

## Original article

# Undifferentiated arthritis: a changing population who did not benefit from enhanced disease-modifying anti-rheumatic drug strategies – results from a 25 year longitudinal inception cohort

Marloes Verstappen<sup>1</sup>, Xanthe M. E. Matthijssen<sup>1</sup> and Annette H. M. van der Helm-van Mil<sup>1,2</sup>

## Abstract

**Objectives.** International guidelines stress timely DMARD initiation in early arthritis as well as when classification criteria are not yet fulfilled. Consequently, undifferentiated arthritis (UA) patients may be increasingly treated with DMARDs. Since UA is a diagnosis of exclusion, the introduction of the 2010 classification criteria presumably decreased the UA population, as former UA patients became regarded as RA. Consequently, the contemporary definition of UA has changed into: no clinical diagnosis and not fulfilling the 1987 nor 2010 RA-criteria. Importantly, placebo-controlled trials on DMARD efficacy in contemporary UA are absent. We aimed to study whether enhanced treatment strategies across the last 25 years improved outcomes in contemporary UA, whereby inclusion period was used as instrumental variable for DMARD treatment.

**Methods.** UA was defined, retrospectively, as clinical arthritis (joint swelling at physical examination) neither fulfilling the 1987 nor 2010 RA-criteria or any other clinical diagnosis. In total, 1132 UA patients consecutively included in the Leiden Early Arthritis Clinic between 1993 and 2019 were divided into five inclusion periods: 1993–1997, 1998–2005, 2006–2010, 2011–2014 and 2015–2019. The frequency of DMARD initiation was compared across the inclusion periods, as were the following outcomes: 28-joint DAS with CRP (DAS28-CRP) and the HAQ Disability Index (HAQ-DI) during follow-up, prevalence of DMARD-free-status within 10 years (DFS; spontaneous remission or sustained remission after DMARD stop) and progression to RA (according 1987/2010 criteria).

**Results.** The contemporary UA population is mainly autoantibody negative, with a median swollen joint count of 2, tender joint count of 3 and HAQ score of 0.6. These characteristics were similar across the inclusion periods. DMARD treatment increased from 17% (1993–1997) to 52% (2015–2019) and methotrexate became more common. The DAS28-CRP during follow-up improved from 2011 onwards (–0.18 to –0.25 DAS units;  $P < 0.05$ ). Disability scores during follow-up did not significantly improve. DFS prevalence also remained similar: 58%, 57% and 61% for 1993–1997, 1998–2005 and 2006–2010, respectively ( $P = 0.77$ ). Likewise, the percentages of RA development did not decrease (14%, 21%, 26%, 18% and 27%, respectively).

**Conclusion.** Although intensified DMARD treatment slightly improved disease activity scores, physical functioning and long-term outcomes did not improve. This suggests overtreatment in the contemporary UA population and underlines the importance of developing stratification methods suitable for this patient-population.

**Key words:** undifferentiated arthritis, early arthritis, DMARDs, outcome, guidelines, remission, physical functioning, RA

<sup>1</sup>Department of Rheumatology, Leiden University Medical Center, Leiden and <sup>2</sup>Department of Rheumatology, Erasmus Medical Center, Rotterdam, The Netherlands

Submitted 24 August 2021; accepted 19 November 2021

Correspondence to: Marloes Verstappen, Department of Rheumatology, Leiden University Medical Center, P.O. Box 9600, 2300 RC Leiden, The Netherlands. E-mail: m.verstappen@lumc.nl

**Rheumatology key messages**

- The contemporary undifferentiated arthritis population is largely autoantibody negative and has a median of 2 swollen joints.
- Increased DMARD use in this population did not result in improved outcomes (physical functioning, DMARD-free remission, RA development).
- Adequate risk stratification methodology for this contemporary UA population needs to be developed to prevent overtreatment.

**Introduction**

Management of undifferentiated arthritis (UA) patients in clinical practice is challenging since the outcome of this group is highly variable, ranging from spontaneous remission to persistent and destructive RA [1, 2]. This prognostic uncertainty can influence decisions to initiate DMARDs. The readiness of rheumatologists to start DMARDs has also changed over time, especially since early DMARD initiation in early arthritis is advocated by international guidelines [3]. Importantly, the characterization of UA has shifted in the previous decade due to the introduction of the 2010 classification criteria for RA [4]. With this, patients who were formerly classified as UA now fulfil the 2010 criteria for RA, especially ACPA-positive patients. Also, some of the ACPA-negative patients who fulfilled the 1987 criteria for RA do not concomitantly fulfil the 2010 criteria, as  $\geq 10$  involved joints are required in the absence of autoantibodies. Considering both sets of criteria, UA in its strict form is currently defined as not fulfilling the 2010 nor the 1987 criteria for RA or any other distinct diagnosis.

There are no clinical trials or studies on the efficacy of DMARDs in this contemporary UA population since all previous trials in UA were completed before 2010 and therefore only included patients with 1987 UA (not fulfilling the 1987 criteria or any other diagnosis) [5–10]. Importantly, these trials did not demonstrate an evident beneficial effect of DMARD treatment on clinical remission, patient-reported outcomes, radiographic progression or progression to RA, and a part of the patients in these trials will today be regarded as RA. Moreover, the EULAR recommendation to initiate DMARD treatment in early arthritis patients with autoantibodies, erosions or high disease activity, even in the absence of fulfilling classification criteria, is based on the definition of UA (UA-1987) that was used before the introduction of the 2010 criteria [11]. Thus scientific evidence cannot substantiate clinical treatment decisions in the contemporary UA population. Also, the increasing use of DMARDs in these UA patients may hamper the effectuation of future placebo-controlled clinical trials in UA.

The primary aim of this study was to explore the efficacy of DMARD treatment in the contemporary UA population using 25 years of observational data in which UA was uniformly defined and the inclusion period was used as an instrumental variable for changes in

treatment strategies. We hypothesized that if DMARD treatment has a disease-modifying effect in UA and DMARD treatment has increased over time, this will have resulted in lower disease activity, improved physical functioning, increased frequency of prolonged DMARD-free status (DFS) and less progression to RA. We ensured a homogeneous population of contemporary UA by selecting, in retrospect, all consecutive patients who did not full the 1987 or 2010 classification criteria for RA or any other distinct clinical diagnosis. In this way, changes in the rheumatologist's clinical perspective on UA identity did not affect the patient population. Our second aim was to substantiate whether the clinical concept of UA has indeed changed over time. Therefore baseline characteristics of clinically diagnosed UA patients were also compared across the last 25 years.

**Patients and methods****Study population**

The Leiden Early Arthritis Clinic (EAC) cohort is a population-based inception cohort including all consecutive patients newly presenting with recent-onset arthritis from 1993 onwards, and has been previously described [12]. In short, all patients presenting with recent-onset arthritis (symptom duration  $< 2$  years), defined as the presence of clinical synovitis at physical examination by a rheumatologist, are included in the Leiden EAC. Since our primary aim was to study whether increased DMARD treatment resulted in improved (long-term) outcomes in UA, we retrospectively selected all consecutive patients included between February 1993 and October 2019 without a distinct clinical diagnosis at baseline and who did not fulfil the 1987 or 2010 classification criteria for RA. This retrospective selection, using the contemporary UA definition, ensured homogeneity of the patient population across the different inclusion periods. UA patients who concomitantly participated in a clinical trial (and were thus not routinely treated) were excluded.

Subsequently we studied whether the rheumatologists' clinical perspective of the identity of UA, irrespective of classification criteria, changed in the last decades. The term UA was first mentioned in the scientific literature around 1987 [13–15], although in a slightly different

concept, whereas the continuum of UA with RA appeared in the literature around 1998 [16–20]. Therefore we studied the baseline characteristics of all consecutive patients included in the Leiden EAC between January 1998 and October 2019 with a clinical diagnosis of UA defined by their treating rheumatologist in the first weeks after presentation (after laboratory measures were done and radiographs of the hands and feet were made). This population was selected regardless of classification criteria, which is different from the main study population described above.

### Treatment

Initiation and choice of DMARD treatment was determined by the patients' treating rheumatologists, who practiced according to treatment guidelines and clinical standards of that specific time period. To study changes in DMARD treatment over time, patients were categorized based on the inclusion period (1993–1997, 1998–2005, 2006–2010, 2011–2014 and 2015–2019). These cut-offs were partly based on developments that may have influenced DMARD treatment in UA [publication of early arthritis guidelines recommending (early) DMARD therapy in early arthritis patients not fulfilling classification criteria in 2006 and 2016; the 2010 classification criteria for RA] and partly arbitrarily chosen to create approximately similar patient numbers per group [3, 4, 11]. Per inclusion period, the frequency of DMARD treatment [and the use of systemic glucocorticosteroids (GCs)] and the type of DMARD within the first year after inclusion were studied.

### Follow-up

Research visits took place at baseline, after 4 months and annually afterwards. During these visits, joint counts were performed, laboratory measurements were done and questionnaires were filled out [among others, the HAQ, visual analogue scale (VAS) pain and VAS fatigue]. At baseline, RF (in-house ELISA, as described previously [21]), and from 2006 onwards ACPA [2006–2008; anti-CCP2 (Euro-Diagnostica), positive if  $\geq 25$  U/ml; 2009–2019; EliA (Phadia), positive if  $\geq 10$  U/ml] were measured. In patients included before 2006, ACPA status was assessed retrospectively in stored baseline serum samples using the Euro-Diagnostica assay. Protocolized visits were performed as long as patients were treated at the outpatient clinic. Follow-up ended in case of release from care due to prolonged DMARD-free remission, death, migration to another area or withdrawal of informed consent while remaining treated.

### Outcome

Four outcomes were compared between the different inclusion periods. First, disease activity scores [28-joint DAS with CRP (DAS28-CRP)] during follow-up, reflecting the direct effect of treatment, were compared across the different inclusion periods. Second, functional disability, measured using the HAQ Disability Index (HAQ-

DI), was compared [22]. Third, prolonged DFS was studied, defined as either spontaneous remission or sustained remission after discontinuation of DMARDs (including GCs). In this case, sustained remission was defined as the sustained absence of synovitis for a minimum of 1 year after discontinuation of all DMARD treatment and for the entire follow-up thereafter. Deliberately, both spontaneous remission and sustained DMARD-free remission were collectively considered as DFS for reasons of comparability, since increased DMARD treatment over time may have led to less spontaneous remission and presumably to more DMARD-free remission. Medical files were studied on the occurrence of DFS until October 2020. Furthermore, the promptness of DMARD tapering and cessation became more common over the last decade compared with earlier time periods, affecting the disease duration after which DFS was obtained. To allow fair comparisons, the development of DFS was primarily assessed within a time span of 10 years of follow-up. In addition, in a sub-analysis, DFS was also evaluated after 5 years of follow-up. This allowed us to include a more recent inclusion period (2011–2014), but may also underestimate the DFS prevalence in the older inclusion periods. Finally, progression to RA was studied, in which RA was defined as a clinical diagnosis of RA and the fulfilment of the 1987 and/or 2010 criteria within the first year after inclusion [23, 24].

### Sensitivity analyses

Previous research has demonstrated significant treatment differences between ACPA-positive and ACPA-negative RA [25], suggesting differences in underlying pathophysiology between these subsets of disease. Since treatment effects may also differ between ACPA-positive and ACPA-negative UA, analyses on disease activity, physical functioning and prolonged DMARD-free status were stratified for ACPA status.

### Statistical analyses

With inclusion period as the instrumental variable for treatment strategy, outcomes were compared between each inclusion period and the reference period (1993–1997). DAS28-CRP and HAQ scores over time were analysed using linear mixed models. Since a non-linear relation was observed between DAS28-CRP/HAQ scores and time, two separate models were used to analyse the first year and the subsequent follow-up thereafter. The model for the first year included a random intercept and an identity covariance matrix. The model for the remainder of the follow-up included a random intercept, a random slope and an independent covariance matrix. Follow-up duration was truncated to 5 years of follow-up. Estimated marginal means were calculated. Missing data on DAS28-CRP (data missing for 0% of patients at baseline and 20% of patients at follow-up visits) and HAQ-DI (data missing for 39% at baseline and 33% at annual follow-up visits) were

imputed using multivariate multiple imputation with predictive mean matching (100 cycles, 30 datasets). Results were pooled using Rubin's rules [26].

Percentages of patients achieving DFS within 10 and 5 years of follow-up and percentages of patients progressing to RA after 1 year of follow-up were visualized per inclusion period. A non-parametric test for trend (Cuzick's test) was used to compare the percentages of DFS and RA development [27]. Stata version 16 (StataCorp, College Station, TX, USA) was used for all analyses.

## Results

### The clinical concept of UA over time

To study whether the rheumatologists' perspective of the identity of UA has changed over time, baseline characteristics of all early arthritis patients with a clinical diagnosis of UA at baseline were compared ( $n=1626$ ). Thus this patient population was not based on classification criteria. As depicted in Fig. 1, over time, UA patients became more frequently autoantibody negative, had fewer affected joints and had less often elevated acute phase reactants (Supplementary Table S1, available at *Rheumatology* online). The observation that the clinical identity of UA from a rheumatologists' perspective changed over time underlines the need to select a homogeneous UA population for this study. Thus, in the subsequent analyses, we studied a criteria-based UA population, i.e. not fulfilling the 1987 and/or 2010 criteria for RA or any other distinct clinical diagnosis, to create a constant patient population meeting the current UA definition.

### UA population according to its contemporary definition

In order to ensure a homogeneous population, meeting the contemporary definition of UA, patients who did not fulfil the 1987 or 2010 criteria for RA (or had any other diagnosis) were retrospectively selected from all consecutively included early arthritis patients between February 1993 and October 2019 ( $n=1259$ ). Patients who were treated in a clinical trial ( $n=127$ ), and therefore not routinely treated, were excluded (Supplementary Table S2, available at *Rheumatology* online). Baseline characteristics of the 1132 UA patients are presented in Table 1 and Fig. 1 and were uniform across the different inclusion periods. Overall, UA patients were mostly autoantibody negative and presented with relatively mild disease at baseline (median SJC of 2, median TJC of 3, median HAQ of 0.6), which is in line with the current clinical characterization of UA.

### DMARD treatment over time

DMARD initiation within the first year of follow-up increased from 17% in 1993–1997 to 52% in 2015–

2019, as methotrexate became more common in the last decade (Fig. 2).

### DAS28-CRP during follow-up

Baseline mean DAS28-CRP in the reference period (1993–1997) was 3.35 (95% CI 2.54, 2.84) and did not significantly differ in later inclusion periods (Fig. 3; Supplementary Table S3, available at *Rheumatology* online). In the reference period, disease activity scores decreased to 2.63 (95% CI 2.54, 2.84) after the first year of follow-up. A similar decrease in DAS28-CRP in the first year was seen in the other inclusion periods. However, compared with the reference period, DASs after the first year of follow-up became significantly lower from 2011 onwards; on average, 0.18 DAS units lower (95% CI  $-0.36$ ,  $-0.00$ ;  $P=0.047$ ) in 2011–2014 and 0.25 DAS units lower (95% CI  $-0.44$ ,  $-0.07$ ;  $P=0.07$ ) in 2015–2019 compared with 1993–1997. Stratification for the different DAS components demonstrated that predominantly SJC and CRP improved from 2011 onwards, in contrast to the TJC and VAS components (Supplementary Table S4, available at *Rheumatology* online).

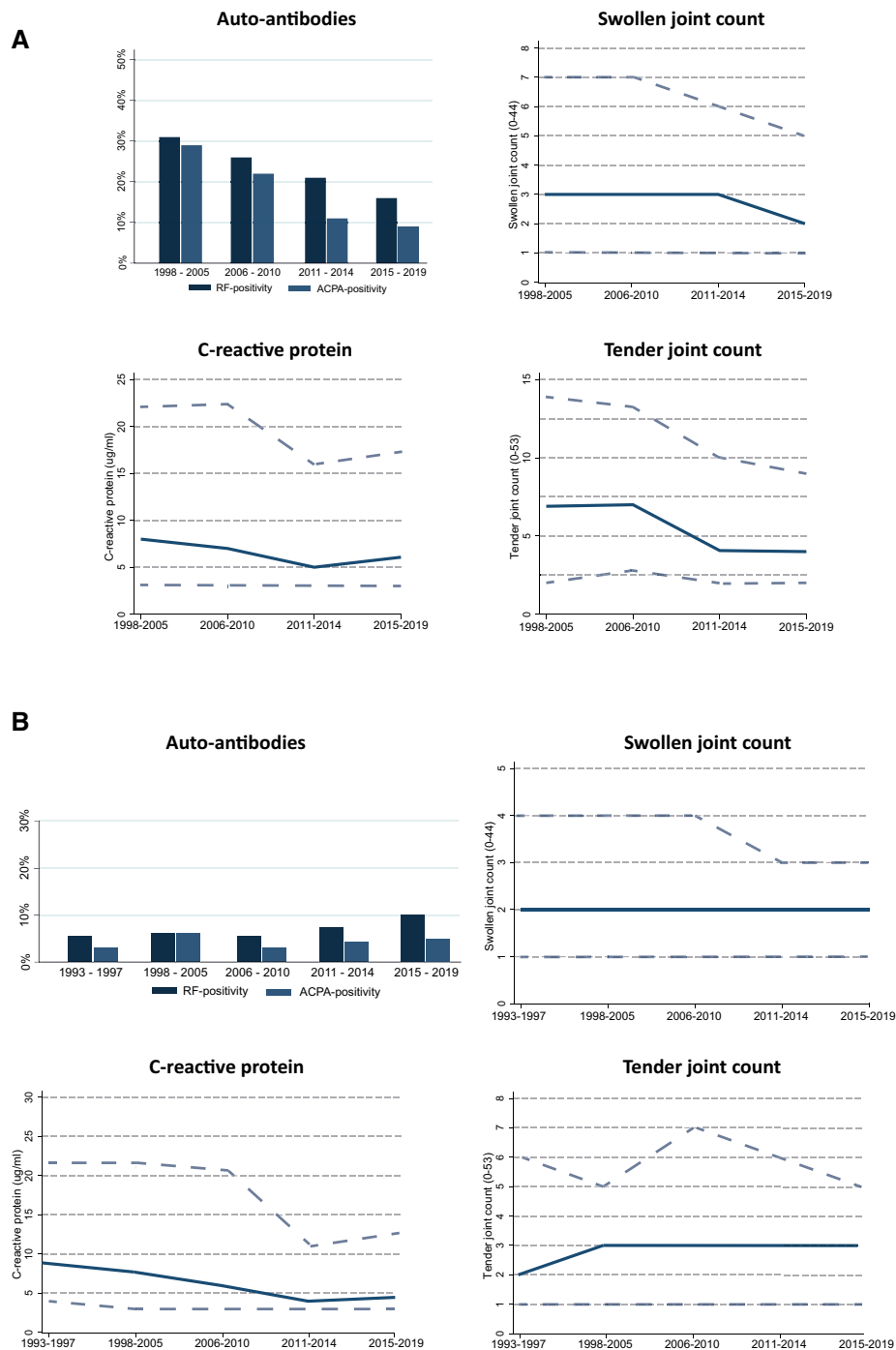
### Functional disability during follow-up

The mean HAQ score at baseline was 0.62 (95% CI 0.53, 0.73) in 1993–1997 and did not significantly differ across the other inclusion periods (Fig. 4, Supplementary Table S5, available at *Rheumatology* online). In the reference period, HAQ scores decreased to 0.44 (95% CI 0.37, 0.52) after the first year of follow-up. A similar decrease in HAQ scores in the first year of follow-up was present in the other inclusion periods. Likewise, HAQ scores after the first year of follow-up did not significantly improve compared with 1993–1997. As patient-reported outcomes are often correlated, pain and fatigue (measured on a VAS) were also studied. Again, no improvement was observed compared with 1993–1997 (Supplementary Fig. S1, available at *Rheumatology* online).

### Prolonged DFS during follow-up

Prolonged DFS was defined as either spontaneous remission or sustained remission after DMARD stop within 10 years of follow-up. This was achieved by 58% of the UA patients included between 1993–1997; in the subsequent inclusion periods this was 57% and 61% (1998–2005 and 2006–2010, respectively;  $P=0.77$ ). When separating DFS into its components, spontaneous remission became less frequent over time (52%, 43% and 40% in 1993–1997, 1998–2005 and 2006–2010, respectively). The median time until spontaneous remission was recorded was 0.4 years [interquartile range (IQR) 0.2–0.9]. DMARD-free remission became more frequent (6%, 14% and 21% in the respective time periods; Fig. 5A), which is in line with the observed increase in frequency of DMARD use in UA. DMARD-free remission was achieved after a median of 3.5 years (IQR 2.1–5.7), after

**Fig. 1** Baseline characteristics over time of patients clinically identified as UA by their treating rheumatologist (**A**) and (**B**) the contemporary UA population



**(A)** Baseline characteristics of patients with a clinical diagnosis of UA, according to their treating rheumatologist. Over time, UA patients became more frequently autoantibody negative, had fewer affected joints and less often had elevated acute phase reactants. The solid line indicates the median of the measured baseline characteristics, with the dashed lines indicating the upper and lower IQRs. **(B)** Baseline characteristics of contemporary UA patients who retrospectively did not fulfil the 1987 or 2010 criteria for RA or any other distinct diagnosis. The characteristics of this primary study population were stable between the inclusion periods, providing homogeneous groups that could be compared over time. The solid line indicates the median of the measured baseline characteristics, with the dashed lines indicating the upper and lower IQRs.



**TABLE 1** Baseline characteristics of a criteria-based UA population; patients not fulfilling the 1987 or 2010 RA criteria or a distinct diagnosis

Characteristics	Total (N = 1132)	1993–1997 (n = 222)	1998–2005 (n = 216)	2006–2010 (n = 214)	2011–2014 (n = 247)	2015–2019 (n = 233)
Age, years, mean (s.d.)	52 (17)	49 (17)	49 (18)	52 (17)	53 (16)	56 (16)
Female, %	59	50	59	62	66	56
Symptom duration, weeks, median (IQR)	11 (4–26)	10 (3–22)	15 (6–33)	16 (6–31)	8 (3–24)	9 (5–20)
ACPA positive, %	5	3	7	3	4	5
RF positive, %	7	5	7	6	8	10
SJC (0–44), median (IQR)	2 (1–4)	2 (1–4)	2 (1–4)	2 (1–4)	2 (1–3)	2 (1–3)
TJC (0–53), median (IQR)	3 (1–6)	2 (1–6)	3 (1–5)	3 (1–7)	3 (1–6)	3 (1–5)
CRP ( $\mu\text{g/mL}$ ), median (IQR)	6 (3–17)	9 (4–22)	8 (3–22)	6 (3–21)	4 (3–11)	5 (3–13)
ESR (mm/h), median (IQR)	14 (6–33)	18 (9–38)	16 (7–34)	14 (6–33)	11 (6–25)	13 (6–31)
VAS (0–100), median (IQR)	30 (14–50)	30 (14–50)	32 (14–51)	26 (11–50)	30 (18–53)	30 (20–50)
DAS28-CRP, median (IQR)	3.29 (2.67–3.93)	3.32 (2.62–3.96)	3.37 (2.69–4.00)	3.34 (2.67–4.08)	3.15 (2.68–3.80)	3.34 (2.67–3.85)
HAQ-DI, median (IQR)	0.50 (0.13–0.91)	0.47 (0.10–0.90)	0.50 (0.14–0.93)	0.51 (0.13–0.93)	0.58 (0.15–0.88)	0.62 (0.13–0.99)

Baseline characteristics of the criteria-based UA population were comparable across the different inclusion periods. Missing values were imputed and used for the baseline table. Values represent means of the imputed medians calculated using Rubin's rules.

which patients were followed for a subsequent median 5.8 years (IQR 2.6–10.6) during which no flare occurred, ensuring the sustainability of this outcome.

A subanalysis on DFS development after 5 years of follow-up also demonstrated no statistically significant increase over time ( $P=0.10$ ; [Supplementary Fig. S2](#), available at *Rheumatology* online). Although this analysis allowed us to include data from a more recent inclusion period, DFS prevalence in 1993–1997, 1998–2005 and 2006–2010 were underestimated compared with the main analysis with 10 years of follow-up.

### Progression to RA

Progression to RA after 1 year of follow-up, according to the 1987 and/or 2010 criteria, varied per inclusion period but did not decrease over time ([Fig. 5B](#)). Between 1993 and 1997, ~14% of UA patients developed RA, whereas in 2015–2019 this was 27%. The test for trend was statistically significant ( $P < 0.01$ ), although an increasing trend in progression to RA was observed and not a decrease.

### Sensitivity analyses

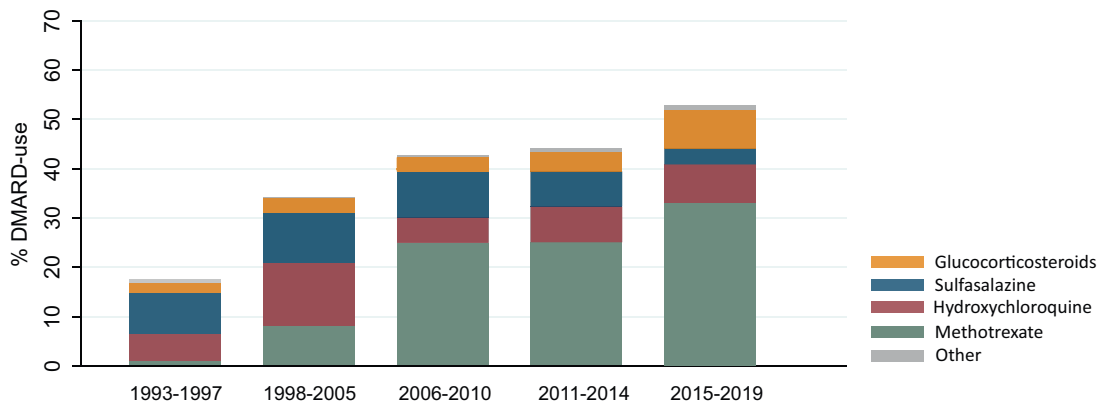
Although the number of ACPA-positive UA patients was relatively low ( $n=45$ ), prohibiting achievement of statistical significance, a tendency towards improved outcomes was seen ([Supplementary Fig. S3](#), available at *Rheumatology* online). The improvements in DASs after 1 year of follow-up seemed to be accompanied by an improvement in physical functioning and an increase in achievement of prolonged DMARD-free remission within the 10 years of follow-up (0% in 1993–1997, 23% in 1998–2005, up to 60% in 2006–2010). Results within ACPA-negative RA were similar for the total UA population ([Supplementary Fig. S3](#), available at *Rheumatology* online).

## Discussion

The characterization of UA has changed in recent decades. Because UA is defined as not having RA or any other distinct diagnosis, this 'remainder' group changed with the introduction of the 2010 RA criteria. Indeed, we observed that the patients who were diagnosed with UA by their treating rheumatologists changed over time and that UA in its contemporary definition is rather mild, mostly autoantibody negative and has on average two swollen joints and moderate functional impairment (HAQ 0.6).

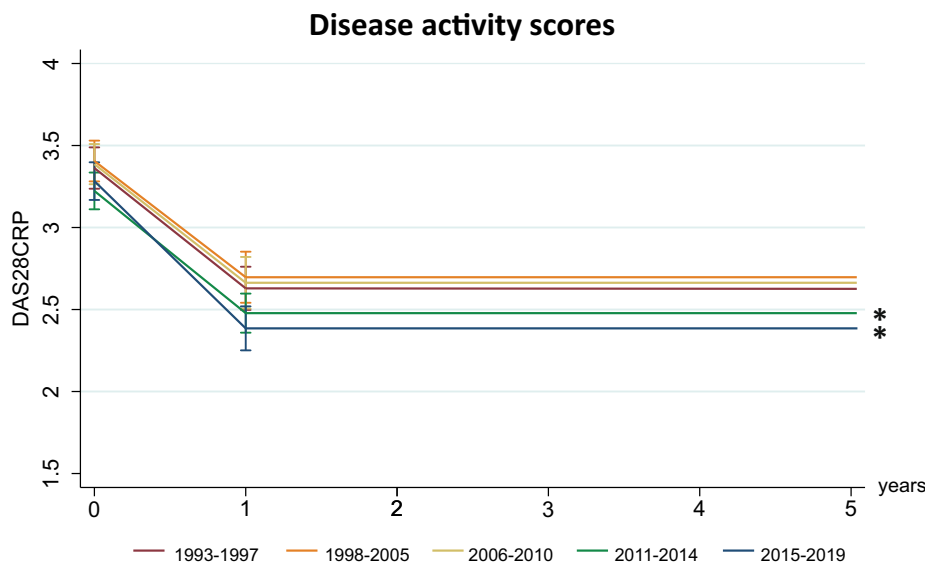
Unfortunately, all trials on DMARD efficacy in UA were done with the former definition of UA (not fulfilling the 1987 criteria for RA). UA patients in these trials were mostly autoantibody positive, regularly (up to 50%) presented with polyarthritis and were more severely impaired (HAQ scores 0.8–1.2) [5, 7, 9]. Part of this UA population will today be regarded as RA. Yet there are no intervention trials in which the UA

**Fig. 2** Frequency of DMARD treatment in contemporary UA (not fulfilling 1987 or 2010 RA criteria or any other diagnosis) during the last 25 years



Initiation of DMARD treatment in UA significantly increased over time from 17% in 1993–1997 to 52% in 2015–2019 ( $P < 0.01$ ). In the last decade, the use of methotrexate became more common. ‘Other’ included leflunomide, azathioprine and rituximab.

**Fig. 3** Disease activity score slightly improved from 2011 onward compared with 1993–1997

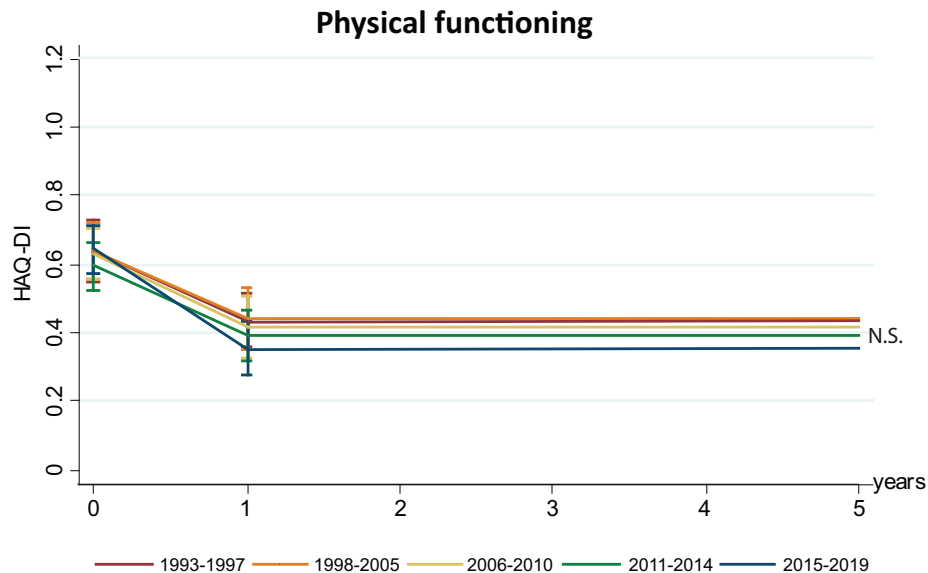


DASs during follow-up per inclusion period compared with the earliest inclusion period (1993–1997). DASs after 1 year of follow-up improved from 2011 onwards. DASs were modelled based on the estimated marginal means resulting from the age- and gender-corrected linear mixed models.  $*P < 0.05$ .

population is defined according to the current definition and presents with milder disease.

Despite the lack of placebo-controlled evidence for DMARD efficacy in contemporary UA, international guidelines (first published in 2006) stress timely initiation of DMARD treatment in early arthritis, also when classification criteria are not yet fulfilled [3, 11]. The increasing tendency to treat UA hampers the effectuation of placebo-controlled randomized clinical trials in contemporary UA in the near future. In order to explore the efficacy of DMARD treatment, we used 25 years of observational data in which UA was retrospectively

defined by its contemporary definition. By this retrospective selection we identified a homogeneous group of UA patients in which we observed that DMARD treatment indeed increased over time to  $>50\%$ . While using the inclusion period as an instrumental variable for DMARD treatment, we studied whether enhanced DMARD strategies resulted in improved disease outcomes. Despite the fact that DASs slightly improved from 2011 onwards (approximately  $-0.20$  DAS units), functional disability, prevalence of prolonged DFS and progression to RA did not concomitantly improve compared with 1993–1997. Notably, the observed

**Fig. 4** Physical functioning during follow-up did not improve with enhanced DMARD strategies in UA

Functional disability during follow-up per inclusion period compared with the earliest inclusion period (1993–1997). HAQ-DI scores during follow-up did not improve compared with 1993–1997. HAQ-DI scores were modelled based on the estimated marginal means resulting from the age- and gender-corrected linear mixed models. \*N.S.: not significant ( $P > 0.05$ ).

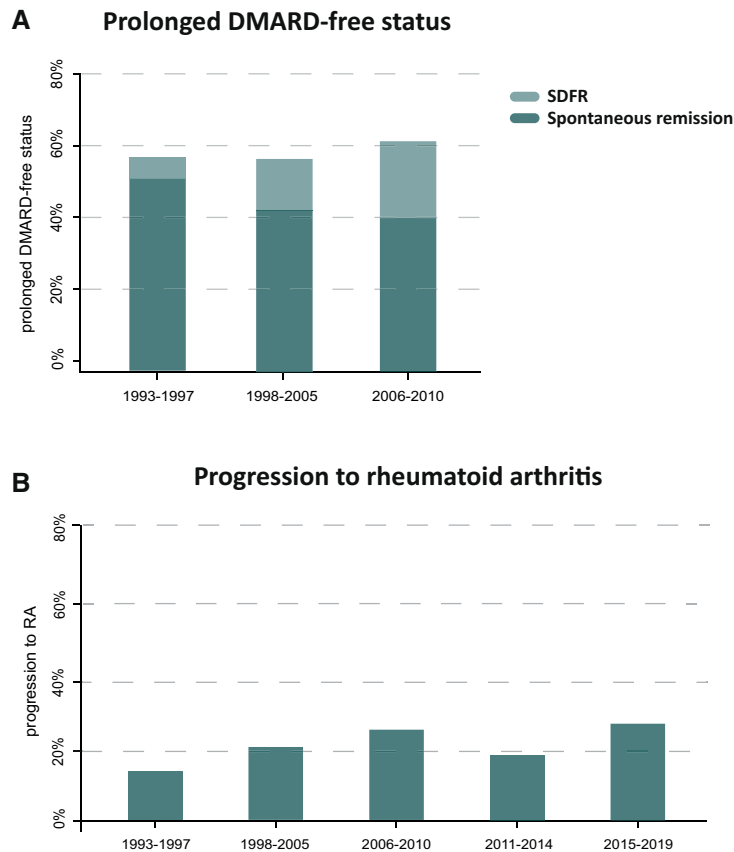
improvement in DAS28-CRP from 2011 onwards does not exceed the minimal clinically important difference of 1.0 [28].

Although we cannot exclude that increased DMARD treatment improved outcomes in the ACPA-positive subset (Supplementary Fig. S3, available at *Rheumatology* online), similar to RA [25], this subgroup is rather small (5% of contemporary UA). The vast majority are ACPA negative and the findings for this group contrast with the situation in RA, where early and more frequent DMARD initiation improved long-term disease outcomes [29], effects that are presumably mediated by a low disease activity and thus treat-to-target treatment strategies [25]. In UA, there is no proof that treat-to-target is valuable, and despite the decrease in DAS during follow-up in the more recent inclusion periods, this was not paralleled by improved long-term outcomes. Our results may therefore indicate that when all UA patients are treated with DMARDs, a significant proportion of patients are being overtreated. Importantly, UA is a diagnosis of exclusion and the introduction of novel classification criteria resulted in the remaining group becoming smaller and more homogeneous. The mild disease characteristics and the large proportion of patients not needing long-lasting DMARD treatment implies that the pathogenic mechanisms in a large proportion of patients might be different from RA. If so, this would imply that treatment guidelines for RA cannot be directly extrapolated to this group of UA patients.

Since our results demonstrate that it is not necessary to treat all UA patients, the development of stratification methods to identify UA patients that progress to RA and/or will benefit from DMARD treatment are needed. In the past, validated prediction models for UA were developed [30–34]. However, these models were also derived before the introduction of the 2010 criteria and are not reflective of the current UA population. So far there is little guidance from the literature on prediction models for the ‘contemporary UA population’, especially since this group is autoantibody negative and autoantibodies contributed significantly to the accuracy of former prediction models. Also, the recommendations for management strategies in current guidelines [11] (erosions, autoantibodies, high disease activity) rely on risk factors that were derived previously and are infrequent in contemporary UA.

Unexpectedly, we observed a significant increase in the percentages of patients developing RA in more recent inclusion periods. This finding is unexplained and as UA is now mostly autoantibody negative, this may relate to the observation that the incidence of ACPA-negative RA has also increased in past decades [35]. A limitation of our study is that although we did our best to create homogeneous patient groups by retrospectively using the same criteria in an inception-based cohort, unknown or unmeasured selection bias cannot be excluded. This is inherent to the observational design of the study [36]. Therefore randomized clinical trials will



**Fig. 5** Prevalence of prolonged DFS and progression to RA did not improve over time

**(A)** Percentage of patients achieving prolonged DFS within the 10 years of follow-up (defined as spontaneous remission of sustained remission after discontinuation of all DMARDs) did not improve. No significant increasing trend was observed among the consecutive time periods ( $P = 0.78$ ). Over time, the proportion of patients achieving spontaneous remission decreased and the proportion of patients achieving spontaneous disease-free remission increased, as would be expected with enhanced DMARD treatment. **(B)** The percentage of patients with progression to RA, according to 1987 and/or 2010 criteria after 1 year of follow-up. Progression to RA after 1 year did not decrease compared with 1993–1997, but rather demonstrated an increasing trend ( $P < 0.01$ ).

be needed to definitely conclude whether DMARD use in contemporary UA is valuable.

A strength of our study is that we assessed several long-term outcomes. The outcome for RA was defined as fulfilment of classification criteria (either the 1987 or 2010 criteria); this is the population from which all scientific data on the efficacy of DMARDs is derived. We acknowledge that other outcomes, such as persistency, are useful. We therefore studied spontaneous remission and sustained DMARD-free remission. The frequency of DMARD use increased over time. Likewise, decisions to taper and stop DMARDs were non-protocolized and DMARD tapering occurred less readily in earlier time periods. In order to create comparability across the inclusion periods despite these two issues, we used a long follow-up duration of 10 years to ascertain whether prolonged DFS was obtained and combined the DMARD-free remission group with the spontaneous remission group. In this way, time-dependent changes in DMARD start, DMARD tapering and DMARD stop were

overcome. Interestingly, although the division between the two components changed over time, the total percentage of this outcome was similar for the different inclusion periods. From patients' perspectives, patient-reported outcomes are relevant, and in our study the HAQ was used to this end. It could be argued that a slight trend towards physical improvement in our study could be observed in 2015–2019 (HAQ  $-0.09$ ), however, such an improvement would not be considered a clinically important difference [37]. Importantly, for all these different outcomes, similar findings were seen, showing the robustness of the results.

In conclusion, this study demonstrated that the clinical characterization of UA has changed over time and that the contemporary UA population is mostly autoantibody negative and presents with relatively mild disease. Although evidence from placebo-controlled trials is needed, our data do not support the use of DMARDs in the broad, contemporary UA population with the perspective of improving long-term outcomes for these

patients. Dedicated stratification methods are needed to identify the subgroup of patients who will develop RA or benefit from DMARDs. Moreover, in the absence of placebo-controlled trials and adequate prediction models for the contemporary UA population, it can be debated whether the international recommendations for treatment of early arthritis should be carefully rephrased.

## Acknowledgements

All authors contributed to the conception and study design. M.V. contributed to acquisition of the data and analysed the data. All authors contributed to interpretation of the data and the development of the manuscript. All authors approved the final version of the manuscript. Patient partners were involved in the design of the Leiden EAC. The Commissie Medische Ethiek of the Leiden University Medical Centre approved this study (B19.008).

**Funding:** The research leading to these results has received funding from the Dutch Arthritis Foundation and the European Research Council under the European Union's Horizon 2020 research and innovation program (starting grant agreement 714312). The funding source had no role in the design and conduct of the study; collection, management, analysis and interpretation of the data; preparation, review or approval of the manuscript; or decision to submit the manuscript for publication.

**Disclosure statement:** The authors have declared no conflicts of interest.

## Data availability statement

All data relevant to the study are included in the article or are uploaded as [supplementary information](#). Additional data are available upon reasonable request.

## Supplementary data

[Supplementary data](#) are available at *Rheumatology* online.

## References

- Wevers-de Boer KV, Heimans L, Huizinga TW *et al*. Drug therapy in undifferentiated arthritis: a systematic literature review. *Ann Rheum Dis* 2013;72:1436–44.
- Olivieri I, Sarzi-Puttini P, Bugatti S *et al*. Early treatment in early undifferentiated arthritis. *Autoimmun Rev* 2012; 11:589–92.
- Combe B, Landewe R, Lukas C *et al*. EULAR recommendations for the management of early arthritis: report of a task force of the European Standing Committee for International Clinical Studies Including Therapeutics (ESCISIT). *Ann Rheum Dis* 2007;66:34–45.
- Aletaha D, Neogi T, Silman AJ *et al*. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Ann Rheum Dis* 2010;69:1580–8.
- Machold KP, Landewé R, Smolen JS *et al*. The Stop Arthritis Very Early (SAVE) trial, an international multicentre, randomised, double-blind, placebo-controlled trial on glucocorticoids in very early arthritis. *Ann Rheum Dis* 2010;69:495–502.
- van Dongen H, van Aken J, Lard LR *et al*. Efficacy of methotrexate treatment in patients with probable rheumatoid arthritis: a double-blind, randomized, placebo-controlled trial. *Arthritis Rheum* 2007;56: 1424–32.
- Saleem B, Mackie S, Quinn M *et al*. Does the use of tumour necrosis factor antagonist therapy in poor prognosis, undifferentiated arthritis prevent progression to rheumatoid arthritis? *Ann Rheum Dis* 2008;67: 1178–80.
- Emery P, Durez P, Dougados M *et al*. Impact of T-cell costimulation modulation in patients with undifferentiated inflammatory arthritis or very early rheumatoid arthritis: a clinical and imaging study of abatacept (the ADJUST trial). *Ann Rheum Dis* 2010;69:510–6.
- Verstappen SMM, McCoy MJ, Roberts C *et al*. Beneficial effects of a 3-week course of intramuscular glucocorticoid injections in patients with very early inflammatory polyarthritis: results of the STIVEA trial. *Ann Rheum Dis* 2010;69:503–9.
- Durez P, de Bellefon LM, Depresseux G, Toukap AN, Lauwerys B, Houssiau FA. Infliximab versus placebo in adult patients with ACPA positive undifferentiated arthritis. Abstract 435. 2011 ACR/ARHP Annual Scientific Meeting, 5–9 November 2011.
- Combe B, Landewe R, Daien CI *et al*. 2016 update of the EULAR recommendations for the management of early arthritis. *Ann Rheum Dis* 2017;76:948–59.
- de Rooy DP, van der Linden MP, Knevel R *et al*. Predicting arthritis outcomes—what can be learned from the Leiden Early Arthritis Clinic? *Rheumatology (Oxford)* 2011;50:93–100.
- Zeidler H. Undifferentiated arthritis and spondylarthropathy as a major problem of diagnosis and classification. *Scand J Rheumatol Suppl* 1987;65:54–62.
- Zeidler H, Werdier D, Klauder A *et al*. Undifferentiated arthritis and spondylarthropathy as a challenge for prospective follow-up. *Clin Rheumatol* 1987;6(Suppl 2): 112–20.
- Zeidler H, Hülsemann JL. Benign polyarthritis and undifferentiated arthritis an epidemiological terra incognita. *Scand J Rheumatol Suppl* 1989;79:13–20.
- Hülsemann JL, Zeidler H. Diagnostic evaluation of classification criteria for rheumatoid arthritis and reactive arthritis in an early synovitis outpatient clinic. *Ann Rheum Dis* 1999;58:278–80.
- Morel J, Legouffe MC, Bozonat MC *et al*. Outcomes in patients with incipient undifferentiated arthritis. *Joint Bone Spine* 2000;67:49–53.

- 18 Wiles N, Symmons DPM, Harrison B *et al.* Estimating the incidence of rheumatoid arthritis: trying to hit a moving target? *Arthritis Rheum* 1999;42:1339–46.
- 19 Machold KP, Eberl G, Leeb BF *et al.* Early arthritis therapy: rationale and current approach. *J Rheumatol Suppl* 1998;53:13–9.
- 20 van der Horst-Bruinsma IE, Speyer I, Visser H *et al.* Diagnosis and course of early-onset arthritis: results of a special early arthritis clinic compared to routine patient care. *Br J Rheumatol* 1998;37:1084–8.
- 21 van der Linden MPM, Batstra MR, Bakker-Jonges LE *et al.* Toward a data-driven evaluation of the 2010 American College of Rheumatology/European League Against Rheumatism criteria for rheumatoid arthritis: is it sensible to look at levels of rheumatoid factor? *Arthritis Rheum* 2011;63:1190–9.
- 22 Bruce B, Fries JF. The Health Assessment Questionnaire (HAQ). *Clin Exp Rheumatol* 2005;23(5 Suppl 39):S14–8.
- 23 Arnett FC, Edworthy SM, Bloch DA *et al.* The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum* 1988;31:315–24.
- 24 Aletaha D, Neogi T, Silman AJ *et al.* 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Arthritis Rheum* 2010;62:2569–81.
- 25 Matthijssen XME, Niemantsverdriet E, Huizinga TWJ *et al.* Enhanced treatment strategies and distinct disease outcomes among autoantibody-positive and -negative rheumatoid arthritis patients over 25 years: a longitudinal cohort study in the Netherlands. *PLoS Med* 2020;17:e1003296.
- 26 Rubin DB. Multiple imputation for nonresponse in surveys. New York: Wiley, 1987.
- 27 Cuzick J. A Wilcoxon-type test for trend. *Stat Med* 1985; 4:87–90.
- 28 Ward MM, Guthrie LC, Alba MI. Clinically important changes in individual and composite measures of rheumatoid arthritis activity: thresholds applicable in clinical trials. *Ann Rheum Dis* 2015;74:1691–6.
- 29 Smolen JS, Landewé RBM, Bijlsma JWJ *et al.* EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2019 update. *Ann Rheum Dis* 2020; 79:685–99.
- 30 van der Helm-van Mil AH, le Cessie S, van Dongen H *et al.* A prediction rule for disease outcome in patients with recent-onset undifferentiated arthritis: how to guide individual treatment decisions. *Arthritis Rheum* 2007;56:433–40.
- 31 van der Helm-van Mil AH, Detert J, le Cessie S *et al.* Validation of a prediction rule for disease outcome in patients with recent-onset undifferentiated arthritis: moving toward individualized treatment decision-making. *Arthritis Rheum* 2008;58:2241–7.
- 32 Kuriya B, Cheng CK, Chen HM *et al.* Validation of a prediction rule for development of rheumatoid arthritis in patients with early undifferentiated arthritis. *Ann Rheum Dis* 2009;68:1482–5.
- 33 Li C, Zhang Y, Song H *et al.* Anti-cyclic citrullinated peptide antibody predicts the development of rheumatoid arthritis in patients with undifferentiated arthritis. *Chin Med J (Engl)* 2019;132:2899–904.
- 34 McNally E, Keogh C, Galvin R *et al.* Diagnostic accuracy of a clinical prediction rule (CPR) for identifying patients with recent-onset undifferentiated arthritis who are at a high risk of developing rheumatoid arthritis: a systematic review and meta-analysis. *Semin Arthritis Rheum* 2014; 43:498–507.
- 35 Matthijssen XME, Huizinga TWJ, van der Helm-van Mil AHM. Increasing incidence of autoantibody-negative RA is replicated and is partly explained by an aging population. *Ann Rheum Dis* 2020;doi: 10.1136/annrheumdis-2020-217609.
- 36 Landewé R, van der Heijde D. Primer: challenges in randomized and observational studies. *Nat Clin Pract Rheumatol* 2007;3:661–6.
- 37 Kitchen H, Hansen BB, Abetz L, Højbjerg L, Strandberg-Larsen M. Patient-reported outcome measures for rheumatoid arthritis: minimal important differences review. Abstract 2268. 2013 ACR/ARHP Annual Meeting, 25–30 October 2013.