate DA, and complex indices were used to define the patient´s DA status. Third, demographic characteristics known to impact disease outcomes, such as gender and age (37, 38), also differed between studies; patients in the ULTRAPRO study were primarily female (up to 90%), as already described in patients from the Latin-America region (39), and younger than patients from the ARTIC and TaSER studies. Fourth, the patients' follow-up was also different between the ULTRAPRO and the TaSER and ARTIC studies, which were conducted with a tight follow-up that may have produced ideal results and a ceiling effect, as highlighted by D´Agostino *et al*. (40). In addition, the ULTRAPRO study compared 6-month follow-up outcomes between the intervention and the control group, while the TaSER and ARTIC studies described/compared 18-month follow-ups and a combination of 16- and 24-month follow-ups, respectively. Finally, in the 3 studies, MUS assess a different number joints and joint locations.

Second, MUS findings were in accordance with the clinical remission status, although up to 37% of the patients included had discordant results between the clinical and the ultrasound assessments; also, grey-scale synovitis was observed in the totality of our patients with remission as previously described (41), meanwhile tenosynovitis was less frequently detected (42). The scientific literature is consistent with the observation that MUS provides a more accurate method for assessing disease activity, when compared to physician evaluation, although there is a need to counterbalance the expanded scientific literature on the generalised benefits of ultrasound in RA management with appropriated strategy trials (40).

Third, we found that in accordance with PROs findings, the intervention strategy did not prevent worsening physician-reported outcomes. We particularly investigated flares as patients included were in remission status, and in such a clinical context, flares represent a clinical deterioration and are easy to identify at follow-up. Flares also add valuable information, as they translate into fluctuations of DA, increase the risk of joint damage (35) and impact orthopedic and hand surgery outcomes (43).

Finally, our results favour the hypothesis that MUS-derived information complements the clinical assessment and drives treatment, although short term follow-up PROs were not impacted; all the patients assigned to the intervention group with discordant findings were recommended a treatment increase; in them, the primary rheumatologist relied upon MUS information over clinical information. Importantly, no serious treatment-related adverse events were identified at the six-month follow-up (data not shown). Meanwhile, treatment from patients assigned to the intervention group with concordant findings, remained unchanged; it could be argued that these patients may had been considered suitable to a treatment decrease; nonetheless, there is a lack of clear recommendations and conflicting advice regarding treatment decrease/withdrawal in patients in remission status; in addition, the remission status is recommended to be sustained and there is no consensus in how to define sustained remission (44).

Limitations of the study need to be addressed. First, the study was developed in a tertiary care-level centre; patients included, although in remission, may represent those with a more aggressive disease. In addition, all the patients were Hispanics, who are recognised to have particular demographic and disease characteristics (38, 39), and the results may not be generalisable. Second, at the 6-month follow-up, up to 11% of the patients were lost; the sample size was calculated considering 20% losses, but we cannot discard the possibility that favourable/unfavourable PROs may have biased the results. Third, MUS evaluated a limited number of joints that may have impacted the ultrasound definition of DA. Fourth, fatigue is important from the patient's perspective (35) and was not assessed; meanwhile, the PROs assessed, such as the PtGA-VAS, are also not without limitations, which include the lack of a gold standard for wording/phrasing and assessment period and being affected by patient´s factors unrelated to RA itself, such as the level of education, psychological distress and comorbidities (45). Fifth, the remission definition used deviated from the recommended remission definition (46, 47); however, we repeated the analysis in the patients with a baseline DAS28 ≤3.2, which was deemed a convenient target, considering that patients had substantial disease duration; the results obtained showed similar trends (data not shown). Sixth, only short-term follow-up PROs are currently reported, but these may be relevant for patients. Finally, a large number of rheumatologists were involved in the study with variable degree of experience, which may have affected decisions related to treatment modifications.

Conclusions

ULTRAPRO study is the first randomised controlled trial that focuses on PROs and supports previous results where the addition of MUS information to complement clinical assessments, drives treatment but do not impact on short-term follow-up PROs in RA patients in remission.

Acknowledgement

C. Sifuentes-Cantu, thanks the CONACyT for the scholarship during his MSc degree.

Significance and innovations

- The addition of MUS to complement clinical assessments drives RA-related treatment but does not impact on short-term follow-up PROs in RA patients in clinical remission.
- In particular, the intervention does not prevent either worsening PROs or flares.

References

1. ANDERSON JK, ZIMMERMAN L, CAPLAN L, MICHAUD K: Measures of rheumatoid arthritis disease activity. *Arthritis Care Res* (Hoboken). 2011; 63: S14-36.
2. SMOLEN JS, BREEDVELD FC, BURMESTER GR *et al*.: Treating rheumatoid arthritis to target: 2014 update of the recommendations of an international task force. *Ann Rheum Dis* 2016; 75: 3-15.
3. LILLEGRAVEN S, PRINCE FH, SHADICK NA *et al*.: Remission and radiographic outcome in rheumatoid arthritis: application of the 2011 ACR/EULAR remission criteria in an observational cohort. *Ann Rheum Dis* 2012; 71: 681-6.
4. WOLFE F, MICHAUD K: Severe rheumatoid arthritis (RA), worse outcomes, comorbid illness, and sociodemographic disadvantage characterize RA patients with fibromyalgia. *J Rheumatol* 2004; 31: 695-700.
5. BROWN AK, CONAGHAN PG, KARIM Z *et al*.: An explanation for the apparent dissociation between clinical remission and continued structural deterioration in rheumatoid arthritis. *Arthritis Rheum* 2008; 58: 2958-67.
6. HASSAN S: Overview of musculoskeletal ultrasound for the clinical rheumatologist. *Clin Exp Rheumatol* 2018; 36 (Suppl. 114): S3-9.
7. ZABOTTI A, MANDL P, ZAMPOGNA G, DE-JACO C, IAGNOCCO A: One year in review 2018: ultrasonography in rheumatoid arthritis and psoriatic arthritis. *Clin Exp Rheumatol* 2018; 36: 519-25.
8. SIFUENTES-CANTÚ C, CONTRERAS-YÁÑEZ I, SALDARRIAGA L, LOZADA AC, GUTIÉRREZ M, PASCUAL-RAMOS V: The added value of musculoskeletal ultrasound to clinical evaluation in the treatment decision of rheumatoid arthritis outpatients: Physician experience matters. *BMC Musculoskelet Disord* 2017; 18: 390.
9. DÍAZ-TORNÉ C, MORAGUES C, TONIOLO E *et al*.: Impact of ultrasonography on treatment decision in rheumatoid arthritis: the IMPULSAR study. *Rheumatol Int* 2017; 37: 891-96.
10. DALE J, PURVES D, MCCONNACHIE A, MCINNES I, PORTER D: Tightening up? Impact of musculoskeletal ultrasound disease activity assessment on early rheumatoid arthritis patients treated using a treat to target strategy. *Arthritis Care Res* (Hoboken) 2014; 66: 19-26.
11. 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* 2016; 75: 1043-50.
12. HAAVARSDSHOLM EA, AGA AB, OLSEN IC *et al*.: Ultrasound in management of rheumatoid arthritis: ARTIC randomised controlled strategy trial. *BMJ* 2016; 354: i42051.
13. VAN TUYL HD, BOERS M: Patient-reported outcomes in core domain sets for rheumatic diseases. *Nat Rev Rheumatol* 2015; 11: 705-12.
14. VAN TUYL LHD, MICHAUD K: Patient-reported outcomes in rheumatoid arthritis. *Rheum Dis Clin North Am* 2016; 42: 219-37.
15. DESHPANDE PR, RAJAN S, SUDEEPTHI BL, ABDUL NAZIR CP: Patient-reported outcomes: A new era in clinical research. *Perspect Clin Res* 2011; 2: 137-44.
16. SOKKA T, HAKKINEN A, KRISHNAN E, HANNONEN P: Similar prediction of mortality by the health assessment questionnaire in patients with rheumatoid arthritis and the general population. *Ann Rheum Dis* 2004; 63: 494-97.
17. U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES: Food and Drug Administration, Center for Drug Evaluation and Research, Center for Biologics Evaluation and Research, Center for Devices and Radiological Health. Guidance for industry. Patient-reported outcome measures: use in medical product development to support labeling claims. 2009. http://www.fda.gov/ downloads/ Drugs/ GuidanceComplianceRegulatoryInformation/Guidances/ UCM193282.pdf.
18. ORBAI AM, BINGHAM CO, III: Patient reported outcomes in rheumatoid arthritis clinical trials. *Curr Rheumatol Rep* 2015; 17: 28.
19. HENDRIKX J, DE JONGE MJ, FRANSEN J, KIEVIT W, VAN RIEL PL: Systematic review of patient-reported outcome measures (PROMs) for assessing disease activity in rheumatoid arthritis. *RMD Open* 2016; 2: e000202.
20. WARD MM, GUTHRIE LC, ALBA MI: Clinically important changes in short form 36 health survey scales for use in rheumatoid arthritis clinical trials: the impact of low responsiveness. *Arthritis Care Res* (Hoboken) 2014; 66: 1783-9.
21. WARD M, GUTHRIE LC, ALBA MI: Clinically important changes in individual and composite measures of rheumatoid arthritis activity: thresholds applicable in clinical trials. *Ann Rheum Dis* 2015; 74: 1691-96.
22. AHMED S, MAGAN T, VARGAS M, HARRISON A, SOFAT N: Use of the painDETECT tool in rheumatoid arthritis suggests neuropathic and sensitization component in pain reporting. *J Pain Res* 2014; 7: 579-88.
23. YANG M, GUO M: Goals for rheumatoid arthritis: treating to target or treating to prevent. *Open Access Rheumatol* 2012; 4: 81-86.
24. KOSINSKY M, ZHAO S, DEDHIYA S, OSTERHAUS J, WARE JE: Determining minimally important changes in generic and disease-specific health-related quality of life questionnaires in clinical trials of rheumatoid arthritis. *Arthritis Rheum* 2000; 43: 1478-87.
25. BACKHAUS M, OHRNDORF S, KELLNER H *et al*.: Evaluation of a novel 7-joint ultrasound score in daily rheumatologic practice: a pilot project. *Arthritis Rheum* 2009; 61: 1194-201.
26. 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* 2013; 72: 1163-9.
27. D'AGOSTINO MA, TERSLEV L, AEGERT P *et al*.: Scoring ultrasound synovitis in rheumatoid arthritis: a EULAR-OMERACT ultrasound taskforce – Part 1: definition and development of a standardized, consensus-based scoring system. *RMD Open* 2017; 3: e000428
28. KITCHEN J, KANE D: Greyscale and power Doppler ultrasonographic evaluation of normal synovial joints: correlation with pro- and anti-inflammatory cytokines and angiogenic factors. *Rheumatology* (Oxford) 2015; 54: 458-62.
29. ARNETT FC, EDWORTHY SM, BLOCH DA *et al*.: The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum* 1988; 31: 315-2.
30. SINGH JA, SAAG KG, BRIDGES SL *et al*.: 2015 American Colege of Rheumatology guideline for the treatment of rheumatoid arthritis. *Arthritis Rheumatol* 2016; 68: 1-26.
31. SMOLEN JS, LANDEWÉ R, BIJLSMA *et al*.: EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2016 update. *Ann Rheum Dis* 2017; 76: 960-77.
32. PINCUS T, CASTREJON I: Evidence that the strategy is more important than the agent to treat rheumatoid arthritis, data from clinical trials of combinations of non-biologic DMARDs, with protocol-driven intensification of therapy for tight control or treat-to-target. *Bull Hosp Jt Sis* 2013; 71 (Suppl. 1): S33-40.
33. D'AGOSTINO MA, TERSLEV L, WAKEFIELD R *et al*.: Novel algorithms for the pragmatic use of ultrasound in the management of patients with rheumatoid arthritis: from diagnosis to remission. *Ann Rheum Dis* 2016; 75: 1902-8.
34. EL MIEDANY Y, EL GAAFARY M, PALMER D: Assessment of the utility of visual feedback in the treatment of early rheumatoid arthritis patients: a pilot study. *Rheumatol Int* 2012; 32: 3061-8.
35. FAUTREL B, ALTEN R, KIRKHAM B *et al*.: Call for action: how to improve use of patient-reported outcomes to guide clinical decision making in rheumatoid arthritis. *Rheumatol Int* 2018; 38: 935-47.
36. WELSING PM, LANDEWÉ RB, VAN RIEL PL *et al*.: The relationship between diseases activity and radiologic progression in patients with rheumatoid arthritis: a longitudinal analysis. *Arthritis Rheum* 2004; 50: 2082-93.
37. AURRECOECHEA E, LLORCA DÍAZ J, DIEZ LIZUAIN ML, MCGWIN G JR, CALVO-ALEN J: Gender-associated comorbidities in rheumatoid arthritis and their impact on outcome: data from GENIRA. *Rheumatol Int* 2017; 37: 479-85.
38. CONTRERAS-YÁÑEZ I, PASCUAL-RAMOS V: Window of opportunity to achieve major outcomes in early rheumatoid arthritis patients: how persistence with therapy matters. *Arthritis Res Ther* 2015; 17: 177.
39. MODY GM, CARDIEL MH: Challenges in the management of rheumatoid arthritis in developing countries. *Best Pract Res Clin Rheumatol* 2008; 22: 621-41.
40. D'AGOSTINO MA, BOERS M, WAKEFIELD RJ, EMERY P, CONAGHAN PG: Is it time to revisit the role of ultrasound in rheumatoid arthritis management? *Ann Rheum Dis* 2017; 76: 7-8.
41. GÄRTNER M, MANDL P, RADNER H *et al*.: Sonographic joint assessment in rheumatoid arthritis. Associations with clinical joint assessment during a state of remission. *Arthritis*

Rheum 2013; 65: 2005-14.
42. NAREDO E, D'AGOSTINO MA, WAKEFIELD R *et al*.: Reliability of a consensus-based ultrasound score for tenosynovitis in rheumatoid arthritis. *Ann Rheum Dis* 2013; 72: 1328-34.
43. CONTRERAS-YÁÑEZ I, GUARACHA-BASÁÑEZ G, DÍAZ-BORJÓN E, IGLESIAS M. PASCUAL-RAMOS VP: Early referral and control of disease´s flares prevent Orthopedic and Hand Surgery indication (OHSI) in a dynamic cohort of Hispanic early rheumatoid arthritis patients. *BMC Musculoskelet Disord* 2018; 19: 378.
44. EINARSSON JT, WILLIM M, ERNESTAM S, SAXNE T, GEBOREK P, KAPETANOVIC MC: Prevalence of sustained remission in rheumatoid arthritis: impact of criteria sets and disease duration, a Nationwide Study in Sweden. *Rheumatology* (Oxford) 2019; 58: 227-36.
45. NIKIPHOROU E, RADNER H, CHATZIDIONYSIOU K *et al*.: Patient global assessment in measuring disease activity in rheumatoid: a review of the literature. *Arthritis Res Ther* 2016; 18: 251.
46. CAPORALI R, SMOLEN JS: Back to the future: forget ultrasound and focus on clinical assessment in rheumatoid arthritis management. *Ann Rheum Dis* 2018; 77: 18-20.
47. BOERS M: Let's stop fooling ourselves. In RA, only ACR/EULAR criteria define remission and equate with absence of disease. *Ann Rheum Dis* 2016; 75: e68.