

## **Supplementary files**

S1. Baseline characteristics of early arthritis patients with a clinical diagnosis of UA changed over time

S2. Baseline characteristics UA-patients excluded from the study population

S3. Disease activity scores during follow-up, compared to 1993-1997

S4. Components of the disease activity scores during follow-up, compared to 1993-1997

S5. Functional disability during follow-up, compared to 1993-1997

S6. Fatigue and pain during follow-up, compared to 1993-1997

S7. prolonged DMARD-free status within 5-years of follow-up

S8. Sensitivity analyses stratified for ACPA-status

**S1. Baseline characteristics of early arthritis patients with a clinical diagnosis of UA changed over time**

	<b>Total</b> N = 1626	<b>1998-2005</b> N = 479	<b>2006-2010</b> N = 447	<b>2011-2014</b> N = 393	<b>2015-2019</b> N = 307
<b>Age (years), mean (SD)</b>	51 (17)	51 (17)	53 (16)	54 (16)	58 (16)
<b>Gender (female), %</b>	55	62	66	68	59
<b>Symptom duration, median (IQR)</b>	14 (6-29)	19 (9-35)	17 (7-31)	10 (4-27)	11 (6-21)
<b>ACPA-positive, %</b>	19	29	22	11	9
<b>RF-positive, %</b>	24	31	26	21	16
<b>SJC (0-44), median (IQR)</b>	3 (1-6)	3 (1-7)	3 (1-7)	3 (1-6)	2 (1-5)
<b>TJC (0-53), median (IQR)</b>	5 (1-12)	7 (2-14)	7 (3-13)	4 (2-10)	4 (2-9)
<b>CRP, median (IQR)</b>	7 (3-20)	8 (3-22)	7 (3-22)	5 (3-16)	6 (3-17)
<b>VAS (0-100), median (IQR)</b>	34 (18-55)	34 (17-54)	31 (12-53)	40 (20-60)	30 (20-50)
<b>DAS28CRP, median (IQR)</b>	3.8 (3.0-4.6)	3.9 (3.1-4.7)	3.8 (3.0-4.7)	3.7 (3.0-4.4)	3.6 (2.8-4.3)
<b>HAQ-DI, median (IQR)</b>	0.63 (0.25-1.13)	0.63 (0.25-1.08)	0.72 (0.25-1.17)	0.63 (0.25-1.15)	0.65 (0.25-1.25)

**Legend:** Baseline characteristics of all UA-patients with a clinical diagnosis of UA, defined by their treating rheumatologist. Over time, UA-patients became more frequently autoantibody-negative, had less affected joints and less often elevated acute-phase-reactants. Missing values were imputed and used for baseline table; values represent mean of the imputed medians calculated using Rubin's rules.

*ACPA: anticitrullinated protein antibody, CRP: C-reactive protein, DAS: disease activity score, ESR: estimated sedimentation rate, HAQ-DI: health assessment questionnaire disability index, SJC: swollen joint count, TJC: tender joint count, VAS: visual analogue scale.*

## S2. Baseline characteristics of UA-patients excluded from the study population

	Study population N = 1132	Excluded UA-patients N = 127
Age (years), mean (SD)	52 (17)	51 (16)
Gender (female), %	59	55
Sympt. dur. (weeks), median (IQR)	11 (4-26)	20 (9-32)*
ACPA-positive, %	5	15*
RF-positive, %	7	16*
SJC (0-28), median (IQR)	1 (1-3)	2 (1-3)
TJC (0-28), median (IQR)	2 (1-4)	4 (2-5)*
CRP, median (IQR)	6 (3-17)	8 (3-24)
VAS (0-100), median (IQR)	30 (14-50)	30 (15-53)
DAS28CRP, median (IQR)	3.3 (2.7-3.9)	3.6 (3.1-4.2)*
HAQ-DI, median (IQR)	0.50 (0.13-0.91)	0.9 (0.4-1.3)*

**Legend:** Baseline characteristics of UA-patients excluded from the study population due to trial participation (n=127), compared to baseline characteristics of the study population. Baseline characteristics were compared using ANOVA or Kruskal Wallis, as appropriate. The excluded UA-patients were more often auto-antibody positive and had higher disease activity at baseline, which would be expected due to trial inclusion criteria. \* p<0.05

*ACPA: anticitrullinated protein antibody, CRP: C-reactive protein, DAS: disease activity score, ESR: estimated sedimentation rate, HAQ-DI: health assessment questionnaire disability index, SJC: swollen joint count, TJC: tender joint count, VAS: visual analogue scale.*

### S3. Disease activity scores during follow-up, compared to 1993-1997

Inclusion period	DAS28CRP at baseline	p-value	Decline in first year	p-value	DAS28CRP after 1 year	p-value
1993 – 1997	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
1998 – 2005	+0.05 (-0.13, +0.22)	0.606	+0.12 (-0.17, +0.40)	0.426	+0.07 (-0.14, +0.26)	0.539
2006 – 2010	+0.01 (-0.16, +0.18)	0.875	-0.07 (-0.38, +0.24)	0.653	+0.02 (-0.19, +0.24)	0.838
2011 - 2014	-0.13 (-0.29, +0.03)	0.112	-0.07 (-0.32, +0.18)	0.564	<b>-0.18 (-0.36, -0.00)</b>	<b>0.047</b>
2015 – 2019	-0.10 (-0.26, +0.06)	0.227	-0.13 (-0.38, +0.13)	0.322	<b>-0.25 (-0.44, -0.07)</b>	<b>0.007</b>

**Legend:** Differences in baseline DAS28CRP, decline in the first year in DAS28CRP and levels for the subsequent follow-up afterwards, compared to the earliest inclusion period (1993-1997), corrected for age and gender, visualized as estimated marginal means (95%CI). Estimated marginal mean for DAS28CRP in the period 1993-1997 at baseline is 3.35 (95% CI, 3.22-3.48) and 2.63 (95% CI, 2.54-2.84) for the subsequent follow-up afterwards.

*CRP: C-reactive protein, DAS: disease activity score*

#### S4. Components of the disease activity scores during follow-up, compared to 1993-1997

Inclusion period	SJC28 at baseline	p-value	Decline in first year	p-value	SJC28 after 1 year	p-value
1993 – 1997	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
1999 – 2005	-0.06 (-0.48, +0.37)	0.796	+0.41 (-0.28, +1.10)	0.245	+0.12 (-0.34, +0.59)	0.602
2006 – 2010	-0.13 (-0.58, +0.32)	0.576	-0.26 (-1.00, +0.49)	0.504	-0.32 (-0.82, +0.18)	0.206
2011 - 2014	<b>-0.45 (-0.83, -0.06)</b>	<b>0.023</b>	-0.24 (-0.85, +0.37)	0.439	<b>-0.77 (-1.18, -0.37)</b>	<b>&lt;0.01</b>
2015 – 2019	<b>-0.47 (-0.88, -0.06)</b>	<b>0.025</b>	-0.46 (-1.08, +0.16)	0.149	<b>-0.92 (-1.36, -0.48)</b>	<b>&lt;0.01</b>

Inclusion period	TJC28 at baseline	p-value	Decline in first year	p-value	TJC28 after 1 year	p-value
1993 – 1997	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
1999 – 2005	+0.06 (-0.62, +0.74)	0.361	+0.64 (-0.37, +1.64)	0.213	+0.45 (-0.34, +1.23)	0.267
2006 – 2010	+0.10 (-0.56, +0.78)	0.511	-0.32 (-0.74, +1.38)	0.556	+0.50 (-0.30, +1.29)	0.219
2011 - 2014	-0.17 (-0.77, +0.40)	0.421	-0.14 (-1.00, +0.72)	0.747	-0.42 (-1.07, +0.22)	0.197
2015 – 2019	-0.00 (-0.58, +0.62)	0.405	-0.58 (-1.47, +0.31)	0.202	-0.58 (-1.27, +0.11)	0.102

Inclusion period	CRP at baseline	p-value	Decline in first year	p-value	CRP after 1 year	p-value
1993 – 1997	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
1999 – 2005	1.5 (-2.4, +5.3)	0.445	-0.9 (-6.9, +5.1)	0.796	-0.9 (-4.7, +3.0)	0.658
2006 – 2010	+2.1 (-1.7, +5.8)	0.280	-4.5 (-11.9, +2.8)	0.224	-2.3 (-6.3, +1.6)	0.251
2011 - 2014	<b>-6.0 (-9.6, -2.5)</b>	<b>0.001</b>	+2.2 (-2.9, +7.4)	0.397	<b>-4.0 (-7.5, -0.6)</b>	<b>0.022</b>
2015 – 2019	<b>-7.0 (-10.7, -3.3)</b>	<b>&lt;0.001</b>	+5.2 (-0.2, +10.6)	0.058	<b>-4.0 (-7.7, -0.3)</b>	<b>0.036</b>

Inclusion period	VAS at baseline	p-value	Decline in first year	p-value	VAS after 1 year	p-value
1993 – 1997	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
1999 – 2005	+0.32 (-4.29, +4.93)	0.891	+1.37 (-6.66, +9.41)	0.737	-0.28 (-5.57, +5.02)	0.918
2006 – 2010	-4.36 (-8.75, +0.02)	0.051	-3.23 (-4.92, +11.39)	0.436	+1.20 (-3.91, +6.13)	0.644
2011 - 2014	-0.03 (-4.12, +4.05)	0.987	-1.58 (-5.00, +8.17)	0.637	+3.47 (-0.94, +7.88)	0.123
2015 – 2019	-2.47 (-6.65, +1.70)	0.245	-5.09 (-1.70, +11.88)	0.141	+1.60 (-3.14, +6.34)	0.507

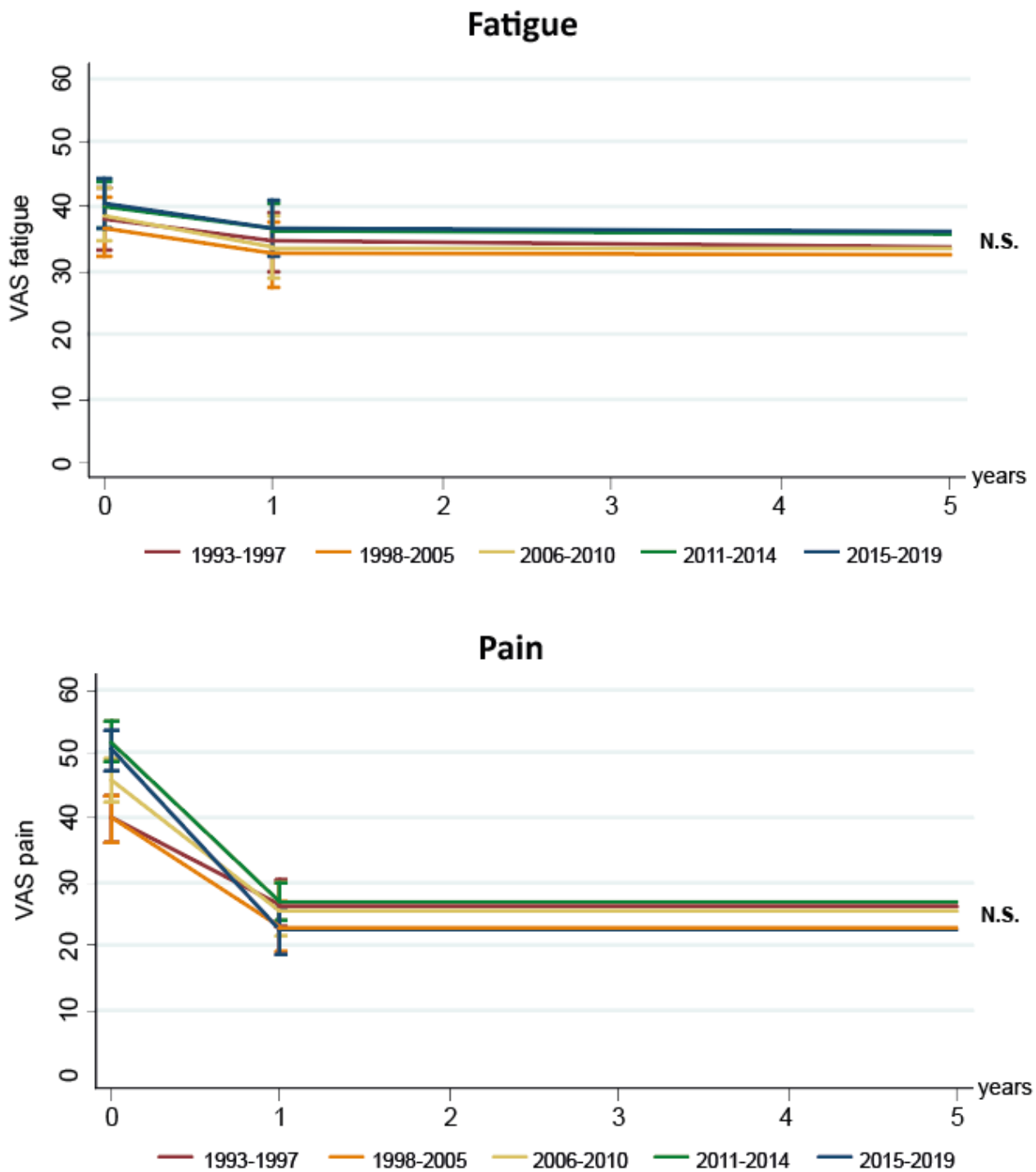
## S5. Functional disability during follow-up, compared to 1993-1997

Inclusion period	HAQ at baseline	p-value	Decline in first year	p-value	HAQ after 1 year	p-value
1993 – 1997	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
1998 – 2005	+0.01 (-0.11, +0.13)	0.848	+0.00 (-0.18, +0.19)	0.973	-0.01 (-0.12, +0.11)	0.917
2006 – 2010	-0.01 (-0.12, +0.10)	0.852	-0.05 (-0.24, +0.14)	0.620	-0.03 (-0.14, +0.08)	0.628
2011 - 2014	-0.04 (-0.14, +0.07)	0.472	-0.02 (-0.17, +0.14)	0.840	-0.05 (-0.14, +0.05)	0.332
2015 – 2019	-0.01 (-0.11, +0.10)	0.886	-0.09 (-0.25, +0.07)	0.263	-0.09 (-0.20, +0.01)	0.086

**Legend:** Differences in baseline HAQ, decline in the first year in HAQ-DI and levels for the subsequent follow-up afterwards, compared to the earliest inclusion period (1993-1997), corrected for age and gender, visualized as estimated marginal means (95%CI). Estimated marginal means for HAQ-DI in 1993-1997 at baseline were 0.62 (95% CI; 0.53-0.73) and 0.44 (95%CI; 0.37-0.52) for the subsequent follow-up afterwards.

*HAQ-DI: health assessment questionnaire disability index*

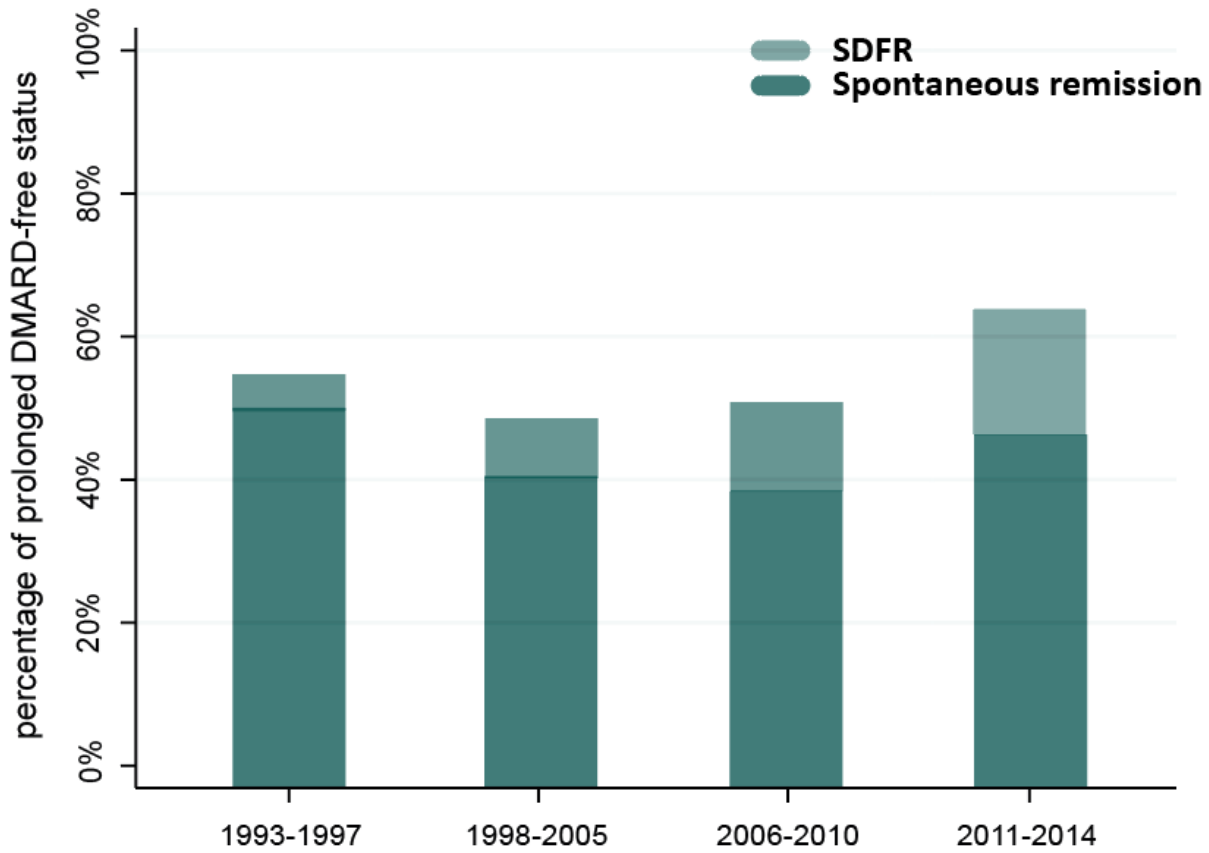
S6. Fatigue and pain during follow-up, compared to 1993-1997



**Legend:** Fatigue and pain (measured using Visual Analogue Scale) did not improve compared to 1993-1997. Fatigue and pain VAS-scores were modelled based on the estimated marginals means which results from the age- and gender-corrected linear mixed models.

N.S.: not significant ( $p > 0.05$ )

### S7. Prolonged DMARD-free status within 5-years of follow-up



**Legend:** Percentages of prolonged DMARD-free status within 5-years of follow-up, i.e. also including the inclusion period 2011-2014. Test for trend demonstrated no significant increase in DFS-prevalence over time ( $p=0.28$ ).

*DFS: prolonged DMARD-free status, SDFR: sustained DMARD-free remission*





**Legend:** Disease activity scores over time, physical functioning over time and prolonged DMARD-free status and fulfillment of classification-criteria for RA stratified for ACPA-positive UA-patients **(A)** and ACPA-negative UA-patients **(B)**. In both ACPA-positive and ACPA-negative UA-patients disease activity scores improved over time. Among ACPA-negative UA-patients, physical functioning, prevalence of prolonged DMARD-free status and the percentage of fulfillment of classification-criteria for RA did not concomitantly improve with improved disease activity scores. Among ACPA-positive UA, physical functioning and prolonged DMARD-free status did tend to concomitantly improve over time.

DAS28CRP and HAQ-DI scores were modelled based on the estimated marginal means which results from the age- and gender-corrected linear mixed models. Due to the limited amount of follow-up data among patients included between 2006-2010, this inclusion period could not be included in the linear mixed model analysis of DAS28CRP and HAQ-DI.

*CRP: C-reactive protein, DAS: disease activity score, HAQ-DI: health assessment questionnaire disability index, DMARD: disease modifying antirheumatic drugs, RA: rheumatoid arthritis*