



T-cell subset biomarkers across the rheumatoid arthritis disease continuum: from clinical utility to adoption in daily practice

Innocent C. Anioke^{1,} , Fatih Tastekin^{1,a,} , Gulay Alp^{1,b,} , Helen Ng¹, Preveena Ravi¹, Isobel Parker¹, Min Lou^{1,} , Hanna Gul¹, Laurence Duquenne^{1,} , Edith Villeneuve¹, Jacqueline L. Nam^{2,} , Sana Sharrack^{1,3,} , Maya H. Buch^{1,4,} , Philip G. Conaghan^{1,3}, Kulveer Mankia^{1,} , Paul Emery^{1,3,} , Frederique Ponchel^{1,5,*}

¹Leeds Institute of Rheumatic and Musculoskeletal Medicine, The University of Leeds, Leeds, UK

²Leeds Teaching Hospitals NHS Trust, Leeds, UK

³NIHR Leeds Biomedical Research Centre, Chapel Allerton Hospital, Leeds, UK

⁴Division of Musculoskeletal & Dermatological Science, University of Manchester, Manchester, UK

⁵Département de Rhumatologie, Université Jean Monnet Saint-Etienne, CHU Saint-Etienne, Mines Saint-Etienne, INSERM, SAINBIOSE U1059, Saint-Etienne, France

*Correspondence to: Frederique Ponchel, Faculte de Medecine, Batiment recherche, Universite Jean Monet, 42055, Saint Etienne, France.

E-mail: frederique.ponchel@univ-st-etienne.fr

^aPresent address: Department of Internal Medicine, Division of Rheumatology, Izmir Bayrakli City Hospital, Izmir, Turkey.

^bPresent address: Department of Internal Medicine, Division of Rheumatology, Uşak University Faculty of Medicine, Uşak, Turkey.

Abstract

Objective: The biomarker potential of CD4⁺ T-cell subsets [naive, regulatory (Treg), inflammation-related cells (IRC)] in patients with rheumatoid arthritis (RA) has been described.

This article investigates the dynamic changes in these biomarkers across the RA disease continuum from at-risk to drug-induced remission.

Methods: T-cell subset biomarker data were acquired using flow cytometry. Multiple group comparisons were performed using ANOVA test (with Bonferroni correction).

Results: In individuals at risk of RA, longitudinal analysis showed that IRC frequencies increased just prior to onset of clinical synovitis, while naive T-cell frequencies reduced in those progressing to clinical synovitis, but increased in non-progressors. The use of naive/IRC data improved the accuracy of RA classification, especially in ACPA-negative patients. A distinct T-cell biomarker signature was observed in late-onset RA (>60 years old vs <59). In untreated RA, the predictive value of naive T-cell frequencies for methotrexate response was confirmed. For patients on methotrexate, naive T cells increased only between 6 and 12 months and only when in remission. IRC and Treg showed no consistent change. In patients treated with TNF inhibitors, naive T-cell frequency increased independently of response, whilst sustained IRC reductions and Treg increases were seen in remission. Once in stable clinical remission, only naive frequencies increased with the length of remission on conventional synthetic disease-modifying anti-rheumatic drugs (csDMARDs), whilst only Treg increased over time in TNF inhibitor-induced remission.

Conclusion: These studies validated that T-cell subset measurements are independent from other currently used biomarkers, highlighting differences in the impact of drug modes of action on the three T-cell subsets. There is still an unmet need for biomarkers to predict response to TNF inhibition in early RA.

Keywords CD4⁺ T-cell subsets, rheumatoid arthritis, first-line treatment, remission

Rheumatology key messages

- T-cell subsets have predictive value for each outcome of the different stages of the inflammatory arthritis continuum.
- Confounding factors such as concomitant osteoarthritis and older age needs to be considered when using these biomarkers.
- These biomarkers can be obtained from hospital services, allowing easy implementation.

Introduction

Rheumatoid arthritis (RA) is a chronic inflammatory disease that primarily affects synovial joints and causes damage to bone and cartilage, leading to pain, swelling and potential lifelong disability if untreated. The diagnosis relies on clinical signs, symptoms and autoantibodies [1], although these are not present in all patients at presentation, highlighting early disease heterogeneity. A pre-clinical RA phase also occurs [2, 3], characterized by systemic autoimmunity [presence of anti-citrullinated peptide antibodies (ACPA)] without clinical synovitis [4]. Given the destructive nature of RA, early remission is key to its management [5]. Thus, novel biomarkers are crucial for predicting progression in at-risk individuals and managing disease progression in people with RA in the short and long term.

In early, drug-naïve RA, our group previously identified changes in the frequencies of CD4+ naïve T cells and regulatory T cells (Treg). We also identified a sub-population of naïve T cells which correlated with the C-reactive protein (CRP) that we called inflammation-related cells (IRC) [6, 7]. We further reported on the potential of these 3 CD4+ T-cell subsets as novel biomarkers for the management of RA patients.

In the pre-clinical phase of RA, we observed disturbances in naïve, Treg and IRC frequencies [8]. Around 40% of at-risk cases progressed to undifferentiated and/or rheumatoid arthritis (UA/RA). Various prediction models using clinical and imaging data have been developed to predict progression [9–13]. Individual T-cell biomarkers were also predictive of progression to inflammatory arthritis, and their value when combined with clinical data was established in predictive models in 150 individuals [14]. After the onset of clinical synovitis, naïve cells had potential for predicting RA classification in patients from an early arthritis clinic (EAC) [14]. A regression analysis suggested that only three variables [naïve T cells, disease activity score (DAS) and age] contributed significantly to the model and achieved 80% accuracy.

Following RA classification, treatment initiation typically involves csDMARDs, most commonly methotrexate (MTX). Naïve T-cell frequency was independently associated with achieving remission on MTX in early RA [15] and a predictive model was developed, selecting the three most predictive biomarkers (naïve T cells, smoking status, DAS28 scores) in 70 patients [14]. Clinical trials have suggested the superiority of anti-tumour necrosis factor agents (TNFi) in achieving better long-term outcomes in early RA [16], but it remains unclear whether T-cell biomarkers can predict response to biological-DMARDs (bDMARDs), although data from nine patients suggested no association with outcome [15]. T-cell biomarkers were widely distributed in RA patients in clinical remission on either csDMARDs or bDMARDs, despite low disease activity (DAS28 < 2.6) [17]. The

ability to sustain remission over 12 months once achieved on csDMARDs was associated with higher naïve cells [14]. Successful tapering of csDMARDs was predicted by IRC in a model using ultrasound power Doppler and the Rheumatoid Arthritis Quality of Life questionnaire (RAQoL) [18]. Treg predicted flare in patients discontinuing a bDMARDs [19] and the successful tapering of bDMARDs was also predicted by Treg and IRC with CRP [20].

Following these initial reports, understanding the dynamic changes in these biomarkers across the early stages of RA and establishing whether they are impacted by demographic or clinical correlates, notably ACPA positivity, remains to be established. Here, data from 1727 participants across >2500 visits at distinct stages of the RA disease continuum are presented. While validating our prior work (detailed in the [supplementary materials](#)), our main goal was to understand these biomarkers' dynamics and their relationships with other current RA biomarkers across the RA disease continuum, from the at-risk stage to drug-induced remission.

Methods

Patients

Ethical approval was obtained from National Research Ethics Committees at different phases of the RA continuum (National Research Ethics Service, West Yorkshire Ethics Committee: REC09/H1307/98, REC10/H1307/138, NCT02433184, REC06/Q1205/169). All participants provided informed consent prior to recruitment. Participants included in this analysis were selected based on availability of T cells subset data. Healthy controls ($n=174$) were used to establish age-associations for naïve and Treg measurements as previously described [14, 21]. Groups included ACPA+ at risk of RA, a group previously extensively described [8, 9], EAC patients evolving to RA/alternative classifications over time, early RA treated with MTX or with MTX +TNFi from the drug arm of two clinical trials [22, 23], and patients achieving DAS28 < 2.6 remission. These groups are described in more detail in the [supplementary materials](#).

Clinical data

Standard data were recorded, including (although not for all groups) age, sex, autoantibody status (ACPA/RF), smoking (ever/never), duration of symptoms, disease activity or remission, HLA-DR status, early morning stiffness (EMS), tender joint count (78 joints), swollen/tender joint count (28 joints), C-reactive protein (CRP), erythrocyte sedimentation rate (ESR) and global health (GH). DAS28 was calculated where relevant.

Cell staining and flow cytometry strategies

Subset quantification for the three subsets was performed by NHS hospital services, according to standardized protocols as previously described [14] between 2013 and 2020, using fresh blood. Gating strategies were previously described [14]. Subset frequencies were reported as % of CD4+ T cells and normalized as previously described for naive and Treg [14]. A reminder of the overall methodology is provided in the [supplementary materials](#). Treg measurements were affected by the shortage of one antibody for several months.

To produce this data, raw data files were retrieved from NHS servers and analysis was performed by an experienced scientist (ICA). Issues were encountered, stemming from human or technical errors, or compromised blood quality. A detailed analysis of the acceptability of data leading to recommendations for implementing such biomarkers in daily practice is presented in the [supplementary materials](#).

Statistical analysis

Continuous variables were described using median, IQR and range. Nominal variables were described as *n* (%). Non-parametric tests were used throughout. For univariate analysis, continuous variable measures were compared between outcomes using Mann–Whitney *U* (MWU) tests and nominal variables with Chi². Correlations were investigated with Spearman tests and described using rho-values. Multiple group analyses were performed using ANOVA with post-hoc 2-group analysis (MWU), using Bonferroni corrections. Analyses were conducted using SPSS 29. Significance for *P*-value was set at <0.05.

Multivariate analysis for prediction models is described in [supplementary materials](#). Predictions were described with odds ratio (OR) and area under the curve (AUC). To evaluate the robustness of our predictive models, we used internal validation using a bootstrapping procedure (500 permutations) to correct for optimism in model performance (generating 95% CI AUC). All study cohorts contained missing data. The pattern of missingness was consistent with data missing at random (Little's MCAR test *P* < 0.0001). Regression-based imputation generated unbiased parameter estimates defining appropriate boundaries for each variable and the 'Random Number Generators' function, followed by the 'Impute Missing Data Values' (five iterations, in SPSS). Model estimates were combined ('pooling' function) to generate overall multiple-imputation estimates. The pooled dataset was compared with the original dataset (odds ratios remained stable) and data distribution checks confirmed no significant differences from non-imputed data. The pooled dataset was used for modelling.

Results

Participants (*n* = 1727) at various stages of the RA continuum are described in [Table 1](#) and [Supplementary Fig. S1](#).

Flow cytometry results

The gating to report data (% of CD4+ T cells) uses simple classic 2D strategies ([Supplementary Fig. S2](#)) as previously reported

[14]. We also present a more advanced analysis (UMAP/tSNE maps) for multiple marker expression, a cluster analysis for the presence of subcategories of cells and a SPADE analysis to highlight the ontology of subsets in [Supplementary Fig. S3](#).

Validation of previous model replicating the predictive value of T-cell biomarkers

We used a forward regression strategy to reduce the number of significantly contributing predictors, directly comparing models with/without adding the T-cell biomarkers.

Outcomes predicted:

- progression to IA [Supplementary Table S1](#);
- RA classification [Supplementary Table S2](#);
- MTX-induced remission [Supplementary Table S3](#).

Results are summarized in [Fig. 1](#), presenting the gain in performance indexes when adding T-cell biomarkers (black) over clinical data only (white) to each model. All details of this analysis are presented and discussed in the [supplementary materials](#). AUC are presented in [Supplementary Fig. S4](#). A summary figure of the predictive value of T-cell subsets is also included after the models.

Progression to inflammatory arthritis in ACPA± at-risk individuals without clinical synovitis

All three biomarkers were significantly different compared with healthy controls (HC, *n* = 174), progressors (Pr, *n* = 198), non-progressors (NP, *n* = 244) and data at onset of clinical synovitis ([Fig. 2a](#), ANOVA *P* < 0.0001). There was a significant difference between Pr and NP (*P* = 3.8×10^{-7}), but not between Pr and onset of clinical synovitis (except for IRC *P* = 0.006) nor between NP and HC.

We did not observe any association between T-cell biomarkers and demographic/clinical data, although trends were seen between age and naive/Treg data ([Supplementary Fig. S5](#)).

Results from a citrullinated peptide (CCP) research test for a repertoire of citrullinated peptides particularly associated with the risk of progression were available [24] in 278 patients. A total of 117/278 (42%) were negative and provided a relevant comparison group. Only 21/117 (18%) of patients progressed in the negative group compared with 100/121 (82%) in the positive group (χ^2 , *P* = 2.3×10^{-13}). The distribution of T-cell biomarkers between positive/negative Pr and NP was significantly different ([Fig. 2b](#), ANOVA *P* < 0.0002). Pr showed no significant difference associated with CCP-research positivity. A difference between NP-ACPA± was observed only for IRC, suggesting that the T-cell subset association is indeed with progression and not with CCP research test results. To validate this conclusion, we tested the added value of the T-cell biomarkers for predicting progression over that of, or in combination with, the CCP-Second Gen test. Logistic regression selected Treg as the best predictor, followed by CCP-Second Gen, then naive T cells, and finally IRC with an overall 75.5% accuracy, AUC = 0.841 ([Supplementary Fig. S4](#)). This was better than the CCP-Second Gen alone (65% accuracy, AUC = 0.720) or the 3 T cell biomarkers together (73% accuracy,

Table 1 Demographic and clinical data for all participating groups.

Cohorts/ Variables	HC	Arthralgia	Arthralgia Onset of IA	Early RA	First MTX treatment	First Anti-TNF treatment	Remission synthetic	Remission biologic	UA/ others	PsA
<i>n</i>	174	442	77	360	221	78	418	265	82	44
[§] Age (years)	43 (24) [19–71]	51 (18) [18–82]		51 (19) [21–87]	57 (17) [19–87]	49 (21) [25–80]	59 (19) [19–87]	56 (22) [22–85]	48 (26) [23–83]	45 (18) [19–66]
Sex M	(41)	130 (29.6)		101 (28.0)	70 (31.7)	25 (32.5)	150 (36.4)	82 (31.1)	25 (29.6)	18 (41.9)
F	(58)	311 (70.4)		259 (72.0)	151 (68.3)	53 (67.5)	268 (63.6)	182 (68.9)	57 (70.4)	26 (58.1)
[§] Duration of symptom/ disease ^b	NA	64 (124) [0–1830]		24 (25) [1–104]	24 (27) [2–105]	19.2 (16) [3–80]	48 (38) [1–260]	74 (119) [3–430]	23 (34) [6–72]	30 (67) [8–176]
[§] Duration of remission (months)	NA	NA	NA	NA	NA	NA	11 (2 796) [1–128]	9 (22) [1–153]	NA	NA
RF Pos		145 (42.3)		152 (63.1)	122 (57.5)	52 (67.5)	192 (58.0)	138 (62.2)	6 (11.5)	
Neg	NA	198 (57.7)		89 (36.9)	90 (42.5)	25 (32.5)	139 (42.0)	84 (37.8)	46 (88.5)	NA
ACPA Pos				190 (55.0)	133 (60.2)	62 (80.5)	235 (70.6)	179 (80.6)	6 (10.9)	
Neg	NA	NA		155 (45.0)	86 (38.9)	15 (19.5)	98 (29.4)	43 (19.4)	49 (89.1)	NA
Smoking ever		258 (58.6)		153 (43)	127 (58)	43 (55.1)	176 (54.5)	111 (51.4)	27 (56.3)	10 (41.7)
Never		182 (41.4)		202 (57)	92 (42)	35 (44.9)	147 (45.5)	105 (48.6)	21 (43.8)	14 (58.3)
HLA-DR SE Pos		176 (44)		>200						
Neg	NA	222 (56)		missing data	NA	NA	Not known	Not known	NA	NA
[§] TJC		1 ^a (2) [0–27]	5 (5) [1–11]	9 (11) [0–28]	9 (12) [0–31]	11 (12) [0–28]	0 (0) [0–5]	00 (0) [0–5]	5 (9) [0–25]	7 (14) [0–44]
[§] SJC			3 (3) [1–13]	5 (7) [0–22]	5 (8) [0–22]	4 (7) [0–19]	0 (0) [0–5]	0 (0) [0–2]	1 (4) [0–13]	4 (6) [0–11]
[§] CRP (mg/L)		4.3 (6) [0–67]	4 (26) [1–77]	9 (24) [1–228]	11 (23) [0–228]	8 (27) [0–122]	0 (0) [0–34]	0 (0) [0–30]	4 (13) [0–163]	1 (7) [0–135]
[§] DAS28(4) –CRP				4.6 (2) [1–8]	4.7 (2) [1–8]	4.8 (2) [2–8]	1.6 (1) [1–3]	1.5 (1) [1–3]	3.5 (2) [1–7]	3.2 (3) [2–5]

Data are presented as Numbers (%) or [§]median (IQR) [min to max]. Note, there are missing data for most variables.

^a Tender joint count 78 was used in the arthralgia cohort.

^b Duration of symptoms in weeks in early IA but months for pre-RA. Duration of disease in years for remission. Duration of disease in years for OA.

AUC=0.807), confirming the additional independent predictive value of combined data (+11% accuracy/+0.164 AUC).

Variations in T-cell biomarkers could occur over time, and we analysed this using annual repeat (AR) visits or at time of onset of clinical synovitis. The distributions of values over time did not change for Treg, while we observed an increase over time for IRC (Fig. 2c, plotting data backward with respect to onset, $\rho=+0.225$, $P=0.0007$) and a reduction trend for naive T cells ($\rho=-0.155$, ns). Paired samples were analysed to assess whether variations in T-cell subsets happen consistently over time (Fig. 2d). Changes were calculated between pairs at inclusion/annual repeat (IN=>AR, $n=57$), annual repeat/onset (AR=>On, $n=24$), and inclusion/onset (In=>On, $n=57$). More reductions were observed for naive T cells [90/138 (65%)], notably over longer periods of time [inclusion-onset 40/57 (70%)]. Despite normalization over time of some high IRC values at inclusion, an increase was seen for IRC in pairs the closest to onset [AR=>On, 17/24 (71%)] compared with during follow-up (45%).

For Treg, no trend was observed, confirming the cross-sectional data (~50% increase/reduction). These results suggest that naive T-cell frequency may continue to reduce over time in Pr, while IRC peak closer to the time of onset of clinical synovitis. In contrast, in 57 In=>AR pairs in NP, an increase over time was seen for naive in 56/77 (73%) cases and for Treg [33/58 (57%)], and a reduction of IRC in 32/57 (57%) cases was seen, with normalization of all high values at inclusion.

RA classification in an early arthritis clinic (EAC)

T-cell biomarkers were also different between EAC outcomes (Fig. 2a, ANOVA, $P<0.006$). A reduction in naive T-cell frequencies in the RA and UA=>RA groups was observed. Higher frequencies of IRC were associated with PsA, but all disease groups had higher IRC than HC, though insignificantly. Treg were

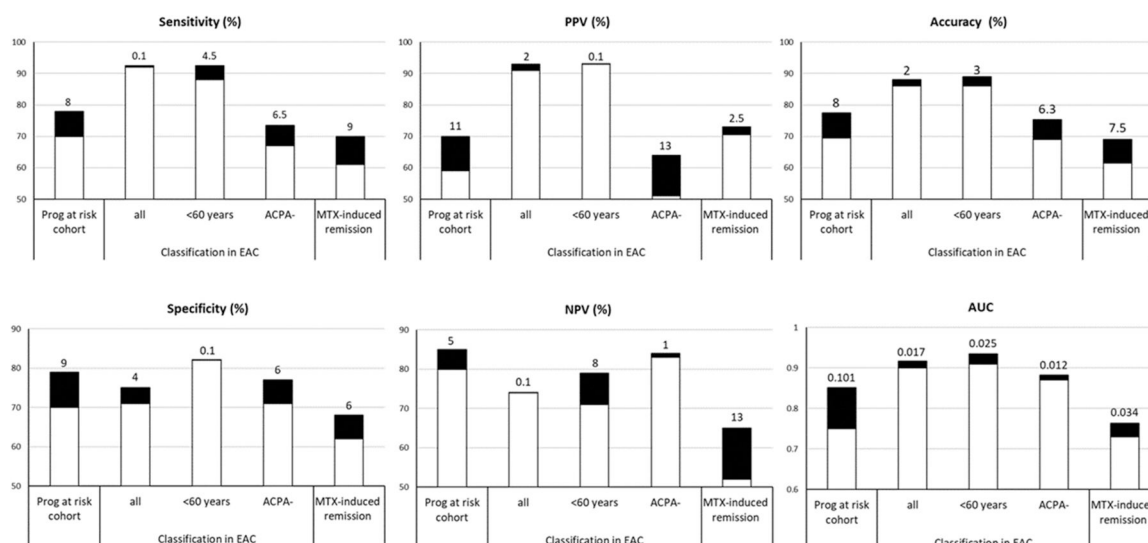


Figure 1 Summary of gain in performance indexes for logistic regression models predicting outcomes across the RA continuum. Sensitivity, specificity, accuracy, positive and negative predictive values (PPV/NPV) were calculated from the model performance matrices and from individual patient probability calculated by the model for AUC. Open bars represent the index values from the regression for clinical data-only models and the black bar above indicates the gain in performance (+ value as %) of these indexes for the model with clinical + T-cell biomarkers. Each stage of the RA continuum is described alongside the x-axis

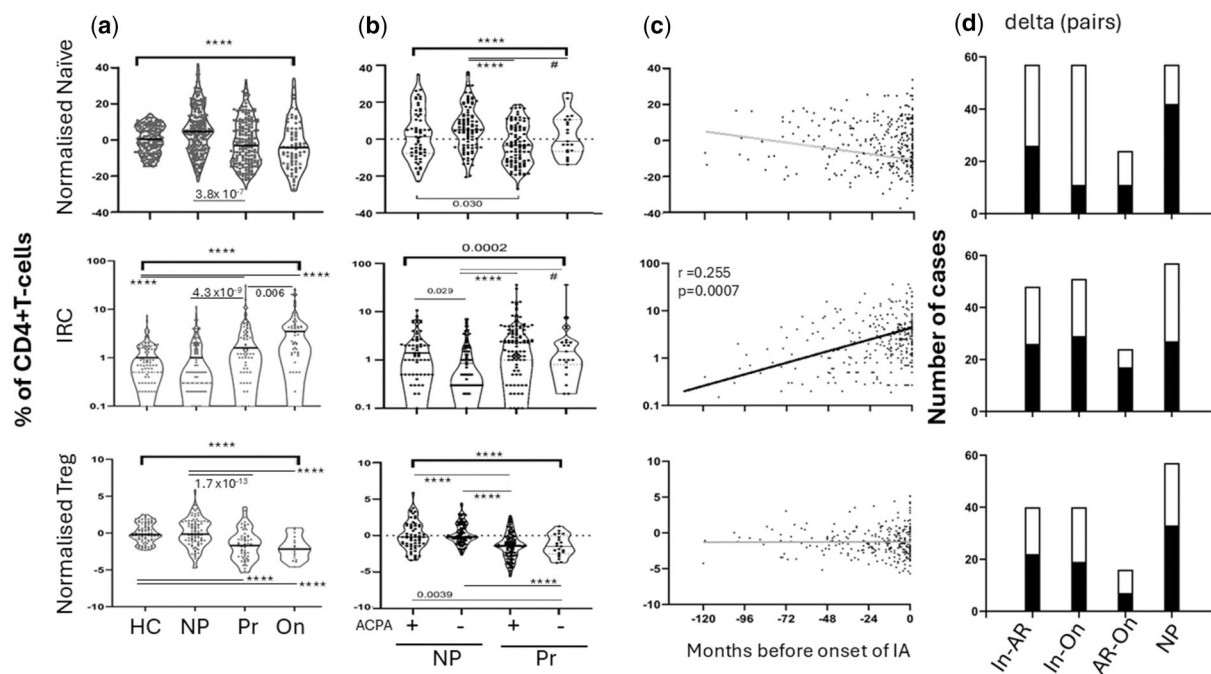


Figure 2 T-cell biomarker data for naive (normalized), IRC and Treg (normalized) in the at-risk of RA, ACPA+ cohort. **a)** T-cell biomarker data displayed in violin plots comparing healthy control (HC, $n = 174$ for naive/IRC and $n = 120$ for Treg), progressor (Pr, $n = 198$) and non-progressor (NP, $n = 244$) ACPA+ at-risk participants and in Pr at time of onset of disease ($n = 77$). **b)** A total of 278 at-risk participants were tested using a research CCP test. T-cell biomarker data are displayed comparing PR to NP, either positive or negative for the CCP test. **c)** Distribution of T-cell biomarkers are displayed in Pr only ($n = 198$) over time, retrospectively calculated with respect to onset of inflammatory arthritis (IA). Correlations were only significant for IRC with a trend for naive cells. **d)** Changes in T-cell biomarker data in paired samples [number of patients showing increase (black bar) and reduction (open) in T-cell biomarker] between inclusion/annual-repeat (In-AR, $n = 57$), annual-repeat/onset of disease (AR-On, $n = 24$), and inclusion/onset (In-On, $n = 57$) as well as in NP ($n = 53$). Thick lines indicate ANOVA and thin lines MWU following Bonferroni correction with P -values (**** $P < 0.0001$). Spearman correlations were calculated (best line of fit are displayed), with rho coefficient and P -value

reduced in UA=>RA/RA. Comparing RA (Fig. 2b, $n=302+77$) to non-RA ($n=126$), only naive (MWU, $P=0.0009$) were significantly reduced in RA while IRC/Treg were borderline.

We investigated data in the RA group ($n=302$ from EAC and 77 at onset of clinical synovitis) for association between T-cell biomarkers and other parameters (Supplementary Fig. S5). There was no association with gender, smoking status or with classification parameters (RF, CRP, joint counts), although lower naive T cells were observed in ACPA+ RA (unadjusted MWU $P=0.012$). Symptom duration did not affect T-cell biomarkers (Fig. 3c), suggesting that disturbances in T-cell subsets are likely to be established during the pre-clinical phase of RA. In contrast, trends for correlations between age and the naive or Treg biomarkers were observed (Fig. 3d, $\text{Rho}=+0.301$, $P=8.1 \times 10^{-9}$ for naive T cell and $\text{rho}=-0.325$, $P=5.2 \times 10^{-9}$ for Treg).

Both correlations appeared driven by a group of patients aged >60 years, not mirrored by a decline in younger patients. We separated RA developing in older patients as late-onset RA (Fig. 3e, >60 years, $n=113$, median age 68 years, range 60–82) vs younger RA patients ($n=266$, median age 47, range 21–59). Naive T cells were lower in the younger RA group (median 0.2%) than in the late-onset RA group (+7.2%, $P<0.0001$). This was

also observed for Treg ($P=0.0001$) where late-onset RA patients had lower Treg (MWU $P=0.001$). IRC were higher in the younger RA group ($P=0.071$). The age relationships with time disappeared, suggesting they indeed represent different groups. RF positivity in late-onset RA patients (43%) was less frequent than in younger RA patients (68%, χ^2 $P=0.00011$) as previously reported [25]. In contrast, ACPA positivity was 80% in the younger patients and 74% in late-onset RA. Other parameters (CRP/SJC-TJC/duration/smoking status) were not different. This is suggesting possible over-classification based on the presence of ACPA autoantibodies in >60-year-old patients.

Of note, regression models for RA classification in >60 and <60-year-old participants are described in Supplementary Table S2. Only ACPA and SJC were predictive of late-onset RA (with no value for T-cell biomarkers), while the model in younger patients performed better than in the overall group, while selecting the same variables (Fig. 1).

Response to MTX±TNFi

We combined patients from the two clinical trial arms ($n=78$) of whom 48/78 (61%) achieved remission at 6 months. There was

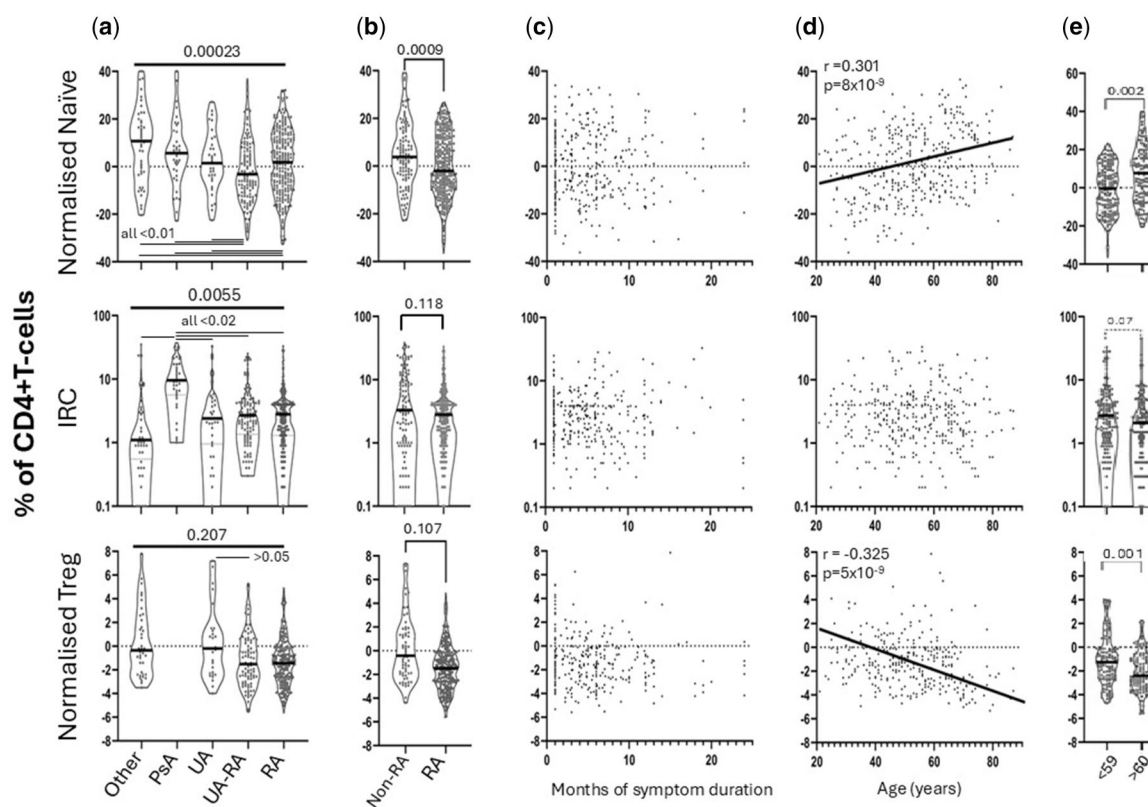


Figure 3 T cell biomarker data for naive (normalized), IRC and Treg (normalized) in an early arthritis clinic cohort. **a)** T-cell biomarker data are displayed comparing various outcomes after 24 months follow-up: 238 patients had RA at baseline and 64 evolved to RA over time, 37 remained undifferentiated arthritis, 44 developed psoriatic arthritis, 45 had others diagnoses (non-persistent inflammation, reactive arthritis, gout). **b)** T-cell biomarker data are displayed comparing RA & UA-RA ($n=302$) versus all others as non-RA ($n=126$). **c)** Distribution of T-cell biomarkers is displayed over duration of symptoms at the time of assessment in RA patients ($n=283$ due to missing data). No correlations were observed. **d)** Distribution of T-cell biomarkers is displayed with respect to age (years) in RA patients only ($n=302$). Correlations were only significant for naive cells and Treg. **e)** T-cell biomarker data are displayed comparing RA patients below (<59, $n=189$) and above (>60, $n=113$) the age of 60 years old. Thick lines indicate ANOVA and thin lines MWU following Bonferroni correction with P -values. Spearman correlations were calculated (best line of fit are displayed), with rho coefficient and P -value

no difference in naive, Treg and IRC frequencies between outcomes (Fig. 4a), despite a lower Treg median in remission. No clinical variable was associated with achieving remission (Supplementary Table S4) and no model could be computed, confirming the still unmet need for biomarkers to predict response to TNFi.

MTX effect on T-cell subsets

The effect of MTX treatment on T-cell biomarkers was analysed over time ($n=94$), separating patients based on achieving DAS remission at 6 months. For patients in remission, treatment remained stable until 12 months ($n=42$). In the non-remission group ($n=52$), patients were escalated to various combinations of csDMARDs after 6 months, some with samples at 12 months ($n=27$). We analysed paired samples by calculating changes (delta) in T-cell biomarkers at 6 or 12 months (Fig. 4b). No clear direction of change for naive T cells was seen in the non-remission group at 6 months (24 increases/20 decreases). For patients achieving remission, naive T cells showed more reduction initially (11 increases/26 decreases, $P=0.025$ compared with patients not achieving remission); however, the tendency was inverted at 12 months (19 increases/13 decreases, $P=0.014$ compared with 6 m) suggesting MTX may need time to affect

naive T cells. IRC continued to increase in both groups at 6 months ($P=0.250$) and did not change direction after 12 months in the remission group. No consistent direction of change was observed for Treg ($P > 0.200$) at 6 or 12 months.

MTX±TNFi effect on T-cell subsets

A similar analysis was performed of patients receiving MTX +TNFi. In the non-responder group [$n=27/70$ (38%)], eight patients achieved remission late at 12 months and many withdrew from the trial (no samples). In the remission group at 6 months [44/70 (62%)], four flared before 12 months. Analysis of the change over time (3 m, 6 m) showed consistent increases in naive T cells in both remission [Fig. 4c, 35/42 (83%)] and non-remission [19/27 (70%), $P=0.570$] at 6 m. IRC clearly showed more reductions in both groups (>70%). We observed more increases in Treg in remission [27/44 (61%) at 6 months] than in non-remission [13/26 (50%)], but not significant ($P=0.355$), with no noticeable change at 12 months.

Importantly, 22 patients who failed MTX at 6 months received TNFi and all went into remission at 12 months. An increase in naive T cells was observed in 13/22 (60%) (Fig. 4b, Sw-12m, grey histogram). In contrast, IRC were no longer being reduced by

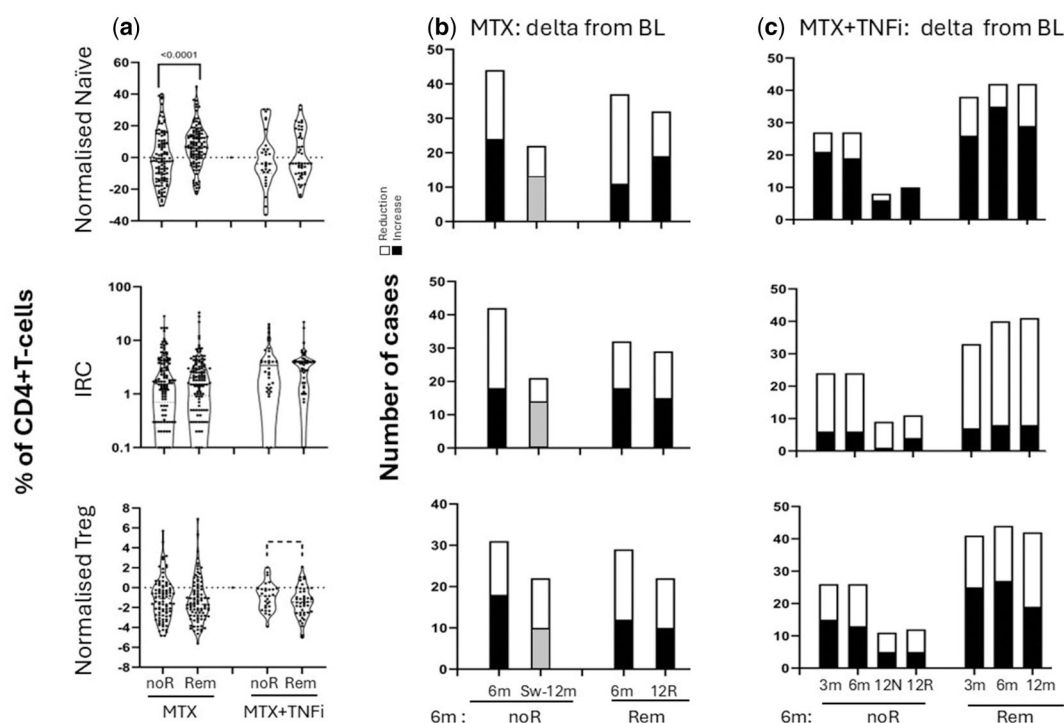


Figure 4 T-cell biomarker data for naive (normalized), IRC and Treg (normalized) in early RA patients treated with methotrexate (MTX) or a TNF-inhibitor (TNFi). **a)** T-cell biomarker data are displayed comparing two outcomes [no remission (noR) and remission (rem) after 24 weeks] in early RA patient treated with MTX ($n=221$) or MTX+TNFi ($n=78$). **b)** Number of patients showing increase (open bar) and reduction (black) in T-cell biomarker data in early RA patients treated with MTX compared 6 months to inclusion, stratifying patients by outcome at 6 months (6m) and achieving remission (Rem, $n=42$) versus no-remission (noR, $n=52$) defined by DAS28 <2.6. After 6 months, some patients not achieving remission were switched to TNFi and increase/reduction were then calculated at 12 months compared to baseline (Sw-12m). In patients achieving remission at 6 months, measures were repeated at 12 months and patients separated based on remaining in remission (12R) or not (12N). **c)** Number of patients showing increase (open bar) and reduction (black) in T-cell biomarker data in early RA patients treated with MTX+TNFi comparing data at inclusion and 3, 6, 9 or 12 months later, stratifying patients by outcome at 6 months (6m) as achieving remission (Rem, $n=48$) versus no-remission (noR, $n=30$) defined by DAS28 <2.6. Of note, most but not all patients had data/samples at all the follow-up time points.

TNFi and continued to rise (67% increase). Treg showed no longer a clear increase either (55% increases).

Established disease patients achieving remission (on drug)

T-cell biomarkers were first compared between remission on either cs- or bDMARDs.

Higher naive cells were observed in csDMARDs remission (Fig. 4a, $n=418$, median +10%) compared with remission on biologic ($n=265$, +1%, $P=0.0003$). This was associated with a possible trend for increase in naive cell over the duration of remission (Fig. 5b, $\rho=0.200$, $P=0.0014$) from first achieving DAS28 < 2.6 to stable remission for several years. In contrast, no effect on naive T cells was seen over time in bDMARD remission (Fig. 5c). There was no difference in IRC frequencies between drugs; however, there was a reduction of IRC when comparing first time in remission to stable/long-term remission on csDMARDs (data not shown, $P=0.0031$), not seen directly as a correlation and not seen in remission on bDMARDs. Finally, Treg were higher in bDMARD remission ($P=0.00005$), in a relationship over time ($\rho=+0.219$, $P=0.042$) and this was not seen in csDMARDs.

Discussion

Our main objective was to understand the dynamics and relationships of circulating T-cell biomarkers compared with those currently in use in clinical practice in a large number of participants across the RA continuum. We described the dynamics of the three T-cell biomarkers over time and established their relationship with demographic/clinical parameters in ACPA+ at-risk Pr and over RA classification (with respect to research CCP-Second Gen ACPA tests and with ACPA-negative early IA). We demonstrated their independence from other biomarkers currently used to predict RA outcomes and validated their added value (discussed in detail in the [supplementary materials](#)). We also established that naive and Treg cells change in patients above the age of 60, possibly due to concomitant osteoarthritis (OA), allowing us to refine the population in which these biomarkers should be used.

At-risk phase

All three biomarkers were highly disturbed and had important predictive value for progression. Only age appeared to influence naive/Treg while not interfering with the prediction of progression. While not significantly decreasing over time as a group,

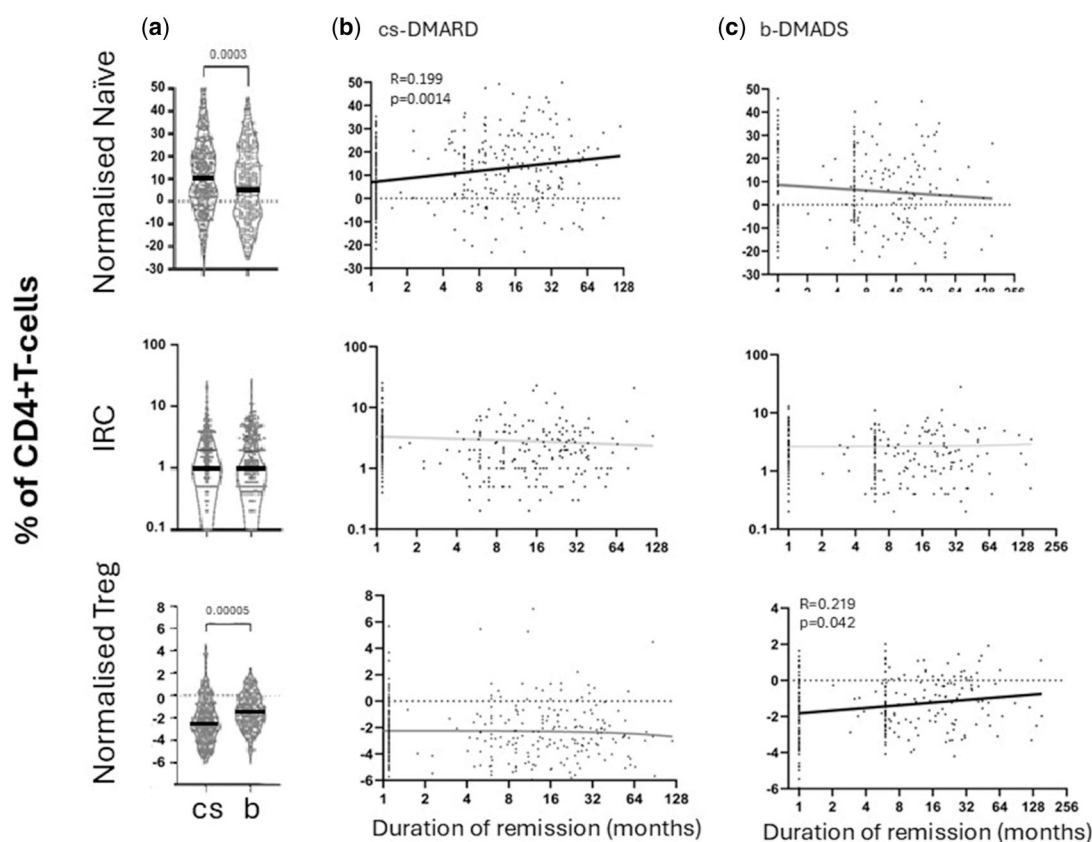


Figure 5 T-cell biomarker data for naive (normalized), IRC and Treg (normalized) in a cohort of RA patients ($n=683$) achieving clinical remission based on DAS28 < 2.6. **a)** T-cell biomarker data are displayed comparing patients achieving remission defined by DASD < 2.6 on classic-synthetic (cs) DMARDs ($n=418$) or on biologic (b) DMARDs ($n=265$). MWU P -values are displayed on the graph. **b)** Distribution of T-cell biomarker data are displayed in RA patients over time, prospectively calculated from first time in remission and treated only with cs-DMARDs ($n=418$). **c)** Distribution of T-cell biomarker data is displayed in RA patients over time, prospectively calculated from first time in remission and treated only with b-DMARDs ($n=265$). Lines indicate MWU test with P -values. Spearman correlations were calculated (best line of fit are displayed), with rho coefficient and P -value

naive cells in paired samples decreased in Pr (56% of pairs) while they increased in NP (83%). This suggested that dynamic changes continue over the at-risk phase in the RA continuum for naive T cells. Switching from low- to high-risk categories over time (based on 80% specificity cut-off as previously defined [14]) therefore justifies repeat analysis (annually at least). Treg decreased in Pr, but only in 42% of pairs (no trend in NP) with no relationship with time. IRC were previously associated with progression >12 months [21]. Here, a cross-sectional analysis confirmed that IRC increase closer to the onset of IA and overall, in 65% of paired samples, whilst not in NP. Monitoring IRC over time in patients with a high risk could provide further stratification toward intervention studies, where patient selection based on an imminent rather than overall risk could reduce the need for follow-up time.

Classification of RA in patients with early arthritis

We confirmed that T-cell biomarkers are independent predictors of fulfilling RA classification, particularly useful in ACPA patients, and are not correlated with other clinical parameters. However, age affects how these biomarkers can be used. While the term 'late-onset RA' (LO-RA) refers to patients who developed RA after the age of 60, other differences with RA (occurring in mid-life) exist. The female-to-male ratio tends to be less pronounced in LO-RA (in our cohort, 61% compared with 75% in the younger group). There is a lower proportion of RF+ in LO-RA individuals (43% in LO-RA vs 68%). Joint count records do not detail large/small joints while LO-RA usually involves more large joints. T-cell biomarkers may provide new insights into difference in pathophysiology between diseases, as naive T cells are clearly higher than expected in LO-RA, Treg are reduced and IRC are not particularly raised. The normalization procedure that we applied allows the effects of age to be accounted for (increase/reduction above expected normalized values), allowing for other effects to be addressed directly. Raised frequencies of naive CD4+ T cells and loss of Treg were previously observed in OA patients [26]. In older individuals, positivity for RF also increases with age [27, 28], while ACPA positivity is also observed in ~6% of OA patients [29, 30]. This suggests that OA rather than age-related immunological changes is a confounding factor for RA classification in ACPA+ patients >60 years old. In either case, the mechanism behind the increase in naive T cells in OA remains to be explained when thymic activity clearly decreases with age [6, 7].

Predicting response

Previous results for the prediction of MTX-induced remission (naive CD4+ T cells only) were replicated and the lack of value for TNFi response was clearly established. This provides a rationale to stratify early, drug-naive RA patients unlikely to respond to MTX for initial treatment with TNFi. This will require randomized control trials to establish a clear gain in preventing long-term effects of RA when treated more aggressively with TNFi.

Drug effect on the biomarkers

In early RA patients achieving MTX-induced remission, naive T cells initially continued to decrease, in contrast to the non-remission group, where changes did not have a particular direction. With time, MTX tended to reverse this tendency and naive cells increased. MTX treatment is not able to effectively control IRC, regardless of disease activity, and Treg appears to be a relatively stable biomarker.

In contrast, over the course of TNFi treatment, naive T cells increased irrespective of remission, suggesting a direct drug effect. The effect of TNF inhibition on the activity of the thymus was not directly investigated here, but there are reports that TNF-alpha is constitutively expressed in the thymus [31, 32]. The thymus, however, should also produce new Treg, but this effect was not seen. The pool of circulating Treg consists of both thymic Treg (t-Treg) and inducible Treg (i-Treg) generated in the periphery. Gut dysbiosis present in pre-clinical RA [33–37] may in part be responsible for altering the profile of immuno-regulatory metabolites notably responsible for the generation of i-Treg [38–41]. TNFi is believed to restore a 'healthier' gut microbiota [42], and it would be relevant to include a marker such as Helios (only expressed in t-Treg) to examine this issue further.

Importantly, the loss of TNFi-mediated effect on IRC, when used after failing to respond to MTX-alone, suggests that a major opportunity to control inflammation-related effects on T cells is missed when the use of TNFi is delayed. This should also be considered when rationalizing the use of T-cell biomarkers to decide between MTX alone or MTX+TNFi.

RA remission

Patients in remission on csDMARDs had higher naive cells compared with those on bDMARDs, possibly due to MTX only being able to produce remission in the less dysregulated individuals. Naive cells, however, increased over time in csDMARD remission, suggesting that long-term control of disease activity may allow recovery of naive cells. Thymic activity is likely responsible for this slow effect, especially in light of previously reported data on naive T-cell recovery in csDMARD remission [43, 44] and the stability of the remission stages [14]. The effect of TNFi on increasing naive T cells during the establishment of remission indicates further stabilization over time in remission.

IRC in remission are clearly lower compared with active disease; however, this also occurs slowly on csDMARDs (higher at time of first achievement of remission), while rapidly in bDMARD-treated patients. Once in stable remission, patients often desire to reduce their treatment although the risk of relapse is significant and requires careful management [45]. Relapse in remission was associated with higher IRC [14, 18, 43] and the strongest predictor of flare in patients tapering csDMARDs [18]. This suggests that tapering should not be offered until IRC are clearly decreased.

IRC are quickly reduced by bDMARDs; however, they showed no predictive value for flare when patients remain on treatment [20], while they were predictive of flare when tapering [20]. The

value of IRC when offering tapering is therefore independent of the type of treatment.

Treg were higher on bDMARDs, and this was seen from initiation of TNFi-induced remission and continued over the course of achieving stable remission. This is specific to bDMARDs and was previously associated with ability to sustain remission when tapering bDMARDs (best predictor) [20]. Monitoring Treg therefore also has important biomarker value for recommending tapering of bDMARDs.

Biological mechanisms

Patients ACPA+ at-risk who later developed RA showed CD4+ T subset disturbances (increasing the closest to disease onset for IRC), while there appear to be limited changes later, once disease has initiated, suggesting that such dysregulations are acquired during the pre-clinical stage and rather a stable event while associated with response to MTX. Whether disturbances of these subsets are merely biomarkers or whether they participate to the progression towards RA remains an open question. We reported dysfunctional features in naive CD4+ T cells in early, untreated RA, notably in relation to a decline in thymic activity, the differentiation of IRC (under a putative IL-6 drive), while response to antigenic challenges suggested maintenance of the naiveite in both IRC and naive cells [6], while the naive TCR repertoire is shaped by HLA alleles [46]. Progression to RA in at-risk individuals has been associated with systemic inflammation (IL-6/TNF, chemokines) and transcriptional change in naive T cells, suggesting activation while remaining naive, and potentially autoreactive naive CD4+ [47]. These data reinforced the concept of a non-immunological/non-antigenic challenge affecting naive CD4+ T cells. Inflammation is a well-known trigger of epigenetic changes, suggesting a mechanistic mean to enable the changes observed in naive T cells. We showed that epigenetic remodeling specific to naive CD4+ T cells occurs in early RA, centered on a network of modifications driven by IL-6/TNF signalling, in a sub-population of naive cells resembling IRCs [48]. Recently, the accumulation of epigenetic changes in naive CD4+ T cells in at-risk individuals over the course of progression to clinical onset was described [49]. Notably, CD62L was involved which is the main marker differentiating naive CD4+ T cells from IRC [48] and their increased resistance to apoptosis [43]. Our data, combined with the effectiveness of Abatacept in preventing RA development [50–51], strongly support the concept that T cells are the central orchestrator of the events leading to RA, while how pre-clinical inflammation develops remain to be elucidated. Microbiota dysbiosis observed preclinically in the mouth [52] or in the gut [53–54] (possibly increased gut permeability/systemic inflammation) offers possible clues to explore, while this may also be related to the loss of Treg via lowering nutrition-generated inducible Treg and/or increasing Th17 differentiation [55] as also seen epigenetically in early RA [48] and in at-risk individuals [49]. More work is needed to identify the origin of the defect in Treg, differentiating between loss of naturally occurring Treg, or inducible Treg or tissue-resident Treg.

Limitation

Although the models we built suggest the added value of including T-cell subsets at each step of the RA continuum, there are limitations in using regressions. This is usually associated with

handling of missing data, using multiple imputation, a recommended approach that reduces imprecision and bias compared with complete-case analysis or single-value imputation methods (even using multiple imputation) [56]. By drawing on the statistical properties of the dataset (distributions and inter-variable relationships), multiple imputations produce more robust estimates and reduce the likelihood of biased or misleading conclusions. Furthermore, we were careful to not use clinical variables with a significant amount of missing data (GH, HLA-DR-SE). For T-cell biomarkers, missing data affected Treg more than naive/IRC while the numbers of cases for each outcome were sufficient to allow for adequate imputations.

These biomarkers remain research tools at this point in time. Prospective studies are needed to demonstrate outcomes of (i) using TNFi in patients predicted to fail response to MTX; (ii) preventing imminent progression by stratifying at-risk individuals for intervention with anti-T-cell co-stimulation, for example; and (iii) offering tapering in patients on bDMARDs achieving a low-risk state. Such findings would support the value of these biomarkers for improved disease management and provide a rationale for economy in drug usage/tapering. Full external validation in an independent cohort remains the next critical step, while replication efforts in biomarker research are often challenged by heterogeneity in cohort selection criteria and/or outcome definitions worldwide.

Furthermore, to adopt this more widely, we developed a dry-tube (DT) approach to staining cells, which would facilitate transferring the technology to standard flow cytometry services. We performed a validation of these antibody pre-coated/bar-coded tubes in [supplementary material](#). We demonstrated the feasibility of the DT technology for two key CD4+ T-cell subsets (naive and IRC), benchmarked against conventional wet-tube (WT), showing equivalent distribution and no bias for naive T-cell enumeration (see [supplementary data](#)). IRC values showed bias at high frequencies, primarily in the clinically relevant range reflecting actual improved subset separation in the DT panel (e.g. superior resolution), which is advantageous for routine NHS laboratory workflows. The lack of signal in DT technology for Treg panel suggest incompatibility of the drying chemistry with intracellular FoxP3 staining. Acceptable detection of 'putative' Treg by surface markers only was nonetheless verified. These features collectively support DT suitability for daily clinical practice and underscore its applicability for routine NHS laboratory use.

Conclusion

This study highlights the potential of CD4+ T-cell subsets as biomarkers for predicting RA progression, developing prevention strategies, classifying early RA, predicting MTX treatment response and rationalizing the use of TNF inhibition early to avoid damage, and counselling patients on tapering bDMARDs. The findings also suggest that coexisting conditions like OA can influence T-cell subset profiles, and confirmed that different DMARDs have distinct effects on these immune cell populations, shaping the immunological state of patients with RA.

Supplementary material

[Supplementary material](#) is available at *Rheumatology* online.

Data availability

Fully anonymized data are available via the corresponding authors based on reasonable request and subject to agreement with the University of Leeds.

Funding

T-cell subset data were collected since 2008 in Dr Ponchel's group supported by an ARUK grant (P0595, 2005–10), a NIHR-RISK grant 'The immune-stethoscope for RA' RC-PC-407–10054 (2007–09), an EU funded FP7-integrated project Masterswitch (No. 223404, 2008–13), an EU/MI funded project BeTheCure (No 115142–2, 2011–17) and the NIHR-LMBRU grants since 2017.

Disclosure statement: No conflict of interest.

Acknowledgements

We are particularly grateful to the staff of the CCP clinic, early arthritis clinic and register, remission clinic for recruiting patients and to the staff of the NHS immunology services for performing the flow cytometry data acquisition. The research is supported by the National Institute for Health Research (NIHR) Leeds Biomedical Research Centre. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

References

- Aletaha D, Eberl G, Nell VP *et al.* Practical progress in realisation of early diagnosis and treatment of patients with suspected rheumatoid arthritis: results from two matched questionnaires within three years. *Ann Rheum Dis* 2002; 61:630–4.
- Gerlag DM, Raza K, van Baarsen LG *et al.* EULAR recommendations for terminology and research in individuals at risk of rheumatoid arthritis: report from the study group for risk factors for rheumatoid arthritis. *Ann Rheum Dis* 2012; 71:638–41.
- Emery P, Mankia K, Nam JL. Individuals at risk of rheumatoid arthritis - The evolving story. *Best Pract Res Clin Rheumatol* 2017;31:1–2.
- Nielen M, van Schaardenburg D, Reesink H *et al.* Specific autoantibodies precede the symptoms of rheumatoid arthritis: a study of serial measurements in blood donors. *Arthritis Rheum* 2004;50:380–6.
- Smolen J, 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.
- Ponchel F, Morgan A, Bingham S *et al.* Dysregulated lymphocyte proliferation and differentiation in patients with rheumatoid arthritis. *Blood* 2002;100:4550–6.
- Lawson CA, Brown AK, Bejarano V *et al.* Isaacs, early rheumatoid arthritis is associated with a deficit in the CD4+CD25high regulatory T-cell population in peripheral blood. *Rheumatology (Oxford)* 2006;45:1210–7.
- Hunt L, Hensor E, Nam J *et al.* T-cell subsets: an immunological biomarker to predict progression to clinical arthritis in ACPA-positive individuals. *Ann Rheum Dis* 2016;75:1884–9.
- Rakieh C, Nam J, Hunt L *et al.* Grainger, Predicting the development of clinical arthritis in anti-CCP positive individuals with non-specific musculoskeletal symptoms: a prospective observational cohort study. *Ann Rheum Dis* 2015; 74:1659–66.
- van de Stadt LA, Witte BI, Bos WH *et al.* A prediction rule for the development of arthritis in seropositive arthralgia patients. *Ann Rheum Dis* 2013;72:1920–6.
- van Steenbergen HW, Aletaha D, Beaart-van de Voorde LJ *et al.* EULAR definition of arthralgia suspicious for progression to rheumatoid arthritis. *Ann Rheum Dis* 2017;76:491–6. 2017
- Boer AC, Wouters F, Dakkak YJ *et al.* Improving the feasibility of MRI in clinically suspect arthralgia for prediction of rheumatoid arthritis by omitting scanning of the feet. *Rheumatology (Oxford)* 2020;59:1247–52.
- Duquenne L, Hensor EM, Wilson M *et al.* Predicting inflammatory arthritis in at-risk persons: development of scores for risk stratification. *Ann Intern Med* 2023;176:1027–36.
- 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.
- Ponchel F, Goëb V, Parmar R *et al.* An immunological biomarker to predict MTX response in early RA. *Ann Rheum Dis* 2014;73:2047–53.
- Ma X, Xu S. TNF inhibitor therapy for rheumatoid arthritis. *Biomed Rep* 2013;1:177–84.
- 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.
- Gul HL, Di Matteo A, Mankia K *et al.* Can biomarkers predict successful tapering of conventional disease-modifying therapy in rheumatoid arthritis patients in stable remission? *Clin Exp Rheumatol* 2023;41:126–36.
- 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.
- Gul H, Di Matteo A, Anioke I *et al.* Predicting flare in patients with rheumatoid arthritis in biologic induced remission, on tapering, and on stable therapy. *ACR Open Rheumatol* 2024; 6:294–303.
- Anioke I, Duquenne L, Parmar R *et al.* Lymphocyte subset phenotyping for the prediction of progression to inflammatory arthritis in anti-citrullinated-peptide antibody-positive at-risk individuals. *Rheumatology (Oxford)* 2024;63:1720–32.
- Emery P, Horton S, Dumitru RB *et al.* Pragmatic randomised controlled trial of very early etanercept and MTX versus MTX with delayed etanercept in RA: the VEDERA trial. *Ann Rheum Dis* 2020;79:464–71.
- Nam JL, Villeneuve E, Hensor EM *et al.* A randomised controlled trial of etanercept and methotrexate to induce remission in early inflammatory arthritis: the EMPIRE trial. *Ann Rheum Dis* 2014;73:1027–36.
- Ponchel F, Duquenne L, Xie X *et al.* Added value of multiple autoantibody testing for predicting progression to

- inflammatory arthritis in at-risk individuals. *RMD Open* 2022;8:e002512.
25. Takanashi S, Takeuchi T, Kaneko Y. Effects of aging on rheumatoid factor and anticyclic citrullinated peptide antibody positivity in patients with rheumatoid arthritis. *J Rheumatol* 2023;50:330–4.
 26. Ponchel F, Burska AN, Hensor EMA *et al*. Changes in peripheral blood immune cell composition in osteoarthritis. *Osteoarthritis Cartilage* 2015;23:1870–8.
 27. van Schaardenburg D, Lagaay AM, Otten HG *et al*. The relation between class-specific serum rheumatoid factors and age in the general population. *Br J Rheumatol* 1993;32:546–9.
 28. Dequeker J, Van Noyen R, Vandepitte J. Age-related rheumatoid factors. Incidence and characteristics. *Ann Rheum Dis* 1969;28:431–6.
 29. Du H, Masuko-Hongo K, Nakamura H *et al*. The prevalence of autoantibodies against cartilage intermediate layer protein, YKL-39, Osteopontin, and cyclic citrullinated peptide in patients with early-stage knee osteoarthritis: evidence of a variety of autoimmune processes. *Rheumatol Int* 2005;26:35–41.
 30. Xie X, van Delft MAM, Shuweihi F *et al*. Auto-antibodies to post-translationally modified proteins in osteoarthritis. *Osteoarthritis Cartilage* 2021;29:924–33.
 31. Giroir BP, Brown T, Beutler B. Constitutive synthesis of tumor necrosis factor in the thymus. *Proc Natl Acad Sci U S A* 1992;89:4864–8.
 32. Pérez AR, Berbert LR, Lepletier A *et al*. TNF- α is involved in the abnormal thymocyte migration during experimental *Trypanosoma cruzi* infection and favors the export of immature cells. *PLoS One* 2012;7:e34360.
 33. Rooney CM, Mankia K, Emery P. The role of the microbiome in driving RA-related autoimmunity. *Front Cell Dev Biol* 2020;8:538130.
 34. Rooney CM, Mankia K, Mitra S *et al*. Perturbations of the gut microbiome in anti-CCP positive individuals at risk of developing rheumatoid arthritis. *Rheumatology* 2021;60:3380–7.
 35. Wells PM, Adebayo AS, Bowyer RC *et al*. Associations between gut microbiota and genetic risk for rheumatoid arthritis in the absence of disease: a cross-sectional study. *Lancet Rheumatol* 2020;2:e418–e27.
 36. Alpizar-Rodriguez D, Lesker TR, Gronow A *et al*. *Prevotella copri* in individuals at risk for rheumatoid arthritis. *Ann Rheum Dis* 2019;78:590–3.
 37. Wang Y, Wei J, Zhang W *et al*. Gut dysbiosis in rheumatic diseases: a systematic review and meta-analysis of 92 observational studies. *EBioMedicine* 2022;80:104055.
 38. Xu T, Stewart KM, Wang X *et al*. Metabolic control of TH17 and induced Treg cell balance by an epigenetic mechanism. *Nature* 2017;548:228–33.
 39. Kempkes RWM, Joosten I, Koenen H *et al*. Metabolic pathways involved in regulatory T-cell functionality. *Front Immunol* 2019;10:2839.
 40. Romero-Figueroa MDS, Ramírez-Durán N, Montiel-Jarquín AJ *et al*. Gut-joint axis: gut dysbiosis can contribute to the onset of rheumatoid arthritis via multiple pathways. *Front Cell Infect Microbiol* 2023;13:1092118.
 41. Lin L, Zhang K, Xiong Q *et al*. Gut microbiota in pre-clinical rheumatoid arthritis: from pathogenesis to preventing progression. *J Autoimmun* 2023;141:103001.
 42. Picchianti-Diamanti A, Panebianco C, Salemi S *et al*. Analysis of gut microbiota in rheumatoid arthritis patients: disease-related dysbiosis and modifications induced by Etanercept. *Int J Mol Sci* 2018;19:2938.
 43. Burgoyne CH, Field SL, Brown AK *et al*. Abnormal T-cell differentiation persists in patients with rheumatoid arthritis in clinical remission and predicts relapse. *Ann Rheum Dis* 2008;67:750–7.
 44. Churchman SM, El-Jawhari JJ, Burska AN *et al*. Modulation of peripheral T-cell function by interleukin-7 in rheumatoid arthritis. *Arthritis Res Ther* 2014;16:511.
 45. Mangoni AA, Al Okaily F, Almoallim H *et al*. Relapse rates after elective discontinuation of anti-TNF therapy in rheumatoid arthritis: a meta-analysis and review of literature. *BMC Rheumatol* 2019;3:10.
 46. Nagafuchi Y, Ota M, Hatano H *et al*. Control of naive and effector CD4 T cell receptor repertoires by rheumatoid arthritis-risk HLA alleles. *J Autoimmun* 2022;133:102907.
 47. He Z, Glass MC, Venkatesan P *et al*. Progression to rheumatoid arthritis in at-risk individuals is defined by systemic inflammation and by T and B cell dysregulation. *Sci Transl Med* 2025;17:eadt7214.
 48. Pitaksalee R, Burska AN, Ajaib S *et al*. Differential CpG DNA methylation in peripheral naïve CD4+ T-cells in early rheumatoid arthritis patients. *Clin Epigenetics* 2020;12:54.
 49. Inamo J, Bylinska A, Smith M *et al*. Longitudinal peripheral blood multi-omic profiling in seropositive individuals identifies immune endotypes and predictive models for future rheumatoid arthritis conversion. *medRxiv* 2026;17. <https://doi.org/10.64898/2026.02.12.26346058>
 50. Cope AP, Jasenecova M, Vasconcelos JC *et al*; APIPPRA study investigators. Abatacept in individuals at high risk of rheumatoid arthritis (APIPPRA): a randomised, double-blind, multicentre, parallel, placebo-controlled, phase 2b clinical trial. *Lancet* 2024;403:838–49.
 51. Rech J, Tascilar K, Hagen M *et al*. Abatacept inhibits inflammation and onset of rheumatoid arthritis in individuals at high risk (ARIAA): a randomised, international, multicentre, double-blind, placebo-controlled trial. *Lancet* 2024;403:850–9.
 52. Hashimoto M, Yamazaki T, Hamaguchi M *et al*. Periodontitis and *Porphyromonas gingivalis* in preclinical stage of arthritis patients. *PloS one* 2015;10:e0122121.
 53. Gilbert BTP, Tadeo RYT, Lamacchia C *et al*. Gut microbiome and intestinal inflammation in preclinical stages of rheumatoid arthritis. *RMD Open* 2024;10:e003589.
 54. Rooney C, Jeffery I, Mankia K *et al*. Dynamics of the gut microbiome in individuals at risk of rheumatoid arthritis: a cross-sectional and longitudinal observational study. *Ann Rheum Dis* 2024;84:562–74.
 55. Sprouse ML, Bates NA, Felix KM *et al*. Impact of gut microbiota on gut-distal autoimmunity: a focus on T cells. *Immunology* 2019;156:305–18.
 56. Woods AD, Gerasimova D, Van Dusen B *et al*. Best practices for addressing missing data through multiple imputation. *Infant Child Dev* 2024;33:e2407