EXTENDED REPORT

Investigation of polymorphisms in the *PADI4* gene in determining severity of inflammatory polyarthritis

A Barton, J Bowes, S Eyre, D Symmons, J Worthington, A Silman

Ann Rheum Dis 2005;**64**:1311–1315. doi: 10.1136/ard.2004.034173

See end of article for authors' affiliations

Correspondence to: Dr Anne Barton, ARC-EU, Stopford Building, University of Manchester, Manchester M13 9PT, UK; anne.barton@manchester.ac.uk

Accepted 7 February 2005
Published Online First 24 February 2005

Background: A functional haplotype of the peptidylarginine deiminase 4 (*PADI4*) gene has recently been identified as a rheumatoid arthritis susceptibility gene in a Japanese but not in a UK population. One possible explanation for this disparity is that the gene determines severity of rather than susceptibility to inflammatory polyarthritis (IP) and that the UK and Japanese cohorts differed in terms of outcome.
Aim: To examine the association between individual *PADI4* single nucleotide polymorphisms (SNPs) and haplotypes, with the development and severity of erosions by five years in patients with IP.
Methods: 438 patients from the NOAR inception cohort of patients with IP were *x* rayed five years after presentation with early IP. Association with four exonic SNPs (padi4_89*G/A, padi4_90*T/C, padi4_92*G/C, and padi4_104*T/C), mapping to the *PADI4* gene and defining a haplotype previously reported to be associated with rheumatoid arthritis, was investigated. Patients were compared for the presence, extent, and progression of erosions by five years and the presence of antibodies to citrullinated peptide (anti-CCP antibodies).
Results: There was no association between individual *PADI4* SNPs or haplotypes and the development or extent of erosions by five years. Restricting analysis to patients who satisfied ACR criteria for rheumatoid arthritis by five years did not alter the conclusions. No association with presence of anti-CCP antibodies was detected.
Conclusions: No evidence was found for association of the *PADI4* gene with severity as assessed by erosive outcome at five years or with presence of anti-CCP antibodies in patients with IP.

Recently, an association of rheumatoid arthritis with the peptidylarginine deiminase 4 (*PADI4*) gene (NM_012387) has been described in a Japanese population.[1] The gene encodes a member of a family of enzymes that post-translationally convert arginine residues to citrulline. It had previously been noted that anticyclic citrullinated peptide (anti-CCP) antibodies appear to be specific for rheumatoid arthritis, predict the development of the disease in patients with undifferentiated inflammatory arthritis, and arise early in the disease course, suggesting that they may have a pathogenic role.[2 3] A susceptibility haplotype of the *PADI4* gene was shown to affect stability of mRNA transcripts and was associated with levels of antibody to anti-CCP antibodies in the sera of Japanese patients with rheumatoid arthritis.[1] Hence it was postulated that the presence of the susceptibility haplotype results in greater transcript stability, in turn leading to more citrullination of epitopes, which then become targets for anti-CCP antibodies.[1] However, in previous work, neither we nor others have shown association of this susceptibility haplotype in white European rheumatoid arthritis populations.[4 5]

One explanation could be that the Japanese and European cohorts differed in terms of severity. In support of this, the presence of anti-CCP antibodies has previously been shown to predict erosive outcome in patients with early inflammatory polyarthritis (IP), suggesting that they may be severity rather than susceptibility markers and, likewise, polymorphism within the gene may be associated with severity rather than susceptibility.[6–8] In order to separate these effects, it is necessary to investigate an inception cohort of patients all followed prospectively to quantify disease outcome. The Norfolk Arthritis Register (NOAR) provides such a resource. It aims to collect information on all incident cases of inflammatory polyarthritis (IP) (defined as swelling of two or more joints lasting for four or more weeks, after exclusion of other causes) presenting to a primary care physician in a geographically defined region, and is a true primary care based inception cohort. Data collection is not restricted to patients fulfilling ACR criteria for rheumatoid arthritis at presentation because the proportion of patients satisfying the criteria changes over time. It is therefore preferable, when studying early disease, to study the whole population of patients with IP and to stratify at different time points for whether or not patients satisfy criteria for rheumatoid arthritis.[9] The aim of the present study was thus to investigate association of the *PADI4* SNPs and haplotypes with the presence and extent of radiographic erosions by five years and with presence of anti-CCP antibodies.

METHODS

Study design

A large prospective inception cohort study was carried out to investigate whether the presence of polymorphisms of the *PADI4* gene played a role in determining erosive outcome in patients with IP (of which rheumatoid arthritis is a major subset). Association of four *PADI4* gene single nucleotide polymorphisms (SNPs), which together define a functional haplotype previously associated with rheumatoid arthritis in a Japanese population, was tested with the presence and extent of radiological erosions by five years and with the presence of anti-CCP antibodies. In addition, association with markers of disease activity and disability at baseline and at five years was investigated.

Abbreviations: ACR, American College of Rheumatology; anti-CCP, anticyclic citrullinated peptide antibody; DMARD, disease modifying antirheumatic drug; EM, expectation maximisation; HAQ, health assessment questionnaire; IP, inflammatory polyarthritis; NOAR, Norfolk Arthritis Register; SNP, single nucleotide polymorphism

Table 1 Analysis of erosive outcome (erosions, yes/no) at five years for the whole dataset and in a subset satisfying ACR criteria for rheumatoid arthritis by five years

	Inflammatory arthritis			ACR criteria positive RA patients		
SNP genotype	Erosive (%), n = 194	Non-erosive (%), n = 244	OR (95% CI)†	Erosive (%), n = 174	Non-erosive (%), n = 143	OR (95% CI)†
Padi4_89						
GG	70 (37.8)	79 (33.9)		63 (36.6)	50 (35.7)	
GA	86 (46.5)	115 (49.4)	0.84 (0.5 to 1.3)	80 (46.5)	64 (45.7)	0.99 (0.6 to 1.7)
AA	29 (15.7)	39 (16.7)	0.84 (0.5 to 1.6)	29 (16.9)	26 (18.6)	0.86 (0.4 to 1.8)
Comparison	p = 0.71			p = 0.93		
Padi4_90						
TT	63 (35.6)	77 (33.6)		59 (35.5)	48 (35.0)	
TC	83 (46.9)	111 (48.5)	0.91 (0.6 to 1.5)	77 (46.4)	63 (46.0)	0.99 (0.6 to 1.7)
CC	31 (17.5)	41 (17.9)	0.92 (0.5 to 1.7)	30 (18.1)	26 (19.0)	0.94 (0.5 to 1.9)
Comparison	p = 0.92			p = 0.98		
Padi4_92						
GG	57 (31.2)	73 (31.6)		52 (30.4)	46 (32.9)	
GC	91 (49.7)	113 (48.9)	1.03 (0.6 to 1.6)	86 (50.3)	66 (47.1)	1.15 (0.7 to 2.0)
CC	35 (19.1)	45 (19.5)	0.99 (0.5 to 1.8)	33 (19.3)	28 (20.0)	1.02 (0.5 to 2.1)
Comparison	p = 0.99			p = 0.83		
Padi4_104						
TT	85 (45.5)	112 (47.7)		77 (44.3)	67 (46.9)	
TC	83 (44.4)	96 (40.8)	1.14 (0.7 to 1.7)	78 (44.8)	59 (41.2)	1.15 (0.7 to 1.9)
CC	19 (10.1)	27 (11.5)	0.93 (0.5 to 1.9)	19 (10.9)	17 (11.9)	0.97 (0.4 to 2.2)
Comparison	p = 0.75			p = 0.81		
Haplotype*						
1/1	83 (45.9)	111 (48.3)		76 (45.0)	65 (46.8)	
½	79 (43.6)	92 (40.0)	1.15 (0.7 to 1.8)	74 (43.8)	57 (41.0)	1.11 (0.7 to 1.8)
2/2	19 (10.5)	27 (11.7)	0.94 (0.5 to 1.9)	19 (11.2)	17 (12.2)	0.96 (0.4 to 2.1)
Comparison	p = 0.75			p = 0.88		

*The *PADI4* susceptibility haplotype, as defined in the Japanese population, was referred to as 2 and all other haplotypes as 1.
†Odds ratio and 95% confidence intervals of erosive change by five years compared with wild type genotype.
CI, confidence interval; OR, odds ratio; RA, rheumatoid arthritis; SNP, single nucleotide polymorphism.

Subjects

Patients were recruited from the NOAR. The case ascertainment procedure has been described previously[10] but, briefly, all cases with IP (defined as swelling of two or more joints lasting for four or more weeks) within the region formerly known as the Norwich Health Authority are assessed by a research nurse using a standard questionnaire and examination. Baseline clinical data are recorded and blood taken for RF measurement and DNA extraction. Patients are reviewed annually and at each assessment they are scored as to whether ACR classification criteria for rheumatoid arthritis are satisfied, using a cumulative evaluation as described previously.[11]

All patients had radiographs of the hands and feet five years after registration. Briefly, *x* rays of the hands and feet were scored by two observers using the Larsen method. A third observer arbitrated in cases of disagreement. A subset of patients who had three or more ACR criteria had *x* rays at one year. For the purposes of this analysis the first 438 consecutive subjects who were successfully followed up for five years and had provided a blood sample for DNA extraction were studied.

All patients were of UK white (Europid) ethnic origin, provided written informed consent, and were recruited with ethical committee approval (LREC 2003-075).

Genotyping methods

Four *PADI4* SNPs (padi4_89*G/A, padi4_90*T/C, padi4_92*G/C (rs874881), and padi4_104*T/C (rs1748033)) were genotyped as described previously.[4]

Anti-CCP antibody

As part of a separate study, the first 380 patients recruited to NOAR with *x* rays available by two years were tested for the presence of anti-CCP antibody. This was an in house first generation test carried out using an enzyme linked immunosorbent assay (ELISA) (done in collaboration with the Leiden University Medical Centre, Netherlands[8]). Two hundred and fourteen of these patients were included in the current cohort.

Statistical analysis

Patients were defined as having erosive arthritis if they had a Larsen score of $\geqslant 2$ in any hand or foot joint. Both the presence of any erosion and the actual Larsen score at five years were used as outcome measures to assess the role of the *PADI4* polymorphisms and the Japanese susceptibility haplotype. Genotype frequencies of each SNP were compared between cases who developed erosions by one and five years and those without, using Fisher's exact test. Further analyses were restricted to cases who satisfied ACR criteria for rheumatoid arthritis by five years, or who were rheumatoid factor (RF) positive at baseline. The two sample (Mann–Whitney) test was used to test for an influence of allele frequency on the following disease features: Larsen score at 1 and 5 years; the number of tender and swollen joints at baseline and 5 years; and the health assessment questionnaire (HAQ) score at baseline and five years. Subgroup analysis was also undertaken restricting analysis to those who satisfied ACR criteria for rheumatoid arthritis by five years.

As mentioned above, in the Japanese population, a haplotype defined by the four SNPs genotyped in the current study was found to affect mRNA stability.[1] The analysis package SNPHAP (http://www-gene.cimr.cam.ac.uk/clayton/software/) was therefore used to estimate and assign haplotypes. This program uses the expectation maximisation (EM) algorithm and the software assigns haplotypes to

Table 2 Association of *PADI4* single nucleotide polymorphisms and haplotype with markers of disease activity and severity

SNP allele	Median (IQR) p value (P_c)* Larsen 5† (n=438)	SJC0† (n=406)	SJC5† (n=176)	TJC0† (n=411)	TJC5† (n=194)	HAQ0† (n=378)	HAQ5† (n=363)
Padi4_89							
G	6 (0 to 20)	6 (2 to 13)	0 (0 to 2)	7 (3 to 17)	0 (0 to 4)	0.75 (0.25 to 1.38)	0.88 (0.25 to 1.63)
A	5 (0 to 22)	6 (2 to 13)	0 (0 to 2)	8 (4 to 18)	0 (0 to 3)	0.875 (0.38 to 1.5)	0.75 (0.25 to 1.75)
Comparison	p=0.57	p=0.33	p=0.36	p=0.41	p=0.22	p=0.03 (P_c=0.60)	p=0.28
Padi4_90							
T	5.5 (0 to 19)	6 (2 to 13)	0 (0 to 3)	7 (3 to 17)	0 (0 to 4)	0.75 (0.25 to 1.38)	0.815 (0.25 to 1.5)
C	5 (0 to 25)	6 (2 to 13)	0 (0 to 2)	8 (4 to 19)	0 (0 to 3)	0.875 (0.38 to 1.5)	0.815 (0.25 to 1.75)
Comparison	p=0.18	p=0.52	p=0.39	p=0.20	p=0.17	p=0.01	p=0.18
Padi4_92							
G	5 (0 to 19.5)	6 (2 to 14)	0 (0 to 2)	8 (3 to 17)	0 (0 to 4)	0.75 (0.25 to 1.38)	0.88 (0.25 to 1.5)
C	5 (0 to 24)	6 (2 to 13)	0 (0 to 2)	8 (4 to 19)	0 (0 to 4)	0.875 (0.38 to 1.5)	0.75 (0.25 to 1.75)
Comparison	p=0.24	p=0.28	p=0.93	p=0.26	p=0.46	p=0.04 (P_c=0.80)	p=0.48
Padi4_104							
T	5 (0 to 20)	6 (2 to 13)	0 (0 to 2)	8 (3 to 17)	0 (0 to 4)	0.75 (0.25 to 1.38)	0.88 (0.25 to 1.5)
C	5 (0 to 27)	6 (2 to 14)	0 (0 to 2)	8 (4 to 19)	0 (0 to 4)	0.875 (0.38 to 1.5)	0.88 (0.25 to 1.88)
Comparison	p=0.46	p=0.58	p=0.76	p=0.36	p=0.44	p=0.04 (P_c=0.80)	p=0.21
Haplotype‡							
1	5 (0 to 19)	6 (2 to 13)	0 (0 to 2)	7 (3 to 17)	0 (0 to 4)	0.75 (0.25 to 1.38)	0.88 (0.25 to 1.5)
2	5 (0 to 27)	6 (2 to 14)	0 (0 to 2)	9 (4 to 19)	0 (0 to 4)	0.875 (0.38 to 1.5)	0.88 (0.25 to 1.88)
Comparison	p=0.34	p=0.64	p=0.77	p=0.22	p=0.53	p=0.09	p=0.21

Values are odds ratios and 95% confidence intervals.
*p Value calculated using Mann–Whitney test; P_c = p value corrected for multiple testing.
†Analysis adjusted using DMARD treatment by 5 years as a co-factor
‡The *PADI4* susceptibility haplotype, as defined in the Japanese population, was referred to as 2 and all other haplotypes as 1.
HAQ0, HAQ score at baseline; HAQ5, HAQ score at five years; IQR, interquartile range; Larsen 5, Larsen score by five years; SJC0, swollen joint count at baseline; SJC5, swollen joint count at five years; SNP, single nucleotide polymorphism; TJC0, tender joint count at baseline; TJC5, tender joint count at five years.

individuals, provides information about the probability of that assignment, and also allows some missing data—for example, as a result of PCR failure. Two haplotypes are assigned to each individual and individuals in whom haplotypes were assigned with a probability of less than 90% were excluded from further analysis in order to ensure stringency. The ''susceptibility'' haplotype, as defined in the Japanese population, was coded as 2 and all other haplotypes as 1, and then frequencies were compared as for the individual SNP markers outlined above.

As the Larsen score is a count, negative binomial regression was used to model the influence of the Japanese susceptibility *PADI4* haplotype on Larsen score at one and five years. Negative binomial regression was used rather than Poisson regression, as the scores obtained display greater variation than a Poisson distribution. Hence Poisson regression underestimates the standard errors. The outcome from the negative binomial regression is a multiplier for each group—for example, a coefficient of 1.2 implies that the Larsen score in that group is, on average, 20% higher than the referent group.

Finally, as treatment with disease modifying antirheumatic drugs (DMARDs) may mask genetic associations, all analyses were also carried out after adjustment for ever using a DMARD.

Power

Sample sizes were calculated based on published allele frequencies (minor allele frequencies of 33–40%) so that each of the association studies had 80% power to detect a genotypic relative risk of 2.0 (the susceptibility effect reported for the padi4_104*T/C SNP in the Japanese population) at the 5% significance level for development of erosions by five years.

Correction for multiple testing

Association with four SNPs and one haplotype were investigated in this study and, in addition, the data were stratified as described above to fully explore whether the *PADI4* gene is associated with severity of IP. As the SNPs are tightly linked, it is difficult to determine accurately the Bonferroni correction that should be applied, as the test assumes independence of loci. In this study, we used a correction factor of 20. The uncorrected probability (p) values are presented, as well as the corrected p values where the uncorrected value achieved statistical significance.

RESULTS

Patients

In all there were 438 IP cases with *x* ray data at five years. Their mean (SD) age was 53.3 (14.3) years; 286 (65.3%) were

Table 3 Results of negative binomial regression for Larsen score at one and five years

	PADI4 haplotype 1	2
Coefficient (95% CI) of Larsen score at one year*	1.0 (referent)	0.80 (0.59 to 1.08); p=0.15
Coefficient (95% CI) of Larsen score at five years*	1.0 (referent)	1.05 (0.83 to 1.31); p=0.70

*Analysis adjusted using DMARD treatment by five years as a cofactor.
CI, confidence interval.

Table 4 Association of PADI4 single nucleotide polymorphisms with satisfying ACR classification criteria for rheumatoid arthritis by five years, rheumatoid factor (RF) positivity at baseline, and anti-CCP antibody positivity at baseline

SNP allele	RA by 5 years (%)	Non-RA by 5 years (%)	RF+ at baseline (%)	RF− at baseline (%)	CCP+ at baseline (%)	CCP− at baseline (%)
Padi4_89						
G	370 (59.3)	129 (60.8)	124 (59.0)	373 (60.0)	44 (56.4)	216 (65.9)
A	254 (40.7)	83 (39.2)	86 (41.0)	249 (40.0)	34 (43.6)	112 (34.1)
Comparison	p=0.75		p=0.87		p=0.15	
Padi4_90						
T	354 (58.4)	120 (58.3)	116 (58.6)	355 (58.8)	45 (56.3)	203 (63.8)
C	252 (41.6)	86 (41.7)	82 (41.4)	249 (41.2)	35 (43.7)	115 (36.2)
Comparison	p=1.0		p=0.74		p=0.25	
Padi4_92						
G	348 (55.9)	116 (56.3)	116 (56.3)	347 (56.3)	42 (55.3)	201 (61.7)
C	274 (44.1)	90 (43.7)	90 (43.7)	269 (43.7)	34 (44.7)	125 (38.3)
Comparison	p=0.94		p=1.0		p=0.36	
Padi4_104						
T	425 (67.0)	148 (70.5)	133 (63.9)	435 (68.8)	50 (62.5)	234 (72.2)
C	209 (33.0)	62 (29.5)	75 (36.1)	197 (31.2)	30 (37.5)	90 (27.8)
Comparison	p=0.39		p=0.20		p=0.10	
Haplotype*						
1	413 (67.0)	146 (70.9)	129 (63.9)	425 (69.0)	49 (62.8)	233 (72.4)
2	203 (33.0)	60 (29.1)	73 (36.1)	191 (31.0)	29 (37.2)	89 (27.6)
Comparison	p=0.34		p=0.19		p=0.10	

*The *PADI4* susceptibility haplotype, as defined in the Japanese population, was referred to as 2 and all other haplotypes as 1.
CCP, cyclic citrullinated peptide; RA, rheumatoid arthritis; RF, rheumatoid factor; SNP, single nucleotide polymorphism.

female; 114 (25.8%) were RF positive at baseline and 166 (37.9%) by five years; and 197 (45.0%) satisfied ACR criteria for rheumatoid arthritis at baseline and 317 (72.4%) by five years. Carriage of shared epitope alleles was as follows: 39.5% carried no alleles, 46.6% one allele, and 13.9% two alleles. One hundred and ninety four (44%) were erosive, with Larsen scores ranging from 2 to 138. A subset of patients (n = 289) also had radiograph data at one year after presentation, of whom 97 (34%) had erosive changes. The subset with *x* rays at one year had more severe disease than those who did not have *x* rays. Anti-CCP antibody data were available for 214 patients, 44 (19.2%) of whom were positive at baseline.

Polymorphisms

All SNPs were in Hardy–Weinberg equilibrium, with allele frequencies close to those reported previously in both the UK and Japanese populations.[1 4]

Neither the SNPs tested individually nor the ''susceptibility'' haplotype defined in the Japanese population were associated with development of erosions by one year (data not shown) or five years (table 1), or with Larsen score at one year (data not shown) or five years (table 2). Restricting analysis to those fulfilling ACR criteria for rheumatoid arthritis by five years (table 1) or those positive for RF at baseline (data not shown) gave similar findings. When the data were analysed using negative binomial regression adjusting for DMARD use by five years, patients carrying the ''susceptibility'' haplotype had Larsen scores at five years that were 5% higher than those in the referent group, but at one year their Larsen scores were 20% lower (table 3). Neither result was statistically significant.

Association with various markers of disease activity was tested but none was found with swollen joint count or tender joint count either at baseline or at five years (table 2). Restricting analysis to those patients satisfying ACR criteria for rheumatoid arthritis by five years did not alter the conclusions (data not shown). Association with HAQ score at baseline but not at five years was detected with the four SNPs individually but not by haplotype (uncorrected p value = 0.10) (table 2). It should be noted that these associations did not remain significant after a Bonferroni correction was applied.

By contrast with a previous report in a Japanese population, the *PADI4* SNPs and ''susceptibility'' haplotype were not associated with presence of anti-CCP antibodies at baseline in this UK cohort of IP patients (table 4). Similarly, no association with presence of RF at baseline or satisfaction of ACR criteria for rheumatoid arthritis by five years was detected (table 4).

DISCUSSION

This is the first study to examine in detail the possible link between polymorphism in the *PADI4* gene and disease severity in IP and rheumatoid arthritis. No association of *PADI4* polymorphism with severity was observed. This was the case whether measures of disease severity or activity at either baseline or by five years were considered.

One of the strengths of the current study is that analysis was not restricted to ACR defined rheumatoid arthritis cases with established disease, because the ACR criteria perform poorly in early disease and it is in early disease that the identification of genes that determine prognosis is likely to have the greatest impact by allowing better targeting of aggressive treatments.[12] However, the majority (75%) of these patients subsequently satisfied ACR criteria for rheumatoid arthritis by five years, and restricting analysis to patients satisfying ACR criteria for rheumatoid arthritis by five years did not alter the findings substantially. The primary outcome measure was the development of erosions; this is accepted as an objective and reliable outcome measure and the presence of erosions is used in many studies as a surrogate marker for quantifying disease severity.[13 14] However, as erosions are part of the criteria for rheumatoid arthritis, studying prevalent cases with variable disease duration would have made separation of susceptibility from severity factors difficult. Therefore, by investigating association in an inception cohort of patients who were all followed prospectively and were

unselected for the presence of erosions at study entry, any bias which might have arisen from this cause should have been eliminated.

The effects of treatment on disease severity may obscure the investigation of genetic factors. Patients with severe disease at baseline who receive treatment early are expected to have milder disease over time than if they had remained untreated. We have tried to account for this possible confounding by treatment allocation using an adjustment in the analysis (whether patients had ever received a DMARD) but this adjustment may be imperfect.

Although the study was adequately powered to detect an associated SNP or haplotype conferring an odds ratio of 2.0 for the development of erosions by five years, if the effect size is small much larger sample sizes will be required. It should be noted, however, that previously identified predictors of erosions—such as carriage of shared epitope alleles, presence of RF at baseline, and presence of anti-CCP antibodies at baseline—were all strongly predictive of erosions by five years in this IP cohort ($p = 2.6\times10^{-6}$, 3.0×10^{-8}, and $p = 7.4\times10^{-5}$, respectively).[8 15 16]

A previous study in a Japanese population found that the haplotype was associated with presence of anti-CCP antibodies, although data were only available for 123 patients.[1] In our study, anti-CCP antibody data were available for 214 patients. We found no association of *PADI4* SNPs or haplotype with the presence of anti-CCP antibodies, although there was a trend towards association for the haplotype (odds ratio = 1.5 (95% confidence interval, 0.9 to 2.07), $p = 0.10$). In order to confirm an association of the haplotype with anti-CCP antibodies, a much larger sample size would be required. In both the Japanese study and our own, haplotypes were estimated using the EM algorithm. Previous work has shown very similar frequencies for estimated and directly sequenced haplotypes in the white population.[17]

The results presented herein suggest that polymorphism within the *PADI4* gene itself does not play a major role in determining erosive or severe outcome in IP in the United Kingdom population. Hence genotyping patients for these polymorphisms at presentation will not help in targeting treatments to patients likely to develop more severe disease.

ACKNOWLEDGEMENTS

This study was supported by the UK Arthritis Research Campaign. AB is in receipt of a Wellcome Advanced Fellowship. We are grateful to Diane Bunn and the Norfolk Arthritis Register research team for the clinical data collection and to the rheumatologists and GPs in the Norwich area for referring patients. We are also grateful to Henk Visser and colleagues at the Leiden Medical Centre, Netherlands, for providing the information on anti-CCP antibodies.

Authors' affiliations
A Barton, J Bowes, S Eyre, D Symmons, J Worthington, A Silman, ARC-EU, University of Manchester, Manchester, UK

REFERENCES

1 **Suzuki A**, Yamada R, Chang X, Tokuhiro S, Sawada T, Suzuki M, *et al*. Functional haplotypes of PADI4, encoding citrullinating enzyme peptidylarginine deiminase 4, are associated with rheumatoid arthritis. *Nat Genet* 2003;**34**:395–402.
2 **Zhou Z**, Menard HA. Autoantigenic posttranslational modifications of proteins: does it apply to rheumatoid arthritis? *Curr Opin Rheumatol* 2002;**14**:250–3.
3 **van Gaalen FA**, Linn-Rasker SP, van Venrooij WJ, de Jong BA, Breedveld FC, Verweij CL, *et al*. Autoantibodies to cyclic citrullinated peptides predict progression to rheumatoid arthritis in patients with undifferentiated arthritis: a prospective cohort study. *Arthritis Rheum* 2004;**50**:709–15.
4 **Barton A**, Bowes J, Eyre S, Spreckley K, Hinks A, John S, *et al*. A functional haplotype of the PADI4 gene associated with rheumatoid arthritis in a Japanese population is not associated in a United Kingdom population. *Arthritis Rheum* 2004;**50**:1117–21.
5 **Caponi L**, Petit-Teixeira E, Sebbag M, Bongiorni F, Moscato S, Pratesi F, *et al*. A family-based study shows no association between rheumatoid arthritis and the PADI4 gene in a French caucasian population. *Ann Rheum Dis* 2005;**64**:587–93.
6 **Bas S**, Genevay S, Meyer O, Gabay C. Anti-cyclic citrullinated peptide antibodies, IgM and IgA rheumatoid factors in the diagnosis and prognosis of rheumatoid arthritis. *Rheumatology (Oxford)* 2003;**42**:677–80.
7 **Meyer O**, Labarre C, Dougados M, Goupille P, Cantagrel A, Dubois A, *et al*. Anticitrullinated protein/peptide antibody assays in early rheumatoid arthritis for predicting five year radiographic damage. *Ann Rheum Dis* 2003;**62**:120–6.
8 **Visser H**, le Cessie S, Vos K, Breedveld FC, Hazes JM. How to diagnose rheumatoid arthritis early: a prediction model for persistent (erosive) arthritis. *Arthritis Rheum* 2002;**46**:357–65.
9 **Symmons DP**, Hazes JM, Silman AJ. Cases of early inflammatory polyarthritis should not be classified as having rheumatoid arthritis. *J Rheumatol* 2003;**30**:902–4.
10 **Symmons DP**, Barrett EM, Bankhead CR, Scott DG, Silman AJ. The incidence of rheumatoid arthritis in the United Kingdom: results from the Norfolk Arthritis Register. *Br J Rheumatol* 1994;**33**:735–9.
11 **MacGregor AJ**, Bamder S, Silman AJ. A comparison of the performance of different methods of disease classification for rheumatoid arthritis. Results from an analysis from a nationwide twin study. *J Rheumatol* 1994;**21**:1420–6.
12 **Dugowson CE**, Nelson JL, Koepsell TD. Evaluation of the 1987 revised criteria for rheumatoid arthritis in a cohort of newly diagnosed female patients. *Arthritis Rheum* 1990;**33**:1042–6.
13 **van der Heijde DM**. Radiographic imaging: the ''gold standard'' for assessment of disease progression in rheumatoid arthritis. *Rheumatology (Oxford)* 2000;**39**(suppl 1):9–16.
14 **Sharp JT**. Radiologic assessment as an outcome measure in rheumatoid arthritis. *Arthritis Rheum* 1989;**32**:221–9.
15 **Bukhari M**, Lunt M, Harrison BJ, Scott DG, Symmons DP, Silman AJ. Rheumatoid factor is the major predictor of increasing severity of radiographic erosions in rheumatoid arthritis: results from the Norfolk Arthritis Register Study, a large inception cohort. *Arthritis Rheum* 2002;**46**:906–12.
16 **Harrison B**, Thomson W, Symmons D, Ollier B, Wiles N, Payton T, *et al*. The influence of HLA-DRB1 alleles and rheumatoid factor on disease outcome in an inception cohort of patients with early inflammatory arthritis. *Arthritis Rheum* 1999;**42**:2174–83.
17 **Hoppe B**, Heymann GA, Tolou F, Kiesewetter H, Doerner T, Salama A. High variability of peptidylarginine deiminase 4 (PADI4) in a healthy white population: characterization of six new variants of PADI4 exons 2–4 by a novel haplotype-specific sequencing-based approach. *J Mol Med* 2004;**82**:762–7.