Chapter 3

Quantitative heritability of ACPA-positive and ACPA-negative rheumatoid arthritis

Diane van der Woude, Jeanine J. Houwing-Duistermaat, René E.M. Toes, Tom W.J. Huizinga, Wendy Thomson, Jane Worthington, Annette H.M. van der Helm-van Mil and René R.P. de Vries

Arthritis and Rheumatism. 2009 April; 60(4): 916-23

ABSTRACT

Objective The majority of genetic risk factors for rheumatoid arthritis (RA) are associated with anti-citrullinated protein antibody (ACPA)-positive RA, while far fewer genetic risk factors have been identified for ACPA-negative RA. To quantify the contribution of genetic risk factors in general and of the predisposing HLA-DRB1 shared epitope (SE) alleles in particular to these two subsets of RA, we computed the heritability and assessed the contribution of HLA SE alleles.

Methods 148 RA twin pairs, of which at least one twin had RA, were tested for ACPA and typed for HLA-DRB1 genotypes. Heritability was assessed by including a bivariate normally distributed random effect representing the contribution of unobserved genetic factors to RA susceptibility in a logistic regression model with the correlation between the random effects fixed according to twin zygosity. The contribution of HLA SE alleles to the genetic variance was assessed by a similar model using genotype-specific population prevalences.

Results The heritability of RA was 66% (95% confidence interval (CI) 44-75%). For ACPA-positive RA the heritability was 68% (CI: 55-79%) and for ACPA-negative RA 66% (CI: 21-82%). The HLA SE alleles explained 18% (CI: 16-19%) of the genetic variance of ACPA-positive RA, and only 2.4% (CI: 1.6-10%) of ACPA-negative RA.

Conclusion The heritability of ACPA-positive and ACPA-negative RA is comparable. These data indicate that genetic predisposition plays an important role in the pathogenesis of ACPA-negative RA, for which most individual genetic risk factors remain to be identified.

INTRODUCTION

Several new genetic risk factors for rheumatoid arthritis (RA) have been identified in the past few years. Candidate gene studies and new techniques such as whole genome scans have led to the discovery of various new genetic variants which predispose to RA, in for example the genomic regions encoding PTPN22, STAT4 and TRAF1/C5 [1-3]. Interestingly, the majority of the genetic factors which predispose for RA have been found to preferentially confer risk to autoantibody-positive disease. The importance of distinguishing autoantibody-positive from autoantibody-negative RA is increasingly being recognized due to the recent findings concerning anti-citrullinated protein antibodies (ACPA).

ACPA were first described as anti-perinuclear factor and anti-keratin antibodies, that both bind to citrullinated fillagrin [4-7]. Since then, ACPA have been found to be the most predictive factor for the future development of RA [8-10]. Due to its superior specificity and predictive characteristics as compared to IgM RF, ACPA are increasingly being used for diagnostic purposes [11]. Approximately two-thirds of patients with established RA are ACPA-positive, and these patients are characterized by a destructive phenotype with higher rates of joint destruction as compared to ACPA-negative RA patients [12].

ACPA-positive and ACPA-negative disease differ not only with regard to clinical disease phenotype, but are also associated with different genetic and environmental risk factors [13, 14]. The HLA-DRB1 shared epitope (SE) alleles for example, which confer the highest risk of all known genetic risk factors, predispose specifically to the development of an anti-citrullinated protein immune response, rather than to the development of RA [15]. PTPN22 gene variants, as well as the C5-TRAF locus, have been found to be predominantly associated with ACPA-positive RA, although an association with ACPA-negative RA cannot be excluded [2, 16]. It is possible that studies which include a larger number of ACPA-negative patients will also find an association between these genetic risk factors and ACPA-negative RA, such as has been shown for STAT4 [17]. At this moment however, the number of genetic risk factors which has been described to be predominantly associated with ACPA-negative RA is more limited, and includes HLA-DR3 haplotypes and interferon regulatory factor 5 (IRF5) [14, 18].

Not only genetic, but also environmental risk factors differ between ACPA-positive and ACPA-negative RA. Recent data have shown an interaction between smoking and the SE alleles in conferring risk for ACPA-positive RA [19, 20]. The differences in clinical phenotype and underlying risk factors have lead to the concept that ACPA-positive and ACPA-negative RA constitute distinct entities with different pathophysiological mechanisms of disease [21]. The extent to which genetic factors contribute to disease development may differ between these two subsets, and thus far, more genetic factors have been found to be associated with ACPA-positive than with ACPA-negative RA.

To investigate the contribution of genetic factors to disease development, twins are a valuable source of information. Twin studies provide two measures for the influence of genetic factors on disease development: the concordance and the heritability. The concordance represents the frequency with which the twin brothers or sisters of diseased individuals are also affected. While the concordance is influenced by the prevalence of the disease (or disease subset), heritability is not. The heritability supplies a quantitative measure for the amount of variation in disease susceptibility which can be explained by genetic factors. Another advantage of heritability estimates compared to concordance rates, is that the heritability is based on the combined data of both mono- and dizygotic twins. This makes heritability modeling a more powerful method, especially when using small populations such as RA twin cohorts. For RA, the heritability has been assessed at approximately 60% [22]. There are no data available however for the heritability of autoantibody-positive versus autoantibody-negative RA.

In this study, we determined the heritability of ACPA-positive and ACPA-negative RA by applying quantitative genetic methods to data from a large twin cohort. In addition, we assessed the contribution of the predisposing HLA-DRB1 shared epitope (SE) alleles to the genetic variance of both groups.

METHODS

Study population

A nationwide study of twin pairs of which one or both had RA was performed in the United Kingdom in 1989 [23]. Twin pairs were recruited by two parallel strategies. All UK rheumatologists were contacted and requested to ask all patients with RA whether they were a twin. Secondly, there was a simultaneous multi-media campaign, inviting patients who had been diagnosed with RA and who had a living twin to contact a study centre. Both members of each twin pair were visited at home by trained research nurses who recorded detailed history and demographic characteristics and performed joint examinations. Blood samples were collected of which serum was stored and lymphocytes were separated for HLA typing. High resolution HLA-DRB1 genotyping was performed enabling the assignment of RA-associated HLA SE alleles [24].

RA was diagnosed according to the 1987 revised ACR criteria [25]. Twin pairs were only eligible for inclusion if at least one twin satisfied these criteria. Zygosity status was verified by DNA fingerprinting on all same-sexed pairs using non-radioactive Southern blotting with a multilocus probe [26].

ACPA (Anti-CCP2) assays

Total IgG anti-CCP2 was measured in serum samples by enzyme-linked immunosorbent assay (ELISA) (Immunoscan RA Mark 2; Eurodiagnostica, Arnhem, The Netherlands). Samples with a value above 25 units/ml were considered positive according to the manufacturer's instructions.

Statistical analysis

Summary statistics were generated to investigate the prevalence of ACPA-positive and ACPA-negative RA and the occurrence of the HLA SE alleles. Concordance figures were calculated by dividing the number of pairs in which both twins were affected by the total number of twin pairs.

To calculate the heritability of RA, as well as of ACPA-positive RA and ACPA-negative RA, we used a logistic regression model with a normally distributed random effect [27-29]. In this case, the normally distributed random effect represents the contribution of unobserved genetic factors to RA susceptibility. The correlation between the outcomes of twins due to genetic factors is modeled according to the extent to which a twin pair shares this random effect. Monozygotic twin pairs share all their genome and therefore they also share all their random effect. Dizygotic twins share half of their genome, hence they also share half of the random effect in the model. The variance of the random effect (genetic variance) is a measurement for the total contribution of genetic factors to the outcome. The heritability is defined as the genetic variance divided by the total variance of the outcome. The variance of the random effect with 95% confidence intervals was estimated using profile likelihood, i.e. for a grid of genetic variances the likelihood was computed and the value for which the likelihood curve obtained its maximum is the maximum likelihood estimate of the genetic variance. Since the twin pairs were selected for having at least one affected twin, the prevalences of RA, ACPA-positive RA and ACPA-negative RA were fixed to their population prevalences (see Table 3 for the values used) (see Appendix A for further details of the heritability modeling approach).

Subsequently, we computed the contribution of SE to the total genetic variance of RA, as well as of ACPA-positive RA and of ACPA negative RA. This measure can be obtained by using a method similar to the method for computing the heritability. Instead of using the general population prevalence, now the genotypic specific population prevalences are used (see Table 3). For this model the variance of the random effect represents the genetic factors other than the genotype under study (see Appendix B for further details of the calculation of the contribution of SE to the genetic variance).

All computations were performed in the freely available software environment for statistical computing R.

RESULTS

Concordance

The characteristics of the study cohort, which consisted of 64 monozygotic and 84 dizygotic twin pairs as described previously [22], are displayed in Table 1. The patient characteristics of the monozygotic and dizygotic twins are comparable with regard to demographic factors, autoantibody status and radiographic severity of RA.

Table 1. Patient characteristics

Patient characteristic †	Monozygotic twins (64 pairs with 74 RA patients)	Dizygotic twins (84 pairs with 87 RA patients)
Age in years, mean (SD)	41 (14)	37 (13)
Female gender, n (%)	65 (88%)	71 (82%)
Rheumatoid factor IgM positive, n (%)	58 (78%)	74 (85%)
Anti-CCP2 positive, n (%)	54 (73%)	71 (82%)
Erosive disease, n (%)	55 (74%)	69 (79%)

† For dichotomous variables, the number and the percentage of patients are listed, relative to the total number of RA patients for whom information about the characteristic under investigation was available.

The concordance of RA, which represents the frequency with which the twin brothers or sisters of RA patients are also affected, was calculated for the monozygotic and dizygotic twin pairs (Table 2). For monozygotic twins the concordance is 15.6% versus 3.6% for dizygotic twins. Based on the results of the anti-CCP2 measurements, the concordance of the ACPA-positive and ACPA-negative subsets was also determined. In this cohort, twins who were concordant for RA all had the same ACPA status, meaning that there was no twin pair consisting of two affected individuals in which one individual had ACPA-positive disease and one individual had ACPA-negative disease. The concordance of ACPA-positive RA is comparable to the concordance of RA, while the concordance of ACPA-negative RA is lower than that of RA and ACPA-positive disease.

Table 2. Concordance of total RA, ACPA-positive and ACPA-negative disease.

	Monozygotic twins 64 pairs			Dizygotic twins 84 pairs		
	concordant	discordant	concordance (%)	concordant	discordant	concordance (%)
RA	10	54	15.6	3	81	3.6
ACPA+ RA	8	38	17.4	3	65	4.4
ACPA- RA	2	16	11.1	0	16	0

Table 3. Population prevalences used for heritability modeling

Characteristic	Prevalence	Reference
Prevalence of RA	1%	32
Prevalence of ACPA-positive RA	0.67%	11
Prevalence of ACPA-negative RA	0.33%	11
Prevalence of ACPA in healthy controls	1%	33
Prevalence of SE in ACPA-positive RA patients	84%	30
Prevalence of SE in ACPA-negative RA patients	59%*	*
Prevalence of SE in healthy controls	44%	31

The prevalence for each genotype was fixed to the specific population prevalence as reported in the United Kingdom [30,31] or, if not available from the UK, in studies from other Caucasian populations [11,32,33]. * personal communication from the Welcome Trust Case Control Consortium, Jane Worthington

Heritability

Heritabilities were estimated to quantify the extent to which genetic variation contributes to the development of disease, independent of the disease prevalence. For the logistic regression model as described in appendix A, the population prevalence of the disease or disease subsets was fixed to the specific population prevalence as reported in the United Kingdom [30, 31] or, if not available, in studies from other Caucasian populations [11, 32, 33] (Table 3).

The overall heritability of RA is 66% with a 95% confidence interval (CI) of 44-75%. The same approach was used to model the outcome of ACPA-positive and ACPA-negative RA. As shown in Table 4, the heritability estimates for ACPA-positive and ACPA-negative RA are similar to the heritability of total RA. The heritability of ACPA-positive RA is 68% (95% CI: 55-79) and the heritability of ACPA-negative RA 66% (95% CI: 21-82%).

Table 4. Heritability

Disease (subset)	twin pairs (n)	Heritability
Total RA	148	66% (95% CI: 44-75%)
ACPA-positive RA	114	68% (95% CI: 55-79%)
ACPA-negative RA	34	66% (95% CI: 21-82%)

Contribution of HLA SE alleles

The HLA SE alleles have recently been shown to be specifically associated with ACPA-positive RA [34]. Therefore we set out to investigate if there is a difference in the extent to which the HLA SE alleles contribute to the genetic variance of ACPA-positive versus ACPA-negative RA.

HLA-DRB1 typing was available for 142 twin pairs (96%). The HLA SE alleles were found to be preferentially associated with ACPA-positive RA. Specifically, 109 (109/123

= 89%) of patients with ACPA-positive RA carried one or two SE alleles, compared to 15 (15/32 = 47%) of patients with ACPA-negative RA. The contribution of the HLA SE to RA in general, as well as to ACPA-positive and to ACPA-negative RA was estimated by including this genetic risk factor in the logistic regression model. The prevalence for each genotype was fixed to the specific population prevalence as reported in the United Kingdom [30, 31] or, if not available, in studies from other Caucasian populations [11, 32, 33] (Table 3).

This analysis revealed that the contribution of HLA SE alleles to the total genetic variance of RA is 11% (CI 10-12%) (Table 5). In the ACPA-positive disease subset the HLA SE alleles contribute 18% (CI 16-19%), whereas they only contribute 2.4% (CI: 1.6-10%) to the heritability of ACPA-negative RA.

Table 5. Contribution of HLA SE alleles to genetic variance

Disease (subset)	twin pairs (n)	Contribution of HLA SE alleles
Total RA	142	11% (95% CI: 10-12%)
ACPA-positive RA	113	18% (95% CI: 16-19%)
ACPA-negative RA	29	2.4% (95% CI: 1.6-10%)

DISCUSSION

The present study investigated the heritability of ACPA-positive and ACPA-negative RA. Data from a large RA twin cohort revealed a heritability for both disease subsets of approximately 66%. Furthermore, the contribution of the predisposing HLA SE alleles to ACPA-positive and ACPA-negative RA was assessed. While the HLA SE alleles contribute 18% to the genetic variance of ACPA-positive RA, they contribute only 2.4% to the genetic variance of ACPA-negative RA.

The heritability of ACPA-positive and ACPA-negative RA is similar despite the fact that the concordance of these two disease subsets differs. Both among mono- and dizygotic twins the concordance of ACPA-negative RA is lower than that of ACPA-positive RA. This can be explained by the fact that the concordance, which indicates how often a twin brother or sister of a patient is affected as well, is influenced by the prevalence of the disease in the population, whereas heritability is not. Concordance therefore reflects the prevalence of the disease in the population. Due to the fact that ACPA-negative RA is less prevalent than ACPA-positive RA the concordance of this disease subset is lower, despite similar heritability.

The fact that the prevalence of the disease affects the concordance but not the heritability also serves to explain the seeming discrepancy between the apparently modest monozygotic concordance of 16% and the heritability of 66% for RA. The concordance figure of 16% needs to be interpreted in light of the prevalence of RA which is only 1% [32].

This indicates that genetic factors are important determinants for disease development and is compatible with a high heritability.

The estimate of 66% (95% CI: 44-75%) for the overall heritability of RA is consistent with the results of a previous study that reported an overall heritability of 65% (95% CI: 50-77%) [22]. The previous report used the same twin cohort but a different modeling approach. Both models quantify the genetic and environmental contribution to a dichotomous variable by assuming that there is a continuous underlying liability to disease and that a threshold of liability divides subjects in affected with RA and unaffected. The previous study used a normal distribution for the liability while we used a logistic distribution. The advantage of a logistic distribution is that this model corresponds to the logistic regression model which is most commonly used for binary outcomes in epidemiology.

In twin studies, the variance of continuous traits, such as the underlying liability to disease, is often divided into variance due to environmental and due to genetic effects. In this paper, we estimated only additive genetic variances. Variance components for dominance genetic effects or for shared environmental effects can be added to this model, though it must be kept in mind that in twins these two effects are statistically confounded and cannot be added simultaneously to the same model. When a dominance genetic effect or a shared environmental effect was added to the model, we found that this determined less than 1% of the total variance of RA and of the ACPA-positive and ACPA-negative disease subsets. According to these data, dominance genetic effects or shared environmental factors do not appear to have a substantial effect on the development of RA.

The modeling approach employed in this study uses the population prevalences of the genetic risk factors of interest to estimate their contribution to the genetic variance. To ensure the validity of our estimates we therefore used prevalence figures of the HLA SE alleles reported in large studies of RA patients. While the prevalence of HLA SE alleles in ACPA-positive RA as reported in the literature (84%) was in line with the number observed in the twin study population (89%), the prevalence of HLA SE alleles in the ACPA-negative twins was found to be lower than the figure observed in a large UK RA study (47% versus 59%). To assess the impact of this discrepancy, the modeling was also performed with other assumed population prevalences, all of which consistently revealed a minor contribution of the HLA SE alleles to ACPA-negative RA. The results provided by the current modeling approach therefore appear to be valid irrespective of small uncertainties of prevalence figures.

The HLA SE alleles are, to our current knowledge, the genetic risk factors which confer the highest risk for disease development [13]. In the past, HLA genes have been estimated to contribute 37% to the overall inherited risk of developing RA [35]. Our result that the contribution of the HLA SE alleles to the genetic variance is only 11% (95% CI: 10-12%) for RA and 18% (95% CI: 16-19%) for ACPA-positive RA may therefore be considered surprising. This difference may be due to the fact that the methods used to calculate the

contribution of the HLA SE alleles may not be completely comparable. Assuming that they are measuring similar parameters, a possible explanation for the difference could be that the contribution of SE was overestimated in the past due to the use of selected literature data for the calculation and weaknesses of the method of Rotter & Landaw, as indicated by Deighton et al. in the original publication [35]. On the other hand, we cannot exclude that our figure of 11% contribution of HLA SE alleles to RA may represent an underestimate of the true value, due to the fact that the genotype-specific population prevalences which were required for our modeling, may vary among populations as discussed in the previous paragraph. Probably the most important explanation is however, the fact that a substantial part of the contribution of HLA to the genetic variance could be due to protective genetic factors [36] in addition to the predisposing SE alleles. While we deliberately only considered the predisposing SE alleles, Deighton et al used HLA haplotype sharing for their calculation, and thereby included the effect of protective HLA alleles, which may have resulted in a higher estimate. It is likely that further characterization of genetic risk factors, in particular the protective HLA-DRB1 alleles, and elucidation of interactions will be necessary to fully understand the role of genetic risk factors in disease development.

Recent genome wide association studies and candidate gene approaches have identified several new genetic risk factors for RA [2, 3]. The majority of these genetic risk factors, such as PTPN22 and C5-TRAF1 have been shown to be predominantly associated with ACPA-positive RA just like the HLA SE alleles [34, 37]. This may partly be due to the fact that some of the larger studies have included mainly ACPA-positive patients. One cannot exclude that newer studies which incorporate a more extensive investigation of ACPA-negative RA will show that these risk factors also predispose to ACPA-negative RA [17].

At this moment however, the number of risk factors described to be preferentially associated with ACPA-negative RA is smaller and to our knowledge restricted to the HLA-DR3 haplotype and IRF5 [14, 18]**.** The latter risk factors however do not confer as high a risk for disease development as the HLA SE alleles. The observation that the heritability of ACPA-negative RA is similar to the heritability of ACPA-positive RA is therefore intriguing. This suggests that most genetic risk factors for ACPA-negative RA are still unknown and remain to be discovered. One reason why the description of genetic risk factors for ACPA-negative RA has not advanced at the same pace as for ACPA-positive RA, could be that ACPA-negative RA may be a more heterogeneous disease entity than ACPA-positive RA, with different phenotypes being associated with different genetic risk factors. Another explanation is that many of the cohorts which have been used for large-scale genetics studies have consisted solely of patients with ACPA-positive RA. Because ACPA-negative disease is less prevalent, it is also more challenging to find sufficient numbers of patients to perform a well-powered study.

For the current study, the limited number of ACPA-negative twins also presented a statistical restriction. This resulted in a larger confidence interval for the heritability estimate of

ACPA-negative RA as compared to that of ACPA-positive RA. Confirmation of these results would therefore be helpful to reach a definitive conclusion regarding the heritability of ACPA-negative RA.

In conclusion, ACPA-positive and ACPA-negative RA have a similar heritability of approximately 66%. This means that genetic predisposition also plays an important role in the pathogenesis of ACPA-negative RA, for which most individual genetic risk factors remain to be identified.

APPENDIX A

Logistic regression model used to compute heritability.

We used a logistic regression model with a random effect to model the outcomes of twin pairs. With u corresponding to the random effect which represents all genetic factors that contribute to the outcome and y being 1 if a twin is affected and 0 if it is healthy, the probability to be affected for a single individual is given by the following equation:

$$P(Y=1\,|\,u)=\frac{\exp(\alpha+u)}{1+\exp(\alpha+u)} \tag{1}$$

Given the population prevalence p of the outcome and the genetic variance τ^2 (= the variance of the random effect u), the parameter α can be derived from the following equation :

$$p=\int \frac{\exp(\alpha+u)}{1+\exp(\alpha+u)}\, dF_{\tau}(u) \tag{2}$$

with F the normal distribution of u with mean zero and variance τ^2. Since u is not observed we integrate over its distribution F. Note that due to the non-linear relationship between the prevalence p and the genetic variance τ^2, the absolute value of the parameter α increases with the genetic variance τ^2. This attenuation effect is well known from regression models which take into account errors in variables [28].

Now given τ^2 and α, the log likelihood of the data is computed. Here (y_1,y_2) and (u_1, u_2) represent the pairs of outcomes and of random effects of a twin pair. The correlation of the random effects (u_1,u_2) is fixed at 1 if the twins are monozygotic and 1/2 if they are dizygotic. The log likelihood is then given by

$$l(y_1y_2\mid \alpha,\tau^2)=\log \int \frac{\exp(y_1(\alpha+u_1))}{1+\exp(\alpha+u_1)}\,\frac{\exp(y_2(\alpha+u_2))}{1+\exp(\alpha+u_2)}\, dF_{\tau}(u_1,u_2), \tag{3}$$

with F the bivariate normal distribution of (u_1,u_2). The above steps are carried out for a grid of values of τ^2 and the value τ^2 for which the curve has its maximum is the maximum likelihood estimate of the genetic variance. The heritability is now given by the genetic

variance (τ^2) divided by the total variance of the outcome ($\tau^2 + 3$). Here the value 3 corresponds to the variance of the logistic distribution [38].

APPENDIX B

Model used to compute the contribution of HLA SE to the total genetic variance.

The contribution of the HLA SE alleles to the total genetic variance of RA, of ACPA-positive RA and of ACPA-negative RA is computed as follows. In an extension of the approach explained in appendix A, we used a logistic regression model which now not only estimates the genetic variance but also the amount of this variance which is explained by the HLA SE alleles. To take into account that the risk provided by the HLA SE alleles differs for the development of RA, ACPA-positive RA, and ACPA-negative RA, these varying genotype specific population prevalences were derived from Table 3 using Bayes' theorem. The prevalences were subsequently entered into equation 2 to obtain the genotype specific α (α_{SE+} and α_{SE-}) of the regression model. The variance of the contribution of SE to the outcome equals $(\alpha_{SE+}-\alpha_{SE-})^2 P_{SE} (1-P_{SE})$, and this variance was then included in the logistic regression model alongside with the parameter for the genetic variance. In the logistic regression model including HLA SE, the variance of the random effect (v^2) now measures the contribution of all remaining genetic factors besides HLA SE. Hence the part of the genetic variance explained by SE is given by the variance of the contribution of SE divided by the total genetic variance (τ^2):

$$\frac{(\alpha_{SE+}-\alpha_{SE-})^2 P_{SE}(1-P_{SE})}{(\alpha_{SE+}-\alpha_{SE-})^2 P_{SE}(1-P_{SE})+v^2} = \frac{(\alpha_{SE+}-\alpha_{SE-})^2 P_{SE}(1-P_{SE})}{\tau^2}$$

REFERENCES

1. Begovich AB, Carlton VE, Honigberg LA et al. A missense single-nucleotide polymorphism in a gene encoding a protein tyrosine phosphatase (PTPN22) is associated with rheumatoid arthritis. Am J Hum Genet 2004;75(2):330-337.
2. Kurreeman FA, Padyukov L, Marques RB et al. A candidate gene approach identifies the TRAF1/C5 region as a risk factor for rheumatoid arthritis. PLoS Med 2007;4(9):e278.
3. Remmers EF, Plenge RM, Lee AT et al. STAT4 and the risk of rheumatoid arthritis and systemic lupus erythematosus. N Engl J Med 2007;357(10):977-986.
4. Nienhuis RL, Mandema E. A new serum factor in patients with rheumatoid arthritis; the anti-perinuclear factor. Ann Rheum Dis 1964;23:302-305.
5. Schellekens GA, de Jong BA, van den Hoogen FH et al. Citrulline is an essential constituent of antigenic determinants recognized by rheumatoid arthritis-specific autoantibodies. J Clin Invest 1998;101(1):273-281.
6. Simon M, Girbal E, Sebbag M et al. The cytokeratin filament-aggregating protein filaggrin is the target of the so-called "antikeratin antibodies," autoantibodies specific for rheumatoid arthritis. J Clin Invest 1993;92(3):1387-1393.
7. Young BJ, Mallya RK, Leslie RD et al. Anti-keratin antibodies in rheumatoid arthritis. Br Med J 1979;2(6182):97-99.
8. 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(10):2741-2749.
9. van der Helm-van Mil AH, le Cessie S, van Dongen H et al. A prediction rule for disease outcome in patients with recent-onset undifferentiated arthritis: how to guide individual treatment decisions. Arthritis Rheum 2007;56(2):433-440.
10. van Gaalen FA, Linn-Rasker SP, van Venrooij WJ 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(3):709-715.
11. Nishimura K, Sugiyama D, Kogata Y et al. Meta-analysis: diagnostic accuracy of anti-cyclic citrullinated peptide antibody and rheumatoid factor for rheumatoid arthritis. Ann Intern Med 2007;146(11):797-808.
12. van der Helm-van Mil AH, Verpoort KN, Breedveld FC et al. Antibodies to citrullinated proteins and differences in clinical progression of rheumatoid arthritis. Arthritis Res Ther 2005;7(5):R949-R958.
13. Kallberg H, Padyukov L, Plenge RM et al. Gene-gene and gene-environment interactions involving HLA-DRB1, PTPN22, and smoking in two subsets of rheumatoid arthritis. Am J Hum Genet 2007;80(5):867-875.
14. Verpoort KN, van Gaalen FA, van der Helm-van Mil AH et al. Association of HLA-DR3 with anti-cyclic citrullinated peptide antibody-negative rheumatoid arthritis. Arthritis Rheum 2005;52(10):3058-3062.
15. van der Helm-van Mil AH, Verpoort KN, Breedveld FC et al. The HLA-DRB1 shared epitope alleles are primarily a risk factor for anti-cyclic citrullinated peptide antibodies and are not an independent risk factor for development of rheumatoid arthritis. Arthritis Rheum 2006;54(4):1117-1121.

16. Feitsma AL, Toes RE, Begovich AB et al. Risk of progression from undifferentiated arthritis to rheumatoid arthritis: the effect of the PTPN22 1858T-allele in anti-citrullinated peptide antibody positive patients. Rheumatology (Oxford) 2007;46(7):1092-5.
17. Orozco G, Alizadeh BZ, Gado-Vega AM et al. Association of STAT4 with rheumatoid arthritis: a replication study in three European populations. Arthritis Rheum 2008;58(7):1974-1980.
18. Sigurdsson S, Padyukov L, Kurreeman FA et al. Association of a haplotype in the promoter region of the interferon regulatory factor 5 gene with rheumatoid arthritis. Arthritis Rheum 2007;56(7):2202-2210.
19. 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(1):38-46.
20. 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(3):366-371.
21. Klareskog L, Ronnelid J, Lundberg K et al. Immunity to citrullinated proteins in rheumatoid arthritis. Annu Rev Immunol 2008;26:651-675.
22. Macgregor AJ, Snieder H, Rigby AS et al. Characterizing the quantitative genetic contribution to rheumatoid arthritis using data from twins. Arthritis Rheum 2000;43(1):30-37.
23. Silman AJ, Macgregor AJ, Thomson W et al. Twin concordance rates for rheumatoid arthritis: results from a nationwide study. Br J Rheumatol 1993;32(10):903-907.
24. 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(11):1205-1213.
25. 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(3):315-324.
26. Jeffreys AJ, Wilson V, Thein SL. Individual-specific 'fingerprints' of human DNA. Nature 1985;316(6023):76-79.
27. Bijkerk C, Houwing-Duistermaat JJ, Valkenburg HA et al. Heritabilities of radiologic osteoarthritis in peripheral joints and of disc degeneration of the spine. Arthritis Rheum 1999;42(8):1729-1735.
28. Carroll RJ RDSLCC. Measurement error in nonlinear models: A modern Perspective. 2nd edition ed. Chapman & Hall; 2006.
29. Houwing-Duistermaat JJ, van Houwelingen HC, de Winter JP. Estimation of individual genetic effects from binary observations on relatives applied to a family history of respiratory illnesses and chronic lung disease of newborns. Biometrics 2000;56(3):808-814.
30. Mewar D, Coote A, Moore DJ et al. Independent associations of anti-cyclic citrullinated peptide antibodies and rheumatoid factor with radiographic severity of rheumatoid arthritis. Arthritis Res Ther 2006;8(4):R128.
31. Thomson W, Harrison B, Ollier B et al. Quantifying the exact role of HLA-DRB1 alleles in susceptibility to inflammatory polyarthritis: results from a large, population-based study. Arthritis Rheum 1999;42(4):757-762.
32. Gabriel SE. The epidemiology of rheumatoid arthritis. Rheum Dis Clin North Am 2001;27(2):269-281.
33. van Venrooij WJ, Zendman AJ, Pruijn GJ. Autoantibodies to citrullinated antigens in (early) rheumatoid arthritis. Autoimmun Rev 2006;6(1):37-41.

34. 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(11):3433-3438.
35. Deighton CM, Walker DJ, Griffiths ID et al. The contribution of HLA to rheumatoid arthritis. Clin Genet 1989;36(3):178-182.
36. van der Helm-van Mil AH, Huizinga TW, Schreuder GM et al. An independent role of protective HLA class II alleles in rheumatoid arthritis severity and susceptibility. Arthritis Rheum 2005;52(9):2637-2644.
37. Wesoly J, van der Helm-van Mil AH, Toes RE et al. Association of the PTPN22 C1858T single-nucleotide polymorphism with rheumatoid arthritis phenotypes in an inception cohort. Arthritis Rheum 2005;52(9):2948-2950.
38. Bickel PJ, Doksum KA. Mathematical Statistics. Holden-Day, Oakland, California; 1977.