RHEUMATOLOGY

Rheumatology 2021;60:5549–5559
doi:10.1093/rheumatology/keab276
Advance Access publication 20 March 2021

Original article

Doppler ultrasound predicts successful discontinuation of biological DMARDs in rheumatoid arthritis patients in clinical remission

Lene Terslev [1,2,*], Cecilie Heegaard Brahe[1,*], Merete Lund Hetland[1,2,3], Stylianos Georgiadis[1], Karen Ellegaard[4], Lars Juul[5], Tuan Huynh[1,6], Uffe Møller Døhn[1], Viktoria Fana[1], Torsten Møller[1], Simon Krabbe[1], Lykke Midtbøll Ørnbjerg[1], Daniel Glinatsi[1,7], Henrik Røgind[1], Anette Hansen[5], Jesper Nørregaard[1], Søren Jacobsen[2,8], Dorte Vendelbo Jensen[3,5], Natalia Manilo[9], Karsten Asmussen[9], Mikael Boesen[10], Zoreh Rastiemadabadi[10], Lone Morsel-Carlsen[10], Jakob Møllenbach Møller [11], Niels Steen Krogh[12] and Mikkel Østergaard[1,2]

Abstract

Objective. To assess the ability of ultrasound to predict successful tapering and successful discontinuation of biological DMARDs (bDMARDs) at the 2-year follow-up in RA patients in sustained remission.

Methods. Patients in sustained remission (DAS28-CRP ≤ 2.6) and with no radiographic progression the previous year tapered bDMARDs according to a standardized regime. A total of 119 of these patients were included in this ultrasound substudy. At baseline, clinical assessment, MRI, X-ray and ultrasound of 24 joints were performed. Ultrasound-detected synovitis was defined and scored 0–3 using the OMERACT scoring system at the joint level for both grey-scale and Doppler activity. Sum scores for each ultrasound modality were calculated for 24 joints at the patient level. The final state of treatment was assessed after 2 years. The predictive value of ultrasound measures for successful tapering and discontinuation at the 2-year follow-up was assessed via logistic regression analyses.

Results. Negative IgM-RF [odds ratio (OR) = 0.29, 95% CI: 0.10–0.85; $P = 0.024$] and lower Doppler sum score of 24 joints (OR = 0.44, 95% CI: 0.15, 0.87; $P = 0.014$) were independent predictors for successful discontinuation of bDMARDs at the 2-year follow-up. The predictive value of the Doppler sum score was independent of MRI findings. Previous numbers of bDMARDs were predictive of successful tapering (OR = 0.58, 95% CI: 0.35, 0.91; $P = 0.018$), whereas ultrasound was not. Clinical parameters were not predictive of successful tapering/discontinuation.

Conclusion. Doppler sum score was an independent predictor for successful discontinuation of bDMARDs at the 2-year follow-up—the odds for achieving successful discontinuation decreased by 56% per one-unit increase in Doppler sum score. Ultrasound could not predict successful tapering.

Key words: ultrasound, Doppler, clinical remission, tapering, discontinuation, bDMARD, discontinuation, synovial hypertrophy, T2T strategy, ultrasound remission

[1]Center for Rheumatology and Spine Diseases, Center of Head and Orthopaedics, Glostrup, Rigshospitalet, Capital Region, Denmark, [2]Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark, [3]DANBIO Registry, Center for Rheumatology and Spine Diseases, Center of Head and Orthopaedics, Glostrup, Rigshospitalet, Capital Region, Denmark, [4]The Parker Institute, Bispebjerg-Frederiksberg Hospital, Copenhagen, Denmark, [5]Center for Rheumatology and Spine Diseases, Center of Head and Orthopedics, Gentofte, Rigshospitalet, Denmark, [6]Holte Rheumatology Clinic, Holte, Capital Region, Denmark, [7]Department of Rheumatology, Skaraborg Hospital, Skövde, Sweden, [8]Copenhagen Lupus and Vasculitis Clinic, Center for Rheumatology and Spine Diseases, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark, [9]Center for Rheumatology and Spine Diseases, Center of Head and Orthopedics, Rigshospitalet, Frederiksberg,, Denmark, [10]Department of Radiology, Bispebjerg-Frederiksberg Hospital, Copenhagen, Denmark, [11]Department of Radiology, Copenhagen University Hospital Herlev-Gentofte, Herlev, Denmark and [12]Zitelabs Aps, Copenhagen, Denmark

Submitted 21 December 2020; accepted 10 March 2021

Correspondence to: Lene Terslev, Copenhagen Center for Arthritis Research, Center for Rheumatology and Spine Diseases, Copenhagen University Hospital, Rigshospitalet, Valdemar Hansensvej 17, 2600 Glostrup, Copenhagen, Denmark. E-mail: lene.terslev.01@regionh.dk

*Lene Terslev and Cecilie Heegaard Brahe contributed equally to this study.

CLINICAL SCIENCE

Rheumatology key messages

- Doppler sum score was an independent predictor for successful discontinuation of bDMARDs at the 2-year follow-up.
- Ultrasound could not predict successful tapering.

Introduction

EULAR treatment recommendations advise tapering of biological DMARD (bDMARD) therapy in patients with RA who have achieved stable clinical remission [1]. In addition, tapering is relevant to reduce costs and because of safety issues for long-term use of bDMARDs [2, 3]. Several randomized clinical trials have reported successful dose tapering or discontinuation in some, but not all, patients [4–10]. A recent systematic literature review highlighted that discontinuation (but not tapering of bDMARDs) was associated with radiographic progression and increased risk of losing remission [11]. Data are needed to correctly identify those patients who may successfully taper or discontinue their bDMARDs.

Previously, we have reported that the number of previous bDMARDs, gender and baseline MRI combined inflammation and combined damage scores may be valuable in predicting successful tapering [12]. However, ultrasound is more widely used as a clinical tool in rheumatology routine care. Previous studies have shown that the majority of RA patients in clinical remission have subclinical synovitis by ultrasound, independent of applied clinical composite scores [13–16]. This subclinical synovitis has been demonstrated to be related to clinical flare and erosive progression on X-ray [17–22], and follow-up studies of 6–12 months have explored its potential predictive value for successful tapering or discontinuation of bDMARDs [23–26]. Different definitions of ultrasound remission exist, and different numbers of joints have been assessed for detecting subclinical inflammation [13–16].

In the current substudy, the aim was to assess whether ultrasound had a predictive value in addition to (i) clinical and demographic parameters and (ii) clinical, demographic and MRI parameters in a cohort of RA patients in longstanding clinical remission on bDMARD therapy for (a) successful tapering at 2-year follow-up and (b) successful discontinuation at 2-year follow-up.

Methods

The Capital Region of Denmark implemented in 2013 a guideline for mandatory tapering of bDMARDs in all RA patients in stable remission for at least 1 year [A Dose OPTimization of Biological Therapy (ADOPT)] (see reference [12] for details). Patients covered by the tapering guideline were offered participation in the current ultrasound substudy.

Of the 143 patients tapered according to the ADOPT guideline [12], 132 patients accepted participation in the ultrasound substudy, while 11 declined. Of the 132 patients included in the substudy, 13 patients were excluded from the analysis (5 patients had missing baseline ultrasound visits, 5 had a missing 2-year clinical visit and 3 were lost to follow-up). Thus, 119 patients were eligible for analysis.

The ultrasound assessments did not influence the tapering regimen. The ultrasound substudy was a research project approved by the research ethical committee of the Capital Region of Denmark (Protocol number: H-1-2012-127) and the Danish Medicines Agency. All patients gave written informed consent, and the study was carried out following the Declaration of Helsinki and the Good Clinical Practice guidelines. For the main study [12], which was published separately, as this was based on the mandatory clinical treatment guideline, no ethical approval was needed according to Danish law.

ADOPT guideline

In Denmark, bDMARDs are provided for RA patients by public hospitals and are paid for by a tax-based system, with no bDMARD-related expenses for the individual patient. The patients are monitored in the clinical DANBIO registry [27].

The ADOPT guideline was implemented in the Capital Region of Denmark between April 2013 and February 2015. The included patients fulfilled the ACR 1987 criteria and/or ACR/EULAR 2010 classification criteria for RA [28, 29] and had achieved and maintained clinical remission (DAS28-CRP $\leq$ 2.6) for at least 1 year as documented by at least three consecutive clinical visits in the DANBIO registry, while being treated with bDMARDs (adalimumab, etanercept, infliximab, tocilizumab, certolizumab, golimumab or abatacept). Patients who had previously unsuccessfully attempted tapering, patients with erosive progression on X-ray the previous year despite clinical remission, and patients who had received glucocorticoid treatment 6 months prior to baseline were excluded from the mandatory tapering regimen.

Treatment algorithm

According to the ADOPT guideline [12], bDMARDs were tapered stepwise following a predefined algorithm. At the time of inclusion (baseline), the dose was tapered to

two-thirds of the standard dose; at week 16 the dose was tapered to half the standard dose; and at week 32 the bDMARD was discontinued (Supplementary File S1, available at *Rheumatology* online). Dose reduction was only carried out if the patient was still in clinical remission. Flare was defined as a DAS28-CRP $\geq$ 2.6 with ΔDAS28-CRP $\geq$ 1.2 from baseline, and it resulted in bDMARD dose escalation to the previous step, with no further tapering being attempted. The same applied if erosive progression on MRI or conventional radiography was reported by a radiologist during the period of the tapering. The bDMARD dose was escalated, if needed, every 4 months until the patient achieved remission, and the final treatment dose was assessed after 2 years.

Clinical and laboratory assessments

Patient-reported outcomes (HAQ), Visual Analogue Scale score for pain, fatigue and global assessment of disease activity, DAS28 joints using CRP (DAS28-CRP), Simplified Disease Activity Index, Clinical Disease Activity Index, ACR/EULAR Boolean remission together with number of swollen and tender joints, physician global assessment (Visual Analogue Scale) and CRP were assessed at baseline and at weeks 4, 8, 16, 24, 32, 40, 48 and 70. In the case of disease flare, a flare-visit was performed, with subsequent follow-up at 8, 16 and 24 weeks after the flare. All patients (with or without flare) had a final clinical visit as a 2-year follow-up (96 weeks).

Imaging

Ultrasound

GE Logiq® E9 R5 (Milwaukee, Wisconsin, USA) ultrasound machines with a 5–16 ML linear array transducer were used for all examinations. Colour Doppler settings for slow flow were kept unchanged throughout the study. The Doppler frequency was set at 7.5 MHz, pulse repetition frequency at 0.4 kHz, colour priority at 100%, and Doppler gain just below the noise level. The size and position of the colour box was set to go to the top of the image to recognize artefacts caused by vessels above the joint [30].

Signs of synovitis were assessed at baseline prior to tapering. Grey-scale (GS) and Doppler ultrasound were performed in 24 joints [elbow, wrist (radiocarpal, midcarpal, distal radio-ulnar joint, using the highest score as representative for the wrist), MCP joints 2–5, knee, ankle and MTP joints 2–5, bilaterally]. Each joint was scored using the OMERACT-EULAR semi-quantitative scorings system (0–3) for GS synovial hypertrophy and for Doppler activity—separately and in combination using the Global OMERACT-EULAR combined Synovitis Score (GLOESS) [31, 32]. At the patient level, an ultrasound sum score of the 24 joints was made for GS synovial hypertrophy, Doppler activity and the GLOESS—each with a range from 0 to 72. To assess whether evaluating hands-only sufficed, a sum score was subsequently calculated for the hands-only—ranging from 0 to 30.

To assess the impact of different ultrasound remission criteria for predicting successful tapering and discontinuation, we applied three different remission definitions for the 24 assessed joints: Strict ultrasound remission (no joints with a GS score >0 and a Doppler score >0); Semi-strict US remission (no joints with a GS score >1 and no joints with a Doppler score >0); Doppler ultrasound remission (no joints with a Doppler score >0, irrespective of the GS score).

Ultrasound was performed by eight sonographers (from four centres) with longstanding experience in musculoskeletal ultrasound (>10 years) and trained in the OMERACT-EULAR synovitis scoring system. Agreement on scoring was obtained prior to study initiation. No inter- or intra-reader agreement statistics are available.

Ultrasound examinations were performed at the same follow-up visits as the clinical examinations. The examination time was ~20 min per patient.

The rheumatologists performing the ultrasound examinations were not involved in the clinical evaluation or the clinical decisions made in the main study.

MRI and X-ray

Conventional radiographs of the hands, wrists and forefeet and MRI of the dominant wrist and MCP 2-5 joints were acquired at baseline, week 16 (only MRI), week 32 and year 2, and evaluated after each examination for absence/presence of erosive progression [12].

After study completion, radiographs were scored by a trained reader according to the Sharp van der Heijde method [33], and baseline MRIs were scored according to the OMERACT RAMRIS (rheumatoid arthritis MRI scoring system) [34–36] by another experienced reader. For MRI, a combined inflammation score and a combined damage score were calculated (see reference [12] for details). The readers were blinded to patient data and chronologic order.

Statistical analysis

According to a predefined Statistical Analysis Plan, the analyses included the predictive value of the sum scores for 24 joints and for the hands-only. Additionally, the predictive value of ultrasound remission was assessed.

Descriptive statistics were applied for demographics and baseline disease activity measures. The Kruskal–Wallis test, χ^2 test and Fisher's exact test were used, as appropriate, for comparisons at the 2-year follow-up: patients on a reduced dose of bDMARDs *vs* those on a full dose, and patients who had discontinued bDMARD therapy *vs* those who had not. Changes in clinical and ultrasound inflammatory parameters from baseline to the 2-year follow-up were assessed via the Wilcoxon signed-rank test and McNemar's test, as appropriate; $P < 0.05$ was considered statistically significant.

Logistic regression analyses on imputed data were used to identify variables associated with successful tapering and successful discontinuation at the 2-year follow-up. For all models, 13 demographic, clinical and

radiographic variables were included as potential predictors: gender, current smoking status, IgM-RF positivity, anti-CCP positivity and ACR/EULAR remission were tested as categorical variables, while age, time since diagnosis, time in remission before tapering, number of previous bDMARDs, HAQ score, DAS28-CRP and Total Sharp van der Heijde Score (TSS) were tested as numeric variables.

For each of the two dependent variables, two models were considered that relied on the independent ultrasound variables included in the analyses: ultrasound for 24 joints and ultrasound for hands-only. In the main analyses, GS and Doppler sum scores were included as ultrasound inflammatory variables for 24 joints and hands-only assessments. Additional analyses were performed, including MRI variables (combined inflammation and combined damage), and by replacing GS and Doppler scores with GLOESS.

Missing data in independent baseline variables were imputed with multiple imputation by chained equations (40 imputed datasets, i.e. approximately the percentage of patients with incomplete data). Variables with a $P < 0.25$ in univariate analyses were included in the initial multivariate model. Backward selection with a significance level of 0.05 was performed in stacked imputed datasets after applying a fixed weight to all observations, accounting for the average fraction of missing data across all variables under consideration. The potential significance of each baseline variable excluded in univariate analyses was tested by reintroducing these variables one at a time into the multivariate model. The linearity of continuous predictors in the logit scale was checked and, whenever the linearity assumption was violated, non-linear continuous predictors were categorized based on quartiles (when applicable). The statistical significance of interactions between predictors was also tested. Once the final multivariate model was selected, the model was fitted in one of two ways: if missing values were present in predictors, all imputed datasets were used and the results were pooled using Rubin's rules; if there were no missing values in predictors, the non-imputed data were used. Internal validation by bootstrapping was performed for all models, and the area under the receiver operating characteristic curve (AUC) was estimated along with its 95% CI. All models were assessed via likelihood ratio tests. The results of logistic analyses are presented by odds ratio (OR), 95% CI of the OR and *P*-value of the likelihood ratio test.

The statistical analyses were performed using R software version 3.6.1.

Results

Baseline parameters and group differences are shown in Table 1. At the 2-year follow-up, 45 patients (38%) had been re-escalated to the standard dose, 56 patients (47%) had tapered to a lower dose than the standard dose (23 patients to two-thirds dose and 33 patients to one-half dose), and 18 patients (15%) had been able to discontinue bDMARDs.

The GS sum score was median 5 [interquartile range (IQR) 2–9] and the Doppler sum score was 0 (IQR = 0–2). Eight percent of the patients were in strict ultrasound remission, 14% in semi-strict ultrasound remission and 56% in Doppler remission.

Baseline differences in patients in relation to successful tapering and successful discontinuation

In the group of patients who were back on the full dose at the 2-year follow-up, the number of previous bDMARDs ($P = 0.021$) was statistically significantly higher, more were women ($P = 0.012$), and the combined inflammation MRI scores ($P = 0.020$) and combined damage MRI scores ($P = 0.006$) were higher than in the group of patients who had successfully tapered at the 2-year follow-up. There were no statistically significant differences in ultrasound parameters. The number of available observations for each baseline parameter is shown in Supplementary Table S2, available at *Rheumatology* online.

Patients who could not discontinue bDMARDs were statistically significantly more frequently IgM-RF positive ($P = 0.009$) and had higher 24-joint Doppler sum scores ($P = 0.009$) than patients who could discontinue treatment. Significantly more patients were in Doppler remission among those who could discontinue treatment ($P = 0.018$).

Changes over time for clinical and ultrasound measures

At baseline, subclinical synovitis was found in 92% of the patients i.e. only 8% were in strict ultrasound remission, whereas 56% of the patients were in Doppler remission. This did not change at the 2-year follow-up (Table 2). The Doppler 24-joint sum score was unchanged between baseline and follow-up, whereas the GS 24-joint sum score increased significantly from baseline to the 2-year follow-up ($P = 0.001$), as did the GLOESS 24-joint sum score ($P = 0.001$). Though small, a statistically significant increase was also seen in tender and swollen joints, patient pain, physician global and in all clinical composite scores for remission except ACR/EULAR remission—see Table 2. No significant change was seen in the TSS from baseline to 2-year follow-up.

Predictors for successful tapering

Final multivariate logistic regression models for successful tapering at the 2-year follow-up are shown in Table 3. All univariate and multivariate logistic regression analyses in stacked imputed datasets can be seen in Supplementary Tables S3 and S5, available at *Rheumatology* online.

GS and Doppler sum scores for both 24 joints and for the hands-only did not have a predictive value for successful tapering at the 2-year follow-up (Table 3). Nor

Table 1 Demographics and disease activity measures at baseline

	All ($n = 119$)	Successful tapering ($n = 74$)	Full dose ($n = 45$)	*P*-value	Successful discontinuation ($n = 18$)	Therapy not discontinued ($n = 101$)	*P*-value
Demographic characteristics							
Women (%)	67	58	82	0.012*	67	67	1
Age, median (IQR), years	60 (47–68)	60 (47–67)	60 (48–69)	0.518	55 (48–66)	60 (47–68)	0.689
Disease duration, median (IQR), years	11 (7–18)	11 (7–16)	11 (7–20)	0.425	10 (6–14)	11 (7–18)	0.360
Smoking current	19%	19%	19%	1	31%	17%	0.184
Time in remission before tapering, median (IQR), years	2 (1–3)	2 (1–3)	2 (1–3)	0.280	2 (2–3)	2 (1–3)	0.636
Concominant DMARD, %	87	85	91	0.405	94	86	0.463
Number of previous bDMARDs, %				0.021*			0.320
0	63	72	49		83	59	
1	25	23	29		17	27	
2	7	3	13		0	8	
≥3	5	3	9		0	6	
Previous bDMARD, %				0.915			1
Adalimumab	16	19	13		0	17	
Etanercept	20	24	17		33	20	
Infliximab	59	57	61		67	59	
Tocilizumab	2	0	4		0	2	
Abatacept	2	0	4		0	2	
RF positive	68	66	71	0.724	39	73	0.009*
Anti-CCP positive	79	76	84	0.365	72	80	0.530
Clinical measures							
Tender joints count (0–28), median (IQR)	0 (0–0)	0 (0–0)	0 (0–0)	0.232	0 (0–0)	0 (0–0)	0.144
Swollen joints count (0–28), median (IQR)	0 (0–0)	0 (0–0)	0 (0–0)	0.173	0 (0–0)	0 (0–0)	0.461
Patient global (0–100), median (IQR)	12 (4–25)	10 (4–24)	17 (5–26)	0.185	14 (2–28)	12 (4–25)	0.997
Patient pain (0–100), median (IQR)	11 (3–20)	8 (2–19)	12 (4–24)	0.159	6 (2–24)	11 (4–19)	0.578
Physician global (0–100)	0 (0–4)	0 (0–5)	0 (0–3)	0.231	1 (0–4)	0 (0–4)	0.477
CRP (mg/l), median (IQR)	5 (2–6)	5 (3–8)	4 (1–5)	0.049*	4 (3–6)	5 (2–6)	0.898
HAQ (0–3), median (IQR)	0.2 (0.0–0.8)	0.1 (0.0–0.6)	0.5 (0.0–1.0)	0.071	0.1 (0.0–0.4)	0.2 (0.0–0.8)	0.318
DAS28–CRP, median (IQR)	1.8 (1.6–2.1)	1.8 (1.6–2.1)	1.8 (1.7–2.1)	0.845	1.8 (1.6–1.9)	1.8 (1.7–2.1)	0.490
SDAI, median (IQR)	2.2 (1.2–3.5)	2.0 (1.2–3.3)	2.7 (1.4–3.7)	0.373	1.8 (0.8–3.7)	2.3 (1.3–3.5)	0.550
CDAI, median (IQR)	1.7 (0.6–3.0)	1.2 (0.6–2.6)	2.4 (0.8–3.0)	0.241	1.4 (0.4–3.2)	1.7 (0.7–2.9)	0.643
ACR/EULAR Boolean remission	37	42	29	0.219	39	37	1
Radiographic measures							
TSS (0–448), median (IQR)	14 (4–45)	10 (5–33)	20 (3–50)	0.312	9 (1–45)	14 (5–45)	0.405
JSN (0–84), median (IQR)	9 (2–34)	8 (2–22)	14 (2–35)	0.346	8 (0–38)	10 (2–34)	0.483
X-ray erosion, median (IQR)	3 (0–12)	3 (0–12)	6 (1–16)	0.237	2 (0–6)	4 (0–15)	0.357
Presence of X-ray erosion	73	72	75	0.911	72	73	1
MRI measures							
Combined inflammation (0–144), median (IQR)	6 (3–12)	6 (2–12)	10 (5–16)	0.020*	6 (2–10)	7 (3–13)	0.417
Combined damage (0–314), median (IQR)	4 (1–17)	2 (1–11)	7 (3–39)	0.006*	4 (1–14)	4 (1–20)	0.757
Ultrasound inflammatory measures							
Grey-scale sum score hands-only (0–30), median (IQR)	2 (0–5)	2 (0–5)	3 (1–5)	0.417	1 (0–2)	3 (1–5)	0.036*
Grey-scale sum score (0–72), median (IQR)	5 (2–9)	5 (2–9)	5 (3–9)	0.873	3 (2–6)	5 (2–9)	0.142

(continued)

TABLE 1 Continued

	All ($n = 119$)	Successful tapering ($n = 74$)	Full dose ($n = 45$)	*P*-value	Successful discontinuation ($n = 18$)	Therapy not discontinued ($n = 101$)	*P*-value
Doppler sum score hands-only (0–30), median (IQR)	0 (0–1)	0 (0–1)	0 (0–2)	0.274	0 (0–0)	0 (0–1)	0.043*
Doppler sum score (0–72), median (IQR)	0 (0–2)	0 (0–1)	1 (0–2)	0.147	0 (0–0)	0 (0–2)	0.009*
GLOESS hands-only (0–30), median (IQR)	2 (0–5)	2 (0–5)	3 (1–5)	0.363	1 (0–2)	3 (1–5)	0.026*
GLOESS (0–72), median (IQR)	5 (2–9)	5 (2–9)	5 (3–9)	0.811	3 (2–6)	5 (3–9)	0.120
Ultrasound strict remission	8	9	7	0.741	0	10	0.356
Ultrasound semi-strict remission	14	16	11	0.591	11	15	1
Ultrasound Doppler remission	56	62	47	0.144	83	51	0.018*

Successful tapering: patients on less than full dose of bDMARD at 2-year follow-up. Full dose: patients on full dose of bDMARD at 2-year follow-up. Discontinued: patients who discontinued bDMARDs at 2-year follow-up. Not discontinued: patients who did not discontinue bDMARDs at 2-year follow-up. Ultrasound strict remission: Grey-scale sum score = 0 and Doppler sum score = 0. Ultrasound semi-strict remission: Grey-scale sum score ≤ 1 and Doppler sum score = 0. Ultrasound Doppler remission: Doppler sum score equal to 0. *P*-values for differences between treatment groups by Kruskal–Wallis test for numeric variables and χ^2 test or Fisher's exact test, as appropriate, for binary and categorical variables; *$P < 0.05$. [b]DMARD: biological DMARD; CDAI: Clinical Disease Activity Index; GLOESS: Global OMERACT/EULAR ultrasound Synovitis Score; IQR: interquartile range; JSN: joint space narrowing; SDAI: Simple Disease Activity Index; TSS: Total Sharp van der Heijde score.

were GLOESS sum scores independent predictors for successful tapering (results not shown).

The number of previous bDMARDs (OR = 0.58, 95% CI: 0.35, 0.91; $P = 0.018$) and gender (female) (OR = 0.37, 95% CI: 0.14, 0.91; $P = 0.031$) were predictors for successful tapering (Table 3). When both MRI and ultrasound were considered as independent variables in the prediction analyses, the number of previous bDMARDs (OR = 0.47; 95% CI: 0.28, 0.75; $P = 0.001$) and MRI combined damage (OR = 0.99; 95% CI: 0.98, 1.00; $P = 0.025$) were independent predictors, but gender was not (Table 4).

Predictors for successful discontinuation

Independent predictors for successful discontinuation at the 2-year follow-up are shown in Table 3. Univariate and multivariate logistic regression analyses in stacked imputed datasets are presented in Supplementary Tables S4 and S6, available at *Rheumatology* online.

The baseline Doppler 24-joint Doppler sum score was an independent predictor of successful discontinuation at the 2-year follow-up (OR = 0.44, 95% CI: 0.15, 0.87; $P = 0.014$), along with RF positivity (OR = 0.29, 95% CI: 0.10, 0.85; $P = 0.024$). Neither the GS 24-joint sum score nor the GS or Doppler sum scores for the hands-only showed predictive value (Table 3), and neither did any of the GLOESS sum scores (data not shown).

Considering a clinical scenario in which a sensitive Doppler modality was unavailable, we assessed whether GS 24-joint sum score alone had predictive value for successful discontinuation by eliminating the Doppler 24-joint sum score completely from the analyses. A GS 24-joint sum score had no predictive value for successful discontinuation (results not shown).

Doppler ultrasound remission for the 24 joints (but not for strict and semi-strict remission) was tested as a predictor for successful discontinuation, and no independent predictive value was found.

Including MRI variables (combined structural damage score and combined inflammation score) in the model did not affect the predictive value of the Doppler 24-joint sum score (Table 4).

Linearity and interactions in logistic regression models

None of the continuous predictors in the multivariate regression models was detected as non-linear. No statistically significant interactions were found.

Performance of statistical models

The performance of the models was assessed by AUC. The estimated AUCs of all models were considered as acceptable. For successful discontinuation, AUC was 0.73 when including ultrasound measures (either 24 joints or hands-only) as independent variables. For successful tapering, an AUC of 0.69 (respectively 0.67) was estimated when MRI measures were introduced (respectively excluded) in the models.

Table 2 Changes from baseline to 2-year follow-up

	Baseline ($n = 119$)	2-year follow-up ($n = 119$)	*P*-value
Clinical measures			
Tender joints count (0–28), median (IQR)	0 (0–0)	0 (0–0)	<0.001*
Swollen joints count (0–28), median (IQR)	0 (0–0)	0 (0–0)	0.005*
Patient global (0–100), median (IQR)	12 (4–25)	13 (5–26)	0.159
Patient pain (0–100), median (IQR)	11 (3–20)	11 (5–24)	0.001*
Physician global (0–100), median (IQR)	0 (0–4)	2 (0–6)	0.003*
CRP (mg/l), median (IQR)	5 (2–6)	4 (1–7)	0.540
HAQ (0–3), median (IQR)	0.2 (0.0–0.8)	0.4 (0.0–0.9)	0.285
DAS28-CRP, median (IQR)	1.8 (1.6–2.1)	1.9 (1.6–2.3)	0.011*
SDAI, median (IQR)	2.2 (1.2–3.5)	2.4 (1.3–5.0)	0.011*
CDAI, median (IQR)	1.7 (0.6–3.0)	2.0 (0.8–4.5)	0.008*
ACR/EULAR remission	37%	34%	0.735
Radiographic measures			
TSS (0–448), median (IQR)	14 (4–45)	14 (3–47)	0.451
JSN (0–84), median (IQR)	9 (2–34)	9 (2–34)	0.513
X-ray erosion, median (IQR)	3 (0–12)	4 (0–13)	0.402
Presence of X-ray erosion	73%	71%	0.617
Ultrasound inflammatory measures			
Grey-scale sum score hands-only (0–30), median (IQR)	2 (0–5)	2 (1–5)	0.229
Grey-scale sum score (0–72), median (IQR)	5 (2–9)	7 (3–10)	0.001*
Doppler sum score hands-only (0–30), median (IQR)	0 (0–1)	0 (0–1)	0.993
Doppler sum score (0–72), median (IQR)	0 (0–2)	0 (0–2)	0.628
GLOESS hands-only (0–30), median (IQR)	2 (0–5)	3 (1–5)	0.246
GLOESS (0–72), median (IQR)	5 (2–9)	7 (3–10)	0.001*
Ultrasound strict remission, %	8	4	0.267
Ultrasound semi-strict remission, %	14	8	0.383
Ultrasound Doppler remission, %	56	57	1

P-values for differences between visits from modified Wilcoxon signed-rank test (Pratt method to handle ties) for numeric variables and McNemar's test for binary variables; *$P < 0.05$. CDAI: Clinical Disease Activity Index; GLOESS: Global OMERACT/EULAR ultrasound Synovitis Score; IQR: Interquartile Range; JSN: Joint Space Narrowing; SDAI: Simple Disease Activity Index; TSS: Total Sharp van der Heijde Score. Ultrasound strict remission: Grey-scale sum score equal to 0 and Doppler sum score equal to 0. Ultrasound semi-strict remission: Grey-scale sum score less than or equal to 1 and Doppler sum score equal to 0. Ultrasound Doppler remission: Doppler sum score equal to 0.

Discussion

In this prospective study of longstanding RA patients in sustained DAS28-CRP remission on a bDMARD for at least 1 year, a lower Doppler sum score of 24 joints prior to tapering was an independent predictor for successful discontinuation of bDMARDs at the 2-year follow-up. Thus, a one-unit increase in the Doppler 24-joint sum score decreased the odds for achieving successful discontinuation at 2 years by 56%. Ultrasound had no independent predictive value for successful tapering to two-thirds or one-half of the dose. Adding MRI variables along with US variables to the prediction model for successful discontinuation did not affect the predictive value of the Doppler 24-joint sum score.

We found that 92% of the patients had subclinical synovitis, and 56% were in Doppler remission at baseline. Ultrasound-detected residual synovitis is frequent in RA patients in clinical remission and has been shown to predict flare and structural progression [17]. The presence of subclinical synovitis assessed by ultrasound led in general to explore its value for predicting the patients who could successfully taper and discontinue bDMARDs. The first study to address this was unable to find any predictive value of ultrasound, and the only parameter with predictive value for successful discontinuation was short duration of untreated symptoms [26]. Later, two studies found Doppler activity to be related to unsuccessful tapering, both at short-and long-term follow-up [24, 37]. Doppler activity but also GS synovial hypertrophy have in one study been predictive for unsuccessful discontinuation of bDMARDs without prior tapering [25]. In our study, a lower Doppler 24-joint sum score was found to be an independent predictor for discontinuation without flare, but not for tapering to a reduced dose of a bDMARD. Furthermore, the GS 24-joint sum score had no predictive value, not even when eliminating Doppler from our predictive analyses and only considering the GS 24-joint sum score. This finding indicates that ultrasound equipment with insensitive Doppler has limited value when assessing patients in remission. in general

Though our data cannot predict successful tapering, our results alone do not indicate that having some

TABLE 3 Multivariate logistic regression analyses for successful tapering and successful discontinuation, including US variables

Independent variables	Successful tapering			
	Ultrasound[a]		Ultrasound (hands-only)[a]	
	OR (95% CI)	*P*-value[b]	OR (95% CI)	*P*-value[b]
Female	0.37 (0.14–0.91)	0.031	0.37 (0.14–0.91)	0.031
Number of previous bDMARDs	0.58 (0.35–0.91)	0.018	0.58 (0.35–0.91)	0.018
AUC (95% CI)[c]	0.67 (0.57–0.77)		0.67 (0.58–0.77)	
	Successful discontinuation			
	Ultrasound[a]		**Ultrasound (hands-only)[d]**	
Smoking (current)	–	–	4.46 (1.14, 17.91)	0.027
RF positive	0.29 (0.10–0.85)	0.024	0.16 (0.04–0.51)	0.001
Doppler sum score	0.44 (0.15–0.87)	0.014	–	–
AUC (95% CI)[c]	0.73 (0.60–0.83)		0.73 (0.62–0.83)	

[a]Results derived from non-imputed data. Predictors were selected by applying backward selection in stacked data. CIs given as profile likelihood CIs. AUC estimated based on internal validation by bootstrapping with 1000 samples. [b]*P*-values by likelihood ratio tests. [c]The bootstrap 0.632+ estimate was calculated to correct for optimism. [d]Results derived from imputed datasets, where model estimates were pooled based on Rubin's rules. Predictors were selected by applying backward selection in stacked data. Profile likelihood CIs calculated according to the Pseudo-Variance modification of Rubin's rule (PVR). AUC estimated based on internal validation by bootstrapping with 100 samples per imputed dataset. AUC: area under the curve; bDMARD: biological DMARD; OR: odds ratio.

Doppler activity precludes attempted tapering of bDMARDs. All the patients that were on a reduced dose (but had not discontinued) at the 2-year follow-up had experienced a flare during the tapering regimen. We subsequently plan to investigate whether the status before each dose-reduction step can predict flaring in the subsequent period.

Compared with more widespread ultrasound assessment of joints, ultrasound of the hands-only has been shown to detect >90% of subclinical synovitis in patients in remission [38], suggesting that ultrasound of the hands could be sufficient for examination in routine care. However, neither GS, Doppler nor GLOESS sum scores of the hands had any predictive value in our study, which indicates that assessing the hands-only is not sufficient when considering tapering or discontinuation.

There is no consensus on how to define ultrasound remission and hence we assessed three different definitions at baseline. However, only Doppler ultrasound remission was tested as a predictor for successful discontinuation, and no independent predictive value was found. The decision to eliminate strict and semi-strict ultrasound remission as items in the prediction analyses was based on the Doppler 24-joint sum score being an independent predictor while GS sum score was not. Furthermore, only a few patients were in the strict and semi-strict ultrasound remission groups. The fact that Doppler ultrasound remission had no predictive value indicates that when applying high-end ultrasound equipment with a sensitive Doppler modality, a Doppler score equal to 0 might not be the correct cut-off. The presence of Doppler activity has been reported in healthy controls—more frequently in the wrist than in the PIP joints [39, 40], and a cut-off between healthy findings and pathology may prove helpful for determining what constitutes ultrasound remission.

Only demographic parameters but no clinical parameters had predictive value for successful tapering or discontinuation. This contrasts with the study by Naredo *et al.*, in which the DAS28 prior to tapering was predictive for successful tapering [24]. However, our findings are in line with the study by Saleem *et al.*, in which only short symptom duration was of predictive value [26]. The lack of clinical predictors could be related to the sparse clinical residual disease activity in our cohort (only 11 and 3 out of 119 patients had tender and swollen joint counts, respectively). Though the patients were all in DAS28CRP remission at the 2-year follow-up without significant change in TSS, we found a statistically significant increase in tender and swollen joints, patient pain, physician global and in all clinical composite scores for remission except ACR/EULAR remission. Furthermore, an increase in GS 24-joint sum score but not Doppler was found. Whether any of these parameters have a predictive value for flare or persistent remission beyond the 2-year follow-up needs to be established in the future.

The strengths of this study are the homogeneous patient cohort, the long follow-up time, and that the study was done in routine care. We have used the OMERACT-validated synovitis scoring system (GS, Doppler and GLOESS) and assessed multiple joints, providing a

TABLE 4 Multivariate logistic regression analyses for successful tapering and successful discontinuation, including MRI and ultrasound variables

Independent variables	Successful tapering			
	MRI and Ultrasound[a]		MRI and Ultrasound (hands-only)[a]	
	OR (95% CI)	*P*-value[b]	OR (95% CI)	*P*-value[b]
Number of previous bDMARDs	0.47 (0.28–0.75)	0.001	0.47 (0.28–0.75)	0.001
MRI combined structural damage score	0.99 (0.98–1.00)	0.017	0.99 (0.98–1.00)	0.017
AUC (95% CI)[c]	0.69 (0.59–0.79)		0.69 (0.59–0.79)	
	Successful discontinuation			
	MRI and Ultrasound[d]		**MRI and Ultrasound (hands-only)[a]**	
Smoking (current)	–	–	4.46 (1.14, 17.91)	0.027
RF positive	0.29 (0.10–0.85)	0.024	0.16 (0.04–0.51)	0.001
Doppler sum score	0.44 (0.15–0.87)	0.014	–	–
AUC (95% CI)[c]	0.73 (0.61–0.83)		0.73 (0.62–0.83)	

MRI combined damage score refers to the sum of bone erosions (0−10) and joint space narrowing (JSN) (0−4), total range 0 − 314. [a]Results derived from imputed datasets, where model estimates are pooled based on Rubin's rules. Predictors were selected by applying backward selection in stacked data. Profile likelihood CIs calculated according to the Pseudo-Variance modification of Rubin's rule (PVR). AUC estimated based on internal validation by bootstrapping with 100 samples per imputed dataset. [b]*P*-values obtained by likelihood ratio tests. [c]The bootstrap 0.632+ estimate was calculated to correct for optimism. [d]Results derived from non-imputed data. Predictors were selected by applying backward selection in stacked data. CIs given as profile likelihood CIs. AUC estimated based on internal validation by bootstrapping with 1000 samples. AUC: Area Under the Curve; bDMARD: biological DMARD; OR: odds ratio.

comprehensive ultrasound status of the patients. A study limitation is the relatively small patient cohort, albeit larger than in previous ultrasound studies. Furthermore, the ultrasound examinations were performed by eight ultrasonographers. However, all were trained in musculoskeletal ultrasound and calibrated in the applied scoring system.

In conclusion, Doppler sum score was an independent predictor for successful discontinuation of bDMARDs at the 2-year follow-up—the odds for achieving successful discontinuation decreased by 56% per one-unit increase in Doppler sum score. Ultrasound could not predict successful tapering. This study suggests a role for Doppler ultrasound for identifying patients who can discontinue bDMARDs.

Acknowledgements

We wish to thank Hanne Slott Jensen and Per Brown Frandsen for their participation in the project. L.T., C.H.B., M.H., M.Ø., H.R., A.N., J.N. and S.J. have made substantial contributions to the study conception and design. L.T., C.H.B., K.E., L.J., K.H., U.M.D., V.F., T.M., S.K., D.V.J., K.A. and N.M. have participated in acquisition of data. M.B., Z.R., L.M.C. and J.M. have read all X-ray and MRI images. S.G. has performed all statistical analyses. N.S.K. has created all databases. L.T. and M.Ø. have interpreted data and drafted the article. All co-authors have revised the article critically for important intellectual content and final approval of the version of the article to be published.

Funding: No specific funding was received from any bodies in the public, commercial or not-for-profit sectors to carry out the work described in this manuscript.

Disclosure statement: L.T.: speakers fees from AbbVie, Janssen, Roche, Novartis, Pfizer, MSD, BMS and GE. C.H.B.: none. M.L.H.: grants from Bristol-Myers Squibb, AbbVie, Roche and Novartis; grants and personal fees from MSD and Biogen and Pfizer; and personal fees from Eli Lilly, Orion Pharma, CellTrion, Samsung Bioepis, and Janssen Biologics BV. S.G.: none. K.E.: none. L.J.: speakers fees and consultancy fees from AbbVie, Eli Lilly and Novartis. U.M.D.: consultancy fees from Eli Lilly, Roche and Novartis; speakers fee from Roche. T.H.: none. V.F.: none. T.M.: none. S.K.: research support from AbbVie, MSD and Novartis. L.Ø.: none. D.G.: none. H.R.: none. A.H.: speakers fee from Eli Lilly. J.N.: none. S.J.: none. D.V.J.: none. N.M. and K.A.: speakers fees and advisory board membership fees from AbbVie, Cellgene, Pfizer, Novartis, Roche, Berlin Chemie, Lilly and MSD. M.B.: research support, consultancy fees and/or speaker fees from Image Analysis Group, Esaote, AbbVie, Celgene, Eli-Lilly, Janssen, Novartis, Pfizer, UCB, Novo, GSK, Takeda, Geurbet, Biogen, Radiobotics and Chondrometrics. Z.R.: none. L.M.C.: none. J.M.: none. N.S.K.: none. M.Ø.: research support, consultancy fees and/or speaker fees from AbbVie, BMS,

Boehringer-Ingelheim, Celgene, Eli-Lilly, Hospira, Janssen, Merck, Novartis, Novo, Orion, Pfizer, Regeneron, Roche, Sandoz, Sanofi and UCB.

Data availability statement

The data underlying this article will be shared on reasonable request to the corresponding author.

Supplementary data

Supplementary data are available at *Rheumatology* online.

References

1 Smolen JS, Landewé RBM, Bijlsma JWJ *et al.* EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2019 update. Ann Rheum Dis 2020; 79:685–99.

2 Ollendorf DA, Klingman D, Hazard E *et al.* Differences in annual medication costs and rates of dosage increase between tumor necrosis factor-antagonist therapies for rheumatoid arthritis in a managed care population. Clin Ther 2009;31:825–35.

3 Bongartz T, Sutton AJ, Sweeting MJ *et al.* Anti-TNF antibody therapy rheumatoid arthritis and the risk of serious infections and malignancies: systematic review and meta-analysis of rare harmful effects in randomized controlled trials. JAMA 2006;295:2275–85.

4 Smolen JS, Nash P, Durez P *et al.* Maintenance, reduction, or withdrawal of etanercept after treatment with etanercept and methotrexate in patients with moderate rheumatoid arthritis (PRESERVE): a randomised controlled trial. Lancet 2013;381:918–29.

5 Van Vollenhoven RF, Østergaard M, Leirisalo-Repo M *et al.* Full dose, reduced dose or discontinuation of etanercept in rheumatoid arthritis. Ann Rheum Dis 2016; 75:52–8.

6 Van den Broek M, Klarenbeek NB, Dirven L *et al.* Discontinuation of infliximab and potential predictors of persistent low disease activity in patients with early rheumatoid arthritis and disease activity score–steered therapy: subanalysis of the BeSt study. Ann Rheum Dis 2011;70:1389–94.

7 Smolen JS, Emery P, Fleischmann R *et al.* Adjustment of therapy in rheumatoid arthritis on the basis of achievement of stable low disease activity with adalimumab plus methotrexate or methotrexate alone: the randomized controlled OPTIMA trial. Lancet 2014; 383:321–32.

8 Den Broeder AA, van Herwaarden N, van der Maas A *et al.* Dose REduction strategy of subcutaneous TNF inhibitors in rheumatoid arthritis: design of a pragmatic randomised non inferiority trial, the DRESS study. BMC Musculoskelet Disord 2013;14:299.

9 Haschka J, Englbrecht M, Hueber AJ *et al.* Relapse rates in patients with rheumatoid arthritis in stable remission tapering or stopping antirheumatic therapy: interim results from the prospective randomised controlled RETRO study. Ann Rheum Dis 2016;75:45–51.

10 Tanaka Y, Takeuchi T, Mimori T *et al.*; for the RRR study investigators. Discontinuation of infliximab after attaining low disease activity in patients with rheumatoid arthritis: RRR (remission induction by Remicade in RA) study. Ann Rheum Dis 2010;69:1286–91.

11 Henaux S, Ruyssen-Witrand A, Cantagrel A *et al.* Risk of losing remission, low disease activity or radiographic progression in case of bDMARD discontinuation or tapering in rheumatoid arthritis: systematic analysis of the literature and meta-analysis. Ann Rheum Dis 2018; 77:515–22.

12 Brahe CH, Krabbe S, Østergaard M *et al.* Dose tapering and discontinuation of biological therapy in rheumatoid arthritis patients in routine care – 2-year outcomes and predictors. Rheumatology (Oxford) 2019;58:110–9.

13 Saleem B, Brown AK, Keen H *et al.* Disease remission state in patients treated with the combination of tumor necrosis factor blockade and methotrexate or with disease-modifying antirheumatic drugs: a clinical and imaging comparative study. Arthritis Rheum 2009;60: 1915–22.

14 Spinella A, Sandri G, Carpenito G *et al.* The discrepancy between clinical and ultrasonographic remission in rheumatoid arthritis is not related to therapy or autoantibody status. Rheumatol Int 2012;32:3917–21.

15 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 2014;33:1061–6.

16 Cruces M, Al Snih S, Serra-Bonett N *et al.* Subclinical synovitis measured by ultrasound in rheumatoid arthritis patients with clinical remission induced by synthetic and biological modifying disease drugs. Reumatol Clin 2019; 15:218–22.

17 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. Rheumatology (Oxford) 2014;53: 2110–8.

18 Brown AK, Quinn MA, Karim Z *et al.* Presence of significant synovitis in rheumatoid arthritis patients with disease-modifying antirheumatic drug-induced clinical remission: evidence from an imaging study may explain structural progression. Arthritis Rheum 2006;54:3761–73.

19 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.

20 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. Rheumatology (Oxford) 2009;48:1092–7.

21 Saleem B, Brown AK, Quinn M *et al.* Can flare be predicted in DMARD treated RA patients in remission,

and is it important? A cohort study. Ann Rheum Dis 2012;71:1316–21.

22 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.

23 Alivernini S, Peluso G, Fedele AL *et al.* Tapering and discontinuation of TNF-α blockers without disease relapse using ultrasonography as a tool to identify patients with rheumatoid arthritis in clinical and histological remission. Arthritis Res Ther 2016;18:39.

24 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 (Oxford) 2015;54:1408–14.

25 Iwamoto T, Ikeda K, Hosokawa J *et al.* Prediction of relapse after discontinuation of biologic agents by ultrasonographic assessment in patients with rheumatoid arthritis in clinical remission: high predictive values of total gray-scale and power Doppler scores that represent residual synovial inflammation before discontinuation. Arthritis Care Res 2014;66:1576–81.

26 Saleem B, Keen H, Goeb V *et al.* Patients with RA in remission on TNF blockers: when and in whom can TNF blocker therapy be stopped? Ann Rheum Dis 2010;69:1636–42.

27 Ibfelt EH, Jensen DV, Hetland ML. The Danish nationwide clinical register for patients with rheumatoid arthritis: DANBIO. Clin Epidemiol 2016;8:737–42.

28 Arnett FC, Edworthy SM, Bloch DA *et al.* The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988;31:315–24.

29 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.

30 Torp-Pedersen S, Terslev L. Settings and artefacts relevant in colour/power Doppler ultrasound in rheumatology. Ann Rheum Dis 2008;67:143–9.

31 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 standardized, consensus-based scoring system. RMD Open 2017;3:e000428.

32 Terslev L, Naredo E, Aegerter P *et al.* Scoring ultrasound synovitis in rheumatoid arthritis: a EULAR-OMERACT ultrasound taskforce- part 2: reliability and application to multiple joints of a standardised consensus-based scoring system. RMD Open 2017;3:e000427.

33 Van der Heijde D. How to read radiographs according to the Sharp/van der Heijde method. J Rheumatol 1999;26:743–5.

34 Østergaard M, Peterfy C, Conaghan P *et al.* OMERACT rheumatoid arthritis magnetic resonance imaging studies. Core set of MRI acquisitions, joint pathology definitions, and the OMERACT RA-MRI scoring system. J Rheumatol 2003;30:1385–6.

35 Østergaard M, Bøyesen P, Eshed I *et al.* Development and preliminary validation of a magnetic resonance imaging joint space narrowing score for use in rheumatoid arthritis: potential adjunct to the OMERACT RA MRI scoring system. J Rheumatol 2011;38:2045–50.

36 Glinatsi D, Bird P, Gandjbakhch F *et al.* Development and validation of the OMERACT rheumatoid arthritis magnetic resonance tenosynovitis scoring system in a multireader exercise. J Rheumatol 2017;44:1688–93.

37 Valor L, Garrido J, Martínez-Estupiñán L *et al.* Identifying markers of sustained remission in rheumatoid arthritis patients on long-term tapered biological disease-modifying antirheumatic drugs. Rheumatol Int 2018;38:1465–70.

38 Hammer HB, Kvien TK, Terslev L. Ultrasound of the hand is sufficient to detect subclinical inflammation in rheumatoid arthritis remission: a *post hoc* longitudinal study. Arthritis Res Ther 2017;19:221.

39 Terslev L, Torp-Pedersen S, Qvistgaard E, von der Recke P, Bliddal H . Doppler ultrasound findings in healthy wrists and finger joints. Ann Rheum Dis 2004;63:644–8.

40 Padovano I, Costantino F, Breban M *et al.* Prevalence of ultrasound synovial inflammatory findings in healthy subjects. Ann Rheum Dis 2016;75:1819–23.