RHEUMATOLOGY

Rheumatology 2020;59:1556–1565
doi:10.1093/rheumatology/kez460
Advance Access publication 19 October 2019

Original article

Synovitis in rheumatoid arthritis detected by grey scale ultrasound predicts the development of erosions over the next three years

Burkhard Möller [1,*], Daniel Aletaha [2], Michael Andor [3], Andrew Atkinson [1,4], Bérengère Aubry-Rozier [5], Laure Brulhart [6,*], Diana Dan [5], Axel Finckh [7], Véronique Grobéty [8], Peter Mandl [2], Raphael Micheroli [9], Michael John Nissen [7,*], Alexander M. Nydegger [10], Almut Scherer [4], Giorgio Tamborrini [11,*], Hans-Rudolf Ziswiler [12,*] and Pascal Zufferey [5,*], and the SONAR group for musculoskeletal ultrasound in the Swiss Clinical Quality Management (SCQM) Foundation

Abstract

Objectives. To evaluate grey scale US (GSUS) and power Doppler US synovitis (PDUS), separately or in combination (CombUS), to predict joint damage progression in RA.

Methods. In this cohort study nested in the Swiss RA register, all patients with sequential hand radiographs at their first US assessment were included. We analysed the summations of semi-quantitative GSUS, PDUS and CombUS assessments of both wrists and 16 finger joints (maximum 54 points) at their upper limit of normal, their 50th, 75th or 87.5th percentiles for the progression of joint damage (ΔXray). We adjusted for clinical disease activity measures at baseline, the use of biological DMARDs and other confounders.

Results. After a median of 35 months, 69 of 250 patients with CombUS (28%), 73 of 259 patients with PDUS (28%) and 75 of 287 patients with available GSUS data (26%) demonstrated joint damage progression. PDUS beyond upper limit of normal (1/54), GSUS and CombUS each at their 50th (9/54 and 10/54) and their 75th percentiles (14/54 and 15/54) were significantly associated with ΔXray in crude and adjusted models. In subgroup analyses, GSUS beyond 14/54 and CombUS higher than 15/54 remained significantly associated with ΔXray in patients on biological DMARDs, while clinical disease activity measures had no significant prognostic power in this subgroup.

Conclusion. Higher levels of GSUS and CombUS are associated with the development of erosions. GSUS appears to be an essential component of synovitis assessment and an independent predictor of joint damage progression in patients on biological DMARDs.

Key words: rheumatoid arthritis, hand, synovium, ultrasonography, biological therapies

CLINICAL SCIENCE

Rheumatology key messages

- The optimal method to demonstrate active synovitis associated with the development of erosions is controversial.
- Synovitis in grey scale was the most robust predictor of erosions in patients with RA.
- Erosion prediction was demonstrated in RA patients with and without biological DMARD therapy.

[1]Rheumatology, Immunology and Allergy, Inselspital, University Hospital of Bern, Bern, Switzerland, [2]Medical Department III, Rheumatology, Medical University Vienna, Vienna, Austria, [3]Prodorso, Zurich, [4]SCQM statistics group, Zurich, [5]Rheumatology, Lausanne University Hospital, Lausanne, [6]Rheumatology, Hôpital Neuchâtelois, La Chaux-de-Fonds, [7]Rheumatology, University Hospitals of Geneva, Geneva, [8]Victoriaplatz Centre for Rheumatology, Bern, [9]Rheumatology, University Hospital Zurich, Zurich, [10]Rheumatology and Rehabilitation, Klinik Schloss Mammern, Mammern, [11]Basel Institute for Locomotor System Ultrasound, Basel and [12]Osteorheuma, Bern, Switzerland

Submitted 30 May 2019; accepted 30 July 2019

Correspondence to: Burkhard Möller, Inselspital, Universitätsklinik für Rheumatologie, Immunologie und Allergologie, University Hospital Bern, Freiburgstrasse 16P, CH-3010 Bern, Switzerland.
E-mail: Burkhard.moeller@insel.ch

*Members of the SONAR group.

Introduction

Major improvements have been achieved in the past decades in RA outcome, both by the use of novel treatment options, and by implementation of the 'treat to target' strategy aiming at clinical remission [1, 2]. Among the concurrent options, the provisional remission criteria endorsed by the ACR and the EULAR [3, 4] are the most stringent ones. However, joint destruction can still occur in patients whose disease activity appears sufficiently controlled [5, 6]. In patients treated with TNF inhibiting (TNFi) and IL-6 receptor inhibiting (IL-6Ri) biological DMARDs (bDMARDs), an uncoupling between inflammation and structural damage was noted [6–12], which may pose a challenge for the appropriate management of an increasing number of bDMARD-treated RA patients.

A worse radiographic outcome in RA is more likely in the presence of synovitis, which may be detected clinically or by different imaging modalities [13]. Grey scale US (GSUS), colour Doppler and power Doppler (PDUS) are capable of detecting clinically verified, as well as subclinical synovitis. Among these US methods, PDUS emerged as the best predictor of progression of joint damage [14–16]. However, when used as a target for escalating treatment from conventional to more aggressive anti-inflammatory therapy, PDUS did not lead to significant improvement in radiographic outcome in two strategic treat-to-target trials in early RA [17, 18]. MRI-detected bone marrow oedema or osteitis was found to be an even better predictor of joint damage than PDUS in several observational studies [9, 15]. Again, when compared with clinical disease activity, escalation from standard treatment to bDMARD therapy in response to bone marrow oedema did not improve the radiographic outcome in RA patients when starting in clinical remission or low disease activity [19].

In addition to the predefined very low PDUS treatment targets in the above-mentioned imaging studies, we set out to define alternative meaningful US cut-offs from the Swiss Sonography in Arthritis and Rheumatism (SONAR) score in comparison with this synovitis level in PDUS as a predictor of radiographic joint damage progression in a real-life setting. This study aimed to have a specific focus on patients treated with bDMARDs. In addition to re-evaluating PDUS and GSUS separately for another set of the 28-joint status, we also assessed their combination (CombUS) as per the most recent recommendation of the EULAR-OMERACT US Taskforce [20].

Methods

Setting

Data in the Swiss Clinical Quality Management (SCQM) database [21] were extracted from patients with RA based on the clinical diagnosis by the treating rheumatologist. Other inclusion criteria were the availability of a baseline radiograph taken within a time period ranging from 6 months prior to and up to 3 months after the initial US assessment of the hands (detailed below) and a follow-up radiograph taken at least 6 months after the baseline X-ray. Data collection started on 1 January 2010, and ended on 10 August 2018. All patients gave their written consent for use of coded data from the SCQM registry. This study complies with the Declaration of Helsinki and was approved by Swiss ethics committee vote KEK-BE 2018-02331.

Outcome and exposure variables

The primary outcome was radiographically detectable change in joint damage between baseline and follow-up hand X-rays, expressed in % of the Ratingen score of the hands [22]. The Ratingen score is a validated semi-quantitative method to measure the damaged bone surface by typical marginal erosions for RA in quintiles [22]. The range of the Ratingen hands score is 0–140. All plain radiographs are centrally evaluated in SCQM according to the Ratingen score by a single evaluator without knowledge of clinical, sonographic and previous radiographic data [23–25]. The continuous quality management in SCQM for radiographs of hand and feet reports intraclass correlation coefficients of minimum 0.98 (0.91; 0.99) for inter-observer and minimum 0.99 (0.97; 1) for intra-observer comparisons, with a point estimate of scoring difference at zero (–6; 6) (SCQM internal report, unpublished). We defined radiographic progression (ΔXray) as a change beyond the minimally detectable change of 3.5% (ΔXray).

The predictor of interest was synovitis detected by different US modalities. PDUS and GSUS was scored according to the SONAR score reduced to the hands [26]. This score evaluates synovitis in both wrists, summarizing the radiocarpal and midcarpal joints into a single joint, and 16 finger joints, which were the second to fifth MCP and PIP joints. In SONAR, GSUS and PDUS of the wrist scans were both performed at the dorsal aspect. Finger joints were scanned for GSUS according to publicly available reference images at start of US data collection in the SCQM RA registry from a volar view [27, 28]. Finger joints were scanned for PDUS at their dorsal aspects with machine settings as described previously [29]. Joints were scored for GSUS and PDUS from 0 to 3 and summed separately, or calculated in combination of both modalities to CombUS according to the EULAR-OMERACT US Taskforce recommendations [20]. Briefly, CombUS is zero in the absence of any pathologies in GSUS and in PDUS. Minimal synovitis (CombUS grade 1) is considered in case GSUS = 1 and by PDUS $\leqslant$ 1. Moderate synovitis (CombUS grade 2) is defined either by GSUS = 2 in combination with PDUS $\leqslant$ 2, or by GSUS = 1 in conjunction with PDUS = 2. Finally, severe synovitis (CombUS grade 3) is defined by the presence of GSUS = 3 or PDUS = 3. Thus, each score has a maximum of 54. Replaced joints and joints with other constraints to performing an appropriate examination in the standard neutral position were excluded from GSUS and PDUS analyses. Calculation of any US score required at least 80% of data completeness. Assuming that synovitis was similar in unevaluated joints, we performed a linear extrapolation for up to three joints. In order to perform

logistic regression analyses for minimal US pathologies, an external reference for PDUS activity (cut-off $\geqslant 1$) was adopted from ARCTIC [18]. The upper limit of normal $\geqslant 8$ in SONAR-GSUS in healthy volunteers [26] served as an external reference. Additional GSUS, PDUS and CombUS cut-offs were derived based on the 50th, 75th and 87.5th percentiles of all scores.

Sonographers in this study were board-certified rheumatologists with a structured further training in musculoskeletal US (MSUS). All sonographers of this study had participated in an additional half-day hands-on exercise to train in the scoring method, before the US examinations were locally performed in the participating centres of SCQM with different commercially available equipment. One year after training, the kappa value for inter-reader agreement for the entire process from scanning on different US equipment to scoring was 0.64 for PDUS and GSUS, which is in the range of published international data from MSUS experts [30].

Statistics

Results are reported as median and interquartile ranges for quantitative variables. Data for qualitative variables are expressed as absolute frequency and as corresponding percentage. The Kruskal–Wallis test was used for continuous and the Fisher's exact test for categorical baseline variables. Radiographic progression was displayed in cumulative probability plots.

Binomial logistic regression with a logit link function was used to analyse associations between US and the odds for radiographic progression. We defined *a priori* a 0.05% alpha error and obtained *post hoc* power estimates of 86% for GSUS, 81% for PDUS and 75% for the CombUS complete dataset model. Age, sex, BMI, anaemia according to WHO definition [31], smoking status (never, former, current), disease duration, RF or anti-CCP antibody positivity, bDMARD treatment at baseline irrespective of its pharmacological target, number of previous bDMARDs, baseline Ratingen hands score, as well as the following disease activity measures were alternatively included each into one of the different adjusted models: the 28-joint-based clinical disease activity scores $DAS28_{ESR}$ [32] or $DAS28_{CRP}$ [33], the Simplified Disease Activity Index or the Clinical Disease Activity Index [34], $DAS28_{ESR}$ defined low disease activity ($\leqslant 3.2$) or remission (<2.6) [35], and Simplified Disease Activity Index or Boolean method-defined ACR/EULAR-defined remission [4]. Whenever indicated, missing baseline covariate data were replaced by multiple imputation using chained equations with 70 iterations for each dataset. Results from models from each dataset containing imputed values for missing covariates were averaged using Rubin's rule.

Results were defined as consistent in complete baseline covariate and multiply imputed datasets if (i) the point estimate of the covariate of interest with the odds for progression after multiple imputation using chained equations was within the CI of the respective estimate in the complete case analysis, and (ii) the inference in both approaches was consistent in the sense that the *P*-values lead to same conclusion.

Subgroup analyses included the following: (i) only ACR/EULAR 2010 classification criteria-positive patients [36], (ii) only bDMARD-naïve patients, (iii) only bDMARD-treated patients since baseline, and (iv) only bDMARD-treated patients since baseline or starting at any time during observation. To further characterize the overall diagnostic performance to predict X-ray progression, we calculated the sensitivities, specificities and receiver operating curve characteristics of all the tested US progression risk categories.

Results

Descriptive statistics

A total of 259 and 287 patients were included into the PDUS and the GSUS 'complete case' analysis, respectively. For the CombUS composite score, 250 patients were included. Unless explicitly otherwise stated, in the following we present the CombUS dataset. Descriptive baseline statistics of this dataset are summarized in Table 1. The median of all PDUS assessments was 1/54 (Fig. 1). The median in GSUS was 9/54 (Fig. 1). Other relevant cut-off values for later-represented association analyses were GSUS $\geqslant 14/54$ at the 75th percentile, CombUS median $\geqslant 10/54$ and CombUS $\geqslant 15/54$ at the 75th percentile (Fig. 1). Thus, relevant CombUS scores were just one point higher than the corresponding GSUS scores.

Single adjusted analyses

Median time between baseline and first follow-up X-ray was 1.4 (interquartile range 1; 2.1) years. Time to progression in the Ratingen score or last X-ray was 2.9 (1.9; 4.7) years. Baseline PDUS $\geqslant 1/54$, GSUS $\geqslant 9/54$, CombUS $\geqslant 10/54$, GSUS $\geqslant 14/54$ and CombUS $\geqslant 15/54$ (Fig. 2) were each significantly associated with ΔXray in separate logistic regression models (Table 2). Point estimates of the odds ratios (OR) of statistically significant crude associations were between 1.94 and 2.41. Among the covariates, no clinical disease activity parameter or bDMARD therapy, but age was significantly associated with ΔXray [OR = 1.03 (1.01–1.05), $P = 0.02$].

Multivariable adjusted analyses

Multivariable adjusted analyses were performed after imputation of missing baseline covariates. As in the crude analyses, CombUS $\geqslant 15/54$ (Table 3), CombUS $\geqslant 10/54$, PDUS $\geqslant 1/54$, GSUS $\geqslant 14/54$ and GSUS $\geqslant 9/54$ (supplementary Tables S1–S4, available at *Rheumatology* online) were all associated with significantly increased odds for radiographic progression. The point estimates for the ORs between ΔXray and the different US parameters in these models ranged between 2.3 and 2.9, and were mostly dependent on US parameters. As in the crude analyses, no clinical disease activity measure or bDMARD therapy, but age was significantly associated with radiographic progression in adjusted models. Notably, the odds for the different clinical disease activity parameters and

Table 1 Description of the study population

	CombUS <15/54	CombUS ⩾15/54	*P*-value
Number	185	65	
Age (years), median, IQR	54.6, 45.7–63.5	56.9, 45.8–67.6	0.30
Female, *n*	153	53	0.85
Former or currently smoking, *n* (N)	50 (102)	8 (32)	0.06
Disease duration (years), median, IQR	5.5, 2.2–12.2	8, 2.6–16.2	0.20
ACR-EULAR classifiable, *n* (N)	131 (178)	44 (63)	0.62
Anti-CCP positive, *n* (N)	145	55	0.37
RF positive, *n* (N)	131 (184)	43 (65)	0.53
PDUS, median, IQR	1, 0–2	5, 2–11	<0.001
GSUS, median, IQR	7, 4–11	18, 16–23	<0.001
BMI, median, IQR (N)	25.6, 22.8–29.8 (167)	26, 23.5–29.1(60)	0.92
ACR-EULAR remission, *n* (N)	25 (68)	2 (21)	0.03
DAS28 CRP, median, IQR (N)	2.6, 1.9–3.5, (153)	3.5, 2.6–4.6, (53)	<0.001
DAS28 ESR, median, IQR (N)	2.9, 2.3–3.8, (139)	4, 2.9–5.2 (52)	<0.001
SDAI, median, IQR (N)	7.9, 2.8–13.4 (66)	9.8, 7.1–24.1 (21)	0.03
CDAI, median, IQR (N)	6, 2–11, (74)	11, 6–24 (23)	0.005
CRP, median, IQR (N)	3, 1.4–8 (156)	6, 0.5–16 (55)	0.26
ESR, median, IQR (N)	14, 7–26 (141)	20, 10.5–27.5(54)	0.02
SJC28, median, IQR (N)	1, 0–3 (171)	4, 1.5–7 (59)	<0.001
TJC28, median, IQR (N)	1, 0–4 (171)	5, 1–9 (59)	<0.001
Ratingen X-ray hands score	6, 2–13	13, 3–22	0.0044
HAQ-DI, median, IQR (N)	0.4, 0.1–0.9 (101)	0.5, 0.1–1.1 (25)	0.58
Time between US and baseline X-ray (months), median, IQR	0, 0–4.8	0, 0–2.9	0.76
Time between baseline and follow-up X-ray (years), median, IQR	1.4, 1–2	1.5, 1.1–2.2	0.23
Time between baseline and progression or last X-ray (years), median, IQR	3, 1.9–4.7	2.8, 1.9–4.7	0.99
Calendar year, median, IQR	2012 (2011–2013)	2011 (2010–2012)	0.003
On corticosteroid, *n*	67	28	0.37
On csDMARD, *n*	113	40	1
On bDMARD, *n*	85	26	0.47
On TNFi, *n*	47	14	0.62
On tsDMARD, *n*	1	0	1
2nd or 3rd line bDMARD, *n*	52	21	0.41
bDMARD continued, *n*	38	7	0.13
bDMARD stopped after baseline, *n*	42	11	
bDMARD started after baseline, *n*	58	27	
Time until bDMARD start or switch since baseline (years), median, IQR (N)	1.32, 0.31–2.7 (100)	0.5, 0.18–1.76 (45)	0.04

Presented are baseline data, unless otherwise stated, from patients with complete CombUS hand scores, stratified by CombUS at its 75th percentile. Continuous data are presented as median and interquartile range (IQR). *n*: absolute numbers; N: number with data; bDMARD: biological DMARD; CDAI: Clinical Disease Activity Index: CombUS: combined grey scale and power Doppler US; csDMARD: conventional synthetic DMARD; DAS28: DAS based on 28 counts; GSUS: grey scale US; HAQ-DI: HAQ disability index; IQR: interquartile range; PDUS: power Doppler US; SDAI: Simplified Disease Activity Index; SJC: swollen joint count; TJC: tender joint count; TNFi: TNF inhibitor; tsDMARD: targeted synthetic DMARD.

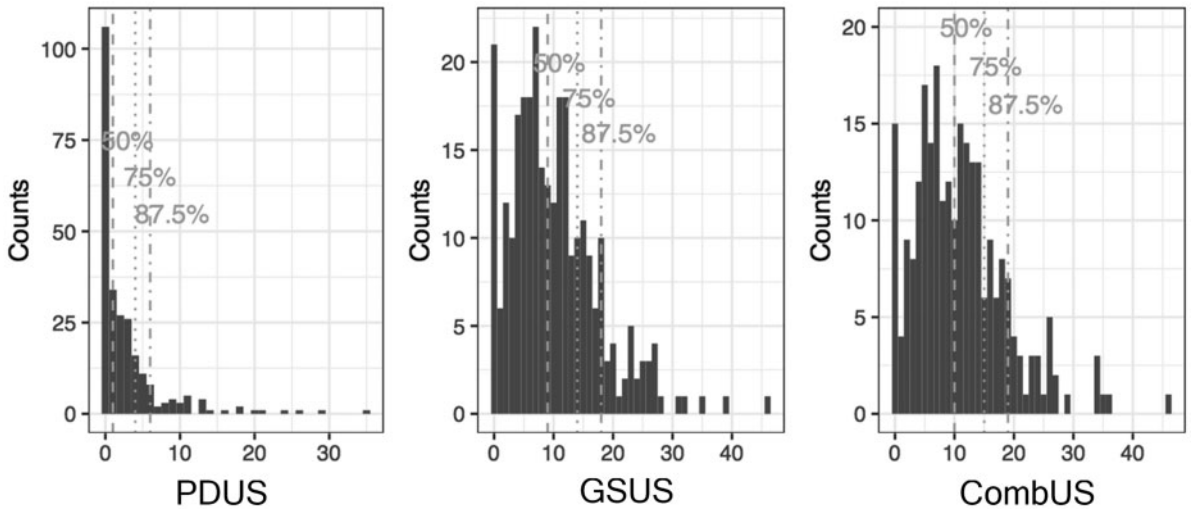
ΔXray tended to be stronger in models with less stringent clinical disease activity parameters. The results after multiple imputation were consistent with the results in analyses with complete baseline covariate data.

Subgroup analyses

Subgroup analyses for CombUS ($n = 200$, 57 events), PDUS ($n = 208$, 61 events) and GSUS ($n = 227$, 61 events) in ACR/EULAR classification criteria-positive patients [36] confirmed the associations observed in the entire datasets. For CombUS ⩾10/54 and PDUS ⩾1/54, we observed a higher OR in ACR/EULAR classifiable patients than in the total study population (supplementary Tables S1 and S2, available at *Rheumatology* online).

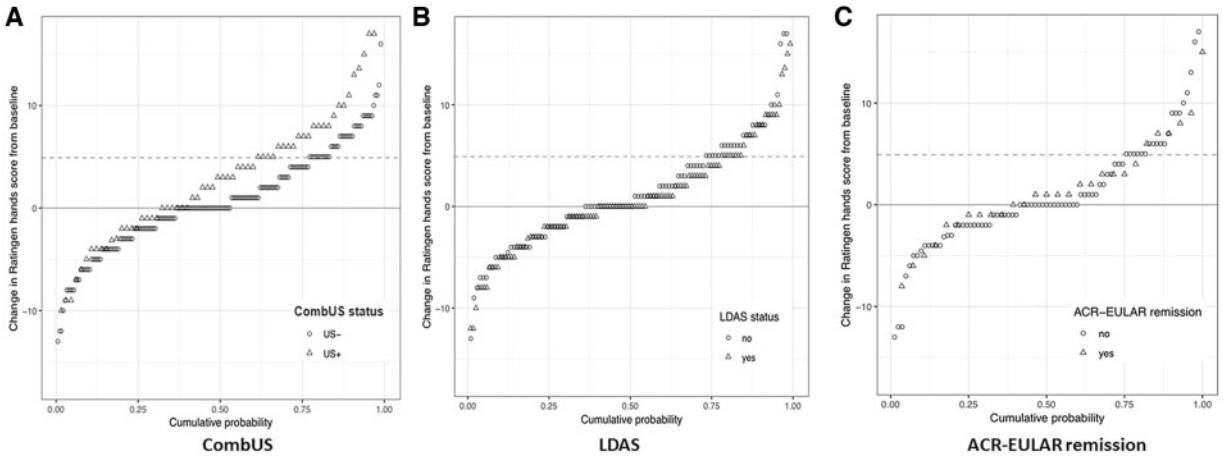
Forty-four percent of patients received bDMARDs at baseline, many of them already as their second or third bDMARD (Table 1). In the bDMARD-treated patients at baseline, we detected statistically significant associations with ΔXray for CombUS ⩾15/54 [$n = 132$, 39 events, OR = 4.14 (1.42–12.09), $P = 0.01$] and GSUS ⩾14/54 [$n = 153$, 45 events, OR 2.82 (1.13–7), $P = 0.03$]. No clinical disease activity parameter was significantly associated with ΔXray in at baseline bDMARD-treated patients.

FIG. 1 Distribution of the three tested synovitis US scores over the range



PDUS (left hand panel) is more right shifted than GSUS (middle) and CombUS (right hand panel). Annotated in the PDUS, GSUS and CombUS score histograms (from left to right) are the cut-off points for 50th, 75th and 87.5th percentiles. PDUS: power Doppler US; GSUS: grey scale US; CombUS: combined grey scale and power Doppler US.

FIG. 2 Association of synovitis imaging and two alternative clinical disease activity measures with joint damage progression



Cumulative probability plot of radiographic damage progression (ΔXray) of the hands are stratified by (**A**) low CombUS <15/54 (US−) *vs* high CombUS ⩾15/54 (US+), (**B**) remission or low disease activity in DAS28 ⩽3.2 (LDAS yes) *vs* active disease with DAS28 >3.2 (LDAS no) and (**C**) ACR/EULAR remission present (yes) or absent (no). Probability plots illustrate the individual mean annual progression in the Ratingen score from baseline to progression or censoring at the last available radiographs. CombUS: combined grey scale and power Doppler US; LDAS: low disease activity score.

Another 35% of patients started bDMARD treatment during follow-up. Type or duration of treatment at baseline was not a formal inclusion criterion, but the present study was performed in an intensively treated RA patient population. Treatment adaptations were not protocol guided, but bDMARDs were more often initiated in patients with clinical and imaging indicators of high disease activity. Furthermore, bDMARDs were more frequently discontinued in subjects with low disease activity indicators. We illustrate this general finding in Table 1. bDMARDs were stopped in 23% of patients with CombUS <15/54 *vs* 17% of patients with CombUS at baseline ⩾15/54. In contrast, bDMARDs were started after baseline in 42% of patients with CombUS ⩾15/54 *vs* 31% of patients with lower CombUS scores at baseline. bDMARD therapy indicators were not significantly associated with ΔXray in crude and in the multivariable adjusted analyses. This finding was reproduced in all the datasets that were obtained after multiple

TABLE 2 Association of US categories and baseline covariates with ΔXray in single adjusted analyses

Parameter	*n/N*	OR	95% CI	*P*-value
$GSUS_{50}$ ⩾9/54	149/286	1.95	1.14, 3.40	0.02
$GSUS_{75}$ ⩾14/54	79/286	2.41	1.37, 4.22	0.002
$GSUS_{87.5}$	43/286	1.11	0.52, 2.25	0.77
$PDUS_{50}$ ⩾1/54	153/259	2.28	1.28, 4.18	0.006
$PDUS_{75}$	66/259	1.68	0.92, 3.04	0.09
$PDUS_{87.5}$	39/259	1.99	0.97, 4.01	0.06
$CombUS_{50}$ ⩾10/54	130/250	1.94	1.11, 3.48	0.02
$CombUS_{75}$ ⩾15/54	65/250	2.2	1.20, 4.02	0.01
$CombUS_{87.5}$	36/250	1.38	0.63, 2.89	0.41
Sex (female reference)	236/286	0.84	0.40, 1.67	0.64
Age (per year)	286	1.03	1.01, 1.05	0.02
BMI (per unit)	259	0.97	0.92, 1.03	0.33
Anaemia	0/228	0.89	0.38, 1.95	0.79
Current *vs* never smoker	29/155	2.14	0.90, 5.00	0.08
Former *vs* never smoker	37/155	0.63	0.17, 1.89	0.44
				0.09[a]
Disease duration (per year)	286	1.02	0.99, 1.04	0.15
Calendar year (per year)	286	0.95	0.80, 1.12	0.55
Anti-CCP positive	195/272	1.08	0.60, 2.00	0.80
RF positive	194/281	1.17	0.66, 2.13	0.60
DAS28 CRP (per unit)	237	1.09	0.87, 1.36	0.44
DAS28 ESR (per unit)	223	1.26	0.47, 1.31	0.57
DAS28 ESR ⩾2.6 *vs* <2.6	223	1.00	0.53, 1.94	1.0
DAS28 ESR >3.2 *vs* ⩽3.2	223	1.30	0.71, 2.39	0.40
SDAI (per unit)	89	1.01	0.97, 1.06	0.51
CDAI (per unit)	99	1.01	0.97, 1.05	0.69
Not in ACR/EULAR remission [4]	81/109	1.26	0.47, 3.80	0.66
Baseline Ratingen score	286	1	0.98, 1.01	0.56
No. of prev. bDMARDs (1 *vs* 0)	39/202	1.6	0.75, 3.32	0.21
No. of prev. bDMARDs (⩾2 *vs* 0)	46/202	1.4	0.68, 2.81	0.35
	258			0.36[a]
bDMARD at baseline	129/286	1.37	0.80, 2.32	0.25

[a]*P*-values for covariates with more than two levels were obtained by likelihood ratio tests. bDMARD: biological DMARD; CDAI: Clinical Disease Activity Index; CombUS: combined grey scale and power Doppler US; DAS28: DAS based on 28 joint counts; GSUS: grey scale US; OR: odds ratio; PDUS: power Doppler US; SDAI: Simplified Disease Activity Index; SJC: swollen joint count; TJC: tender joint count. Subscript 50, 75 and 87.5 denote the 50th, 75th and 87.5th percentiles.

imputation using chained equations (Table 3 and supplementary Tables S1–S4, available at *Rheumatology* online). As before in bDMARD-treated patients at baseline, now in a larger subgroup of patients who were receiving bDMARDs at any time, CombUS ⩾15/54 (Table 3) and GSUS ⩾14/54 at baseline (supplementary Table S3, available at *Rheumatology* online) was significantly associated with ΔXray. Furthermore, PDUS ⩾1/54 [*n* = 206, 63 events, OR 2.15 (95% CI 1.03–4.47), *P* = 0.04] (supplementary Table S2, available at *Rheumatology* online) and GSUS ⩾9/54 [*n* = 225, 45 events, OR 1.94 (95% CI 1–3.76), *P* = 0.05] (supplementary Table S4, available at *Rheumatology* online), but no clinical disease activity measure became significantly associated with ΔXray after inclusion of all patients into analysis with bDMARD therapy since baseline and those starting later.

Overall diagnostic performance

Despite the many significant associations of US-detected synovitis and ΔXray, the formal diagnostic performance of every single US index test at baseline was not satisfactory. However, as we could not observe a stronger association between ΔXray and any of the tested clinical disease activity measures in this study, we consider this a practically relevant finding. We obtained similar areas under the curve in the receiver operating characteristics analyses for PDUS [0.6 (95% CI 0.52; 0.67)], GSUS [0.6 (95% CI 0.53; 0.68)] and CombUS [0.61 (95% CI 0.53; 0.69)]. However, PDUS ⩾1/54 had the best sensitivity [0.73 (95% CI 0.61; 0.82)], but GSUS ⩾14/54 [0.77 (95% CI 0.71; 0.82)] and CombUS ⩾15/54 [0.78 (95% CI 0.72; 0.84)] represented the best testing specificities among the MSUS criteria in significant association with ΔXray.

Discussion

This study shows for the first time in a large registry the value of a single joint US for the risk assessment of radiographic joint damage on a subsequent median 3-year

Table 3 Association of $CombUS_{75}$ in combination with different clinical disease activity covariates and ΔXray

OR for ΔXray by CombUS ⩾15/54	Clinical disease activity covariate	OR for ΔXray by clinical covariate
2.69 (1.28–5.63), $P = 0.0093$	DAS28 CRP	1.16 (0.88–1.53), $P = 0.28$
2.58 (1.23–5.54), $P = 0.01$	DAS28 ESR	1.21 (0.93–1.57), $P = 0.17$
2.72 (1.29–5.74), $P = 0.0092$	SDAI	1.01 (0.98–1.05), $P = 0.44$
2.77 (1.3–5.89), $P = 0.0088$	CDAI	1.01 (0.97–1.05), $P = 0.59$
2.73 (1.32–5.64), $P = 0.0072$	DAS28 ESR >3.2	1.68 (0.85–3.31), $P = 0.14$
2.87 (1.38–5.95), $P = 0.0051$	DAS28 ESR ⩾2.6	1.19 (0.56–2.54), $P = 0.65$
2.85 (1.37–5.95), $P = 0.0055$	Not in ACR/EULAR remission	0.83 (0.36–1.88), $P = 0.65$
Subgroup analyses		
2.7 (1.22–5.98), $P = 0.02$[a]	DAS28 ESR >3.2 *vs* ⩽3.2	1.68 (0.85–3.31), $P = 0.14$
4.14 (1.42–12.09), $P = 0.01$[b]	DAS28 ESR >3.2 *vs* ⩽3.2	0.99 (0.38–2.57), $P = 0.98$
2.72 (1.23–6.02), $P = 0.01$[c]	DAS28 ESR >3.2 *vs* ⩽3.2	1.26 (0.61–2.61), $P = 0.53$
2.74 (0.85–8.77), $P = 0.09$[d]	DAS28 ESR >3.2 *vs* ⩽3.2	2.75 (0.91–8.3), $P = 0.07$

Multiple variable adjusted logistic regression analyses were performed in the complete dataset ($n = 250/69$ events) and different subgroups after MICE. [a]Dataset of only ACR/EULAR classification criteria-positive patients ($n = 200$, 57 events). [b]Patients treated with bDMARD at baseline ($n = 132$, 39 events). [c]bDMARD-treated patients at baseline or started later during follow-up ($n = 197$, 59 events). [d]bDMARD-naïve patients at baseline ($n = 118$, 30 events). bDMARD: biological DMARD; CDAI: Clinical Disease Activity Index; DAS28: DAS based on 28 joint counts; CombUS: combined grey scale and power Doppler US; MICE: multiple imputation in chained equations; OR: odds ratio (margins of 95% CI); SDAI: Simplified Disease Activity Index. Subscript 75 denotes the 75th percentile.

period. The real life setting of this study with many examiners, long observation time and the different equipment in use make this result practically relevant. We identified two candidate cut-off values each in GSUS and CombUS in association with structural damage progression, one close to 20% and another at 30% of maximal possible synovitis pathologies. Furthermore, we found PDUS pathologies at very low levels similar to that which was used in ARCTIC [18], and which in The Targeting Synovitis in Early Rheumatoid Arthritis (TaSER) study [17] were associated with the development of erosions, but with a lower OR than the optimal GSUS and CombUS cut-offs. Thus, quite low grades of synovial pathologies that are not far from the upper level of normal findings [29] can already be of relevance for the structural outcome in RA.

The decision about the selection of the joints and the US technique (volar and dorsal) in the SONAR score to be evaluated in this study was oriented at the 28-joint status and available data of US synovitis imaging before 2010 [27, 28]. The US methods remained unchanged in the Swiss RA cohort over approximately a decade which made this long-term study possible. However, this methodological stability to the costs of some differences to the continuously evolving recommendations [20]. Furthermore, for GSUS and PDUS, SONAR uses the same, universally applied semi-quantitative scoring system ranging from 0–3 [20], applied in most US studies, including ARCTIC and TaSER [17, 18]. Apart from always arbitrary joint selection, the only major difference between the SONAR and other GSUS assessments is the exclusive evaluation from the palmar view. We propose in depth and practically elaborated parameters for further adaptations and testing in future strategic trials.

The results of the present study have to be interpreted in the context of more general potential limitations associated with an observational setting. Patients were followed-up without a protocol, and there was large variation in time intervals between follow-up radiographs. Furthermore, medical treatment was not prescribed per protocol. Furthermore, with the recommended frequency of annual visits in the SCQM register, individual changes in medical treatment could not be tracked back to disease activity. However, with more treatment intensifications in case of higher baseline disease activity, the observed bDMARD treatment adaptations appeared to be in principle in accordance with the treat-to-target concept, which in Switzerland is almost unaffected by restricted access to drugs. Together with the fact that the vast majority of patients was at least temporarily exposed to bDMARDs, this may be a reason that bDMARD treatment was not *per se* associated with ΔXray, although it is evident from many clinical trials that bDMARDs better inhibit structural damage progression than conventional synthetic DMARDs. Furthermore, we decided on a linear imputation of occasionally missing US data in destroyed or otherwise unevaluable joints, which may have affected the precision of the models. As another limitation for adjusted models, we had to impute the often missing patient global disease activity parameters necessary to calculate Clinical Disease Activity Index, Simplified Disease Activity Index and the ACR-EULAR remission rates. Finally, data could not be adjusted for unmeasured potentially confounding genetic data or other information on disease biology [37].

As the SONAR score is hand dominated, we decided *a priori* to focus on US as well as radiographic data only of the hands. This decision limits the generalizability of our data to other joint regions. Especially the MTP joints are frequently affected in RA, and their exclusion may be one of the most important limitations of this study and the SONAR score in its present form. Furthermore, the

applied Ratingen X-ray score has a numerical range of pathologies for hands that is 20 points lower compared with the van der Heijde Sharp score, which goes up to 160 points [38]. As far as is comparable when obtained from different studies, the Ratingen score appears to have 10% lower standardized response means than the van der Heijde Sharp score [39, 40]. However, it could be argued that less sensitivity to change may even strengthen the results in terms of being able to demonstrate tangible differences.

This study population contained many bDMARD-treated patients. It is commonly recognized that the link between damage progression and clinical disease activity may be disconnected in patients on this type of therapy [7, 8, 11, 12]. Taking this into consideration, the finding that clinical disease activity measures were associated with structural outcome only in patients without bDMARD therapy was not unexpected. Furthermore, the complexity of the models was reduced to only baseline exposure data, which does not cover fluctuations in disease activity over time [6, 41, 42]. Nevertheless, we handled clinical and MSUS data in exactly the same way. As far as demonstrated, GSUS and CombUS, obtained at only one occasion, was in itself informative for detecting an increased risk of joint damage progression over such a long period of time. Regarding other covariates, we observed a statistically significant association of age with erosive progression. More severe arthritis and accelerated joint damage progression in elderly RA patients has been observed previously [43]. In contrast to some previous studies, joint damage was not associated with RF or CCP antibody status [44, 45] in this study.

As a consequence of the negative outcome for the imaging arms in ARCTIC, TaSER and Imagine-RA [17–19], arthritis imaging should be restricted to difficult diagnostic situations [46]. Identification of the progression of erosions in RA patients on bDMARD therapy appears to have the most relevance to these challenging situations [7, 8], in which imaging synovitis according to the present data appears to make sense. Patients treated with bDMARDs typically have only small changes in joint damage. Thus, the significance of the amount in ΔXray used in this study might be debatable, but protection of the joint structures is a central aspect of good long-term outcomes. As bDMARDs are increasingly used worldwide, our findings will probably gain even more relevance in the near future. Though PDUS was on its own a predictor of damage progression in several studies [14–16], inclusion of PDUS into CombUS here did not provide a major additional benefit. We have no clear evidence from our own data [30], but we consider technical and methodological issues in daily practice to be a more relevant problem for the reliability and validity of PDUS than for GSUS.

In summary, this large, multicentre, real-life study strongly suggests the usefulness of imaging synovitis to predict the structural disease outcome from a perspective of years, especially in RA patients on bDMARD therapy. Exceeding 20–30% of the maximum synovitis grading in GSUS or in CombUS appeared to be of equivalent relevance. This result does not contradict the importance of synovial hypervascularity or even more of osteitis for joint destruction, which cannot be displayed by synovitis US, nor does this study call into question the current primacy of a treat-to-target concept on the basis of reiterated clinical disease activity assessments. While our observations highlight the importance of synovitis for the destructive course of RA, ongoing technical and methodological improvements in US imaging synovitis appear likewise necessary and achievable.

B.M. planned this study and is responsible for the overall content as guarantor. A.A. and A.S. performed statistical analyses. B.M., P.Z., L.B., G.T., R.M., D.D., V.G., M.J.N., A.F., B.A.-R. and H.-R.Z. contributed to data acquisition. D.A. and P.M. critically revised the manuscript and provided important intellectual input. All authors approved the final manuscript. Physicians and institutions who contributed patients to the SCQM registry without fulfilling the requirements for authorship are listed under www.scqm.ch/institutions.

Acknowledgements

We are grateful to the rheumatologists (listed under www.scqm.ch/institutions) and patients who participated in SCQM, thereby making this study possible. This project was financed in parts by AbbVie, Switzerland, and by AQUIRA 2009, an unrestricted research grant provided by Bristol-Myers-Squibb, Switzerland. The SCQM Foundation is supported by sponsors: www.scqm.ch/sponsors. SCQM sponsors had no role in study design, data collection, data analysis, data interpretation, writing of this manuscript or the decision to submit this manuscript. All members of the SONAR group for musculoskeletal US in the SCQM database are listed as authors.

Funding: This work supported by unrestricted research grants from AbbVie Switzerland and Bristol-Myers Squibb Switzerland.

Disclosure statement: The authors declare no conflicts of interest.

Supplementary data

Supplementary data are available at *Rheumatology* online.

References

1 Smolen JS, Landewé R, Bijlsma J *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.

2 Singh JA, Furst DE, Bharat A *et al*. 2012 update of the 2008 American College of Rheumatology recommendations for the use of disease-modifying antirheumatic drugs and biologic agents in the treatment of rheumatoid arthritis. Arthritis Care Res (Hoboken) 2012;64:625–39.

3 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. Arthritis Rheum 2011;63:573–86.

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 2011;70:404–13.

5 Molenaar ET, Voskuyl AE, Dinant HJ *et al*. Progression of radiologic damage in patients with rheumatoid arthritis in clinical remission. Arthritis Rheum 2004;50:36–42.

6 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.

7 Smolen JS, Han C, Bala M *et al*. Evidence of radiographic benefit of treatment with infliximab plus methotrexate in rheumatoid arthritis patients who had no clinical improvement: a detailed subanalysis of data from the anti-tumor necrosis factor trial in rheumatoid arthritis with concomitant therapy study. Arthritis Rheum 2005;52:1020–30.

8 Smolen JS, Avila JC, Aletaha D. Tocilizumab inhibits progression of joint damage in rheumatoid arthritis irrespective of its anti-inflammatory effects: disassociation of the link between inflammation and destruction. Ann Rheum Dis 2012;71:687–93.

9 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.

10 McQueen F, Naredo E. The 'disconnect' between synovitis and erosion in rheumatoid arthritis: a result of treatment or intrinsic to the disease process itself? Ann Rheum Dis 2011;70:241–4.

11 Landewé R, van der Heijde D, Klareskog L, van Vollenhoven R, Fatenejad S. Disconnect between inflammation and joint destruction after treatment with etanercept plus methotrexate: results from the trial of etanercept and methotrexate with radiographic and patient outcomes. Arthritis Rheum 2006;54:3119–25.

12 Emery P, Genovese MC, van Vollenhoven R *et al*. Less radiographic progression with adalimumab plus methotrexate versus methotrexate monotherapy across the spectrum of clinical response in early rheumatoid arthritis. J Rheumatol 2009;36:1429–41.

13 Dougados M, Devauchelle-Pensec V, Ferlet JF *et al*. The ability of synovitis to predict structural damage in rheumatoid arthritis: a comparative study between clinical examination and ultrasound. Ann Rheum Dis 2013;72:665–71.

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

15 Dohn UM, Ejbjerg B, Boonen A *et al*. No overall progression and occasional repair of erosions despite persistent inflammation in adalimumab-treated rheumatoid arthritis patients: results from a longitudinal comparative MRI, ultrasonography, CT and radiography study. Ann Rheum Dis 2011;70:252–8.

16 Foltz V, Gandjbakhch F, Etchepare F *et al*. Power Doppler ultrasound, but not low-field magnetic resonance imaging, predicts relapse and radiographic disease progression in rheumatoid arthritis patients with low levels of disease activity. Arthritis Rheum 2012;64:67–76.

17 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.

18 Haavardsholm EA, Aga AB, Olsen IC *et al*. Ultrasound in management of rheumatoid arthritis: ARCTIC randomised controlled strategy trial. BMJ 2016;354:i4205.

19 Møller-Bisgaard S, Hørslev-Petersen K, Ejbjerg B *et al*. Effect of Magnetic Resonance Imaging vs Conventional Treat-to-Target Strategies on Disease Activity Remission and Radiographic Progression in Rheumatoid Arthritis: the IMAGINE-RA Randomized Clinical Trial. JAMA 2019;321:461–72.

20 D'Agostino MA, Terslev L, Aegerter P *et al*. Scoring ultrasound synovitis in rheumatoid arthritis: a EULAR-OMERACT ultrasound taskforce-Part 1: definition and development of a standardised, consensus-based scoring system. RMD Open 2017;3:e000428.

21 Uitz E, Fransen J, Langenegger T, Stucki G. Clinical quality management in rheumatoid arthritis: putting theory into practice. Swiss Clinical Quality Management in Rheumatoid Arthritis. Rheumatology (Oxford) 2000;39:542–9.

22 Rau R, Wassenberg S, Herborn G, Stucki G, Gebler A. A new method of scoring radiographic change in rheumatoid arthritis. J Rheumatol 1998;25:2094–107.

23 Kyburz D, Gabay C, Michel BA, Finckh A; physicians of SCQM-RA. The long-term impact of early treatment of rheumatoid arthritis on radiographic progression: a population-based cohort study. Rheumatology (Oxford) 2011;50:1106–10.

24 Finckh A, Möller B, Dudler J *et al*. Evolution of radiographic joint damage in rituximab-treated versus TNF-treated rheumatoid arthritis cases with inadequate response to TNF antagonists. Ann Rheum Dis 2012;71:1680–5.

25 Möller B, Scherer A, Forger F, Villiger PM, Finckh A. Swiss Clinical Quality Management Program for Rheumatic D. Anaemia may add information to standardised disease activity assessment to predict radiographic damage in rheumatoid arthritis: a prospective cohort study. Ann Rheum Dis 2014;73:691–6.

26 Zufferey P, Brulhart L, Tamborrini G *et al*. Ultrasound evaluation of synovitis in RA: correlation with clinical disease activity and sensitivity to change in an observational cohort study. Joint Bone Spine 2014;81:222–7.

27 Scheel AK, Hermann KG, Kahler E *et al*. A novel ultrasonographic synovitis scoring system suitable for analyzing finger joint inflammation in rheumatoid arthritis. Arthritis Rheum 2005;52:733–43.

28 Ziswiler HR, Aeberli D, Villiger PM, Möller B. High-resolution ultrasound confirms reduced synovial hyperplasia following rituximab treatment in rheumatoid arthritis. Rheumatology (Oxford) 2009;48:939–43.

29 Zufferey P, Möller B, Brulhart L *et al*. Persistence of ultrasound synovitis in patients with rheumatoid arthritis fulfilling the DAS28 and/or the new ACR/EULAR RA

remission definitions: results of an observational cohort study. Joint Bone Spine 2014;81:426–32.

30 Brulhart L, Ziswiler HR, Tamborrini G, Zufferey P; SONAR/SCQM programmes. The importance of sonographer experience and machine quality with regards to the role of musculoskeletal ultrasound in routine care of rheumatoid arthritis patients. Clin Exp Rheumatol 2015;33:98–101.

31 DeMaeyer E, Adiels-Tegman M. The prevalence of anaemia in the world. World Health Stat Q 1985;38:302–16.

32 Prevoo ML, van't Hof MA, Kuper HH *et al*. 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 1995;38:44–8.

33 Wells G, Becker JC, Teng J *et al*. Validation of the 28-joint Disease Activity Score (DAS28) and European League Against Rheumatism response criteria based on C-reactive protein against disease progression in patients with rheumatoid arthritis, and comparison with the DAS28 based on erythrocyte sedimentation rate. Ann Rheum Dis 2009;68:954–60.

34 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 2007;21:663–75.

35 Fransen J, van Riel PL. DAS remission cut points. Clin Exp Rheumatol 2006;24(6 Suppl 43):S-29-32.

36 Aletaha D, Neogi T, Silman AJ *et al*. 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheum 2010;62:2569–81.

37 Yarwood A, Huizinga TW, Worthington J. The genetics of rheumatoid arthritis: risk and protection in different stages of the evolution of RA. Rheumatology (Oxford) 2016;55:199–209.

38 van der Heijde D. How to read radiographs according to the Sharp/van der Heijde method. J Rheumatol 2000;27:261–3.

39 Tanaka E, Yamanaka H, Matsuda Y *et al*. Comparison of the Rau method and the Larsen method in the evaluation of radiographic progression in early rheumatoid arthritis. J Rheumatol 2002;29:682–7.

40 Drossaers-Bakker KW, Amesz E, Zwinderman AH, Breedveld FC, Hazes JM. A comparison of three radiologic scoring systems for the long-term assessment of rheumatoid arthritis: findings of an ongoing prospective inception cohort study of 132 women followed up for a median of twelve years. Arthritis Rheum 2000;43:1465–72.

41 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.

42 Aletaha D. Nothing lasts forever - a critical look at sustained remission. Arthritis Res Ther 2012;14:116.

43 Mangnus L, van Steenbergen HW, Lindqvist E *et al*. Studies on ageing and the severity of radiographic joint damage in rheumatoid arthritis. Arthritis Res Ther 2015;17:222.

44 Aletaha D, Alasti F, Smolen JS. Rheumatoid factor determines structural progression of rheumatoid arthritis dependent and independent of disease activity. Ann Rheum Dis 2013;72:875–80.

45 Berglin E, Johansson T, Sundin U *et al*. Radiological outcome in rheumatoid arthritis is predicted by presence of antibodies against cyclic citrullinated peptide before and at disease onset, and by IgA-RF at disease onset. Ann Rheum Dis 2006;65:453–8.

46 Aletaha D, Smolen JS. Achieving clinical remission for patients with rheumatoid arthritis. JAMA 2019;321:457–8.