Can biomarkers predict successful tapering of conventional disease-modifying therapy in rheumatoid arthritis patients in stable remission?

H.L. Gul[1,2], A. Di Matteo[1,2], K. Mankia[1,2], J. Wu[3], F. Ponchel[1], P. Emery[1,2]

[1]Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds;
[2]NIHR Leeds Biomedical Research Centre, Leeds Teaching Hospitals NHS Trust;
[3]Leeds Dental Institute, The University of Leeds, United Kingdom.

Abstract

Objective

Specific guidelines for managing RA patients in clinical remission for ≥6 months on cs-DMARDs are lacking. Tapering of treatment is encouraged, however, without validated biomarkers for success. We aimed to assess the rate of sustained remission after 12 months in patients who either (i) followed structured cs-DMARD tapering or (ii) continued therapy, focusing on the added value of biomarkers as predictors of outcome.

Methods

RA patients fulfilling 3v-DAS28CRP<2.6 for ≥6 months on stable cs-DMARD therapy were included. Patients were offered structured tapering, with 117 accepting tapering and 83 continuing therapy. Clinical, ultrasound, immunological (T-cell subsets) and patient-reported outcome (PRO) data were collected. The primary endpoint was the proportion of patients in sustained remission without relapse after 12 months. Regression analyses were used to identify predictors of sustained remission.

Results

Of those who tapered, 64% remained in clinical remission after 12 months compared with 80% (p=0.018) of patients on stable treatment. In the tapering group, higher levels of CRP, TJC, % inflammation-related T-cell (IRC) and PROs were associated with flare (all p<0.05), with a trend for total PD (p=0.066). A model predicting sustained remission retained RAQoL, total PD and IRC (85% accuracy, AUROC=0.893, p<0.0001). In the non-tapering group, higher CRP, ESR, SJC and shorter disease duration (all p<0.05) were associated with flare, with no parameter able to predict sustained remission.

Conclusion

In the tapering group, the combination of clinical, PRO, US and T-cell parameters demonstrated added value for predicting sustained remission compared with clinical parameters alone. These data may inform best tapering practice.

Key words
rheumatoid arthritis, remission, disease-modifying therapy, tapering, biomarkers

Hanna L. Gul, MBChB, MRCP, MSc
Andrea Di Matteo, MD, PhD
Kulveer Mankia, BMBCh, MA, MRCP, DM
Jianhua Wu, PhD
Frederique Ponchel, PhD
Paul Emery, MD, MA, FRCP

Please address correspondence to:
Paul Emery,
Chapel Allerton Hospital,
Chapeltown Road,
Leeds LS7 4SA, United Kingdom.
E-mail: p.emery@leeds.ac.uk

Received on January 23, 2022; accepted in revised form on April 19, 2022.

© Copyright CLINICAL AND EXPERIMENTAL RHEUMATOLOGY 2023.

Funding: funding for this work was provided by the Leeds Biomedical Research Centre and the Leeds Hospital Charitable Foundation.
An abstract of this work was presented at the EULAR 2021 virtual conference (Ann Rheum Dis 2021; 80 (Suppl. 1): 110).
Competing interests: none declared.

Introduction

Increasing numbers of RA patients reach and maintain a state of clinical remission, especially when following a treat-to-target (T2T) approach (1, 2). Stable long-term remission is now possible in >50% of patients using this strategy (3, 4). However, there is little guidance on how to manage patients once remission has been achieved. A key issue is whether disease-modifying therapy is required indefinitely, with concerns for over-treatment. This can lead to poor patient adherence to therapy, with approximately 15% self-discontinuing treatment when they are in remission, with unpredictable morbidity (5, 6). Although life-changing for RA patients, the benefits of long-term cs-DMARD use in remission must be balanced against patients' wishes to minimise drug use in addition to reducing the risk of potential drug-induced toxicities (7), the cost of therapy, and its monitoring by healthcare providers (8). Current guidelines recommend tapering (with the possibility of discontinuation) in RA patients who achieve sustained remission on stable cs-DMARDs (9-11) however, this practice is not routinely adopted into standard care due to the potential risk of disease flare and the lack of specific guidance on how to perform tapering (12). Data for cs-DMARD tapering/discontinuation are sparse and with small patient numbers (13-18) however, the possibility of successful tapering has been described. The majority of evidence comes from randomised controlled trials (RCTs) for patients on a range of monotherapies (13-18). Flare was more likely in those who discontinued treatment however, individuals successfully re-captured disease control after re-starting treatment. Additional evidence comes from studies in which a step-down approach in tapering was used (notably combination cs-DMARDs to monotherapy) which also reported sustained clinical response (19-23).

To be able to offer safe tapering, clinicians need to be able to identify appropriate patients, ideally using objective biomarkers to predict sustained remission. Several studies have reported associations with sustained remission, however validated biomarkers are yet to be identified and studied prospectively. Specifically, shorter disease duration, longer remission duration, younger age, male gender, low baseline disease activity/PRO scores and early treatment initiation were associated with achieving sustained remission (4, 19, 21, 24-27). Baseline musculoskeletal ultrasound (US) assessment (prior to tapering) has been proposed as an objective biomarker for sustained remission, since patients can show evidence of sub-clinical disease and radiographic progression despite achieving DAS-remission (28) and high baseline power-Doppler (PD) score has been associated with disease flare following tapering (29). Since immune dysregulation is key to RA pathogenesis (30), T-cell subsets have also been studied as a biomarker in RA. Specifically, $CD4^+$ T-cell subset abnormalities have been demonstrated across the RA continuum from at-risk individuals, evolving RA and those patients achieving clinical remission (31-34). They were furthermore associated with disease flare when tapering biologic therapy (35). To date, only the absence of serum rheumatoid factor (RF) and anti-citrullinated protein antibody (ACPA) have been identified to predict successful tapering of disease-modifying therapy (12, 26, 36). A model of deep remission, which requires the achievement of clinical DAS/Boolean remission, imaging (no synovitis/osteitis), serological (normal CRP/ESR) and immunological (negative RF/ACPA) parameters, has been proposed to more precisely define remission, however this needs to be validated prospectively (36). Similarly, we previously explored the concept of multi-dimensional remission (MDR) in RA patients at the time of achieving DAS28-remission (34). We hypothesised that objective measures of inflammation could help define different depths of DAS28 remission. We demonstrated wide variations in patient characteristics including demographics, serology, clinical, US and $CD4^+$ T-cells subset parameters. Achieving deep remission (defined as achieving clinical, imaging and T-cell subset remission) was associated with better PROs, suggesting that

these could be potential biomarkers, allowing a patient-centred dimension to be considered. In a recent study with a similar aim (de-escalation of DMARDs *vs*. continuation) a model based on achieving of DAS28(CRP)<1.82 in the absence of subclinical synovitis on ultrasound was able to predict successful tapering in RA patients in clinical remission on cs-DMARD monotherapy. Flare was nonetheless more common in the de-escalation group (23).

Based on previous work we proposed that objective measures of inflammation and PROs could predict successful tapering of treatment. We conducted a prospective observational study of two different treatment strategies (taper *vs*. continuation) for RA patients in stable clinical remission. The aim was to assess the rate of sustained remission after 12 months (without flare) and to determine baseline predictors of successful tapering, towards developing models to aid risk stratification in clinical practice.

Order of tapering if on dual/triple therapy ▼

DRUG	Baseline Dose	Taper 1	Taper 2	Taper 3	Taper 4	Taper 5
Hydroxychloroquine	200mg bd	200mg od	Stop	-	-	-
Hydroxychloroquine	200mg od	Stop	-	-	-	-
Sulfasalazine	1.5g bd	1g bd	1.5g od	500mg bd	500mg od	Stop
Sulfasalazine	1g bd	1.5g od	500mg bd	500mg od	Stop	-
Sulfasalazine	1.5g od	500mg bd	500mg od	Stop	-	-
Sulfasalazine	500mg bd	500mg od	Stop	-	-	-
Sulfasalazine	500mg od	Stop	-	-	-	-
Methotrexate	20-25mg/wk	15mg/wk	7.5mg/wk	No change	Stop	-
Methotrexate	10-15mg/wk	7.5mg/wk	No change	Stop	-	-
Methotrexate	7.5mg/wk	Stop	-	-	-	-

Fig. 1. cs-DMARD tapering schedule.

Patients and methods

RA patients (ACR/EULAR 2010 classification) were recruited from the Leeds RA remission clinic based on achieving sustained DAS remission using 3 variables (3v-DAS28(CRP)<2.6) for at least six months on stable cs-DMARD doses, and no systemic corticosteroid therapy. Only 6/200 patients received glucocorticoid therapy (intramuscular methylprednisolone) over the 12-month period before study inclusion (>6 months before inclusion). These patients were all in the taper group. PGA score was omitted from the DAS score due to missing data at this is not always clearly documented in the outpatient clinic. Informed, written consent for participation was obtained upon inclusion (ethical approval: Leeds (West) Research Ethics Committee - 09/H1307/98, 15/01/2010).

Patients were offered the option to continue therapy or to taper according to a pre-defined protocol (Fig. 1), in-line with the EULAR/ACR guidance for managing remission in RA (10, 11). Patients were informed of the potential risk of flare on tapering *versus* the risk of over-treatment/potential side-effects of continuing therapy prior to making their decision. They were followed prospectively for 12 months (reviewed every 3 months and at time of flare). The primary endpoint was the proportion of patients still in sustained 3v-DAS28(CRP) remission, without flare after 12 months. Flare was defined as loss of remission (3v-DAS28(CRP)≥2.6) or evidence of at least one new clinically swollen joint and was treated with corticosteroids and/or increasing therapy to the previous effective dose (in the tapering group). No further attempt at tapering was made for individuals who flared during the study. Due to limited clinic capacity and patient availability (notably due to the COVID-19 pandemic) some patients did not have face-to-face follow-up visits at all time -points.

Clinical assessment and investigations

At baseline, demographic details were collected and participants completed patient questionnaires, provided a clinical history of their symptoms and had a systems examination by a rheumatologist which included a joint count. Demographic and clinical data included: age, gender, smoking status, disease duration, remission duration, 28 tender and swollen joint counts TJC/SJC (28), ESR (mm/h)/CRP (mg/l), duration of early morning stiffness (EMS, mins) and autoantibody status (IgM-RF/ACPA). Standardised patient questionnaires for PROs included visual analogue scores (VAS) for patient global assessment of disease (PGA), disease activity (DA), fatigue and pain, HAQ-DI and RAQoL scores. Ultrasound assessments were performed by an independent assessor and recorded grey scale (GS) and power Doppler (PD) signal as presence/absence per joint and total score, graded according to the OMERACT standardised consensus-based scoring system (37). The joints chosen represented a pragmatic and feasible core set, which were most commonly affected in patients with RA (32). T-cell subsets (naive $CD4^+$ cells, T-regulatory cells (Treg) and IRC were measured by flow-cytometry. This technology is routinely used by our hospital services and we transferred our research panels to the NHS immunology services of the Leeds Teaching Hospitals Trust in 2013. Peripheral blood was collected into EDTA (4 ml) tubes. Flow cytometry was performed according to the NHS procedures, using 250 µl of whole blood per panel and red cell lysis. The 3 subsets were quantified as previously described in Figure 1a of the paper by Ponchel *et al*. in 2020 (31). We previously demonstrated the age relationship between naive and Treg cells in health, while IRC frequencies were independent of age. Expected naive and Treg frequencies (% of $CD4^+$T-cells) at a defined age in health were calculated as follows: [expected naive %] = -0.63 x [age] +66.6 (rho=0.850, p<0.0001; [expected Treg %] = +0.061 x [age] +1.83 (rho=0.554, p=0.001) using regression calculated from <120 healthy control data points. Normalised naive and Treg frequencies (% of $CD4^+$T-cells) were calculated as the difference from observed values to

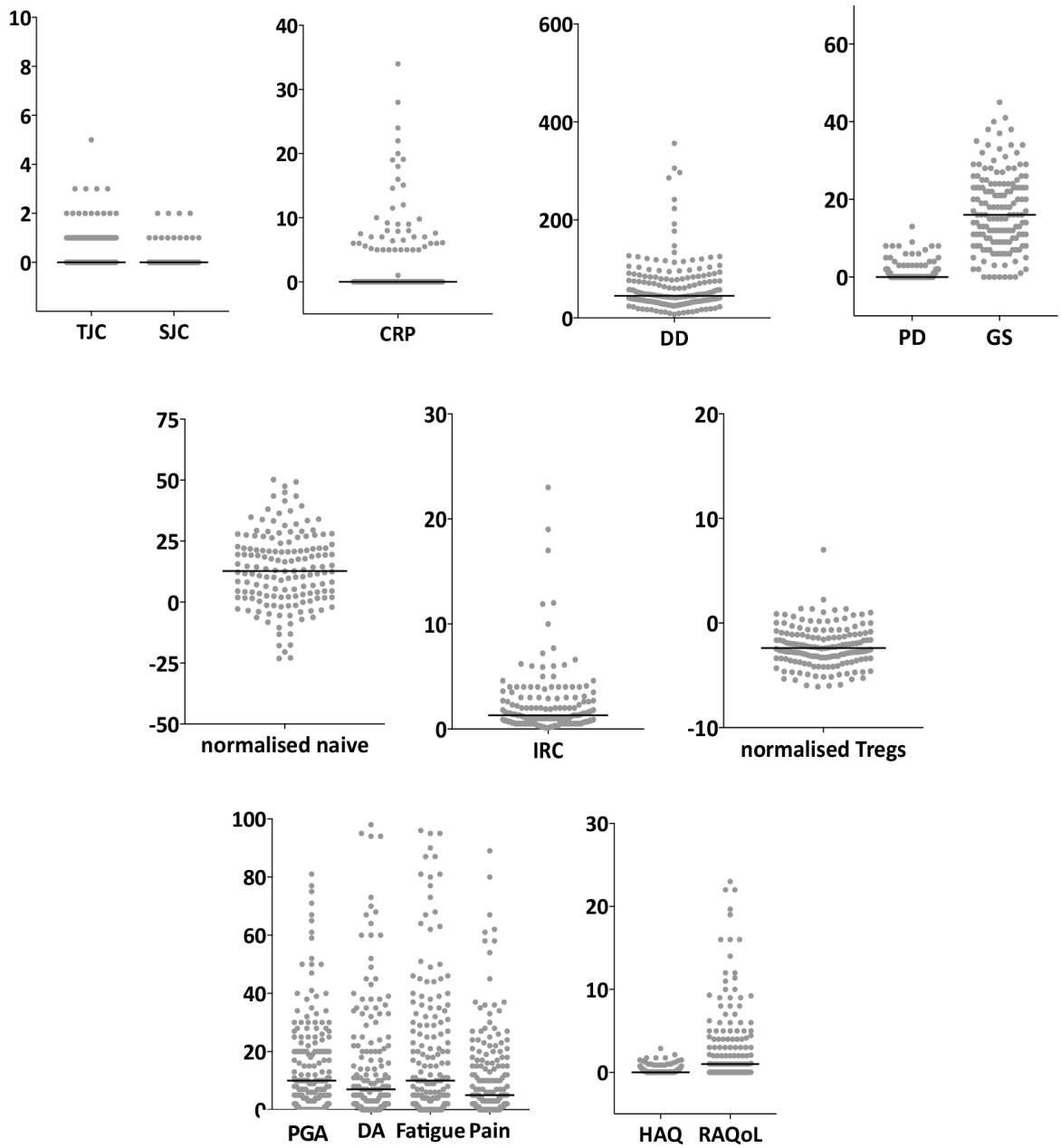


Fig. 2. Baseline characteristics of total cohort (n=200).
TJC/SJC(28): number of tender and swollen joints out of 28; CRP (mg/L) – values plotted at 0 are those deemed below detection (<5mg/L); DD: disease duration (months); PD and GS: total power Doppler and grey scale scores; T-cell subsets (% of $CD4^+$T-cells): naive T-cells; IRC: inflammation related cells; Treg: T-regulatory cells; PGA: Patient Global Assessment; DA: disease activity (fatigue and pain scores all out of 100); HAQ: health assessment questionnaire; RAQoL: RA quality of life score.

expected values. These were reported as positive when the observed values were higher than expected in health or negative when below (31).

Statistical analysis

Baseline data are described using medians ([IQR]) or number and proportion (%). Distribution of data did not verify normality (Kolmogorov-Smirnov test) and continuous measures were therefore explored comparing groups using a Mann-Whitney U-test and nominal measures with Chi-square tests. Corrections for multiple testing were not applied in the descriptive tables. Area under the ROC curve (AUC) was performed for univariate analyses to assess potential predictive values. Every attempt was made to obtain complete data for clinical, imaging, immunological and PRO parameters however, this was not always possible due to the availability of US and laboratory facilities at the time of the visit. Furthermore, some of the recent follow-up visits had to be conducted via telephone due to the COVID-19 pandemic. Before modelling, missing data were imputed (using 5 rounds of data imputation in SPSS). Unadjusted odds ratios (OR) were calculated first. Predicting sustained remission using multiple dimensions was performed using forward binary logistic and Cox regression allowing the model to independently se-

lect the best predictors. Kaplan Meier plots were used to compare probability to flare over time. Analyses were conducted using SPSS 21.1.

Results

Baseline characteristics

Two hundred patients were recruited since 2014. Clinical, imaging and immunological parameters were highly variable (Fig. 2 and Table I), confirming previous data on the heterogeneity of DAS28 remission (34). The majority of patients were taking methotrexate either as monotherapy (n=114) or combination therapy with hydroxychloroquine and or sulfasalazine (n=70).

Compared with the data acquired when patients first achieved remission on cs-DMARDs (34) (Supplementary Fig. S2 and Suppl. Table S1, n=419), this cohort showed a trend for lower scores for most baseline characteristics. Reductions in ranges and IQRs were observed (joint counts/CRP/PD] scores/GS scores)/PROs), while medians remained largely the same. Increased normalised naive T-cells and reduced IRC (but no change for Treg) were observed. This confirms that DAS28-remission is a dynamic state, with some patients continuing to improve over a prolonged period, and highlights that biomarkers may require different cut-offs when remission is stable, compared to when first achieved.

Using our previously reported definition of multi-dimensional remission (34), 89 of 200 (45%) patients were in clinical remission (modified Boolean definition: TJC/SJC and CRP all ≤1), 49/152 (32%) in ultrasound remission (PD=0), while 125/154 (81%) were in T-cell remission (positive normalised naive T-cells). In addition, 32/154 (21%) showed normal IRC and 54/154 (35%) reduced normalised Treg as previously described (31) but not used to define T-cell remission. Achievement of these 3 remission criteria [as defined at 1st visit (34)] occurred in 45/152 (29.6%) patients.

One hundred and seventeen patients chose to taper (Table II). No difference in baseline drug regimens (mono/combination therapy) was observed. Only male gender (p=0.036) and longer length of remission (p=0.015) were associated with the patient's decision to taper (not significant after correction). TJC/SJC/CRP/ESR were within the healthy range, while EMS was lower in the tapering group (p=0.034). US evaluation showed higher total GS in the tapering group (p=0.003). Naive and Treg CD4+cells showed no differences, while IRC were lower in the tapering group (p=0.001, only significant variable after correction). Lower medians for PROs were observed in the tapering group, potentially suggesting a perception of better health status by these patients.

Table I. Baseline characteristics of taper *vs*. non-tapering cohort (total n=200).

	Missing data (% of cases)	Taper cohort (n=117)	Non-taper cohort (n=83)	*p*-value
Female, n (%)	-	63 (54.3%)	56 (69.1%)	0.036
Age (years)*	-	63.5 (14)	60.5 (17.5)	0.146
Disease duration (months)*	6%	46.9 (47.5)	40.2 (44.7)	0.913
Remission duration (months)*	-	18 (21)	12.2 (8.3)	0.015
RF+, n (%)	-	61 (52.6%)	49 (60.5%)	0.271
ACPA+, n (%)	-	82 (70%)	52 (64.2%)	0.461
Smoking (never), n (%)				
Never	1%	54 (45.7%)	59 (50.9%)	0.558
Ever		34 (42%)	45 (55.6%)	
TJC28*	-	0 (0.75)	0 (1)	0.489
SJC28*	-	0 (0)	0 (0)	0.757
CRP (mg/L)*	-	<5 (<5)	<5 (<5)	0.401
ESR (mm/h)*	8.5%	9 (12.7)	11.5 (12)	0.034
EMS (mins)*	7.7%	0 (5)	0 (10)	0.767
VAS PGA*	13.7%	9 (21.7)	10 (27)	0.464
VAS Pain*	15.4%	5.5 (18.8)	10 (18.3)	0.330
VAS DA*	16.2%	5 (19.8)	11 (33.5)	0.130
VAS Fatigue*	19.7%	10 (27.5)	9 (34)	0.854
HAQ-DI*	16.2%	0 (0.38)	0.1 (0.5)	0.963
RaQoL*	19.7%	0.5 (4.07)	1 (4.14)	0.495
Total PD*	17.1%	0 (2)	0 (0 to 0)	0.084
Total GS*	17.1%	18 (15.5)	14.5 (12)	0.003
Normalised naive*	-	12.69 (24.38)	13.61 (19.1)	0.462
Normalised Treg*	18.8%	-1.97 (2.21)	-3.26 (2.69)	0.251
IRC*	18.8%	1.00 (1.5)	2.00 (3)	0.001
Loss of remission (3vDAS28≥2.6)	-	42 (36%)	17 (20.5%)	0.018
Monotherapy, n (%)	-	78/117 (67%)	49/83 (59%)	
Combination therapy, n (%)		39/117 (33%)	34/83 (41%)	0.194

*Median (IQR) *CRP <5 mg/l = lowest detectable limit.
RF: rheumatoid factor; CCP: anti-CCP antibody; TJC/SJC(28): number of tender and swollen joints out of 28; CRP: C-reactive protein, mg/l; ESR: erythrocyte sedimentation rate, mm/h; EMS: early morning stiffness, mins; VAS: visual analogue score; PGA: patient global assessment of disease; DA: disease activity; HAQ-DI: Health Assessment Questionnaire Disability Index; RAQoL: RA quality of life questionnaire score; PD: power Doppler score; GS: grey scale synovial hypertrophy score; T-cell subsets (% of CD4+T-cells): naive T-cells; IRC: inflammation related cells; Treg: T-regulatory cells.

Rate of sustained remission

Seventy-five of 117 (64%) patients remained in sustained remission over 12 months in the tapering group compared to 66/83 (79.5%) in the non-tapering group (p=0.018). 1/117 patients achieved drug-free remission while 11/117 (9%) managed to do so over 15 months, adapting the tapering schedule. Time-to-flare survival analysis comparing the 2 groups is displayed in Supplementary Figure S1. The median time to flare was 277 days (IQR: 222–400) in non-tapering and 238 days (119–343.5) in tapering patients. 114/117 (97%) patients who flared in the tapering group re-captured remission following reinstitution of treatment at the last effective dose, compared to 100% of those who did not taper. 3/6 (50%) patients who received glucocorticoids withing 12 months of starting the study remained in remission.

Characteristics of patients in sustained remission vs. flare

In the non-taper group (Suppl. Table S1), only inflammation-related para-

meters (CRP/ESR, $p<0.003$) suggested that better disease control tended to be associated with the maintenance of remission (with a trend for SJC, $p=0.055$). Only lower ESR showed significant predictive value for sustained remission (AUC=0.844, $p<0.001$). No further modelling was performed with only one variable.

In the taper group (Table II), there was no association with demographic variables and sustained remission. Clinical parameters suggested less well-controlled inflammation in patients who flared, with higher CRP ($p=0.001$) and TJC (28) ($p=0.011$). This was also reflected in higher IRC ($p<0.0001$, after correction) while sub-clinical inflammation detected by PD signal was higher but not significant ($p=0.066$). PROs were also higher (all $p<0.007$), with the exception of PGA.

Modelling the prediction of sustained remission in the tapering group

Only 73/117 (63%) patients had a complete dataset for all variables due to the practical limitation of accessing US and T-cell assessment. Unadjusted odds ratios (ORs, 95% CI) were calculated. Missing data were imputed. Pooled ORs were calculated again and compared to that of the original dataset. Minimal differences were observed, allowing further analysis using the overall group (n=117).

Logistic forward regression modelling was progressed by analysing each dimension when added to a reference (model-1) based only on demographic/clinical variables (Table III). Model-1 provided 69% accuracy in predicting sustained remission (retaining TJC/CRP). The sequential addition of imaging to this model resulted in added prediction (model-2, +3.6% accuracy), as well as when adding PROs (model-3, +3.9%) or immunological (model 4, +12%) dimensions. Similarly, a combination of 2 dimensions (models-5 imaging/PRO, +9.4%; model-6 PRO/immunological, +14.6%; model-7 imaging/immunological, +14.6%) keep increasing performances. The best model (model-8) combined all 4 dimensions and provided the highest added value (85% accuracy, +15%), with

Table II. Baseline characteristics of flare (loss of DAS28 remission) *vs*. sustained remission in the taper cohort (n=117).

	Flare (n=42)	Sustained Rem (n=75)	*p*-value	AUC (95%CI) *p*-value
Demographic variables				
Female, n (%)	26 (62%)	37 (49.3%)	0.146	0.430 (0.321–0.538) $p=0.211$
Age*	63 (14.5)	64.5 (15.3)	0.686	0.523 (0.414–0.632) $p=0.686$
Disease duration*	54.5 (51.8)	44.1 (48.6)	0.300	0.559 (0.447–0.671) $p=0.300$
Remission duration *	17 (20.2)	18 (23)	0.603	0.471 (0.361–0.580) $p=0.603$
RF^+ n (%)	25 (59%)	36 (48%)	0.810	0.435 (0.326–0.544) $p=0.249$
$ACPA^+$ n (%)	32 (76%)	50 (67%)	0.312	0.443 (0.335–0.551) $p=0.312$
Smoking n (%)				
Never	18 (42.9%)	36 (48%)	0.345	0.546 (0.435–0.657)
Ever	24 (57.1%)	35 (46.7%)		$p=0.417$
Clinical variables				
TJC28*	0 (1)	0 (0)	0.001	0.624 (0.510–0.739) $p=0.034$
SJC28*	0 (0)	0 (0)	0.221	0.526 (0.410–0.642) $p=0.660$
CRP*	<5 (<5 to 6.1)	<5 (<5 to <5)	0.011	0.602 (0.490–0.714) $p=0.069$
ESR*	10 (11.8)	9 (13)	0.342	0.565 (0.448–0.682) $p=0.267$
EMS*	0 (15)	0 (0)	0.106	0.621 (0.484–0.758) $p=0.080$
PRO variables				
VAS PGA*	15 (23)	7 (16)	0.061	0.604 (0.489–0.718) $p=0.077$
VAS Pain*	15 (20.6)	3 (10)	0.004	0.653 (0.522–0.783) $p=0.027$
VAS DA*	7.5 (30.4)	2.5 (0 to 11.5)	0.002	0.643 (0.514–0.771) $p=0.039$
VAS Fatigue*	29.5 (35.5)	7.5 (20.8)	0.007	0.650 (0.515–0.785) $p=0.030$
HAQ-DI*	0.25 (0.812)	0 (0.125)	<0.0001	0.706 (0.579–0.833) $p=0.003$
RaQoL*	2.5 (9.17)	0 (3)	<0.0001	0.682 (0.553–0.811) $p=0.009$
Ultrasound variables				
Total PD*	0 (4.25)	0 (1.75)	0.066	0.558 (0.420–0.695) $p=0.403$
Total GS*	18.5 (16.8)	17 (15)	0.207	0.574 (0.437–0.710) $p=0.285$
T-cell variables				
Normalised naive*	14.47 (24.6)	12.60 (21.6)	0.838	0.524 (0.394–0.654) $p=0.714$
Normalised Treg*	-1.57 (2.65)	-1.99 (1.99)	0.940	0.506 (0.370–0.641) $p=0.932$)
IRC*	2.70 (3.3)	0.60 (0.50)	<0.0001	0.860 (0.774–0.946) $p<0.0001$
Drug category				
Monotherapy, n (%)	28/42 (66%)	51/75 (68%%)	0.617	0.523 (0.412–0.634)
Combination therapy n (%)	15/42 (36%)	24/ 75 (61.5%)		$p=0.684$

*Median (IQR)). CRP <5 mg/l = lowest detectable limit,
RF: rheumatoid factor; CCP: anti-CCP antibody; TJC/SJC(28): number of tender and swollen joints out of 28; CRP: C-reactive protein, mg/l; ESR: erythrocyte sedimentation rate, mm/h; EMS: early morning stiffness, mins; VAS: visual analogue score; PGA: patient global assessment of disease; DA: disease activity; HAQ-DI: Health Assessment Questionnaire Disability Index; RAQoL: RA quality of life questionnaire score; PD: power Doppler score; GS: grey scale synovial hypertrophy score; T-cell subsets (% of $CD4^+$T-cells): naive T-cells; IRC: inflammation related cells; Treg: T-regulatory cells.

Table III. Modelling the prediction for sustained remission after 12m in the tapering cohort using logistic regression (n=117), presented as OR (95% CI) and *p*-value.

	Unadjusted	Model 1 Clin only	Model 2 Clin + US	Model 3 Clin + PRO	Model 4 Clin + T-cells	Model 5 Clin + PRO + US	Model 6 Clin + PRO + T-cells	Model 7 Clin + US + T-cells	Model 8 Clin + PRO + US + T-cells
TJC28	0.430 (0.225-0.821) *p*=0.011	0.389 (0.200-0.758) *p*=0.006	0.427 (0.224-0.816) *p*=0.010	0.446 (0.233-0.852) *p*=0.015	0.459 (0.224-0.940) *p*=0.033	0.480 (0.258-0.896) *p*=0.021	Not selected by model	Not selected by model	Not selected by model
CRP	0.914 (0.836-0.999) *p*=0.046	0.898 (0.819-0.985) *p*=0.022	0.911 (0.834-0.994) *p*=0.037	0.912 (0.836-0.996) *p*=0.040	Not selected by model	0.920 (0.845-1.001) *p*=0.052	Not selected by model	Not selected by model	Not selected by model
VAS Pain	0.963 (0.931-0.995) *p*=0.024	NA	NA	Not selected by model	NA	Not selected by model	Not selected by model	NA	Not selected by model
VAS PGA	0.967 (0.941-0.994) *p*=0.018	NA	NA	Not selected by model	NA	Not selected by model	Not selected by model	NA	Not selected by model
VAS DA	0.972 (0.947-0.997) *p*=0.028	NA	NA	Not selected by model	NA	Not selected by model	Not selected by model	NA	Not selected by model
VAS Fatigue	0.971 (0.951-0.992) *p*=0.006	NA	NA	Not selected by model	NA	Not selected by model	Not selected by model	NA	Not selected by model
HAQ-DI	0.194 (0.068-0.555) P=0.002	NA	NA	0.222 (0.076-0.651) P=0.006	NA	0.213 (0.073-0.622) *p*=0.005	Not selected by model	NA	Not selected by model
RaQoL	0.835 (0.746-0.935) *p*=0.002	NA	NA	Not selected by model	NA	Not selected by model	0.837 (0.740-0.948) *p*=0.005	NA	0.855 (0.757-0.477) *p*=0.012
Total PD	0.779 (0.651-0.932) *p*=0.006	NA	0.804 (0.668-0.968) *p*=0.021	NA	NA	0.786 (0.647-0.956) *p*=0.016	NA	0.750 (0.619-0.908) *p*=0.003	0.765 (0.625-0.937) *p*=0.010
IRC	0.277 (0.154-0.500) *p*<0.0001	NA	NA	NA	0.291 (0.161-0.524) *p*<0.0001	NA	0.279 (0.155-0.500) *p*<0.0001	0.250 (0.132-0.473) *p*<0.0001	0.253 (0.134-0.477) *p*<0.0001
Accuracy		**69%**	**71.6%**	**75.9%**	**81%**	**78.4%**	**83.6%**	**83.6%**	**84.5%**
SEN; SPE		64.71%; 69.7%	66.67%; 72.83%	74.07%; 76.40%	82.76%; 80.46%	76.67%; 79.07%	86.67%; 82.56%	63.41%; 90.67%	84.85%; 84.34%
PPV; NPV		26.83%; 92%	39.02%; 89.33%	48.78%; 90.67%	58.54%; 99.33%	56.10%; 90.67%	63.41%; 94.67%	78.79%; 81.93%	68.29%; 93.33%
AUROC (95% CI) *p*-value		0.725 (0.624-0.827) *p*<0.0001	0.780 (0.689-0.871) *p*<0.0001	0.781 (0.690-0.872) *p*<0.0001	0.832 (0.749-0.914) *p*<0.0001	0.825 (0.743-0.908) *p*<0.0001	0.847 (0.775-0.918) *p*<0.0001	0.866 (0.794-0.939) *p*<0.0001	0.893 (0.826-0.946) *p*<0.0001

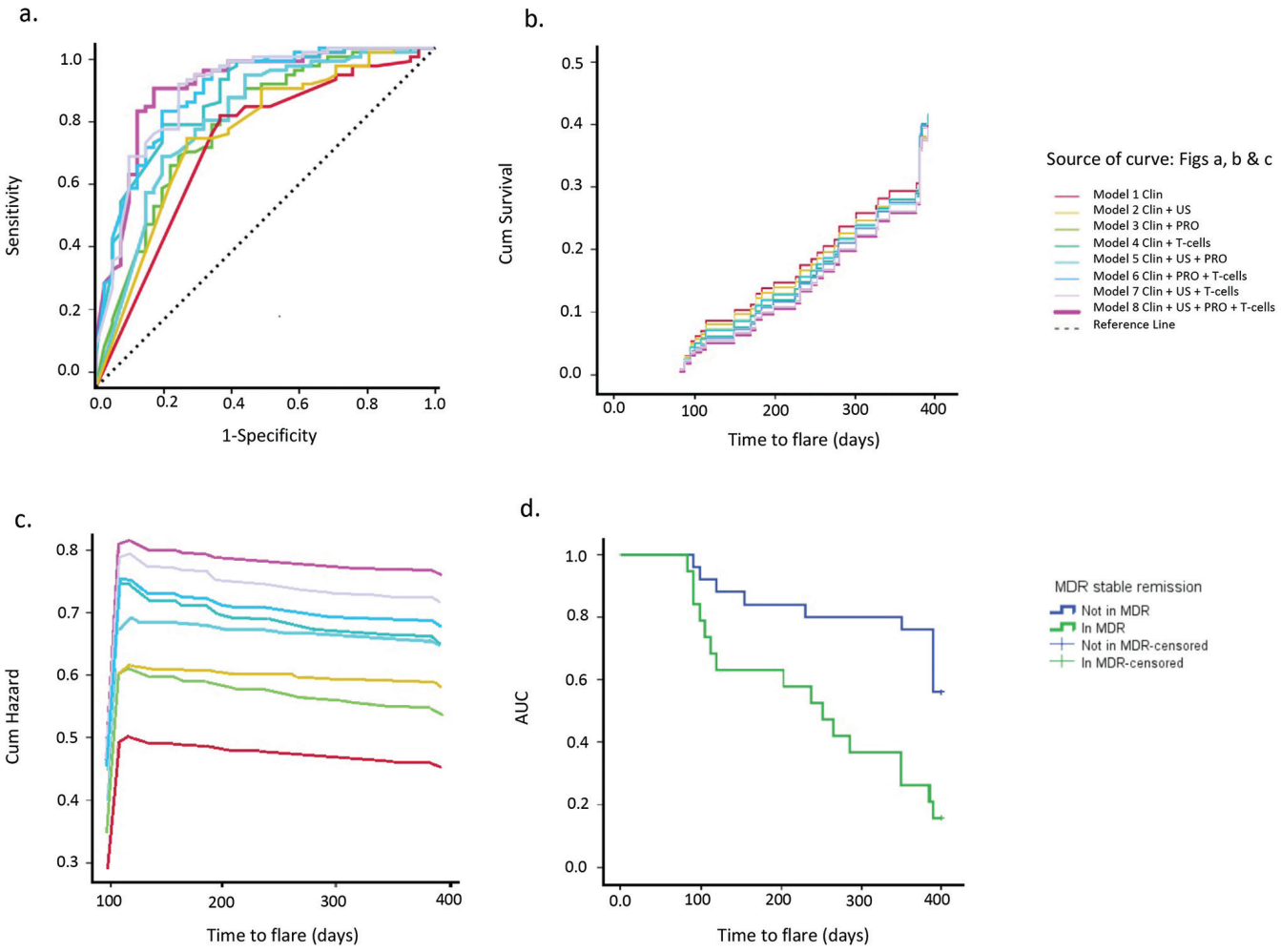
RF: rheumatoid factor; CCP: anti-CCP antibody; TJC/SJC(28): number of tender and swollen joints out of 28; CRP: C-reactive protein, mg/l; ESR: erythrocyte sedimentation rate, mm/h; EMS: early morning stiffness, mins; VAS: visual analogue score; PGA: patient global assessment of disease; DA: disease activity; HAQ-DI: Health Assessment Questionnaire Disability Index; RAQoL: RA quality of life questionnaire; PD: power Doppler score; GS: grey scale synovial hypertrophy score; T-cell subsets (% of $CD4^+$T-cells): naive T-cells; IRC: inflammation related cells; Treg: T-regulatory cells.

anAUROC of 0.893 (Fig. 3a). Further performance measures relating to these models are described in Table III, with for Model-8 85%/84% sensitivity/specificity and 68%/93% PPV/NPV.

Using individual patients' calculated probability of remaining in remission (model-8) and using a high specificity cut-off (80%, probability >0.650 based on AUROC data), the risk of flare was dichotomised: high risk was observed in 40/117 (34%) patients of whom 33/40 (82.5%) flared; low risk in 77/117 (66%) patients with only 8/76 (10.5%) flare. Model-8's probability derived risk for flare could therefore predict flare with 80.5% sensitivity, 91% specificity, an OR=8.62 and 82.5% PPV and 89.5% NPV.

A Cox regression model of time to flare was progressed in a similar way, adding dimensions to the reference model (Suppl. Table S1 and Fig. 3b). We estimated the performance of individual models using AUROC (Fig. 3c), which indicated that all the models had high AUROC between 100–150 days which lowered slowly over time. Overall, model-8 was again the best (AUROC=0.761) confirming that the combination of remission dimensions is highly relevant.

In order to use the concept of MDR in clinical practice for this group of patients, risk categories were defined using the three variables retained in model-8 (Total PD=0, IRC<2% and RaQoL≤1). The individual contribution to flare-free survival of each variable is shown in Suppplementary Figure 2b. MDR (3 variables fulfilling low-risk) was achieved in 43/117 (36.8%) patients, highly associated with the ability to remain in remission ($p<0.0001$). The predictive performances of achieving MDR were calculated: AUROC=0.850, SEN/SPE 52%/90.2% and PPV/NPV 90.7%/50.7%. A sur-



e **Chance of sustained remission (SR) on tapering**

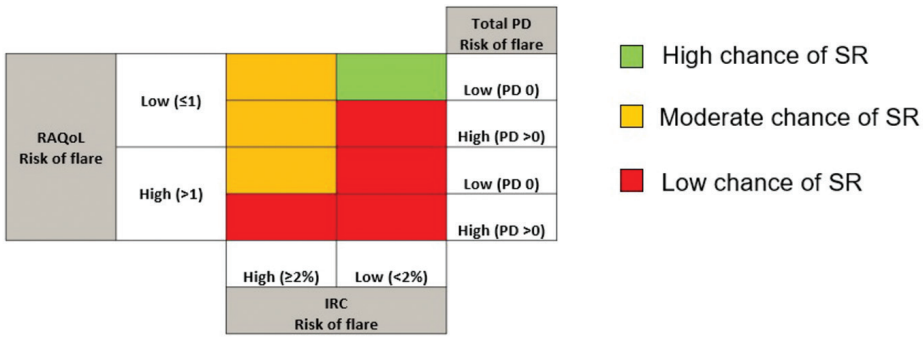


Fig. 3. Tapering cohort modelling.
a: AUROC analysis of the different prediction models (logistic regression)
b: Hazard function at mean of covariates for each COX regression model**.**
c: AUC for each Cox regression model. Legend for a) b) and c) in the side panel.
d: Kaplan Meier plot of time to flare depending on patients achieving a multi-dimensional remission (MDR) (green) *versus* not in MDR (blue).
e: Probability matrix for the chance to remain in sustained remission (high in green, intermediate in orange and low in red) after tapering.

vival analysis showed that patients in MDR (Fig. 3d) had a better cumulative survival and confirmed the cumulative effect of combining the 3 dimensions. Based on the dichotomisation of data used to define MDR, we developed a probability matrix to estimate the chance of sustained remission when offering tapering in clinical practice (Fig. 3e).

Discussion

The effect of two different management strategies (tapering *vs*. continuation) on the rate of sustained remission over 12 months, was compared in RA

patients in stable cs-DMARD-induced remission. This study demonstrated that tapering and perhaps even stopping cs-DMARDs is possible in a proportion of RA patients. It was expected that the rate of sustained remission would be lower in patients choosing to taper and we hypothesised that biomarkers at baseline could predict disease flare. Loss of remission was indeed more frequently observed in the tapering group (36% compared to 20%), while remission was re-captured in 97% of patients (resuming therapy at the previous effective dose). Our study further demonstrated that successful tapering can be predicted using objective measures in addition to those routinely used to define DAS-remission. Furthermore, our study highlights the safety of structured tapering while fulfilling the need for evidence-based predictive biomarkers, to only offer tapering to the patients able to maintain remission, along with rapid re-introduction of treatment if necessary.

Differences in baseline characteristics between patients who chose to taper were identified. Asymptomatic patients who had been in stable remission longer were more likely to be willing to reduce their drug burden. Although not significant, patients who chose tapering also demonstrated generally lower PRO scores. Males were more likely to accept tapering, which may reflect an increase in risk-taking behaviour (38), while alternatively, differences in perceived health status may vary between genders as suggested by non-significant higher PRO scores in females (data not shown). Lower serum levels of inflammation (ESR/%IRC) were also observed in the patients who chose tapering, compatible with longer remission duration.

The study participants demonstrated overall very low levels/absence of clinically apparent and sub-clinical inflammation (median DAS28 of 1.15, no SJC and normal CRP/ESR), which is necessary when considering tapering. Despite this, spontaneous loss of disease control occurred in the non-tapering group, perhaps due to less well-controlled inflammation demonstrated by slightly higher scores for baseline IRC (%) and PD. Furthermore, compared with data acquired when patients first achieve remission (34), patients with stable remission for >6 months showed trends for lower scores for most characteristics at baseline, with reductions in the ranges/IQRs for TJC/SJC/CRP/PD/GS/PROs, while medians remained largely the same (data not shown). In contrast, an increase in naive T-cells and a reduction in IRCs but no change for Treg were demonstrated (data not shown). This confirms that DAS28-remission is a dynamic state, with some patients continuing to improve over time (notably for T-cell subsets), while also highlighting that the duration of remission may allow improvement in certain dimensions over time, enabling to reach cut-offs needed to achieve MDR. An association between shorter symptom duration and seronegativity (for ACPA and IgM-RF) with sustained remission seen in historic studies, was not observed in our study, perhaps as our out-patient population was highly heterogeneous compared to trial patients, notably in term of disease duration (29–82 months).

Differences in baseline characteristics between patients who chose to taper *vs*. not were identified. Although not significant, patients who chose tapering demonstrated generally lower PRO scores. Asymptomatic patients who had been in stable remission longer were more likely to be willing to reduce their drug burden. Males were more likely to accept tapering, which may reflect an increase in risk-taking behaviour (38), while alternatively, differences in perceived health status may vary between genders as suggested by non-significant higher PRO scores in females (data not shown). Lower serum levels of inflammation/sub-clinical disease (ESR/%IRC) were also observed in the patients who chose tapering, compatible with longer duration of remission. Those who tapered also had higher total GS scores, reflecting cumulative synovial thickening, which is also likely to reflect longer overall disease duration.

Modelling was performed to evaluate whether multi-dimensional models combining imaging, immunological and PRO biomarkers to clinical parameters (model-1) routinely collected provided added-value to predict successful tapering of cs-DMARDs. Models-2/3/4 (adding 1 dimension) demonstrated an increase in accuracy, with imaging providing modest improvement, PROs a sizeable increase and then T-cells adding the most value. There was further improvement in predictive accuracy when adding +2 dimensions, and clinical variables were no longer retained by some models but included the need for being in DAS-remission. Interestingly, models 6 and 7 provided the same level of accuracy, suggesting that PROs and US are of similar added-value when combined with clinical + T-cells, while still providing +2.6% improvement. This is an important consideration for clinical practice since obtaining PROs is more feasible and inexpensive compared to US, while T-cells still appear to have the largest added-value. Ultimately, the combination of 4 dimensions (model-8) demonstrated the best predictive accuracy, thus suggesting capacity for US to still add value. Cox regression analysis also confirmed the performance of model 8 as superior to the 3-dimension models.

MDR, as established here for patients with ≥6 months stable remission and defined using dichotomisation of variables retained by model-8, was different to previously reported data (34) and reflect the dynamics of being in remission as defined only by DAS28. MDR was highly associated with the ability to sustain remission and a good AUROC. Dichotomisation of linear data is essential to facilitate risk categorisation in clinical practice and with high specificity and PPV (both >90%), MDR has the potential to select individuals with the best chances of successful tapering. Of those classified as low risk, only 4/43 (9.3%) flared while for those with a high risk, we indeed observed flares in 37/73 (50.6%).

This study had limitations, mainly with missing data. Unfortunately, some patients did not attend pre-arranged US scans and blood tests in addition to some follow-up data not being collected due to having to perform telephone consultations. Absence of data was not associated with lower PROs, suggesting it was indeed random rather than

a result of patients' perception of being in a better health. This nevertheless needs replication in a new group of patients.

Conclusion

This study both identifies and utilises objective predictors of successful tapering of cs-DMARD therapy, in addition to proposing a tool for managing tapering prospectively. Our results reflect real-life practice of tapering in an out-patient clinic and demonstrate that successful tapering of cs-DMARDs is possible in ~60% of patients and that it could be predicted. Several baseline characteristics were individually associated with sustained remission and through statistical modelling, the combination of IRC/PD/RAQoL predicted sustained remission following tapering with an accuracy of 85%. Had tapering only been offered to patients at low risk of flare using MDR-status, only 4/43 (9.3%) would have deteriorated compared to 41/117 (35%) based on only achieving DAS28-remission. From a practical viewpoint, by only offering tapering to patients with a low risk of disease flare using this tool, it may be possible to rationalise their follow-up, *i.e.* less often than would normally be advised, which could free up resources and waiting list times.

These would also have an excellent chance of regaining remission with re-initiating treatment. Further validation is needed, however, this work will allow the design of studies with a limited number of variables to be collected.

References

1. AGA AB, LIE E, UHLIG T *et al*.: Time trends in disease activity, response and remission rates in rheumatoid arthritis during the past decade: results from the NOR-DMARD study 2000-2010. *Ann Rheum Dis* 2015; 74: 381-8. https://doi.org/10.1136/annrheumdis-2013-204020
2. COMBE B, RINCHEVAL N, BENESSIANO J *et al*.: Five-year favorable outcome of patients with early rheumatoid arthritis in the 2000s: data from the ESPOIR cohort. *J Rheumatol* 2013; 40: 1650-7. https://doi.org/10.3899/jrheum.121515
3. LENERT A, LENERT P: Tapering biologics in rheumatoid arthritis: a pragmatic approach for clinical practice. *Clin Rheumatol* 2017; 36: 1-8. https://doi.org/10.1007/s10067-016-3490-8
4. KUIJPER TM, LAMERS-KARNEBEEK FBG, JACOBS JWG, HAZES JMW, LUIME JJ: Flare rate in patients with rheumatoid arthritis in low disease activity or remission when tapering or stopping synthetic or biologic DMARD: a systematic review. *J Rheumatol* 2015; 42: 2012-22. https://doi.org/10.3899/jrheum.141520
5. GRIJALVA CG, CHUNG CP, ARBOGAST PG, STEIN CM, MITCHEL EF JR, GRIFFIN MR: Assessment of adherence to and persistence on disease-modifying antirheumatic drugs (DMARDs) in patients with rheumatoid arthritis. *Med Care* 2007; 45 (10 Suppl. 2): S66-76. https://doi.org/10.1097/mlr.0b013e318041384c
6. HUGHES LD, DONE J, YOUNG A: A 5 item version of the Compliance Questionnaire for Rheumatology (CQR5) successfully identifies low adherence to DMARDs. *BMC Musculoskelet Disord* 2013; 14: 286. https://doi.org/10.1186/1471-2474-14-286
7. LEDINGHAM J, GULLICK N, IRVING K *et al*.: BSR and BHPR guideline for the prescription and monitoring of non-biologic disease-modifying anti-rheumatic drugs. *Rheumatology* (Oxford) 2017; 56: 865-8. https://doi.org/10.1093/rheumatology/kew479
8. SCOTT IC, KINGSLEY GH, SCOTT DL: Can we discontinue synthetic disease-modifying anti-rheumatic drugs in rheumatoid arthritis? *Clin Exp Rheumatol* 2013; 31 (4 Suppl. 78): S4-8.
9. DEIGHTON C, O'MAHONY R, TOSH J, TURNER C, RUDOLF M: Management of rheumatoid arthritis: summary of NICE guidance. *BMJ* 2009; 338: b702. https://doi.org/10.1136/bmj.b702
10. 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. https://doi.org/10.1136/annrheumdis-2015-207524
11. SINGH JA, SAAG KG, BRIDGES SL JR *et al*.: 2015 American College of Rheumatology Guideline for the Treatment of Rheumatoid Arthritis. *Arthritis Rheumatol* 2016; 68: 1-26. https://doi.org/10.1002/acr.22783
12. VAN MULLIGEN E, WEEL AEAM, KUIJPER TM, HAZES JMW, VAN DER HELM-VAN MIL AHM, DE JONG PHP: The impact of a disease flare during tapering of DMARDs on the lives of rheumatoid arthritis patients. *Semin Arthritis Rheum* 2020; 50: 423-31. https://doi.org/10.1016/j.semarthrit.2020.02.011
13. TEN WOLDE S, BREEDVELD FC, HERMANS J *et al*.: Randomised placebo-controlled study of stopping second-line drugs in rheumatoid arthritis. *Lancet* 1996; 347: 347-52. https://doi.org/10.1016/s0140-6736(96)90535-8
14. AHERN MJ, HALL ND, CASE K, MADDISON PJ: D-penicillamine withdrawal in rheumatoid arthritis. *Ann Rheum Dis* 1984; 43: 213-7. https://doi.org/10.1136/ard.43.2.213
15. KREMER JM, RYNES RI, BARTHOLOMEW LE: Severe flare of rheumatoid arthritis after discontinuation of long-term methotrexate therapy. Double-blind study. *Am J Med* 1987; 82: 781-6. https://doi.org/10.1016/0002-9343(87)90015-5
16. DE SILVA M, HAZLEMAN BL: Long-term azathioprine in rheumatoid arthritis: a double-blind study. *Ann Rheum Dis* 1981; 40: 560-3. https://doi.org/10.1136/ard.40.6.560
17. GØTZSCHE PC, HANSEN M, STOLTENBERG M *et al*.: Randomized, placebo controlled trial of withdrawal of slow-acting antirheumatic drugs and of observer bias in rheumatoid arthritis. *Scand J Rheumatol* 1996; 25: 194-9. https://doi.org/10.3109/03009749609069987
18. VAN DER LEEDEN H, DIJKMANS BA, HERMANS J, CATS A: A double-blind study on the effect of discontinuation of gold therapy in patients with rheumatoid arthritis. *Clin Rheumatol* 1986; 5: 56-61. https://doi.org/10.1007/bf02030968
19. LANDEWÉ RB, BOERS M, VERHOEVEN AC *et al*.: COBRA combination therapy in patients with early rheumatoid arthritis: long-term structural benefits of a brief intervention. *Arthritis Rheum* 2002; 46: 347-56. https://doi.org/10.1002/art.10083
20. MOTTONEN T, HANNONEN P, LEIRISALO-REPO M *et al*.: Comparison of combination therapy with single-drug therapy in early rheumatoid arthritis: a randomised trial. FIN-RACo trial group. *Lancet* 1999; 353: 1568-73. https://doi.org/10.1016/s0140-6736(98)08513-4
21. MARCHESONI A, BATTAFARANO N, ARREGHINI M, PANNI B, GALLAZZI M, TOSI S: Radiographic progression in early rheumatoid arthritis: a 12-month randomized controlled study comparing the combination of cyclosporin and methotrexate with methotrexate alone. *Rheumatology* (Oxford). 2003; 42: 1545-9. https://doi.org/10.1093/rheumatology/keg394
22. KLARENBEEK NB, VAN DER KOOIJ SM, GULER-YUKSEL M *et al*.: Discontinuing treatment in patients with rheumatoid arthritis in sustained clinical remission: exploratory analyses from the BeSt study. *Ann Rheum Dis* 2011; 70: 315-9. https://doi.org/10.1136/ard.2010.136556
23. WANG L, GENG Y, HAN J, SUN X, ZHANG Z: A combination model to predict relapse and successful conventional DMARDs de-escalation in rheumatoid arthritis patients with sustained clinical remission. *Clin Exp Rheumatol* 2019; 37: 120-6.
24. MOTTONEN T, HANNONEN P, LEIRISALO-REPO M *et al*.: Comparison of combination therapy with single-drug therapy in early rheumatoid arthritis: a randomised trial. FIN-RACo trial group. *Lancet* 1999; 353: 1568-73. https://doi.org/10.1016/s0140-6736(98)08513-4
25. QUINN MA, EMERY P: Window of opportunity in early rheumatoid arthritis: possibility of altering the disease process with early intervention. *Clin Exp Rheumatol* 2003; 21 (5 Suppl. 31): S154-7.
26. VAN DER WOUDE D, YOUNG A, JAYAKUMAR K *et al*.: Prevalence of and predictive factors for sustained disease-modifying antirheumatic drug-free remission in rheumatoid arthritis: results from two large early arthritis cohorts. *Arthritis Rheum* 2009; 60: 2262-71. https://doi.org/10.1002/art.24661
27. BAKER KF, SKELTON AJ, LENDREM DW *et al*.: Predicting drug-free remission in rheumatoid arthritis: A prospective interventional cohort study. *J Autoimmun* 2019; 105: 102298. https://doi.org/10.1016/j.jaut.2019.06.009

28. 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. https://doi.org/10.1002/art.23945
29. 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. https://doi.org/10.1136/annrheumdis-2011-200548
30. COPE AP: T cells in rheumatoid arthritis. *Arthritis Res Ther* 2008; 10 (Suppl. 1): S1.
31. PONCHEL F, BURSKA AN, HUNT L *et al*.: T-cell subset abnormalities predict progression along the inflammatory arthritis disease continuum: implications for management. *Sci Rep* 2020; 10: 3669. https://doi.org/10.1038/s41598-020-60314-w
32. PONCHEL F, GOEB V, PARMAR R *et al*.: An immunological biomarker to predict MTX response in early RA. *Ann Rheum Dis* 2014; 73: 2047-53. https://doi.org/10.1136/annrheumdis-2013-203566
33. HUNT L HENSOR EM, NAM J *et al*.: T cell subsets: an immunological biomarker to predict progression to clinical arthritis in ACPA-positive individuals. *Ann Rheum Dis* 2015.
34. GUL HL, EUGENIO G, RABIN T *et al*.: Defining remission in rheumatoid arthritis: does it matter to the patient? A comparison of multi-dimensional remission criteria and patient reported outcomes. *Rheumatology* (Oxford) 2020; 59: 613-21. https://doi.org/10.1093/rheumatology/kez330
35. 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. https://doi.org/10.1136/ard.2009.117341
36. SCHETT G, EMERY P, TANAKA Y *et al*.: Tapering biologic and conventional DMARD therapy in rheumatoid arthritis: current evidence and future directions. *Ann Rheum Dis* 2016; 75: 1428-37. https://doi.org/10.1136/annrheumdis-2016-209201
37. 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. https://doi.org/10.1136/rmdopen-2016-000428
38. BYRNES JP, MILLER DC, SCHAFER WD: Gender differences in risk taking: a meta-analysis. *Psychol Bull* 1999;125: 367-83. https://psycnet.apa.org/doi/10.1037/0033-2909.125.3.367