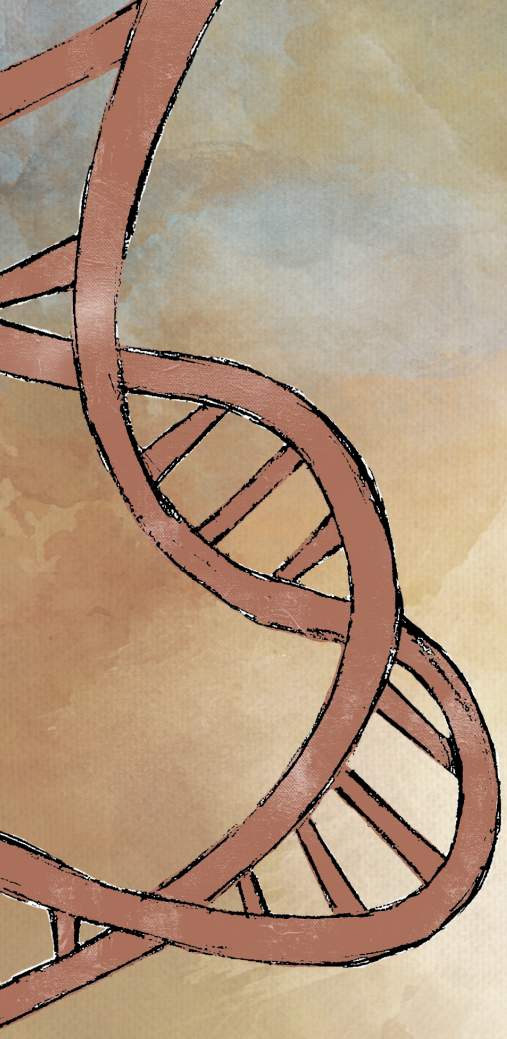
Clinically suspect arthralgia: unraveling the development of rheumatoid arthritis

Wouters, F.



7

Determining in which pre-arthritis stage HLA-shared epitope alleles and smoking exert their effect on the development of rheumatoid arthritis

Fenne Wouters
Marc P. Maurits
Laurette van Boheemen
Marloes Verstappen
Kulveer Mankia
Xanthe M.E. Matthijssen
Annemarie L. Dorjée
Paul Emery
Rachel Knevel
Dirkjan van Schaardenburg
René E.M. Toes
Annette H.M. van der Helm-van Mil

Published as *Original Research* in Annals of the Rheumatic Diseases
2022 Jan;81(1):48-55.
doi:10.1136/annrheumdis-2021-220546.

Abstract

Objectives

The HLA-shared epitope-alleles (HLA-SE) and smoking are the most prominent genetic and environmental risk factors for rheumatoid arthritis (RA). However, at which pre-arthritis stage (asymptomatic/symptomatic) they exert their effect is unknown. We aimed to determine whether HLA-SE and smoking are involved in the onset of autoantibody-positivity, symptoms (Clinically Suspect Arthralgia, CSA) and/or progression to clinical arthritis.

Methods

We performed meta-analyses on results from literature on associations of HLA-SE and smoking with anti-citrullinated protein antibodies (ACPA) in the asymptomatic population. Next, we studied associations of HLA-SE and smoking with autoantibody-positivity at CSA-onset, and with progression to clinical inflammatory arthritis (IA) during follow-up. Associations in ACPA-positive CSA-patients were validated in meta-analyses with other arthralgia-cohorts. Analyses were repeated for rheumatoid factor (RF), anti-carbamylated and anti-acetylated antibodies (anti-CarP, AAPA).

Results

Meta-analyses showed that HLA-SE is not associated with ACPA-positivity in the asymptomatic population (OR 1.06 (95%CI;0.69-1.64)), whereas smoking was associated (OR 1.37 (1.15-1.63)). At CSA-onset, both HLA-SE and smoking associated with ACPA-positivity (OR 2.08 (1.24-3.49), OR 2.41 (1.31-4.43)). During follow-up, HLA-SE associated with IA-development (HR 1.86 (1.23-2.82)), in contrast to smoking. This was confirmed in meta-analyses in ACPA-positive arthralgia (HR 1.52 (1.08-2.15)). HLA-SE and smoking were not associated with RF, anti-CarP or AAPA-positivity at CSA-onset. Longitudinally, AAPA associated with IA-development independent from ACPA and RF (HR 1.79 (1.02-3.16)), anti-CarP did not.

Conclusions

HLA-SE and smoking act at different stages: smoking confers risk for ACPA- and symptom-development, whereas HLA-SE mediates symptom- and IA-development. These data enhance the understanding of the timing of the key risk factors in development of RA.

Introduction

The HLA shared epitope (SE) is the most well-known and strongest genetic risk factor for development of rheumatoid arthritis (RA), especially for anti-citrullinated protein antibody (ACPA)-positive RA.[1-14] Similarly, smoking is the strongest environmental risk factor for autoantibody-positive RA[2,9,10,12,15]; multiple studies have shown this effect is mostly present in people carrying HLA-SE alleles.[1,3,5,6,8,14,16] This knowledge is mostly obtained from case-control studies comparing RA-patients and healthy controls. During the last decade, research attention has shifted to the stages that precede clinical arthritis and RA and several pre-RA stages have been discerned. However, so far it remains undetermined at which stage(s) HLA-SE alleles and smoking exert their effect.

The following stages are distinguished. An asymptomatic stage in which autoimmune responses can develop, resulting in autoantibody-positivity. Then, autoimmune responses can mature and a symptomatic stage develops. The pattern of symptoms that is considered specific for an increased risk of RA is called clinically suspect arthralgia (CSA). Patients with CSA can progress to clinically apparent inflammatory arthritis (IA); the stage when RA is generally diagnosed.[17] This model suggests that genetic factors exert their influence first, followed by smoking with subsequent autoantibody-development.[17,18] However, this time-order has never been shown.

In addition to a nested case-control study,[19] several longitudinal studies assessed genetic factors and/or smoking and provided data either from healthy to IA but not the intermediate stages, or from mixed populations of asymptomatic and symptomatic people.[20-24] These approaches do not allow determination of stage-dependent effects. As for the asymptomatic stage, contrasting findings are reported on associations between HLA-SE alleles and smoking and the presence of ACPA in the general population.[2,14,25-28] To the best of our knowledge only one study evaluated the effect of smoking on the progression from ACPA-positivity to CSA.[29] Furthermore, longitudinal studies within arthralgia are scarce and their findings varied.[30,31] The mentioned studies focused on ACPA, however, HLA-SE and smoking might also interact with other autoantibodies such as rheumatoid factor (RF), anti-carbamylated (anti-CarP) and anti-acetylated (AAPA) protein antibodies, the time-effects of which have not yet been studied.

We aimed to determine at which pre-RA stage HLA-SE and smoking exert their effect by studying both original and previously reported data. More specifically, we performed meta-analyses on literature from the general population, analyzed our own data at CSA-onset and during progression to IA, and finally performed meta-

analyses using data from different longitudinal arthralgia-cohorts. In doing this we focused on fine-staging the effects in the development of ACPA-positive RA. Analyses were repeated for ACPA-negative RA and associations of RF, anti-CarP and AAPA.

Methods

Summarizing literature obtained from the general population

The literature was reviewed on studies reporting associations between HLA-SE and/or smoking with the presence of ACPA in the asymptomatic population, as described supplementary. Results were pooled in meta-analyses. Although these studies were cross-sectional in nature, observed findings were considered to reflect the influence of HLA-SE/smoking on ACPA-development, as this is most likely the first event in the development of ACPA-positive RA.

The symptomatic phase

Associations of HLA-SE and smoking with autoantibodies at CSA-onset were investigated in the Leiden CSA-cohort, we did not identify large cohorts for validation since most arthralgia-cohorts did not include autoantibody-negative patients. Additionally, the role of HLA-SE and smoking in progression from arthralgia to IA was investigated in the Leiden CSA-cohort. Results obtained in the ACPA-positive subgroup were validated in ACPA-positive arthralgia/at-risk-patients from two independent cohorts (Amsterdam, Leeds).

Measurements at CSA-onset

Patients presenting with CSA to the Leiden rheumatology outpatient clinic between April 2012-September 2019 were studied. As described in detail previously,[32] patients had recent-onset (<1 year) arthralgia of small joints and were, according to the clinical expertise and pattern recognition of the rheumatologist, at risk for progression to RA. Patients were excluded if clinical arthritis was already present, or if a different explanation for the joint pain was more likely. At baseline smoking-status (present/past/never) was obtained through questionnaires. Presence of IgM RF (in-house ELISA, cut-off >3.5 IU/mL) and IgG ACPA (anti-CCP2, Phadia, Nieuwegein, the Netherlands, cut-off >7 IU/mL) was determined during routine laboratory measurements in all patients, presence of IgG anti-CarP and IgG AAPA with in-house ELISA in a subset of patients. Detailed methods are described supplementary. The HLA-SE alleles were extracted from whole genome sequencing data; the HLA-region was isolated and imputed using the SNP2HLA software and T1DGC reference panel.[33] HLA-SE positivity was subsequently defined as the presence of 1 or 2 of the HLA-DRB1 alleles *0101, *0102, *0401, *0404, *0405, *0408 and *1001 (see supplementary material).[34]

Measurements on the progression from CSA to IA

Patients in the Leiden CSA-cohort were prospectively followed (median (IQR) 106 weeks (43-114)) for development of IA, which was defined as ≥1 swollen joints at physical examination by a rheumatologist. Treatment with disease-modifying anti-rheumatic drugs (DMARDs, including systemic or intra-articular corticosteroids) was not allowed before IA-development. Analyses evaluating progression to IA were stratified for ACPA-status and results from the ACPA-positive subgroup were studied in meta-analyses with the results from ACPA-positive patients included in the Amsterdam and Leeds cohorts. The Amsterdam cohort included ACPA- and/or RF-positive patients; for this study the data from ACPA-positive arthralgia-patients was obtained and studied.[31] Data on smoking history, presence of HLA-SE, RF, ACPA and anti-CarP were collected previously and are described supplementary. In addition, IgG AAPA was determined in baseline serum samples simultaneous with Leiden CSA-samples. Results on predictive value of HLA-SE and smoking in ACPA-positive patients from the Leeds cohort were obtained from Rakieh et al.,[30] detailed methods are described supplementary. Anti-CarP and AAPA were not determined in the Leeds cohort.

In sub-analyses, the association of HLA-SE and smoking with RA-development was studied using Leiden CSA-data; RA was defined as development of IA plus fulfillment of the 1987 and/or 2010 EULAR/ACR criteria at that time.[35,36]

Statistics

Results from literature on associations of HLA-SE and smoking with ACPA in the asymptomatic population were pooled in inverse-variance weighted meta-analyses.

Associations of HLA-SE and smoking with autoantibody-positivity at CSA-onset were investigated with logistic regression analyses. Results of smoking were also stratified for HLA-SE. Associations of HLA-SE and smoking with ACPA-level in ACPA-positive patients were evaluated with Mann-Whitney U tests and logistic regression.

Associations with IA-development were studied with cox regression, also stratified for ACPA. Results in ACPA-positive arthralgia were summarized in inverse-variance weighted meta-analyses.

Associations of anti-CarP and AAPA with IA-development were corrected for concomitant ACPA- and RF-positivity in multivariable analyses with the autoantibody-negative group as reference in the Leiden data (the Amsterdam cohort did not include autoantibody-negative patients). The additional value of anti-CarP and AAPA to ACPA- and RF-positivity for prediction of IA-development was determined in the ACPA+RF+

subgroup from the Leiden and Amsterdam cohorts.

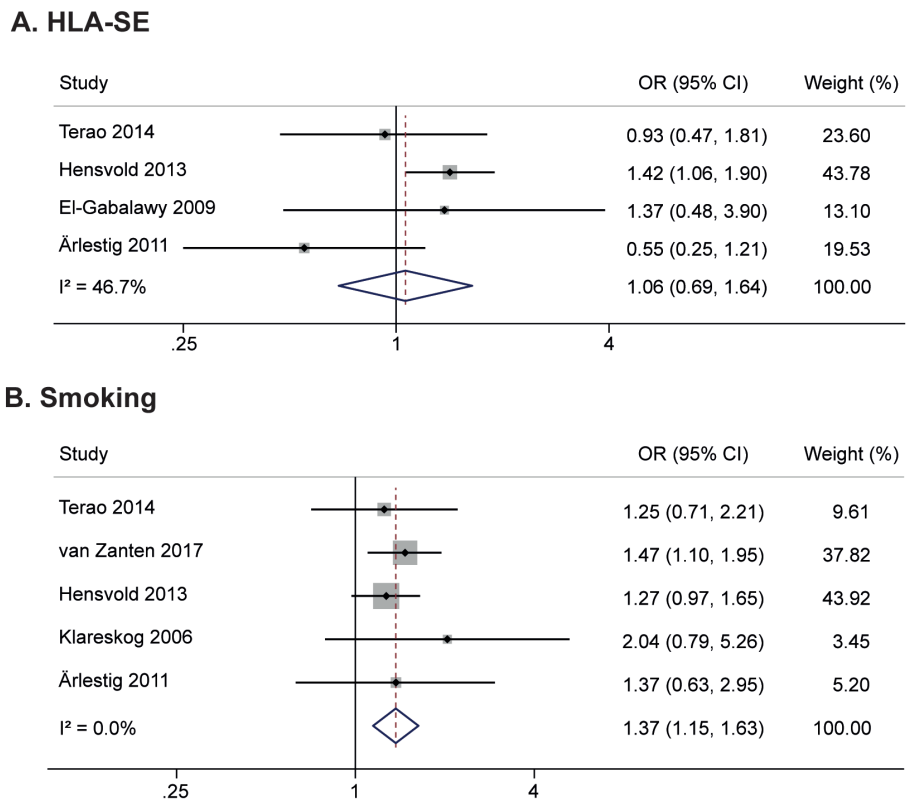
P-values <0.05 were considered statistically significant. IBM SPSS Statistics (V25) and STATA (V16) were used.

Results

Summarizing literature obtained from the asymptomatic stage

Four studies were identified on the association of HLA-SE with ACPA, and five on smoking (Supplementary File 1). Meta-analyses revealed that HLA-SE was not associated with ACPA-positivity (OR 1.06 (95%CI;0.69-1.64)), whereas smoking was associated (OR 1.37 (1.15-1.63)), Figure 1. This suggests that smoking, but not HLA-SE, conferred risk for ACPA-development in the asymptomatic stage.

Figure 1. Meta-analyses on HLA-SE (A) and smoking (B) in asymptomatic healthy individuals and first-degree relatives, showing associations with presence of ACPA for smoking but not for HLA-SE



HLA-SE: shared epitope, ACPA: anti-citrullinated protein antibody, OR: odds ratio, CI: confidence interval

Associations with ACPA at CSA-onset

Characteristics of patients presenting with CSA (n=577) are provided in the supplementary materials. HLA-SE positive CSA-patients were more often ACPA-positive (OR 2.08 (95%CI;1.24-3.49), this relation was dependent on the number of alleles (Table 1). Patients that smoked were also more often ACPA-positive (OR 2.41 (1.31-4.43)), which was also dose-dependent with an higher OR for current smokers than ex-smokers (Table 1). In addition, within smokers, it was dependent on number of packyears, because the odds for being ACPA-positive increased per increase in packyear (OR 1.03 (1.00-1.06)). As it has been reported in RA that the association of smoking is dependent on HLA-SE status, we stratified the analyses of smoking (ever versus never) for HLA-SE; smoking was associated with ACPA-status in both HLA-SE negative and HLA-SE positive CSA-patients (Table 1). The association of HLA-SE and smoking with ACPA-positivity was present for both ACPA double-positivity (ACPA+RF+) and single-positivity (ACPA+RF-), and thus independent from RF (Supplementary Table 2). Studying the levels of ACPA within ACPA-positive patients at CSA-onset revealed that HLA-SE positive patients tended to have higher levels than HLA-SE negative patients (median (IQR) 236 (72-340) versus 144 (32-340), p=0.12), whilst no effect on ACPA-levels was present for smoking (229 (64-340) versus 222 (52-340), p=0.89), see Supplementary Table 3 for results from regression analyses.

Table 1. Associations of HLA-SE and smoking with presence of ACPA in patients newly presenting with CSA

		ACPA positive, n (%)	ACPA negative, n (%)	OR (95% CI)	p-value
All patients					
HLA-SE	Absent	27 (39)	259 (57)	Reference	--
	Present	42 (61)	194 (43)	2.08 (1.24-3.49)	0.006
HLA-SE	0	27 (39)	259 (57)	Reference	--
	1	31 (45)	161 (36)	1.85 (1.06-3.21)	0.029
	2	11 (16)	33 (7)	3.20 (1.45-7.04)	0.004
Smoking	Never	15 (23)	185 (42)	Reference	--
	Ever	49 (77)	251 (58)	2.41 (1.31-4.43)	0.005
Smoking	Never	15 (23)	185 (42)	Reference	--
	Ex-smoker	28 (44)	161 (37)	2.15 (1.12-4.16)	0.024
	Current smoker	21 (33)	90 (21)	2.88 (1.42-5.85)	0.003

Table 1. *Continued*

		ACPA positive, n (%)	ACPA negative, n (%)	OR (95% CI)	p-value
HLA-SE positive subgroup					
Smoking	Never	10 (27)	77 (45)	Reference	--
	Ever	27 (73)	95 (55)	2.19 (1.00-4.80)	0.051
Smoking	Never	10 (27)	77 (45)	Reference	--
	Ex-smoker	13 (35)	57 (33)	1.76 (0.72-4.29)	0.22
	Current smoker	14 (38)	38 (22)	2.84 (1.15-6.98)	0.023
HLA-SE negative subgroup					
Smoking	Never	4 (18)	99 (43)	Reference	--
	Ever	18 (82)	130 (57)	3.43 (1.12-10.45)	0.030
Smoking	Never	4 (18)	99 (43)	Reference	--
	Ex-smoker	11 (50)	89 (39)	3.06 (0.94-9.95)	0.063
	Current smoker	7 (32)	41 (18)	4.23 (1.17-15.22)	0.027

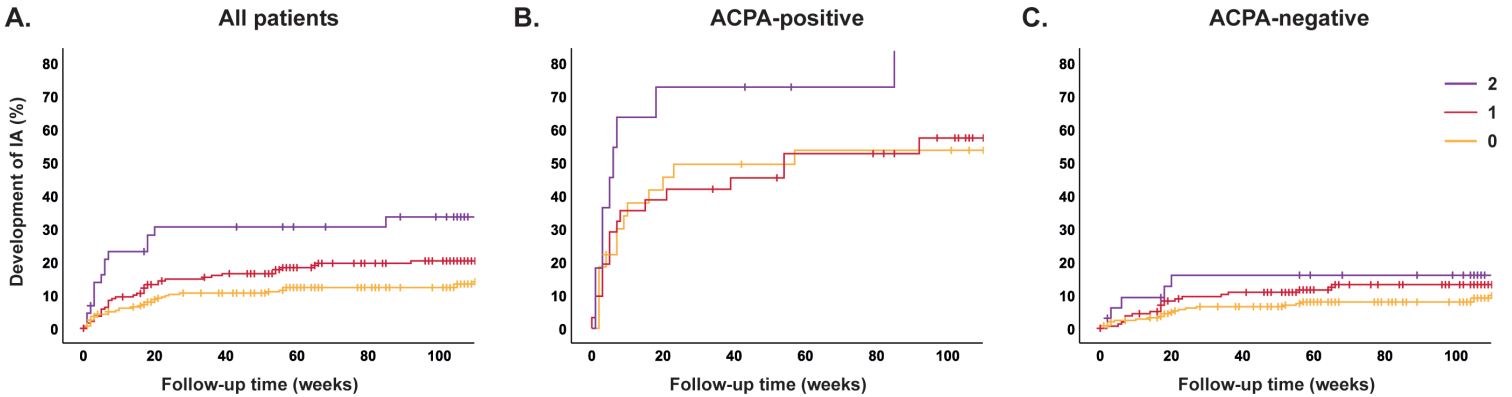
*Numbers on smoking in HLA-SE strata do not add up to numbers in the total CSA-group as some patients with data on smoking have missing data on HLA-SE.
HLA-SE: shared epitope, ACPA: anti-citrullinated protein antibody, CSA: clinically suspect arthralgia, OR: odds ratio, CI: confidence interval

Progression to IA in ACPA-positive CSA

Patients were followed for development of IA; median time till IA was 16 weeks (IQR 3-36), non-progressors were followed for median 109 (62-116) weeks. Presence of HLA-SE was significantly associated with IA-development in all CSA-patients (HR 1.86 (95%CI;1.23-2.82)), also here a dose-response relation was present (Figure 2A, Supplementary Table 4). Within the ACPA-positive subgroup the HR was 1.29 (0.67-2.47, Figure 2B, Supplementary Table 4). Because of the small sample size after stratification and risk of type-II error, we performed meta-analysis including ACPA-positive patients from two other arthralgia-cohorts. This showed that HLA-SE significantly associated with IA-development in ACPA-positive patients (HR 1.52 (1.08-2.15), Figure 4A).

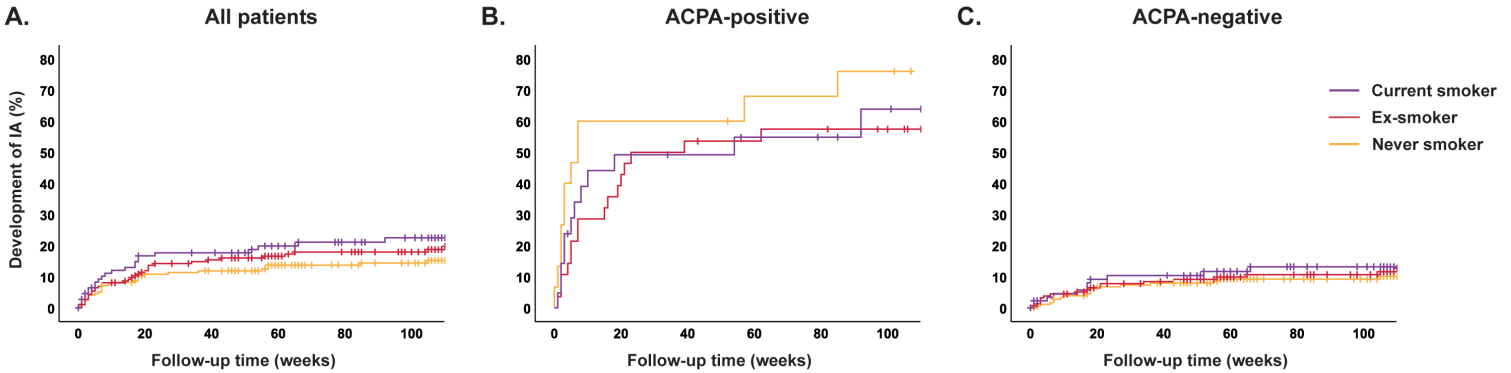
Smoking was not associated with IA-development, neither in the total CSA-population (HR 1.40 (0.90-2.18), Figure 3A, Supplementary Table 5), nor in the ACPA-positive subgroup (HR 0.59 (0.29-1.18), Figure 3B, Supplementary Table 5), nor in meta-analysis including ACPA-positive patients from three cohorts (HR 0.94 (0.67-1.33), Figure 4B).

Figure 2. Associations of number of HLA-SE alleles (0/1/2 alleles present) with progression from CSA to inflammatory arthritis



Corresponding hazard ratios, with 0 HLA-SE alleles as reference category were: (A) HR 1.65 (95% CI 1.06-2.56) and HR 3.03 (1.64-5.61) for 1 and 2 HLA-SE alleles respectively, (B) HR 1.05 (0.52-2.13) and HR 2.32 (1.00-5.41), and (C) HR 1.66 (0.94-2.94) and HR 2.00 (0.76-5.28), see Supplementary Table 4.
HLA-SE: shared epitope, CSA: clinically suspect arthralgia, ACPA: anti-citrullinated protein antibody, IA: inflammatory arthritis, HR: hazard ratio, CI: confidence interval

Figure 3. Associations of smoking with progression from CSA to inflammatory arthritis



Corresponding hazard ratios, with never smoker as reference category were: (A) HR 1.25 (95% CI 0.76-2.06) and HR 1.66 (0.97-2.83) for ex-smoker and current smoker respectively, (B) HR 0.55 (0.26-1.19) and HR 0.64 (0.28-1.45), and (C) HR 1.17 (0.61-2.24) and HR 1.56 (0.76-3.18), see Supplementary Table 5.
CSA: clinically suspect arthralgia, ACPA: anti-citrullinated protein antibody, IA: inflammatory arthritis, HR: hazard ratio, CI: confidence interval

Thus HLA-SE, but not smoking, influenced the risk to progress from ACPA-positive CSA to RA.

Associations of HLA-SE and smoking in ACPA-negative CSA

Presence of HLA-SE was associated with IA-development in ACPA-negative patients (HR 1.71 (0.99-2.96)), although the CI just included 1 (Figure 2C, Supplementary Table 4). Within ACPA-/RF- and ACPA-/RF+ CSA-patients associations of HLA-SE with IA-development were HR 1.64 (0.90-2.99) and HR 2.07 (0.55-7.75), respectively.

The tendency of HLA-SE to associate with IA-development in ACPA-negative patients disappeared in sensitivity analyses with the outcome RA, in contrast to the effect that remained within ACPA-positive patients (Supplementary Figure 3). Hence, HLA-SE was not convincingly associated with progression from symptoms to IA in ACPA-negative patients.

Smoking did also not associate with progression to IA in ACPA-negative patients (HR 1.30 (0.73-2.33)), Figure 3C, Supplementary Table 5.

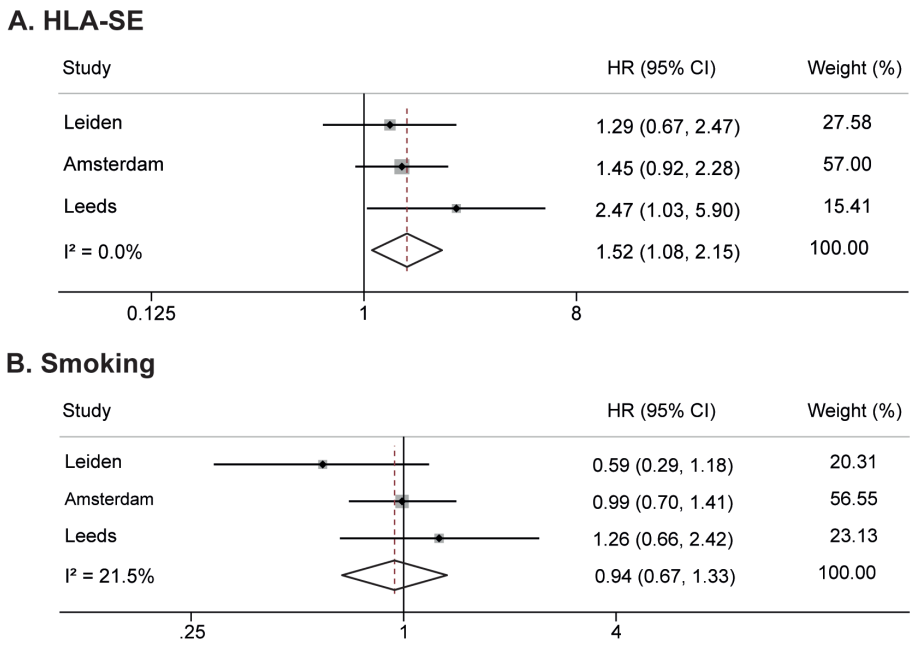
Associations of HLA-SE and smoking with anti-CarP and AAPA at CSA-onset

Neither HLA-SE positivity nor smoking was associated with a higher frequency of RF, anti-CarP or AAPA at presentation with CSA, both in univariable analyses and after correction for concomitant presence of ACPA (Supplementary Table 6).

Associations of anti-CarP and AAPA with IA-development

In univariable analyses, anti-CarP and AAPA were associated with IA-development (Table 2). Correcting for ACPA and RF in the Leiden cohort revealed that AAPA was significantly associated with RA-development, but anti-CarP was not. Similar multivariable analyses were not possible in the Amsterdam cohort because of the lack of an autoantibody-negative reference group. Instead, we studied the association of both AMPA's in the ACPA+/RF+ subgroups. Meta-analyses of data from the two cohorts revealed a significant association for AAPA (HR 1.53 (1.02-2.28)), but not for anti-CarP (HR 1.29 (0.85-1.97), Figure 5).

Figure 4. Meta-analyses on HLA-SE (A) and smoking (B) in three cohorts of ACPA-positive arthralgia patients, showing an association with clinical arthritis development for HLA-SE but not for smoking



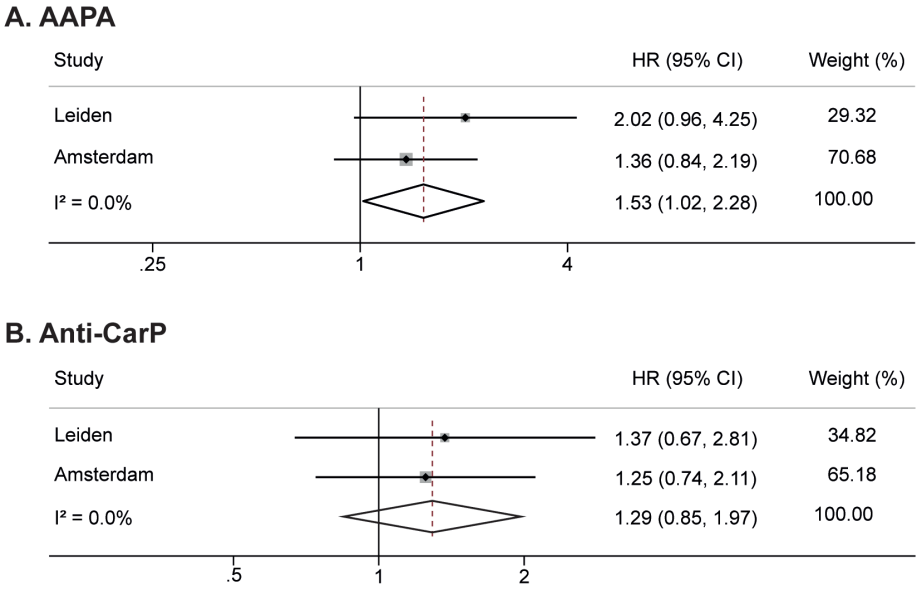
Raw data from ACPA-positive patients from the Amsterdam cohort as described by van de Stadt et al. were obtained and analysed. Results from the Leeds cohort were obtained from Rakieh et al. (Table 2 from reference[30]).
HLA-SE: shared epitope, ACPA: anti-citrullinated protein antibody, HR: hazard ratio, CI: confidence interval, CSA: clinically suspect arthralgia

Table 2. Associations of autoantibodies with development of inflammatory arthritis in patients newly presenting with arthralgia

	Univariable cox regression		**Multivariable cox regression**		**Multivariable cox regression**	
CSA cohort	*HR (95% CI)*	*P-value*	*HR (95% CI)*	*P-value*	*HR (95% CI)*	*P-value*
ACPA IgG	3.29 (2.11-5.13)	<0.001	2.55 (1.44-4.53)	0.001	2.97 (1.73-5.10)	<0.001
RF IgM	1.72 (1.11-2.67)	0.015	1.01 (0.61-1.69)	0.96	0.98 (0.58-1.67)	0.95
AAPA IgG	3.07 (1.90-4.98)	<0.001	1.79 (1.02-3.16)	0.043	--	--
Anti-CarP IgG	2.85 (1.59-5.11)	<0.001	--	--	1.47 (0.75-2.87)	0.26

ACPA: anti-citrullinated protein antibody, RF: rheumatoid factor, AAPA: anti-acetylated protein antibody, anti-CarP: anti-carbamylated protein antibody, HR: hazard ratio, CI: confidence interval

Figure 5. Meta-analyses on AAPA (A) and anti-Carp (B) in two cohorts of ACPA-positive/RF-positive arthralgia patients, showing an association with IA-development for AAPA but not for anti-CarP

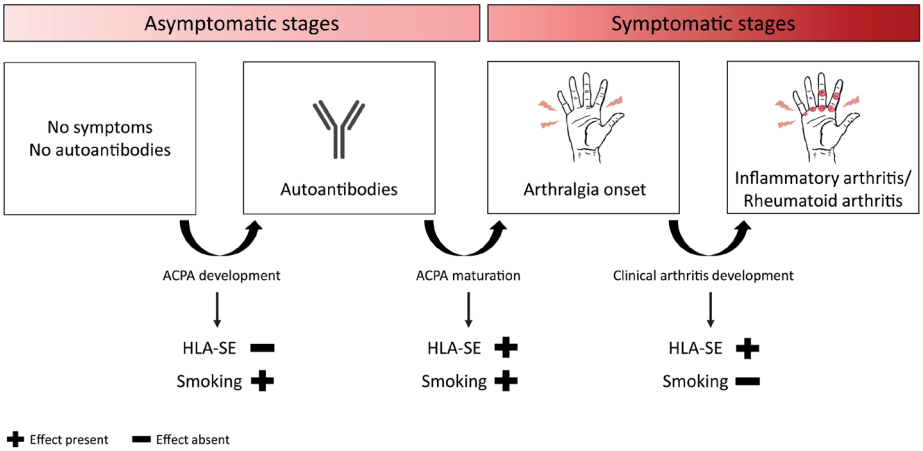


Raw data from ACPA-positive patients from the Amsterdam cohort as described by van de Stadt et al. were obtained and analysed.
AAPA: anti-acetylated protein antibody, anti-CarP: anti-carbamylated protein antibody, RF: rheumatoid factor, ACPA: anti-citrullinated protein antibody, HR: hazard ratio, CI: confidence interval, CSA: clinically suspect arthralgia

Discussion

Although it has been extensively shown that HLA-SE and smoking are risk factors for RA, it was thus far unclear in which pre-arthritis stage these factors exert their effect. We aimed to fine-stage the effects of HLA-SE and smoking, taking advantage of our own cohort data, as well as published data. Results from meta-analyses in people in the asymptomatic stage indicated that smoking, but not HLA-SE, is involved in development of ACPA. At CSA onset, both HLA-SE and smoking were associated with the presence of ACPA, although only HLA-SE associated with progression towards arthritis and RA. Presuming that autoantibody-development as a proxy for the emerging autoimmune response, is the first event, these results imply that smoking is involved in autoantibody-development and possibly symptom-development, but not with further IA-development. In contrast, HLA-SE is not involved in initial autoantibody-development, but rather associated with autoantibody-maturation and symptom-development as implied by results found at CSA-onset. Furthermore, it associates with further progression to clinical disease (Figure 6).

Figure 6. Summary of results on the role of HLA-SE and smoking in the asymptomatic and symptomatic phase of rheumatoid arthritis development



Meta-analyses in the asymptomatic stage indicated that smoking, but not HLA-SE, is involved in development of ACPA. At CSA onset, both HLA-SE and smoking were associated with presence of ACPA. Only HLA-SE further stimulated progression towards arthritis and ACPA-positive RA. Together these data imply that smoking is involved in autoantibody- and symptom-development, HLA-SE plays a role in autoantibody-maturation, symptom-development and progression to clinical disease.

To evaluate the role of HLA-SE and smoking in the asymptomatic phase we reviewed the literature following PRISMA guidelines for systematic literature reviews as much as possible (Supplementary File 1).[37] The results of identified studies performed in asymptomatic populations were combined in meta-analyses. These revealed an effect for smoking and absence of an association of HLA-SE with ACPA-positivity. Recent data in RA-patients indicated that smoking does not associate with ACPA as such, but rather with RF or autoantibodies in general.[6,15,16,38,39] Although not all of the studies included in the meta-analyses contained data on RF, pooled analysis did not identify an association between smoking and RF in the asymptomatic population (supplementary material). Also in CSA-patients no association between RF and smoking was found. All included studies were cross-sectionally performed in the general population. As we presumed that ACPA-positivity is the first event in the development of ACPA-positive RA, we believe the observed findings reflect effects of HLA-SE and smoking on autoantibody-development.

For smoking an association with ACPA was found at the asymptomatic stage and at CSA-onset. Our analyses at CSA-onset were cross-sectional in nature; therefore we cannot definitely conclude whether smoking truly associates with progression from autoantibody-positivity to symptom-development (alternatively, the association found at CSA-onset could be reflective of the association with ACPA-development). However, one longitudinal study evaluated ACPA-positive individuals from the

general population until development of CSA and showed a significant association of smoking with CSA-development.[29] Together with our data this suggests that smoking plays a role in the development of ACPA, further maturation and symptom-development.

The absence of an association of HLA-SE with ACPA in the asymptomatic population, the presence of this association at CSA-onset, and the finding that ACPA-levels tended to be higher in HLA-SE positive CSA-patients (which is in line with a previous study on ACPA-levels in arthralgia[40]), suggest that HLA-SE associates with maturation of the ACPA-response and/or symptom-onset. However, the latter implication is based on deductions from cross-sectional data, longitudinal data from ACPA-positivity to symptom-onset would have been preferable.

Several nested case-control studies have shown that autoantibody-development and the increase in levels can occur years before disease-onset.[41-43] The current study and previous studies on CSA showed that the period between CSA-onset and clinical arthritis development is on average 4-6 months.[44] We recently showed that the autoantibody-response had already matured at CSA-onset and did not mature further towards RA-development.[45] Together these results indicate that autoantibody-response maturation took place before symptom-onset, and was influenced by smoking and HLA-SE. However, although case-control studies have found gene-environment interactions,[6,9,10,14] we found no statistically signification interaction between HLA-SE and smoking for presence of ACPA at CSA-onset (p=0.52). Interestingly, in the asymptomatic phase ACPA-positivity can serorevert to negativity, as is shown in symptom-free relatives of RA-patients.[23] This is in contrast to what is described in the symptomatic phases of CSA and clinical RA,[45-48] where autoantibody-status and -levels were shown to be stable and seroreversion was infrequent. Regarding time-lines, this suggests that the autoimmune response is no longer reversible at symptom-onset. However, disease chronicity is then not yet established; only a proportion of CSA-patients develop RA, and both joint symptoms and subclinical inflammation can resolve spontaneously, also in ACPA-positive patients.[49] The final processes resulting in irreversible ACPA-positive RA remain to be elucidated. However the current data also suggest that this final step is influenced by HLA-SE.

This is not the first longitudinal study on HLA-and smoking and the progression from arthralgia to clinical arthritis. We took advantage of existing data to strengthen the findings and show consistency in the ACPA-positive group. Furthermore, the fact that the Leiden CSA-cohort included patients based on the clinical phenotype and not on autoantibody-status, ensured inclusion of also autoantibody-negative CSA-patients. This served to explore the role of HLA-SE and smoking in ACPA-negative RA.

Although, HLA-SE seemed to promote IA-development in ACPA-negative patients; this effect was not present for RA-development as outcome. Large case-control studies have suggested a role for HLA-SE also in ACPA-negative RA albeit with a smaller effect size than in ACPA-positive RA.[50] The present longitudinal data on ACPA-negative IA- or RA-development were insufficient to support a role for HLA-SE in the symptomatic pre-RA stage.

This study focused on associations of ACPA as measured with anti-CCP2, associations with other ACPA-tests (e.g. anti-CCP3) were not studied. However, in addition to ACPA, we did evaluate other AMPA's. Although different studies have shown cross-reactivity between ACPA and other AMPA's.[51,52] associations with HLA-SE and smoking at CSA-onset seemed to be specific for ACPA as no such associations were found for AAPA and anti-CarP in our patient population. This is in line with findings in RA, where anti-CarP was also not associated with HLA-SE and smoking.[53]

We aimed to fine-stage the effects of HLA-SE and smoking. Identification of predictive markers for IA- or RA-development in CSA was not our primary aim. Nonetheless, we included an exploration and observed that AAPA, but not anti-CarP, associated with IA, independent of ACPA and RF. Further research is needed to ascertain the diagnostic value of these autoantibodies, especially their relevance on top of ACPA and RF that are measured in daily practice.

This study has extended knowledge on the timing of HLA-SE and smoking in the different stages of RA-development. Intriguingly, HLA-SE and smoking exert their effect in partly different phases. Although requiring further biological exploration, it is tempting to speculate that initial autoantibody-development is stimulated by smoking, whereas further expansion of the autoimmune response is promoted differently; by an HLA-SE-restricted T-cell reaction, that drives further ACPA-response maturation. As such, smoking may contribute to development of autoantibodies in general.[6,15,16,38,39] This initial antibody-development does, most likely, require T-cell help as the antibodies are of the IgG isotype and hence the antibody producing B-cells have undergone isotype-switching, a T-cell dependent process. However, as no association with the HLA-system is observed at this stage, these T-cells most likely act in a HLA-SE-independent manner. In contrast, the subsequent expansion of the ACPA-response does associate with HLA-SE, indicating that another, second, T-cell response is involved in the further expansion of the ACPA-response. These T-cells are associated with HLA-SE and, conceivably, recognize other antigens than the ones involved in the T-cell response underlying the "initial" ACPA-response. Thereafter, ACPA-positive persons with HLA-SE are particularly prone for further progression towards RA. These insights in timing of environmental and genetic factors support a

further refinement of the SE-hypothesis; the HLA-SE specific T-cell response may not promote the initial break of tolerance to citrullinated-antigens, but rather promotes the expansion of the (already existing) ACPA-response prior to disease-onset. Conceptually, this would explain why ACPA-positive patients with HLA-SE develop RA more often than ACPA-positive patients without HLA-SE, and why HLA-SE does not associate with the other autoantibodies.

The findings of our study can guide future prevention studies. Prevention often concentrates on health-promoting behaviors. Our results on smoking imply that cessation of smoking might be able to influence the risk of ACPA-development and/or symptom-onset, but also that it may not be effective in reducing the risk of progression from CSA to clinical arthritis. This would mean that trials on smoking cessation might preferably assess the efficacy in disease prevention in the asymptomatic population (primary prevention), rather than in arthralgia-patients (secondary prevention).

To conclude, HLA-SE and smoking act in partly different pre-RA stages. Smoking confers risk for development of ACPA and/or joint symptoms, but does not further associate with IA-development. In contrast, HLA-SE does not associate with ACPA in the general population, but does mediate symptom-development and progression to IA. Even though the underlying time-specific biological pathways need further exploration, these data enhance understanding of timing of key genetic and environmental risk factors in development of RA.

References

1. Linn-Rasker SP, van der Helm-van Mil AH, van Gaalen FA et al. Smoking is a risk factor for anti-CCP antibodies only in rheumatoid arthritis patients who carry HLA-DRB1 shared epitope alleles. Ann Rheum Dis. 2006;65:366-71.

2. Arlestig L, Mullazehi M, Kokkonen H et al. Antibodies against cyclic citrullinated peptides of IgG, IgA and IgM isotype and rheumatoid factor of IgM and IgA isotype are increased in unaffected members of multicase rheumatoid arthritis families from northern Sweden. Ann Rheum Dis. 2012;71:825-9.

3. van der Helm-van Mil AH, Verpoort KN, le Cessie S et al. The HLA-DRB1 shared epitope alleles differ in the interaction with smoking and predisposition to antibodies to cyclic citrullinated peptide. Arthritis Rheum. 2007;56:425-32.

4. Huizinga TW, Amos CI, van der Helm-van Mil AH et al. Refining the complex rheumatoid arthritis phenotype based on specificity of the HLA-DRB1 shared epitope for antibodies to citrullinated proteins. Arthritis Rheum. 2005;52:3433-8.

5. Michou L, Teixeira VH, Pierlot C et al. Associations between genetic factors, tobacco smoking and autoantibodies in familial and sporadic rheumatoid arthritis. Ann Rheum Dis. 2008;67:466-70.

6. Hedstrom AK, Ronnelid J, Klareskog L et al. Complex Relationships of Smoking, HLA-DRB1 Genes, and Serologic Profiles in Patients With Early Rheumatoid Arthritis: Update From a Swedish Population-Based Case-Control Study. Arthritis Rheumatol. 2019;71:1504-11.

7. Lee HS, Irigoyen P, Kern M et al. Interaction between smoking, the shared epitope, and anti-cyclic citrullinated peptide: a mixed picture in three large North American rheumatoid arthritis cohorts. Arthritis Rheum. 2007;56:1745-53.

8. Pedersen M, Jacobsen S, Garred P et al. Strong combined gene-environment effects in anti-cyclic citrullinated peptide-positive rheumatoid arthritis: a nationwide case-control study in Denmark. Arthritis Rheum. 2007;56:1446-53.

9. Too CL, Yahya A, Murad S et al. Smoking interacts with HLA-DRB1 shared epitope in the development of anti-citrullinated protein antibody-positive rheumatoid arthritis: results from the Malaysian Epidemiological Investigation of Rheumatoid Arthritis (MyEIRA). Arthritis Res Ther. 2012;14:R89.

10. Lee YH, Bae SC, Song GG. Gene-environmental interaction between smoking and shared epitope on the development of anti-cyclic citrullinated peptide antibodies in rheumatoid arthritis: a meta-analysis. Int J Rheum Dis. 2014;17:528-35.

11. Svard A, Skogh T, Alfredsson L et al. Associations with smoking and shared epitope differ between IgA- and IgG-class antibodies to cyclic citrullinated peptides in early rheumatoid arthritis. Arthritis Rheumatol. 2015;67:2032-7.

12. Morgan AW, Thomson W, Martin SG et al. Reevaluation of the interaction between HLA-DRB1 shared epitope alleles, PTPN22, and smoking in determining susceptibility to autoantibody-positive and autoantibody-negative rheumatoid arthritis in a large UK Caucasian population. Arthritis Rheum. 2009;60:2565-76.

13. Gregersen PK, Silver J, Winchester RJ. The shared epitope hypothesis. An approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis. Arthritis Rheum. 1987;30:1205-13.

14. Klareskog L, Stolt P, Lundberg K et al. A new model for an etiology of rheumatoid arthritis: smoking may trigger HLA-DR (shared epitope)-restricted immune reactions to autoantigens modified by citrullination. Arthritis Rheum. 2006;54:38-46.

15. van Wesemael TJ, Ajeganova S, Humphreys J et al. Smoking is associated with the concurrent presence of multiple autoantibodies in rheumatoid arthritis rather than with anti-citrullinated protein antibodies per se: a multicenter cohort study. Arthritis Res Ther. 2016;18:285.

16. Ishikawa Y, Ikari K, Hashimoto M et al. Shared epitope defines distinct associations of cigarette smoking with levels of anticitrullinated protein antibody and rheumatoid factor. Ann Rheum Dis.

2019;78:1480-7.

17. 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.

18. Raza K, Holers VM, Gerlag D. Nomenclature for the Phases of the Development of Rheumatoid Arthritis. Clin Ther. 2019;41:1279-85.

19. Kokkonen H, Brink M, Hansson M et al. Associations of antibodies against citrullinated peptides with human leukocyte antigen-shared epitope and smoking prior to the development of rheumatoid arthritis. Arthritis Res Ther. 2015;17:125.

20. Lahiri M, Luben RN, Morgan C et al. Using lifestyle factors to identify individuals at higher risk of inflammatory polyarthritis (results from the European Prospective Investigation of Cancer-Norfolk and the Norfolk Arthritis Register--the EPIC-2-NOAR Study). Ann Rheum Dis. 2014;73:219-26.

21. Ford JA, Liu X, Marshall AA et al. Impact of Cyclic Citrullinated Peptide Antibody Level on Progression to Rheumatoid Arthritis in Clinically Tested Cyclic Citrullinated Peptide Antibody-Positive Patients Without Rheumatoid Arthritis. Arthritis Care Res (Hoboken). 2019;71:1583-92.

22. de Hair MJ, Landewe RB, van de Sande MG et al. Smoking and overweight determine the likelihood of developing rheumatoid arthritis. Ann Rheum Dis. 2013;72:1654-8.

23. Tanner S, Dufault B, Smolik I et al. A Prospective Study of the Development of Inflammatory Arthritis in the Family Members of Indigenous North American People With Rheumatoid Arthritis. Arthritis Rheumatol. 2019;71:1494-503.

24. Sparks JA, Chang SC, Deane KD et al. Associations of Smoking and Age With Inflammatory Joint Signs Among Unaffected First-Degree Relatives of Rheumatoid Arthritis Patients: Results From Studies of the Etiology of Rheumatoid Arthritis. Arthritis Rheumatol. 2016;68:1828-38.

25. Terao C, Ohmura K, Ikari K et al. Effects of smoking and shared epitope on the production of anti-citrullinated peptide antibody in a Japanese adult population. Arthritis Care Res (Hoboken). 2014;66:1818-27.

26. van Zanten A, Arends S, Roozendaal C et al. Presence of anticitrullinated protein antibodies in a large population-based cohort from the Netherlands. Ann Rheum Dis. 2017;76:1184-90.

27. Hensvold AH, Magnusson PK, Joshua V et al. Environmental and genetic factors in the development of anticitrullinated protein antibodies (ACPAs) and ACPA-positive rheumatoid arthritis: an epidemiological investigation in twins. Ann Rheum Dis. 2015;74:375-80.

28. El-Gabalawy HS, Robinson DB, Hart D et al. Immunogenetic risks of anti-cyclical citrullinated peptide antibodies in a North American Native population with rheumatoid arthritis and their first-degree relatives. J Rheumatol. 2009;36:1130-5.

29. Westra J, Brouwer E, Raveling-Eelsing E et al. Arthritis autoantibodies in individuals without rheumatoid arthritis: follow-up data from a Dutch population-based cohort (Lifelines). Rheumatology (Oxford). 2020.

30. Rakieh C, Nam JL, Hunt L et al. 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.

31. 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.

32. van Steenbergen HW, van Nies JA, Huizinga TW et al. Characterising arthralgia in the preclinical phase of rheumatoid arthritis using MRI. Ann Rheum Dis. 2015;74:1225-32.

33. Jia X, Han B, Onengut-Gumuscu S et al. Imputing amino acid polymorphisms in human leukocyte antigens. PLoS One. 2013;8:e64683.

34. van der Woude D, Lie BA, Lundstrom E et al. Protection against anti-citrullinated protein antibody-positive rheumatoid arthritis is predominantly associated with HLA-DRB1*1301: a meta-analysis of HLA-DRB1 associations with anti-citrullinated protein antibody-positive and anti-citrullinated

protein antibody-negative rheumatoid arthritis in four European populations. Arthritis Rheum. 2010;62:1236-45.

35. 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.

36. 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.

37. Moher D, Liberati A, Tetzlaff J et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Int J Surg. 2010;8:336-41.

38. Murphy D, Mattey D, Hutchinson D. Anti-citrullinated protein antibody positive rheumatoid arthritis is primarily determined by rheumatoid factor titre and the shared epitope rather than smoking per se. PLoS One. 2017;12:e0180655.

39. Regueiro C, Rodriguez-Rodriguez L, Lopez-Mejias R et al. A predominant involvement of the triple seropositive patients and others with rheumatoid factor in the association of smoking with rheumatoid arthritis. Sci Rep. 2020;10:3355.

40. Bos WH, Ursum J, de Vries N et al. The role of the shared epitope in arthralgia with anti-cyclic citrullinated peptide antibodies (anti-CCP), and its effect on anti-CCP levels. Ann Rheum Dis. 2008;67:1347-50.

41. Rantapaa-Dahlqvist S, de Jong BA, Berglin E et al. Antibodies against cyclic citrullinated peptide and IgA rheumatoid factor predict the development of rheumatoid arthritis. Arthritis Rheum. 2003;48:2741-9.

42. Nielen MM, van Schaardenburg D, Reesink HW et al. Specific autoantibodies precede the symptoms of rheumatoid arthritis: a study of serial measurements in blood donors. Arthritis Rheum. 2004;50:380-6.

43. Gan RW, Trouw LA, Shi J et al. Anti-carbamylated protein antibodies are present prior to rheumatoid arthritis and are associated with its future diagnosis. J Rheumatol. 2015;42:572-9.

44. van Steenbergen HW, Mangnus L, Reijnierse M et al. Clinical factors, anticitrullinated peptide antibodies and MRI-detected subclinical inflammation in relation to progression from clinically suspect arthralgia to arthritis. Ann Rheum Dis. 2016;75:1824-30.

45. Wouters F, Niemantsverdriet E, Salioska N et al. Do autoantibody-responses mature between presentation with arthralgia suspicious for progression to rheumatoid arthritis and development of clinically apparent inflammatory arthritis? A longitudinal serological study. Ann Rheum Dis. 2020.

46. Ten Brinck RM, van Steenbergen HW, van Delft MAM et al. The risk of individual autoantibodies, autoantibody combinations and levels for arthritis development in clinically suspect arthralgia. Rheumatology (Oxford). 2017;56:2145-53.

47. Boeters DM, Burgers LE, Toes RE et al. Does immunological remission, defined as disappearance of autoantibodies, occur with current treatment strategies? A long-term follow-up study in rheumatoid arthritis patients who achieved sustained DMARD-free status. Ann Rheum Dis. 2019;78:1497-504.

48. Ursum J, Bos WH, van de Stadt RJ et al. Different properties of ACPA and IgM-RF derived from a large dataset: further evidence of two distinct autoantibody systems. Arthritis Res Ther. 2009;11:R75.

49. Ten Brinck RM, Boeters DM, van Steenbergen HW et al. Improvement of symptoms in clinically suspect arthralgia and resolution of subclinical joint inflammation: a longitudinal study in patients that did not progress to clinical arthritis. Arthritis Res Ther. 2020;22:11.

50. Viatte S, Plant D, Bowes J et al. Genetic markers of rheumatoid arthritis susceptibility in anti-citrullinated peptide antibody negative patients. Ann Rheum Dis. 2012;71:1984-90.

51. Kissel T, Reijm S, Slot LM et al. Antibodies and B cells recognising citrullinated proteins display a broad cross-reactivity towards other post-translational modifications. Ann Rheum Dis. 2020;79:472-80.

52. Sahlstrom P, Hansson M, Steen J et al. Different Hierarchies of Anti-Modified Protein Autoantibody Reactivities in Rheumatoid Arthritis. Arthritis Rheumatol. 2020.

53. Jiang X, Trouw LA, van Wesemael TJ et al. Anti-CarP antibodies in two large cohorts of patients with rheumatoid arthritis and their relationship to genetic risk factors, cigarette smoking and other autoantibodies. Ann Rheum Dis. 2014;73:1761-8.

Supplementary File 1 – Literature and meta-analyses

A literature search was performed to investigate associations between HLA-SE, smoking and anti-citrullinated protein antibodies (ACPA) in the asymptomatic general population and first-degree relatives (FDR). See Supplementary Figure 1 for a flowchart of the article selection. PubMed was searched until September 2020 with terms as: "SE"[All Fields] AND ("smoke"[MeSH Terms] OR "smoke"[All Fields] OR "smoke s"[All Fields] OR "smoked"[All Fields] OR "smokes"[All Fields] OR "smoking"[MeSH Terms] OR "smoking"[All Fields] OR "smokings"[All Fields] OR "smoking s"[All Fields]) AND ("autoantibodies"[MeSH Terms] OR "autoantibodies"[All Fields] OR "autoantibody"[All Fields]) AND ("healthies"[All Fields] OR "healthy"[All Fields]), including different combinations of SE, shared epitope, smoking, autoantibodies, ACPA, healthy, asymptomatic, preclinical, general population, population-based and FDR (see below for all combinations used). Additional articles were identified by hand searching reference lists. After removal of reviews, duplicates and articles that did not apply to our research question (based on title and abstract screening), ten articles remained. Subsequently all studies that were either cross-sectional or longitudinal studies and contained information on associations between HLA-SE, smoking and ACPA-development were considered eligible. No longitudinal studies starting the healthy population were identified; all identified studies were cross-sectional in nature. However, since ACPA-development is most likely the first event in development of ACPA-positive disease, we believe the observed findings in cross-sectional studies reflect effects of HLA-SE and smoking on ACPA-development. The ten selected articles were further studied on the description of the population. Studies evaluating populations in which part of subjects had swollen and/or tender joints, and studies in which associations with autoantibodies were not investigated for ACPA (separate from other autoantibodies) were excluded. Finally, six cross-sectional studies were eligible for meta-analyses. Four studies evaluated associations between HLA-SE and ACPA,[1-4], and five studies evaluated associations between smoking and ACPA.[1,2,4-6] Meta-analyses were first performed without a stringent evaluation of study bias and thereafter with assessing this bias as recommended according to the PRISMA guidelines.[7] First, the odds ratios (OR) of these four and five studies respectively were combined in inverse-variance weighted meta-analyses, I^2 was determined to evaluate heterogeneity across the studies. Some studies were truly population based (no selection), these were the studies of van Zanten et al. (2017) and Terao et al. (2014). Other studies included asymptomatic persons whom were selected because of having a relative with RA,[3,4] being a twin,[2] or were blood bank donors.[6] In the latter no odds ratio was given on the association of smoking and ACPA in healthy individuals, however, since all required information was available we calculated the odds ratio from the data presented in the publication. In Terao et al. (2014) results were stratified

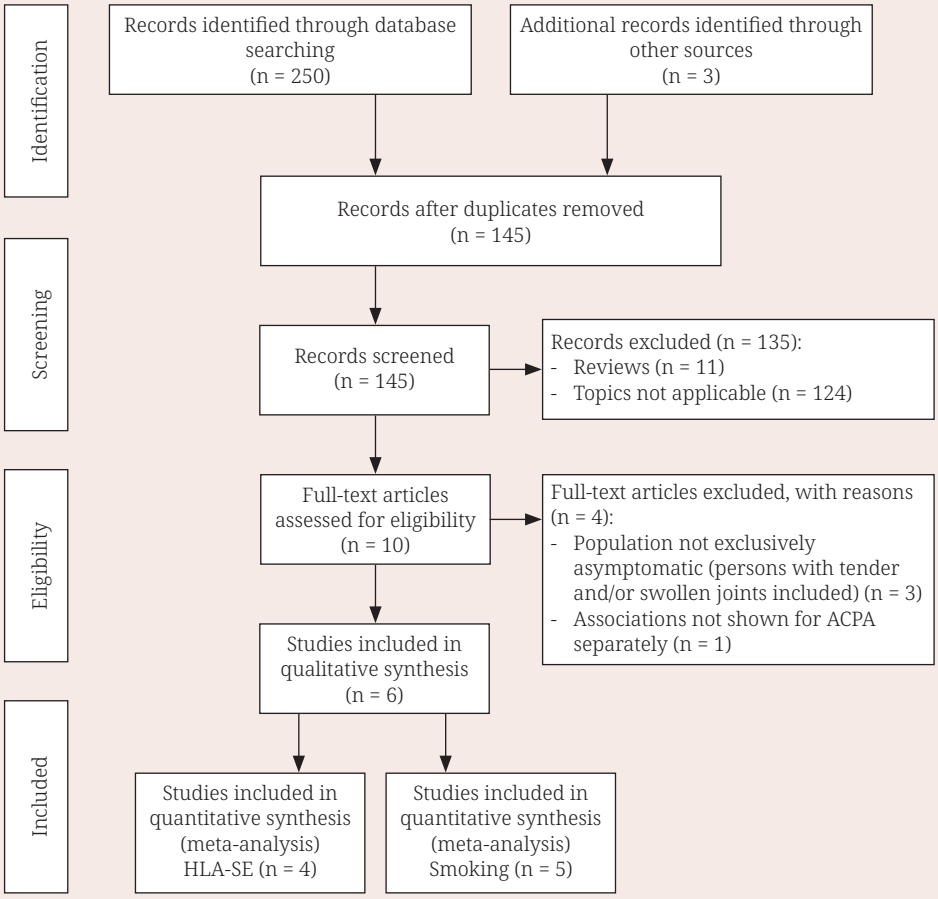
for gender; for the present meta-analyses we included results presented for women, however, results from the meta-analysis was similar when men were included in the analyses (data not shown).

Next, to eliminate possible inclusion bias, only the two studies that evaluated unselected individuals from the general population (van Zanten et al. (2017) and Terao et al. (2014)) were included in the meta-analysis. Then similar results were obtained, see Supplementary Figure 2.

Finally, since recent data in RA-patients indicated that smoking does not associate with ACPA as such, but rather with RF or autoantibodies in general,[8-12] and we wished to evaluate whether this is the same in the general population, we retrieved RF data, if present, from the identified studies, and performed additional meta-analyses. Two studies (Terao et al. (2014) and Ärlestig et al. (2011)) provided data on associations of HLA-SE and smoking with RF. Meta-analyses indicated that was no association was present between HLA-SE and RF (OR 0.97 (95% CI 0.70-1.35)), nor between smoking and RF (OR 0.84 (0.55-1.28); forest plots not shown).

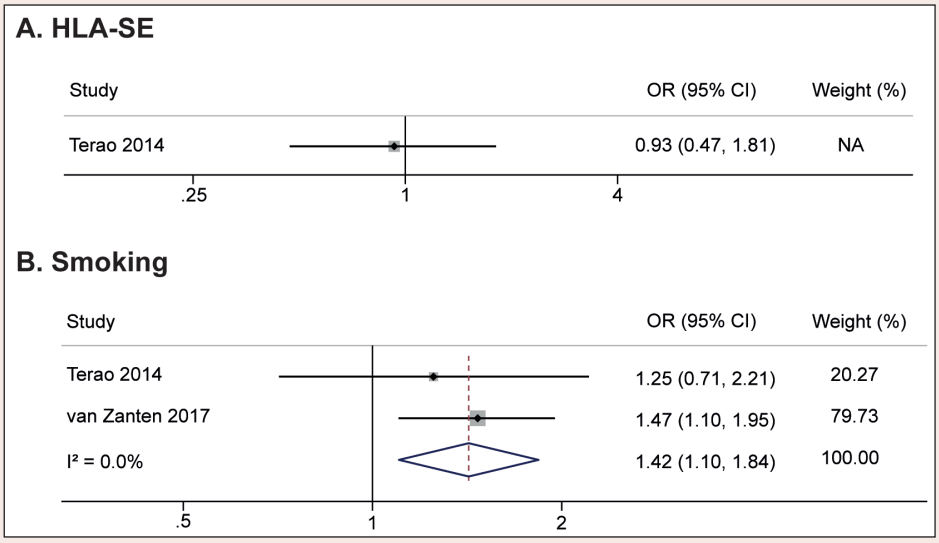
List of search terms that were used:
- SE smoking autoantibodies healthy (15 results)
- SE smoking autoantibodies asymptomatic (0 results)
- SE smoking autoantibodies preclinical (0 results)
- SE smoking ACPA healthy (8 results)
- SE ACPA healthy (32 results)
- shared epitope ACPA healthy (33 results)
- smoking ACPA healthy (32 results)
- smoking ACPA general population (16 results)
- shared epitope ACPA general population (10 results)
- smoking ACPA population-based (25 results)
- shared epitope ACPA population-based (15 results)
- smoking ACPA FDR (2 results)
- SE ACPA FDR (2 results)
- shared epitope ACPA FDR (2 results)
- SE smoking ACPA (58 results)

Supplementary Figure 1. Flowchart of literature review and article selection for meta-analyses on association of HLA-SE and smoking in the asymptomatic population.



Meta-analyses were performed in the 4 (HLA-SE) en 5 (smoking) included studies, respectively, without assessment of inclusion bias. Some of these studies did not select persons from the general asymptomatic population. Therefore sensitivity analyses with the population-based studies were performed, see Supplementary Figure 2.

Supplementary Figure 2. Meta-analyses on HLA-SE (A) and smoking (B) with presence of ACPA in unselected asymptomatic healthy population-based individuals



HLA-SE: shared epitope, ACPA: anti-citrullinated protein antibody, OR: odds ratio, CI: confidence interval

Supplementary File 2 – Detailed description of methods

Leiden CSA cohort

Patients presenting with CSA to the Leiden rheumatology outpatient clinic between April 2012-September 2019 were studied. The cohort is described in detail previously.[13] Patients had recent-onset (<1 year) arthralgia of small joints and were, according to the clinical expertise and pattern recognition of the rheumatologist, at risk for progression to RA. Autoantibody status was largely unknown at inclusion as (in line with Dutch guidelines) general practitioners in the area of Leiden are discouraged to perform autoantibody tests. Inclusion in the CSA-cohort was therefore predominantly based on history taking and physical examination. Patients were excluded if clinical arthritis was already present, or if a different explanation for the joint pain was more likely. Patients were followed for at least 2 years for development of clinically apparent inflammatory arthritis (IA), defined as ≥1 swollen joints at physical examination by a rheumatologist, with scheduled follow-up visits at 4, 12 and 24 months. In case of an increase in symptoms or suspicion of clinical arthritis additional visits were planned. During follow-up treatment with disease-modifying anti-rheumatic drugs (DMARDs, including systemic or intra-articular corticosteroids) was not allowed, therefore patients participating in a randomized controlled trial investigating the effect of methotrexate in the phase of CSA (RCT; treat earlier) were excluded from the current study.

Baseline visit consisted of physical examination, blood sampling, questionnaires (including questions on current and past smoking) and an MRI. Presence of IgM RF (in house ELISA, cut-off >3.5 IU/mL) and IgG ACPA (anti-CCP2, Phadia, Nieuwegein, the Netherlands, cut-off >7 U/mL) was determined during routine laboratory measurements in all patients. Because of limited laboratory capacity not all patients were selected for additional anti-CarP and AAPA measurements. Patient selection for these tests was first based on availability of samples and presence of RF and/or ACPA at baseline. In total 189 patients were included for additional autoantibody measurements. First, patients positive for RF and/or ACPA at baseline with available serum samples were included (n=89, 45 progressing and 44 non-progressing patients). Then, autoantibody-negative patients with available samples that progressed to IA were additionally included (n=37). Finally, from the large group of autoantibody-negative patients that did not progress a random selection was made (n=63). Baseline characteristics of the randomly selected non-progressing autoantibody-negative patients were similar to that of the patients that were not selected (data not shown); suggesting that the selection is representative for this total group. The fact that anti-CarP and AAPA were only measured in a selection of autoantibody-negative patients

limits interpretation of predictive accuracy of these autoantibodies in the entire CSA-population. However, since no selection was made in RF-positive/ACPA-positive patients, analyses were repeated and validated with the Amsterdam cohort in this subgroup.

The study was approved by the local medical ethical committee and all participants gave written informed consent.

HLA-SE

HLA-SE status was derived from whole genome sequencing data. All patients were whole genome sequenced using the Illumina Global Screening Array (GSA). Standard quality control steps were performed using Plink v1.90;[14] individuals with a missingness of more than 2% were removed, as were variants with a missingness of over 2%, a minor allele frequency (MAF) below 0.01 or a Hardy-Weinberg equilibrium p-value below 0.0001. In order to prepare the genotype data for imputation of the HLA region we used the McCarthy program,[15] which checks for ambiguous SNPs with a MAF above 0.4 (only A/T & G/C SNPs), SNPs with a MAF deviating more than 0.2 from the HRC reference panel[16] and SNPs not in the aforementioned reference panel. We subsequently use the SNP2HLA tool with the T1DGC reference to impute the HLA region.[17] From these imputed data we extracted the HLA-SE, defined as the HLA-DRB1 variants *0101, *0102, *0401, *0404, *0405, *0408 and *1001.[18] Patients were deemed HLA-SE positive if 1 or 2 of the SE variants were present.

Anti-carbamylated and anti-acetylated antibodies

In serum, we determined the presence and levels of IgG anti-carbamylated and anti-acetylated protein antibodies (anti-CarP and AAPA, respectively). In-house ELISA was used for all measurements as described previously.[19] Briefly, plates were coated with carbamylated FCS and CCP1 acetylated lysine for measurements of anti-CarP and AAPA, respectively. To determine background signal, plates were additionally coated with non-modified antigens (non-modified FCS and CCP1 norleucine, respectively). Serum samples were diluted 1:50 and incubated. After washing, plates were subsequently incubated with HRP-labeled rabbit-anti-human IgG (Dako). HRP-activity was visualized with ABTS and measurements were expressed in arbitrary units per milliliter (aU/mL). On every plate a dilution standard was included to determine the linear part of the curve; standards from all plates were used in the analyses. The fourth standard, which is expected to be in the middle (and therefore linear part) of the curve, is further diluted and additionally included as a reference sample. Serum of healthy subjects (n=199) was used to determine the cut-off of all autoantibody measurements, which was calculated as the mean plus two times the standard deviation of healthy subjects. When the difference in optical density

(OD) between the non-modified antigens and the modified proteins was <0.1, the measurement was considered non-specific; non-specific measurements with values above the cut-off were considered negative. Inter-assay variation of in-house ELISAs was determined previously by reevaluation of ~10% of samples; measurements were highly correlated (Pearson's r ranges 0.96-0.97) and changes in positivity of the test were infrequent, see Supplementary Figure 4. Intra-assay variability was determined for anti-CarP by measurement of 3 samples 10 times, the mean coefficient of variation (CV, mean % (SD)) was 20.4 (6.8). Of note, although not absolute at the monoclonal- or polyclonal level, cross-reactivity of ACPA towards other post translationally modified proteins have been conclusively shown in different studies,[20,21] and hence should be regarded as anti-modified protein antibody-reactivities.

Amsterdam arthralgia cohort

Analyses evaluating progression to IA were stratified for ACPA-status and results from the ACPA-positive subgroup were studied in meta-analyses with the results from two other ACPA-positive cohorts. The first cohort concerned arthralgia patients from Amsterdam, as described in van de Stadt et al.[22] As this study evaluated ACPA- and/or RF-positive patients, raw data from ACPA-positive individuals was obtained for this study. Data on smoking was obtained by history taking. Presence of HLA-SE was inferred from HLA-DQA1, HLA-DQB1 haplotypes using strong linkage disequilibrium with HLA-DRB1 alleles in Caucasians.[23] Patients were positive for HLA-SE if 1 or 2 copies of the HLA-DRB1*0101, *0102, *0401, *0404, *0405, *0408, *0410 or *1001 alleles were present. Presence of IgM RF and ACPA were determined with in-house ELISA (cut-off >30 IU/mL) and aCCP ELISA (Axis Shield; cut-off >5 aU/mL), respectively. Presence of anti-CarP was determined with in-house ELISA similar to methods used in the CSA cohort.[19] Baseline serum samples were obtained to determine presence of IgG AAPA simultaneously with Leiden CSA-samples.

Leeds ACPA-positive cohort

The second cohort consists of ACPA-positive patients with non-specific musculoskeletal symptoms from the Leeds cohort, as described by Rakieh et al.[24] Since only ACPA-positive patients were studied, results obtained from the literature were sufficient for inclusion in the meta-analyses; raw data was not evaluated. Also here smoking status was obtained through history taking. Presence of HLA-SE was determined by PCR amplification with sequence-specific primers.[25] Patients were positive for HLA-SE if 1 or 2 copies of alleles in the HLA-DRB1*01, *04 and *10 locus were present. Presence of IgM RF (initial cut-off >40 IU/mL, later >20 IU/mL) and ACPA (anti-CCP2, immunocap assay, Phadia; initial cut-off >7 IU/mL, later >2.99 IU/mL) was determined. No serum samples were obtained from this cohort and the presence of anti-CarP and AAPA was not determined.

References

1. Terao C, Ohmura K, Ikari K et al. Effects of smoking and shared epitope on the production of anti-citrullinated peptide antibody in a Japanese adult population. Arthritis Care Res (Hoboken). 2014;66:1818-27.

2. Hensvold AH, Magnusson PK, Joshua V et al. Environmental and genetic factors in the development of anticitrullinated protein antibodies (ACPAs) and ACPA-positive rheumatoid arthritis: an epidemiological investigation in twins. Ann Rheum Dis. 2015;74:375-80.

3. El-Gabalawy HS, Robinson DB, Hart D et al. Immunogenetic risks of anti-cyclical citrullinated peptide antibodies in a North American Native population with rheumatoid arthritis and their first-degree relatives. J Rheumatol. 2009;36:1130-5.

4. Arlestig L, Mullazehi M, Kokkonen H et al. Antibodies against cyclic citrullinated peptides of IgG, IgA and IgM isotype and rheumatoid factor of IgM and IgA isotype are increased in unaffected members of multicase rheumatoid arthritis families from northern Sweden. Ann Rheum Dis. 2012;71:825-9.

5. van Zanten A, Arends S, Roozendaal C et al. Presence of anticitrullinated protein antibodies in a large population-based cohort from the Netherlands. Ann Rheum Dis. 2017;76:1184-90.

6. Klareskog L, Stolt P, Lundberg K et al. A new model for an etiology of rheumatoid arthritis: smoking may trigger HLA-DR (shared epitope)-restricted immune reactions to autoantigens modified by citrullination. Arthritis Rheum. 2006;54:38-46.

7. Moher D, Liberati A, Tetzlaff J et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Int J Surg. 2010;8:336-41.

8. van Wesemael TJ, Ajeganova S, Humphreys J et al. Smoking is associated with the concurrent presence of multiple autoantibodies in rheumatoid arthritis rather than with anti-citrullinated protein antibodies per se: a multicenter cohort study. Arthritis Res Ther. 2016;18:285.

9. Ishikawa Y, Ikari K, Hashimoto M et al. Shared epitope defines distinct associations of cigarette smoking with levels of anticitrullinated protein antibody and rheumatoid factor. Ann Rheum Dis. 2019;78:1480-7.

10. Murphy D, Mattey D, Hutchinson D. Anti-citrullinated protein antibody positive rheumatoid arthritis is primarily determined by rheumatoid factor titre and the shared epitope rather than smoking per se. PLoS One. 2017;12:e0180655.

11. Hedstrom AK, Ronnelid J, Klareskog L et al. Complex Relationships of Smoking, HLA-DRB1 Genes, and Serologic Profiles in Patients With Early Rheumatoid Arthritis: Update From a Swedish Population-Based Case-Control Study. Arthritis Rheumatol. 2019;71:1504-11.

12. Regueiro C, Rodriguez-Rodriguez L, Lopez-Mejias R et al. A predominant involvement of the triple seropositive patients and others with rheumatoid factor in the association of smoking with rheumatoid arthritis. Sci Rep. 2020;10:3355.

13. van Steenbergen HW, van Nies JA, Huizinga TW et al. Characterising arthralgia in the preclinical phase of rheumatoid arthritis using MRI. Ann Rheum Dis. 2015;74:1225-32.

14. Plink v1.90. https://www.cog-genomics.org/plink/.

15. McCarthy Program. https://www.well.ox.ac.uk/~wrayner/tools/.

16. HRC reference panel. http://www.haplotype-reference-consortium.org/site.

17. Jia X, Han B, Onengut-Gumuscu S et al. Imputing amino acid polymorphisms in human leukocyte antigens. PLoS One. 2013;8:e64683.

18. van der Woude D, Lie BA, Lundstrom E et al. Protection against anti-citrullinated protein antibody-positive rheumatoid arthritis is predominantly associated with HLA-DRB1*1301: a meta-analysis of HLA-DRB1 associations with anti-citrullinated protein antibody-positive and anti-citrullinated protein antibody-negative rheumatoid arthritis in four European populations.

Arthritis Rheum. 2010;62:1236-45.

19. Shi J, Knevel R, Suwannalai P et al. Autoantibodies recognizing carbamylated proteins are present in sera of patients with rheumatoid arthritis and predict joint damage. Proc Natl Acad Sci U S A. 2011;108:17372-7.

20. Kissel T, Reijm S, Slot LM et al. Antibodies and B cells recognising citrullinated proteins display a broad cross-reactivity towards other post-translational modifications. Ann Rheum Dis. 2020;79:472-80.

21. Sahlstrom P, Hansson M, Steen J et al. Different Hierarchies of Anti-Modified Protein Autoantibody Reactivities in Rheumatoid Arthritis. Arthritis Rheumatol. 2020.

22. 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.

23. Knipper AJ, Hakenberg P, Enczmann J et al. HLA-DRB1,3,4,5 and -DQB1 allele frequencies and HLA-DR/DQ linkage disequilibrium of 231 German caucasoid patients and their corresponding 821 potential unrelated stem cell transplants. Hum Immunol. 2000;61:605-14.

24. Rakieh C, Nam JL, Hunt L et al. 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.

25. Olerup O, Zetterquist H. HLA-DR typing by PCR amplification with sequence-specific primers (PCR-SSP) in 2 hours: an alternative to serological DR typing in clinical practice including donor-recipient matching in cadaveric transplantation. Tissue Antigens. 1992;39:225-35.

Supplementary Table 1. Baseline characteristics

	CSA cohort Leiden			Arthralgia cohort Amsterdam
	All patients (n=577)	*ACPA-positive (n=77)*	*ACPA-negative (n=500)*	*ACPA-positive (n=244)*
Female, n (%)	451 (78.2)	61 (79.2)	390 (78.0)	186 (76.2)
Age in years, mean (SD)	43.7 (12.6)	48.1 (12.0)	43.0 (12.5)	48.4 (11.4)
Symptom duration in weeks, median (IQR)	20 (9-46)	25 (13-53)	19 (9-43)	52 (28-104)
53-TJC, median (IQR)	5 (2-10)	3 (2-7)	5 (2-10)	0 (0-3)
Increased CRP, n (%)	126 (22.0)	23 (30.3)	103 (20.7)	35 (14.5)
RF positivity, n (%)	113 (19.6)	59 (76.6)	54 (10.8)	108 (44.3)
ACPA positivity, n (%)	77 (13.3)	NA	NA	NA

Data on HLA-SE and smoking was available in the CSA cohort in 522 and 500 patients, respectively. In Amsterdam in 133 and 243 patients, respectively.
CSA: clinically suspect arthralgia, ACPA: anti-citrullinated protein antibody, SD: standard deviation, IQR: interquartile range, TJC: tender joint count, RF: rheumatoid factor, CRP: c-reactive protein

Supplementary Table 2. Associations of HLA-SE and smoking with ACPA-positivity stratified for RF, in patients newly presenting with CSA

		ACPA positive, n (%)	ACPA negative, n (%)	OR (95% CI)	p-value
RF-positive patients					
HLA-SE	Absent	22 (42)	32 (64)	Reference	--
	Present	31 (59)	18 (36)	2.51 (1.13-5.55)	0.024
HLA-SE	0	22 (42)	32 (64)	Reference	--
	1	23 (43)	16 (32)	2.09 (0.91-4.83)	0.084
	2	8 (15)	2 (4)	5.82 (1.13-30.05)	0.036
Smoking	Never	13 (26)	23 (49)	Reference	--
	Ever	37 (74)	24 (51)	2.73 (1.16-6.40)	0.021
Smoking	Never	13 (26)	23 (49)	Reference	--
	Ex-smoker	22 (44)	21 (45)	1.85 (0.75-4.58)	0.18
	Current smoker	15 (30)	3 (6)	8.85 (2.15-36.38)	0.003
RF-negative patients					
HLA-SE	Absent	5 (31)	227 (56)	Reference	--
	Present	11 (69)	176 (44)	2.84 (0.97-8.32)	0.057
HLA-SE	0	5 (31)	227 (56)	Reference	--
	1	8 (50)	145 (36)	2.51 (0.80-7.81)	0.11
	2	3 (19)	31 (8)	4.39 (1.0-19.30)	0.050
Smoking	Never	2 (14)	162 (42)	Reference	--
	Ever	12 (86)	227 (58)	4.28 (0.95-19.39)	0.059
Smoking	Never	2 (14)	162 (42)	Reference	--
	Ex-smoker	6 (43)	140 (36)	3.47 (0.69-17.48)	0.13
	Current smoker	6 (43)	87 (22)	5.59 (1.10-28.27)	0.038

HLA-SE: shared epitope, ACPA: anti-citrullinated protein antibody, RF: rheumatoid factor, CSA: clinically suspect arthralgia, OR: odds ratio, CI: confidence interval

Supplementary Table 3. Associations of HLA-SE and smoking with ACPA-level in ACPA-positive patients evaluated with Mann-Whitney U tests and logistic regression

HLA-SE positive	**HLA-SE negative**	**Mann-Whitney U**	**Logistic regression**
Median ACPA level (IQR)	*Median ACPA level (IQR)*	*p-value*	*OR (95% CI), p-value*
236 (72-340)	144 (32-340)	0.12	1.002 (0.999-1.006), 0.18

Ever smoking	**Never smoking**	**Mann-Whitney U**	**Logistic regression**
Median ACPA level (IQR)	*Median ACPA level (IQR)*	*p-value*	*OR (95% CI), p-value*
229 (64-340)	222 (52-340)	0.89	1.000 (0.996-1.004), 0.99

Odds ratios indicate effect measures for every unit increase in ACPA-level.
HLA-SE: shared epitope, ACPA: anti-citrullinated protein antibody, IQR: interquartile range, OR: odds ratio, CI: confidence interval

Supplementary Table 4. Associations of HLA-SE with development of inflammatory arthritis in patients newly presenting with CSA; hazard ratios corresponding to Figure 2 of the main article

		IA, n (%)	No IA, n (%)	HR (95% CI)	p-value
All patients					
HLA-SE	Absent	39 (42)	247 (58)	Reference	--
	Present	55 (59)	181 (42)	1.86 (1.23-2.82)	0.003
HLA-SE	0	39 (42)	247 (58)	Reference	--
	1	41 (44)	151 (35)	1.65 (1.06-2.56)	0.027
	2	14 (15)	30 (7)	3.03 (1.64-5.61)	<0.001
ACPA positive subgroup					
HLA-SE	Absent	15 (37)	12 (43)	Reference	--
	Present	26 (63)	16 (57)	1.29 (0.67-2.47)	0.44
HLA-SE	0	15 (37)	12 (43)	Reference	--
	1	17 (42)	14 (50)	1.05 (0.52-2.13)	0.90
	2	9 (22)	2 (7)	2.32 (1.00-5.41)	0.051
ACPA negative subgroup					
HLA-SE	Absent	24 (45)	235 (59)	Reference	--
	Present	29 (55)	165 (41)	1.71 (0.99-2.96)	0.055
HLA-SE	0	24 (45)	235 (59)	Reference	--
	1	24 (45)	137 (34)	1.66 (0.94-2.94)	0.083
	2	5 (9)	28 (7)	2.00 (0.76-5.28)	0.16

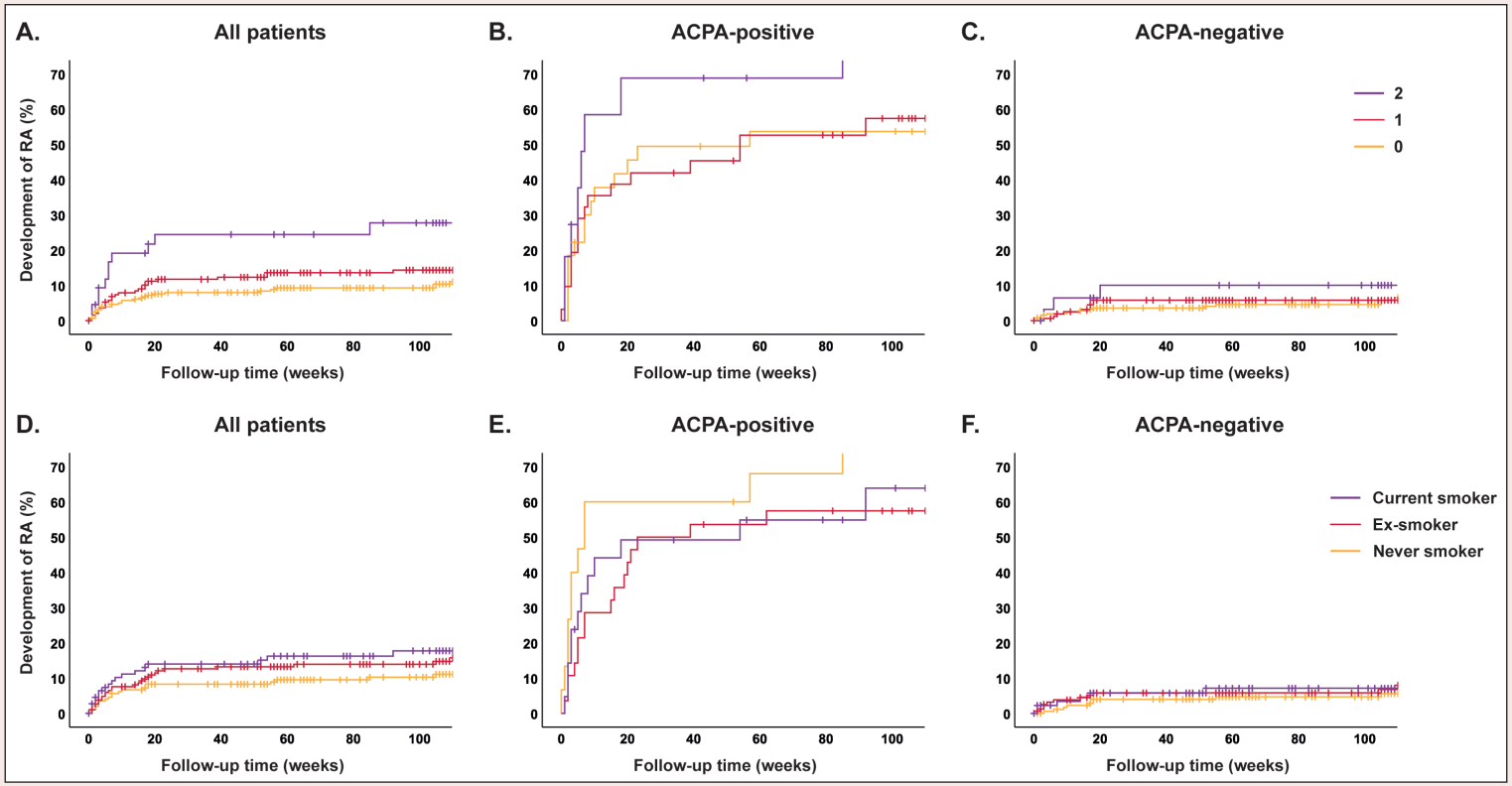
HLA-SE: shared epitope, CSA: clinically suspect arthralgia, IA: inflammatory arthritis, HR: hazard ratio, CI: confidence interval, ACPA: anti-citrullinated protein antibody

Supplementary Table 5. Associations of smoking with development of inflammatory arthritis in patients newly presenting with CSA; hazard ratios corresponding to Figure 3 of the main article

		IA, n (%)	No IA, n (%)	HR (95% CI)	p-value
All patients					
Smoking	Never	29 (32)	171 (42)	Reference	--
	Ever	63 (69)	237 (58)	1.40 (0.90-2.18)	0.14
Smoking	Never	29 (32)	171 (42)	Reference	--
	Ex-smoker	38 (41)	151 (37)	1.25 (0.76-2.06)	0.37
	Current smoker	25 (27)	86 (21)	1.66 (0.97-2.83)	0.065
ACPA positive subgroup					
Smoking	Never	11 (28)	4 (17)	Reference	--
	Ever	29 (73)	20 (83)	0.59 (0.29-1.18)	0.13
Smoking	Never	11 (28)	4 (17)	Reference	--
	Ex-smoker	17 (43)	11 (46)	0.55 (0.26-1.19)	0.13
	Current smoker	12 (30)	9 (38)	0.64 (0.28-1.45)	0.28
ACPA negative subgroup					
Smoking	Never	18 (35)	167 (44)	Reference	--
	Ever	34 (65)	217 (57)	1.30 (0.73-2.33)	0.37
Smoking	Never	18 (35)	167 (44)	Reference	--
	Ex-smoker	21 (40)	140 (37)	1.17 (0.61-2.24)	0.64
	Current smoker	13 (25)	77 (20)	1.56 (0.76-3.18)	0.23

CSA: clinically suspect arthralgia, IA: inflammatory arthritis, HR: hazard ratio, CI: confidence interval, ACPA: anti-citrullinated protein antibody

Supplementary Figure 3. Associations of HLA-SE (A-C) and smoking (D-F) with progression from CSA to RA



RA is defined as fulfilment of the 1987 and/or 2010 criteria at the time of clinically apparent inflammatory arthritis development.
Corresponding hazard ratios, with 0 HLA-SE alleles as reference category were: (A) HR 1.42 (95% CI 0.85-2.38) and HR 2.97 (1.48-5.96) for 1 and 2 HLA-SE alleles respectively, (B) HR 1.05 (0.52-2.13) and HR 2.08 (0.87-5.00), and (C) HR 1.16 (0.53-2.53) and HR 1.87 (0.54-6.51).
Corresponding hazard ratios, with never smoker as reference category were: (D) HR 1.44 (0.81-2.56) and HR 1.81 (0.97-3.39) for ex-smoker and current smoker respectively, (E) HR 0.55 (0.26-1.19) and HR 0.64 (0.28-1.45), and (F) HR 1.42 (0.59-3.43) and HR 1.66 (0.62-4.46).
HLA-SE: shared epitope, CSA: clinically suspect arthralgia, RA: rheumatoid arthritis, ACPA: anti-citrullinated protein antibody, HR: hazard ratio, CI: confidence interval

Supplementary Table 6. Associations of HLA-SE and smoking with presence of RF, anti-CarP and AAPA in patients newly presenting with CSA

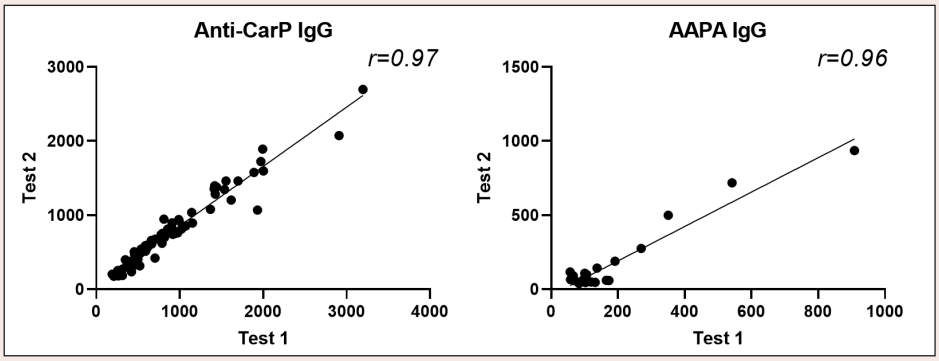
RF	Univariable OR (95% CI)	p-value	Multivariable[a] OR (95% CI)	p-value
HLA-SE present	1.13 (0.73-1.73)	0.59	0.71 (0.41-1.22)	0.21
Ever-smoker	1.16 (0.74-1.84)	0.52	0.70 (0.40-1.23)	0.22

Anti-CarP	Univariable OR (95% CI)	p-value	Multivariable[a] OR (95% CI)	p-value
HLA-SE present	1.63 (0.59-4.50)	0.34	1.31 (0.42-4.07)	0.65
Ever-smoker	1.05 (0.37-2.95)	0.93	0.58 (0.17-2.01)	0.39

AAPA	Univariable OR (95% CI)	p-value	Multivariable[a] OR (95% CI)	p-value
HLA-SE present	1.41 (0.65-3.05)	0.39	1.15 (0.47-2.82)	0.76
Ever-smoker	0.58 (0.26-1.31)	0.19	0.28 (0.098-0.80)	0.018

[a] Corrected for presence of ACPA
HLA-SE: shared epitope, RF: rheumatoid factor, anti-CarP: anti-carbamylated protein antibody, AAPA: anti-acetylated protein antibody, CSA: clinically suspect arthralgia, OR: odds ratio, CI: confidence interval, ACPA: anti-citrullinated protein antibody

Supplementary Figure 4. Inter-assay variation of in-house ELISAs



Inter-assay variation resulted in changes in positivity of the test infrequently: anti-CarP IgG 3.9%, AAPA IgG 4.2%.
anti-CarP: anti-carbamylated protein antibody, AAPA: anti-acetylated protein antibody