The course of clinically suspect arthralgia and early rheumatoid arthritis : clinical features, imaging and genetics

Steenbergen, H.W. van

Part II

Genetic factors and disease outcome in rheumatoid arthritis

Predicting the severity of joint damage in rheumatoid arthritis; the contribution of genetic factors

Hanna van Steenbergen, Roula Tsonaka, Tom Huizinga, Saskia le Cessie, Annette van der Helm-van Mil

8

Ann Rheum Dis 2015;74:876-82

ABSTRACT

Introduction

The severity of radiologic progression is variable between rheumatoid arthritis (RA) patients. Recently, several genetic severity variants have been identified and were replicated, these belong to 12 loci. This study determined the contribution of the identified genetic factors to the explained variance in radiologic progression and whether genetic factors, in addition to traditional risk factors, improve the accuracy of predicting the severity of radiologic progression.

Methods

426 early RA patients with yearly radiologic follow-up were studied. The main outcome measure was the progression in Sharp-van der Heijde score (SHS) over 6 years, assessed as continuous outcome or categorised in no/little, moderate or severe progression. Assessed were improved fit of a linear mixed model analysis on serial radiographs, R^2 using linear regression analyses, C-statistic and the net proportion of patients that was additionally correctly classified when adding genetic risk factors to a model consisting of traditional risk factors.

Results

The genetic factors together explained 12–18%. When added to a model including traditional factors and treatment effects, the genetic factors additionally explained 3–7% of the variance (p-value $_{R^2\ change}$=0.056). The percentage of patients that was correctly classified increased from 56% to 62%; the net proportion of correct reclassifications 6% (95% CI 3 to 10%). The C-statistic increased from 0.78 to 0.82. Sensitivity analyses using imputation of missing radiographs yielded comparable results.

Conclusion

Genetic risk factors together explained 12–18% of the variance in radiologic progression. Adding genetic factors improved the predictive accuracy, but 38% of the patients were still incorrectly classified, limiting the value for use in clinical practice.

INTRODUCTION

The severity of rheumatoid arthritis (RA) is commonly expressed by the extent of damage of hand and feet joints. Joint damage can be measured objectively with validated scoring methods and is associated with long-term functional disability [1]. The severity is highly variable between patients; many patients show mild progression and few severe progression. The processes underlying these differences are partly understood. The observation that the heritability of radiologic progression is 45-58% [2] underlined the notion that genetic factors play a role. Presently, several genetic risk factors for radiologic progression have been identified and replicated. Some of these variants were also associated with differences in mRNA or protein expression [3–6]. Here, we aimed to explore the relevance of currently known genetic risk factors with regards to (1) explaining the interindividual variance in radiologic progression and (2) improving the accuracy of predicting radiologic progression for individual patients.

Known traditional risk factors explain about one-third of the variance in joint damage after 5 years of disease; the majority of these risk factors were related to patient characteristics (age, gender), inflammation (acute phase reactants, swollen joint counts) and the presence of auto-antibodies [7]. The contribution of the genetic risk factors to the explained variance has not been explored.

Prediction of RA severity on the level of individual patients is not yet accurate. Several matrices to predict rapid radiologic progression have been derived, consisting of three or four risk factors. Most of these matrices are not validated in the general RA population, and failed to correctly classify ~50% of patients. In particular, the patients who developed progressive disease were not recognized [8–13]. Consequently, the value of these matrices for clinical practice is still limited. Whether the addition of genetic factors improves prediction is unknown.

This study examined the variance in joint damage progression explained by recently identified genetic risk factors and their value in improving the prediction of the severity of joint damage progression. We assessed traditional performance measures of prediction models and the net proportion of RA patients that is additionally correctly classified when adding risk factors to a prediction model consisting of known risk factors.

PATIENTS AND METHODS

Patients

Between 1993 and 2006, 600 RA patients (1987-ACR-criteria) were included in the Leiden Early Arthritis Clinic (EAC) [7]. Inclusion in the EAC took place when arthritis was confirmed at physical examination and symptom duration was less than 2 years. At first visit, patients and rheumatologists filled questionnaires, 66-swollen and 68-tender joint counts were performed (66-SJC and 68-TJC [14]), and blood samples were taken. Patients were followed yearly. The initial treatment strategy differed for different inclusion periods: patients included in 1993–

1995 were initially treated with NSAIDs, patients included in 1996–1998 were initially treated with hydroxychloroquine or sulphasalazine, and patients included since 1999 were promptly treated with methotrexate. The severity of radiologic progression differed for these three treatment groups; therefore, treatment effects were incorporated in the analyses. The traditional risk factors studied were age, gender, symptom duration at first visit, localisation initial joint symptoms, 66-SJC, presence of anti-citrullinated peptide antibodies (ACPA), presence of rheumatoid factor (RF) and erythrocyte sedimentation rate (ESR).

Selection of genetic risk factors and genotyping

We selected single nucleotide polymorphisms (SNPs) using the following criteria: the SNP was studied in relation to the severity of radiologic progression in several cohorts and the association was independently replicated or found significant in a meta-analysis including all published data. Based on these criteria, we came to a selection of genetic variants that is presented in table 1. Notably, rs4810485 in *CD40* and rs7607479 in *SPAG16* were identified as risk factors for radiologic progression only in ACPA-positive RA. Genotypings in the EAC

Table 1. Genetic variants studied and the R2 of each variant for radiologic progression over six years.

Genetic variant(risk allele)	Located in/nearby gene(s)	Chr.	MAF*	Tested Model*	R^2 $\Delta SHS_{0\text{-}6years}$ (%) in RA (n=239)	R^2 $\Delta SHS_{0\text{-}6years}$ (%) in ACPA-pos RA (n=144)
SE [29]	*HLA-DRB1*	6	0.39	add	4.0	<0.01
rs4810485 (T) [15]	*CD40*	20	0.24	rec	0.1	<0.01
rs7667746 (G) [16]	*IL-15*	4	0.33	rec	2.6	3.9
rs7665842 (G) [16]	*IL-15*	4	0.40	rec	2.7	3.7
rs4371699 (A) [16]	*IL-15*	4	0.19	rec	0.3	1.0
rs6821171 (C) [16]	*IL-15*	4	0.29	rec	0.1	1.4
rs1896368 (G) [4]	*DKK-1*	10	0.47	add	0.3	1.1
rs1896367 (A) [4]	*DKK-1*	10	0.41	add	0.4	0.7
rs1528873 (A) [4]	*DKK-1*	10	0.47	add	2.1	3.0
rs2104286 (C) [18]	*IL2RA*	10	0.24	add	0.3	<0.1
rs8192916 (A) [3]	*GRZB*	14	0.42	rec	0.8	1.4
rs1119132 (A) [17]	*IL-4R*	16	0.13	rec	0.5	1.1
rs7607479 (C) [6]	*SPAG16*	2	0.33	add	0.6	2.5
rs26232 (T) [19]	*C5orf30*	5	0.29	add	0.3	<0.1
rs11908352 (A) [5]	*MMP-9*	20	0.21	add	4.7	1.3
rs451066 (A) [5]	rs1465788	14	0.20	add	1.1	0.2
rs1485305 (T) [30]	*OPG*	8	0.44	add	1.4	0.6

The presented R^2s were based on univariable analyses for each individual risk factor. *The MAFs and tested models are presented as reported in the previous studies. MAF=minor allele frequency; R^2= proportion of explained variance; $\Delta SHS_{0\text{-}6years}$= progression in Sharp-van der Heijde score over six years; add=additive; rec=recessive; SE=shared epitope.

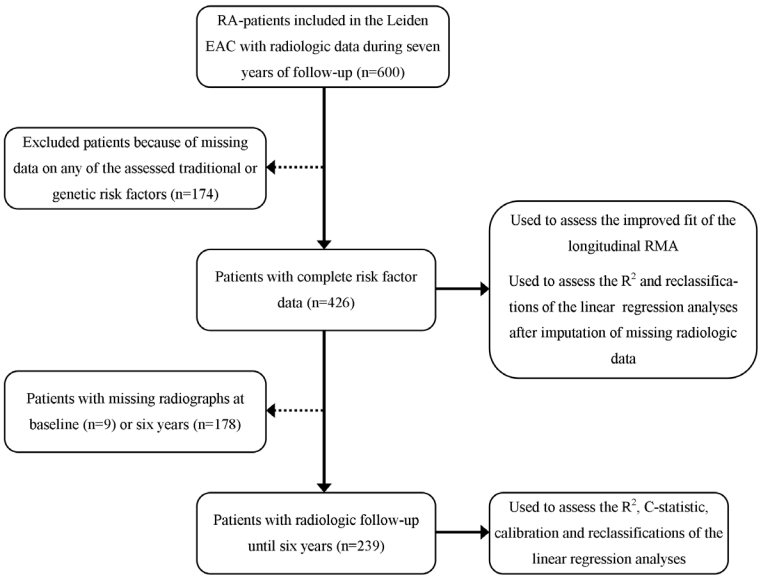


Figure 1. Flow chart of patient selection.
Baseline characteristics of the included (n=426) and excluded patients (n=174) were not different (data not shown). The patients with follow-up until six years (n=239) were younger compared to the patients without complete follow-up until six years (n=187) (mean (SD) 53.9 (14.5) versus 60.0 (15.7) years, p<0.001), had a higher 66-SJC (median (IQR) 9 (5–16) versus 8 (3–13), p=0.009) at baseline and were more frequent ACPA-positive (60.3% vs 44.4%, p=0.001). RMA=repeated measurement analysis; R^2=proportion of explained variance.

were done with allele-specific kinetic PCR analysis [15], Illumina Golden Gate platform [3,4,16,17], Illumina Immunochip [5,18], Sequenom iPLEX [6] and LightSnp (Roche) [19]. Quality control of genotyping was performed as described previously [3–6,15–19]. 426 patients had complete data on all evaluated traditional and genetic risk factors (figure 1).

Radiologic outcome

X-rays were taken at baseline and with yearly intervals. Totally, over 7 years, 2680 X-ray sets of hands and feet of 426 patients were made and scored by one experienced reader using Sharp-van der Heijde scores (SHSs) blinded to any clinical or genetic data (intraclass correlation coefficient 0.91). The numbers of patients with radiologic data at baseline and over 4, 5, 6 and 7 years were, respectively, 321, 286, 239 and 206. The main outcome measure in this study was radiologic progression in the first 6 years after inclusion ($\Delta SHS_{0-6\ years} = SHS_{6\ years} - SHS_{0\ years}$). Although radiologic data was known up till 7 years, the progression over 6 years was chosen as main outcome as fewer patients completed 7 years of follow-up. For some accuracy measures the continuous outcome was categorised in three groups of progression over 6 years: $\Delta SHS_{0-6\ years}$ ≤6, 7–30 and >30 units, indicating no/little, moderate and severe radiologic progression (figure 3A). The first cut-off was chosen as progression of ≤1 SHS-unit per year is minimal; the latter cut-off was chosen because rapid radiologic progression is generally defined as an increase of 5 SHS-units per year [8–10]. In all analyses, the difference in

SHS was log10-transformed to approximate a normal distribution.

Analyses

The performance of prediction models can be evaluated using different aspects, see table 2 [20,21]. Inherent to the method of determining these aspects, the radiologic progression rate over 6 years was studied as a continuous or categorised outcome.

Improved fit - First, a linear mixed model analysis was used with serial log10-transformed SHS over 6 years as response variable and time and risk factors as variables.

Table 2. Different measures to evaluate the performance of prediction models; inherent to the statistical method used, progression over six years was assessed as a continuous or categorical outcome variable.

Aspect and measure	Characteristics	Used model and radiologic outcome
Overall performance: 'goodness-of-fit' of model, quantification of how close predictions are to the observed outcome; captures both aspects of calibration and discrimination		
Improved fit	Improved fit of model after adding additional variables to the model (%). Measured as relative increase in Nagelkerke R^2 (modified version of Cox and Snell's pseudo R2) 23.	Linear mixed model analysis with yearly SHS scored X-rays over six years as outcome. Patients with missing radiographs at a certain time point were included.
R^2	Variance in outcome explained by the included variables. The explained variance can be corrected for the number of variables in the model (adjusted R^2) (%).	Linear regression analysis with radiologic progression between baseline and six years ($\Delta SHS_{0\text{-}6years}$) on a continuous scale as outcome. Analyses are done on patients with complete data (n=239) and on all patients (n=426) when imputing missing radiological data.
Discrimination: ability to discriminate between those with and without the outcome		
C-statistic	Assessing pairs of patients where one has more severe outcome than the other, it reflects the fraction of patients where those with the more severe outcome have higher predictions than those with the less severe outcome [24].	Linear regression analyses where the predicted $\Delta SHS_{0\text{-}6years}$ is compared with the observed $\Delta SHS_{0\text{-}6years}$ categorized in no/little, moderate and severe progression ($\Delta SHS_{0\text{-}6years}$ ≥6, 7-30 and >30).
Calibration: agreement between observed and predicted outcomes		
Calibration	Scatterplot with predicted outcome on the x-axis and observed outcome on the y-axis.	Scatter plot of observed versus predicted progression over six years ($\Delta SHS_{0\text{-}6years}$), both as continuous outcomes.
Reclassification: ability to reclassify patients by adding predictors to the model		
Net correct reclassification	Comparing the predicted classification with the observed classification when using two models; assessed is the net change in the correct direction (correct minus incorrect reclassifications).	Linear regression analyses in which the predicted $\Delta SHS_{0\text{-}6years}$ is calculated. Then both the observed and predicted $\Delta SHS_{0\text{-}6years}$ are categorized in no/little, moderate and severe progression ($\Delta SHS_{0\text{-}6years}$ ≥6, 7-30 and >30). Analyses were done on patients with complete data (n=239) and on all patients (n=426) when imputing missing radiological data.

R^2=proportion of explained variance; $\Delta SHS_{0\text{-}6years}$=progression in Sharp-van der Heijde score over six years

The ARH1 covariance matrix was used as suggested previously by Knevel et al [22]. Valuable of this repeated measurement analysis (RMA) is that it takes advantage of within-patient correlations of serial X-rays and allows the inclusion of patients with missing X-rays at certain time-points (figure 1). The improved fit of the model when adding treatment effects, traditional risk factors, genetic risk factors or combinations of these to a model consisting of only time effect was measured as the relative increase in the Nagelkerke R^2 between the models with the risk factors of interest and with only the time effect [23]. Importantly, this is not a direct measure of the explained variance, which cannot be determined in RMA such as linear mixed model analysis. Therefore, the R^2 was subsequently determined in linear regression analyses.

R^2 - This reflects the absolute proportion of the variance that is explained by the factors in the model and was determined using linear regression analyses with $\Delta SHS_{0-6\ years}$ as outcome. A limitation of this outcome is that only patients with complete follow-up could be studied (figure 1). Regression models were fitted that included treatment effects, traditional risk factors, genetic risk factors or combinations of these. Because adding more variables to a model will increase the fit of a model and thus the R^2, the adjusted R^2 was also calculated. This includes a correction for the number of variables in the regression model.

C-statistic - Harrel's C-statistic was assessed as described in the online supplementary methods [24]. It reflects the accuracy of discriminating patients with and without the outcome and does not reflect the absolute risk on an outcome. For clinical risk prediction it is more relevant that a new model can more accurately stratify individuals into risk categories. Hence, calibration (agreement between observed and predicted outcomes) and reclassification have gained popularity [25,26].

Calibration and Reclassification - First, the observed ΔSHS0–6 years was plotted against the $\Delta SHS_{0-6\ years}$ that was predicted by linear regression models including treatment effects and traditional factors or including treatment effects, traditional and genetic factors (calibration plot). Then the actual observed $\Delta SHS_{0-6\ years}$ and the predicted $\Delta SHS_{0-6\ years}$ were categorised in three severity groups ($\Delta SHS_{0-6\ years}$ ≤6, 7–30 and >30 units). To assess the improvement in predictive performance gained by adding genetic information to the prediction model, the proportion of patients that was correctly reclassified (correct reclassifications minus incorrect reclassifications) was determined. This was done for the total population and for each severity group separately.

Sensitivity analyses

The R^2 depends on the variance of the outcome. Therefore, the R^2 may change in case other follow-up durations are studied. To assess the influence of this effect, the R^2 was also determined for radiologic progression over 4, 5 and 7 years ($\Delta SHS_{0-4\ years}$, $\Delta SHS_{0-5\ years}$ and $\Delta SHS_{0-7\ years}$).

Only 239 of the 426 patients had complete radiologic data till year 6. As missing was not completely at random, we repeated the linear regression analyses with missing radiologic data imputed. We performed single conditional mean imputation by replacing missing values with the values predicted by the RMA with SHSs over 7 years of disease as outcome, and time and all traditional and genetic risk factors as variables. Subsequently, the R^2 and reclassifications were again determined.

Some of the genetic variants were identified as risk factors for radiologic progression in ACPA-positive RA, the more severe subset of RA. Therefore, the analyses of R^2 and reclassifications with $\Delta SHS_{0-6\ years}$ as outcome were repeated in the ACPA-positive subgroup (n=144).

Analyses were performed using SPSS V.20.0 and Stata V.12.

RESULTS

Patients and traditional risk factors

Patient characteristics are presented in table 3. The median SHS at baseline was 5.0 (IQR 2.0–10.0) and at year 6 it was 22.3(IQR 9.0–47.0); the median $SHS_{0-6\ years}$ was 14.0 (IQR 4.5–39.0). Treatment effects explained 7.1% of the variance of radiologic progression over 6 years (ΔSHS0–6 years). The R^2 of the individual traditional risk factors, determined in univariable regression analyses showed the highest values for ACPA and RF (R2 22.8% and 19.4%, respectively, table 3). All traditional risk factors together explained 31.2% of the variation and treatment effects and traditional risk factors combined explained 36.5% (figure 2A, see online supplementary table S2A). The adjusted R^2s were respectively 28.5% and 33.7% (figure 2D).

Genetic risk factors

Improved fit

First, all radiologic data of 426 patients were assessed using RMA. Models without and with genetic risk factors were compared, revealing that the model including the genetic risk factors had a 3.2% better fit in predicting radiologic progression com-pared to a model including only treatment effects and traditional risk factors (see online supplementary table S1). Since this measure is difficult to interpret, we continued with determining the R^2.

R^2

The R^2 of individual genetic risk factors was determined in univariable analyses (table 1). Rs11908352 in MMP-9 and the human leucocyte antigens-shared epitope (HLA-SE) alleles had the largest R^2 (4.7% and 4.0%, respectively). All genetic risk factors together explained 18.1% of the variance in $\Delta SHS_{0-6\ years}$ (figure 2B). The adjusted R^2 was 11.8% (figure 2E). Next, it was studied to what extent the genetic risk factors increased the R^2 compared to a model including treatment effects and traditional risk factors. A model including all factors (treatment effects, traditional and genetic risk factors) resulted in an R^2 of 43.9% and adjusted

Table 3. Characteristics of patients and the R^2 of each individual characteristic for radiologic progression over six years.

	All patients (n=426)	R^2 $\Delta SHS_{0\text{-}6years}$ (%) in RA (n=239)	R^2 $\Delta SHS_{0\text{-}6years}$ (%) ACPA-pos RA (n=144)
Age, mean (sd), years	56.6 (15.3)	<0.1	<0.1
Female gender, n (%)	290 (68.1%)	1.7	0.9
Symptom duration at first visit, median (IQR), months	4.4 (2.4-8.6)	2.3	0.7
Localization initial joint symptoms		<0.1	0.1
Upper extremities, n (%)	204 (47.9%)		
Lower extremities, n (%)	57 (13.4%)		
Upper and lower extremities, n (%)	165 (38.7%)		
66-SJC, median (IQR), n	8 (4-14)	2.0	<0.1
BMI, median (IQR), n	25.4 (23.0-27.6)	3.1	1.8
ACPA-positive, n (%)	227 (53.3%)	22.8	-
IgM-RF positive, n (%)	248 (58.2%)	19.4	0.3
ESR, median (IQR), mm/h	33.0 (18.0-55.0)	2.7	2.0

The presented R^2 s were based on univariable analyses of each individual risk factor. 239 patients of the total included 426 patients completed follow-up until six years, 144 of these patients were ACPA-positive. R^2=proportion of explained variance; $\Delta SHS_{0\text{-}6years}$=progression in Sharp-van der Heijde score over six years.

R^2 of 36.7%. As the R^2 of the model, including treatment and traditional risk factors was 36.5%, the increase in the R^2 by genetic risk factors was 7.4% (p-value $_{R^2\ change}$=0.056, figure 2C, see online supplementary table S2A). When comparing adjusted R^2s, genetic factors increased the R^2 with 3.0% (figure 2F).

C-statistic

The C-statistic increased from 0.78 (95% CI 0.73 to 0.82) for a model with treatment and traditional factors to 0.82 (95% CI 0.77 to 0.86) for a model including treatment, traditional and genetic factors.

Calibration and reclassification

Observed progression rates were plotted against predicted progression rates by a linear regression model with treatment effects and traditional risk factors as variables. When categorising patients in three groups ($\Delta SHS_{0-6\ years}$ ≤6, 7–30 and >30 units) 134 of 239 patients (56.1%) were correctly classified. When genetic factors were added, 148 out of 239 patients (61.9%) were correctly classified by the model (figure 3B,C). Hence in total this concerned a net increase in correctly classified patients (proportion of correct reclassifications) of 5.9% (95% CI 3.2 to 9.6%). Evaluating the reclassifications per severity group, showed no net change for the group with no/little progression, a 5.1% net increase in correctly classified

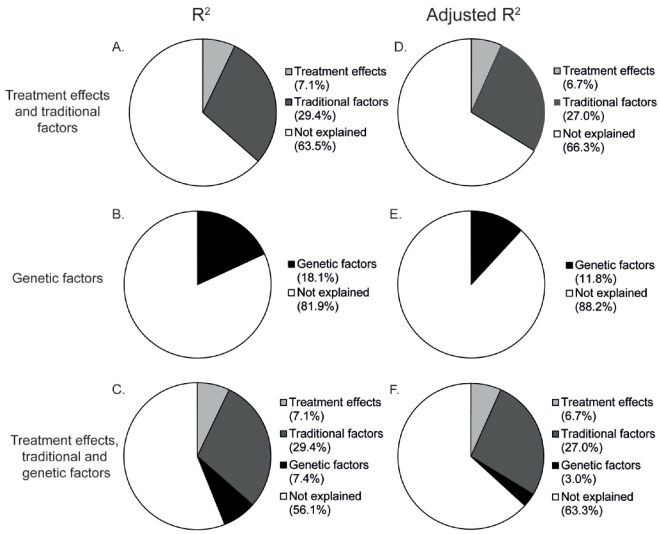


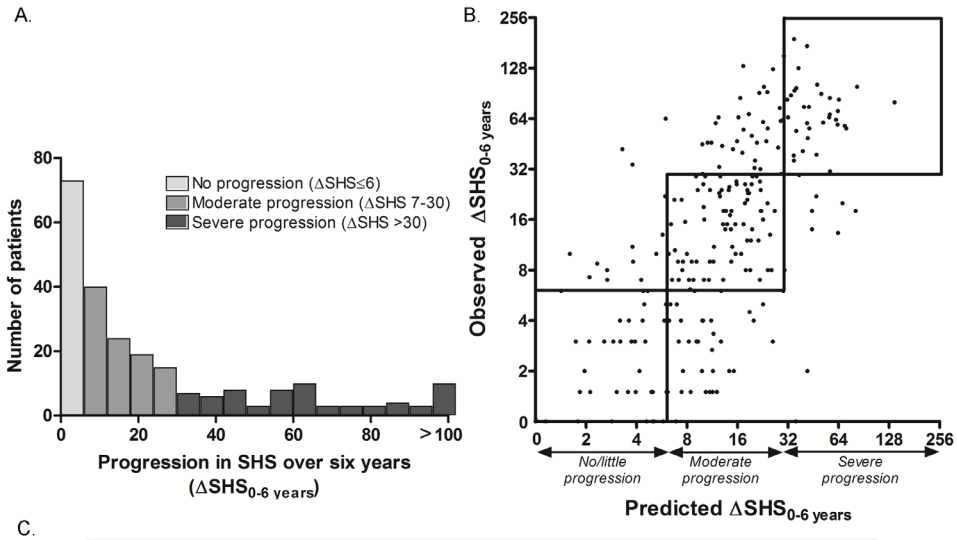
Figure 2. (A-C) Proportions of explained variance (R^2) (D–F) and adjusted explained variance (adjusted R^2) in progression in Sharp-van der Heijde score over six years ($\Delta SHS_{0-6\ years}$) by treatment effects and traditional risk factors (A and D), genetic risk factors (B and E), and treatment effects, traditional and genetic risk factors (C and F). The treatment strategy differed for different inclusion period. Therefore, the effects of treatment were determined before adding traditional and genetic factors (A, C, D and F). The studied traditional risk factors are presented in table 3 and included age, gender, symptom duration at first visit, localisation initial joint symptoms, 66-SJC, BMI, ACPA-positivity, RF-positivity and ESR. The studied genetic risk factors are presented in table 1 and included genetic variants in *HLA-DRB1, CD40, IL-15, DKK-1, IL2RA, GRZB, IL-4R, SPAG16, C5orf30, MMP-9* and *OPG*. The data presented are based on the patients with complete data over six years (n=239). Analyses on all patients after imputation of missing data (n=427) revealed similar results, see online supplementary table S2B. The p-value for change in R^2 after adding genetic factors was 0.056 for patients with complete data.

patients for the group with moderate progression, and a 13.2% net increase in correctly classified patients for the group with severe progression (figure 3C, see online supplementary table S3). Thus, the proportion of patients that was correctly reclassified when adding genetic factors increased in particular in the most severe patient group.

Sensitivity analyses

To check for consistency, $\Delta SHS_{0-4\ years}$, $\Delta SHS_{0-5\ years}$, and $\Delta SHS_{0-7\ years}$ were also assessed as outcomes. Adding genetic risk factors to a model with treatment effects and traditional risk factors yielded an increase in R^2 of 5.5% for $\Delta SHS_{0-4\ years}$ (p-value $_{R^2\ change}$=0.085), 7.1% for $\Delta SHS_{0-5\ years}$ (p-value $_{R^2\ change}$ =0.035) and 9.8% for $\Delta SHS_{0-7\ years}$ (p-value $_{R^2\ change}$=0.026) (see online supplementary table S4).

When missing SHSs were imputed and all 426 patients were studied, the increase in R^2 when adding genetic factors to a model with treatment and traditional risk factors and $\Delta SHS_{0-6\ years}$ as outcome was 5.3% (p-value $_{R^2\ change}$=0.001) (see online supplementary table S2B). The net proportion of patients that was correctly reclassified was 5.4% (95% CI 3.5 to 8.0%); for the groups with no/little, moderate and severe progression these were respectively 1.3%, 6.6% and 9.5%. The proportion of correctly classified patients was 286/426 (67.1%) (see



C.

	Observed progression over 6 years			
	No (ΔSHS ≤6)	Moderate (ΔSHS 7-30)	Severe (ΔSHS >30)	Total
Predicted progression over 6 years by model without genetic risk factors				
No (ΔSHS ≤6)	**34**	15	1	50
Moderate (ΔSHS 7-30)	34	**75**	42	150
Severe (ΔSHS >30)	4	9	**25**	39
Predicted progression over 6 years by model with genetic risk factors				
No (ΔSHS ≤6)	**34**	11	3	48
Moderate (ΔSHS 7-30)	36	**80**	31	147
Severe (ΔSHS >30)	2	8	**34**	44
Total	72	99	68	239

Figure 3. (A) Distribution of observed progression in Sharp-van der Heijde score over six years ($\Delta SHS_{0-6\ years}$), (B) observed versus predicted $\Delta SHS_{0-6\ years}$ by a model consisting of treatment effects, traditional and genetic risk factors and (C) numbers of patients per categorized observed and predicted $\Delta SHS_{0-6\ years}$ by models without and with genetic risk factors, resulting in the net proportion of correct reclassifications. (B) The dots in the boxes represent the 148 of the 239 patients in whom the severity of radiologic progression over six years was correctly predicted by the model, including treatment effects, traditional and genetic risk factors. (C) The model without genetic risk factors correctly classified 134 of 239 patients (56.1%) and the model with genetic risk factors correctly classified 148 of 239 patients (61.9%), resulting in a total net proportion of correct reclassifications of 5.8% (95% CI 3.2 to 9.6%). Evaluating reclassifications per severity group showed, respectively, no net change (5 correct and 5 incorrect reclassifications, 0/72), a 5.1% net increase (10 correct and 5 incorrect reclassifications, 5/99) and a 13.2% net increase (14 correct and 5 incorrect reclassification, 9/68) in correct classifications for the groups with no/little, moderate and severe progression (see also online Supplementary table S3).

online supplementary table S4).

In the subset of ACPA-positive patients, the median SHS at year 6 was 32.5 (IQR 17.3–65.8), and the median $\Delta SHS_{0-6\ years}$ 24.0 (IQR 10.6–57.5). The genetic factors together explained 17.1% of the variance in $\Delta SHS_{0-6\ years}$. Adding genetic factors to a model already including treatment effects and traditional risk factors increased the R^2 with 15.1% (p-value

R^2_{change}=0.11, see online supplementary table S5). The net proportion of correctly reclassified patients was 4.9% (95% CI 2.0 to 9.8%); for three severity groups, these were 0%, 3.1% and 8.6%. The model including all factors classified 91/144 (63.2%) of the ACPA-positive patients correctly (see online supplementary table S6).

DISCUSSION

New genetic risk factors for radiologic progression in RA have been identified recently. This study evaluated how much of the variance in radiologic progression is explained by these genetic factors together and whether these genetic factors improve predicting the severity of the disease course. We observed that genetic risk factors together explained 12–18% of the variance in joint destruction, and that adding the genetic factors to a prediction model already consisting of treatment effects and traditional risk factors resulted in a net increase of correctly classified patients of 6%. This increase was largely due to improved identification of patients with severe progression. Based on the Icelandic RA population, the heritability of radiologic progression was estimated at 45–58% [2]. Our observation that studied genetic factors explained around 18% suggest that part of the heritability is still missing. Several explanations may account for this. Part of the relevant genetic variants may still be unidentified or gene-gene interactions may play a role. The heritability in the Icelandic and Dutch RA population may also be dissimilar, prohibiting a direct comparison of percentages.

Adding genetic factors to a model with known risk factors had a small but independent contribution (3–7%) to the explained variance in radiologic progression. An explanation that this increase is less than the 12–18% of variation found for genetics alone is that part of the genetic factors are associated with traditional risk factors that were already included in the model. Probably these genetic factors relate to the outcome by mediating through these traditional risk factors and, therefore, they do not contribute to the model when the intermediate risk factors are also included. This observation differs from previous observations done for RA susceptibility where identified genetic susceptibility factors did not contribute independently to predicting the development of RA using a model with traditional factors, among which is ACPA [27]. The variants that had the largest independent contribution to the increase in R^2 were rs1528873 (*DKK-1*), rs7607479 (*SPAG16*) and rs11908352 (*MMP-9*) (data not shown). Intriguingly, all these proteins are involved in bone metabolism or cartilage destruction, processes that were not represented by the assessed traditional factors. Notably, due to the strong correlation between ACPA and HLA-SE, adding only HLA-SE to a model already containing ACPA was not helpful (R^2 change 0.1%, p-value $_{R^2_{change}}$=0.63). Conversely, the R^2 change when adding the non-HLA variants to the model including traditional factors was 7.3% (p-value $_{R^2_{change}}$=0.045).

The existing prediction matrices for rapid radiologic progression consist of a few traditional risk factors, were developed in a selected set of severe RA patients, could not adequately classify ~50% of the patients and had difficulties with identifying the patients

with severe progression in particular [8–13]. We evaluated nine traditional factors in a general population of RA patients, and observed that also here, 46% of RA patients were incorrectly classified. When evaluating traditional and genetic factors 62% of RA patients were correctly classified and 38% misclassified. Assuming that clinicians prefer to have at least 80% of the patients correctly predicted, the derived models including genetic variants were still insufficient for use in clinical practice. Importantly, with the help of genetic factors, the correct identification of especially those RA patients with severe radiologic progression increased.

We have chosen to study genetic variants that were replicated in independent studies or found significant in meta-analysis including all published data. Variants that were associated with radiologic progression in only one or two cohorts but not replicated or significant in meta-analyses were not included [28]. Potentially, future research will reveal more severity factors for RA and might increase the predictive accuracy.

Because of the negative implication of our conclusion, we did not seek for external validation or internal validation using cross-validation. The observed R^2 values may have been overestimated as many variables were included. Controlling for overfitting was done by determining the adjusted R^2 (correcting for the number of variables). However, some variables were correlated (for SNPs the correlation coefficients were <0.8) and, consequently, the correction may have been too stringent and the adjusted R^2 values underestimated. Presumably, the actual explained variance lies between the presented R^2 and adjusted R^2 values.

Several sensitivity analyses were done to check for the consistency of the results on the R2. Because missing radiologic data may be due to selection bias, analyses were also repeated after imputation of missing radiographs. Nonetheless, the consistent results in all sensitivity analyses indicate the reliability of our results.

Because some of the genetic risk factors were identified in ACPA-positive RA, we also performed subanalyses on ACPA-positive patients. Compared to the total RA population, the R^2 of the traditional risk factors was smaller (this may be explained by absence of the effect of ACPA) and the increase in R^2 when adding genetic factors was larger. Importantly, the R^2 values between the total and ACPA-positive population cannot be directly compared, as the total variance in $\Delta SHS_{0–6\ years}$ differed. The number of ACPA-positive patients was relatively small, providing another limitation.

In conclusion, all genetic severity factors together explained 12–18% of the variance in radiologic progression. Additional use of genetic factors resulted in increased correct classification of patients in severity risk groups. Nonetheless, 38% of the patients were still not correctly classified. Therefore, we considered the predictive performance of the derived prediction model insufficient for use in clinical practice.

SUPPLEMENTARY DATA

Supplementary data are published on the website of the *Annals of the Rheumatic Diseases*.

REFERENCES

1. Ødegård S, Landewé R, van der Heijde D, et al. Association of early radiographic damage with impaired physical function in rheumatoid arthritis: A ten-year, longitudinal observational study in 238 patients. Arthritis Rheum 2006;54:68–75.
2. Knevel R, Gröndal G, Huizinga TW, et al. Genetic predisposition of the severity of joint destruction in rheumatoid arthritis: a population-based study. Ann Rheum Dis 2012;71:707–9.
3. Knevel R, Krabben A, Wilson AG, et al. A genetic variant in granzyme B is associated with progression of joint destruction in rheumatoid arthritis. Arthritis Rheum 2013;65:582–9.
4. De Rooy DP, Yeremenko NG, Wilson AG, et al. Genetic studies on components of the Wnt signalling pathway and the severity of joint destruction in rheumatoid arthritis. Ann Rheum Dis 2013;72:769–75.
5. De Rooy DP, Zhernakova A, Tsonaka R, et al. A genetic variant in the region of MMP-9 is associated with serum levels and progression of joint damage in rheumatoid arthritis. Ann Rheum Dis 2014;73:1163–9.
6. Knevel R, Klein K, Somers K, et al. Identification of a genetic variant for joint damage progression in autoantibody-positive rheumatoid arthritis. Ann Rheum Dis 2014;73:2038–46.
7. De Rooy DP, van der Linden MP, Knevel R, et al. Predicting arthritis outcomes—what can be learned from the Leiden Early Arthritis Clinic? Rheumatology 2011;50:93–100.
8. Visser K, Goekoop-Ruiterman YP, de Vries-Bouwstra JK, et al. A matrix risk model for the prediction of rapid radiographic progression in patients with rheumatoid arthritis receiving different dynamic treatment strategies: post hoc analyses from the BeSt study. Ann Rheum Dis 2010;69:1333–7.
9. Vastesaeger N, Xu S, Aletaha D, et al. A pilot risk model for the prediction of rapid radiographic progression in rheumatoid arthritis. Rheumatology 2009;48:1114–21.
10. Fautrel B, Granger B, Combe B, et al. Matrix to predict rapid radiographic progression of early rheumatoid arthritis patients from the community treated with methotrexate or leflunomide: results from the ESPOIR cohort. Arthritis Res Ther 2012;14:R249.
11. Syversen SW, Gaarder PI, Goll GL, et al. High anti-cyclic citrullinated peptide levels and an algorithm of four variables predict radiographic progression in patients with rheumatoid arthritis: results from a 10-year longitudinal study. Ann Rheum Dis 2008;67:212–7.
12. Lillegraven S, Paynter N, Prince FH, et al. Performance of Matrix-Based Risk Models for Rapid Radiographic Progression in a Cohort of Patients With Established Rheumatoid Arthritis. Arthritis Care Res 2013;65:526–33.
13. Durnez A, Vanderschueren G, Lateur L, et al. Effectiveness of initial treatment allocation based on expert opinion for prevention of rapid radiographic progression in daily practice of an early RA cohort. Ann Rheum Dis 2011;70:634–7.
14. Van Riel PL. EULAR handbook of clinical assessments in rheumatoid arthritis. 2004. Alphen aan den Rijn: van Zuiden Communications BV: 10–21.
15. Van der Linden MP, Feitsma AL, le Cessie S, et al. Association of a single-nucleotide polymorphism in CD40 with the rate of joint destruction in rheumatoid arthritis. Arthritis Rheum 2009;60:2242–7.
16. Knevel R, Krabben A, Brouwer E, et al. Genetic variants in IL15 associate with progression of joint destruction in rheumatoid arthritis: a multicohort study. Ann Rheum Dis 2012;71:1651–7.
17. Krabben A, Wilson AG, de Rooy DP, et al. Brief Report: Association of Genetic Variants in the IL4 and IL4R Genes With the Severity of Joint Damage in Rheumatoid Arthritis: A Study in Seven Cohorts. Arthritis Rheum 2013;65:3051–7.
18. Knevel R, de Rooy DP, Zhernakova A, et al. Association of Variants in IL2RA With Progression of Joint Destruction in Rheumatoid Arthritis. Arthritis Rheum 2013;65:1684–93.
19. Teare MD, Knevel R, Morgan MD, et al. Allele-Dose Association of the C5orf30 rs26232 Variant With Joint Damage in Rheumatoid Arthritis. Arthritis Rheum 2013;65:2555–61.

20. Cook NR, Ridker PM. Advances in Measuring the Effect of Individual Predictors of Cardiovascular Risk: The Role of Reclassification Measures. Ann Intern Med 2009;150:795–802.
21. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for some traditional and novel measures. Epidemiol Camb Mass 2010;21:128–38.
22. Knevel R, Tsonaka R, le Cessie S, et al. Comparison of methodologies for analysing the progression of joint destruction in rheumatoid arthritis. Scand J Rheumatol 2013;42:182–9.
23. Harrell FE Jr. Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis. New York: : Springer-Verlag 2001. 204.
24. Harrell FE, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med 1996;15:361–87.
25. Pencina MJ, D' Agostino RB, D' Agostino RB, et al. Evaluating the added predictive ability of a new marker: From area under the ROC curve to reclassification and beyond. Stat Med 2008;27:157–72.
26. Cook NR. Use and Misuse of the Receiver Operating Characteristic Curve in Risk Prediction. Circulation 2007;115:928–35.
27. Van der Helm-van Mil AH, Toes RE, Huizinga TW. Genetic variants in the prediction of rheumatoid arthritis. Ann Rheum Dis 2010;69:1694–6.
28. Krabben A, Huizinga TW, van der Helm-van Mil AH. Biomarkers for radiographic progression in rheumatoid arthritis. Curr Pharm Des 2015;21 (2):147–69.
29. 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;52:2637–44.
30. Knevel R, de Rooy DP, Saxne T, et al. A genetic variant in osteoprotegerin is associated with progression of joint destruction in rheumatoid arthritis. Arthritis Res Ther 2014;16:R108.