





A Risk Score to Detect Subclinical Rheumatoid Arthritis–Associated Interstitial Lung Disease

Pierre-Antoine Juge,¹  Benjamin Granger,² Marie-Pierre Debray,³ Esther Ebstein,⁴ Fabienne Louis-Sidney,⁴ Joanna Kedra,⁵  Tracy J. Doyle,⁶ Raphaël Borie,⁷ Arnaud Constantin,⁸ Bernard Combe,⁹ René-Marc Flipo,¹⁰ Xavier Mariette,¹¹ Olivier Vittecoq,¹² Alain Saraux,¹³  Guillermo Carvajal-Alegria,¹³ Jean Sibilia,¹⁴ Francis Berenbaum,¹⁵ Caroline Kannengiesser,¹⁶ Catherine Boileau,¹⁷ Jeffrey A. Sparks,¹⁸  Bruno Crestani,⁷ Bruno Fautrel,⁵ and Philippe Dieudé¹

Objective. Patients at high risk of rheumatoid arthritis–associated interstitial lung disease (RA-ILD) would benefit from being identified before the onset of respiratory symptoms; this can be done by screening patients with the use of chest high-resolution computed tomography (HRCT). Our objective was to develop and validate a risk score for patients who have subclinical RA-ILD.

Methods. Our study included a discovery population and a replication population from 2 prospective RA cohorts (ESPOIR and TRANSLATE2, respectively) without pulmonary symptoms who had received chest HRCT scans. All patients were genotyped for *MUC5B* rs35705950. After multiple logistic regression, a risk score based on independent risk factors for subclinical RA-ILD was developed in the discovery population and tested for validation in the replication population.

Results. The discovery population included 163 patients with RA, and the replication population included 89 patients with RA. The prevalence of subclinical RA-ILD was 19.0% and 16.9%, respectively. In the discovery population, independent risk factors for subclinical RA-ILD were presence of the *MUC5B* rs35705950 T allele (odds ratio [OR] 3.74 [95% confidence interval (95% CI) 1.37, 10.39]), male sex (OR 3.93 [95% CI 1.40, 11.39]), older age at RA onset (for each year, OR 1.10 [95% CI 1.04, 1.16]), and increased mean Disease Activity Score in 28 joints using the erythrocyte sedimentation rate (for each unit, OR 2.03 [95% CI 1.24, 3.42]). We developed and validated a derived risk score with receiver operating characteristic areas under the curve of 0.82 (95% CI 0.70–0.94) for the discovery population and 0.78 (95% CI 0.65–0.92) for the replication population. Excluding *MUC5B* rs35705950 from the model provided a lower goodness of fit (likelihood ratio test, $P = 0.01$).

Conclusion. We developed and validated a risk score that could help identify patients at high risk of subclinical RA-ILD. Our findings support an important contribution of *MUC5B* rs35705950 to subclinical RA-ILD risk.

INTRODUCTION

Interstitial lung disease (ILD) is an extraarticular manifestation of rheumatoid arthritis (RA). Subclinical RA-associated ILD

(RA-ILD) is detected in 20–60% of RA patients when the method of systematic high-resolution computed tomography (HRCT) of the chest is used, and clinically significant RA-ILD presents in almost 10% of RA patients (1–6). The course of subclinical

The content is solely the responsibility of the authors and does not necessarily represent the official views of Harvard University, its affiliated academic health care centers, or the National Institutes of Health.

ESPOIR cohort (NCT03666091) was supported by the French Society of Rheumatology and INSERM, Merck Sharp and Dohme, Pfizer, AbbVie, Lilly, Fresenius, and Galapagos. TRANSLATE2 cohort (NCT04227535) was supported by the French Society of Rheumatology and Bristol Myers Squibb (IM101-647). Dr. Sparks' work was supported by the National Institute of Arthritis and Musculoskeletal and Skin Diseases, NIH (grants R01-AR-077607, P30-AR-070253, and P30-AR-072577), and the R. Bruce and Joan M. Mickey Research Scholar Fund.

¹Pierre-Antoine Juge, MD, PhD, Philippe Dieudé, MD, PhD: Université de Paris, INSERM UMR 1152, F-75018, and Service de Rhumatologie, Hôpital Bichat-Claude Bernard, AP-HP, F-75018, Paris, France; ²Benjamin Granger, MD: Sorbonne Université, Institut Pierre Louis d'Épidémiologie et de Santé Publique Département de Biostatistiques, INSERM UMR 1136, F-75013, and

Santé Publique et Information Médicale, Groupe Hospitalier Pitié-Salpêtrière, AP-HP, F-5013, Paris, France; ³Marie-Pierre Debray, MD: Université de Paris, INSERM UMR 1152, F-75018, and Service de Radiologie, Hôpital Bichat-Claude Bernard, AP-HP, F-75018, Paris, France; ⁴Esther Ebstein, MD, Fabienne Louis-Sidney, MD: Service de Rhumatologie, Hôpital Bichat-Claude Bernard, AP-HP, F-75018, Paris, France; ⁵Joanna Kedra, MD, Bruno Fautrel, MD, PhD: Sorbonne Université, Institut Pierre Louis d'Épidémiologie et de Santé Publique Département de Biostatistiques, INSERM UMR 1136, F-75013, and Service de Rhumatologie, Groupe Hospitalier Pitié-Salpêtrière, AP-HP, F-75013, Paris, France; ⁶Tracy J. Doyle, MD: Division of Pulmonary and Critical Care Medicine, Department of Medicine, Brigham and Women's Hospital, Boston, Massachusetts; ⁷Raphaël Borie, MD, PhD, Bruno Crestani, MD, PhD: Université de Paris, INSERM UMR 1152, F-75018, Service de Pneumologie A, Centre de compétences maladies pulmonaires rares, Hôpital Bichat-Claude Bernard, AP-HP, F-75018, Paris, France; ⁸Arnaud Constantin, MD, PhD: Université

RA-ILD is heterogeneous, with some patients having disease progression to pulmonary symptoms and decreased respiratory function (7,8). Once clinically significant, RA-ILD is associated with high levels of morbidity and mortality, with a median survival ranging from 3 years to 8 years (1,9,10).

For patients in whom a diagnosis of RA-ILD is clinically suspected, HRCT of the chest is a useful screening tool to confirm the diagnosis of ILD and to assess ILD both qualitatively (by reviewing the pattern of interstitial pneumonia) and quantitatively (by determining the extent of ILD) (11–14). However, for patients who lack respiratory symptoms and have a particularly high risk of disease, screening for ILD is challenging (11,15). Therefore, improving ways to determine risk of progression in patients with subclinical RA-ILD could be of great value, especially because recently developed therapeutic interventions could help reduce the decline of lung function in patients with progressive disease (16).

To date, RA-ILD-associated risk factors have been investigated in patients with clinically significant disease in retrospective case-control or register studies (17–19). Clinical RA-ILD-associated risk factors consistently observed across many studies include male sex, older age, tobacco smoking, high RA activity, extraarticular features, and longer RA disease duration (1,2,20–23). Positivity for RA autoantibodies (rheumatoid factor [RF] and/or anti-citrullinated protein antibodies [ACPAs]) remains controversial because no consistent association was observed in recent large studies (23–25). The *MUC5B* rs35705950 genetic variant has been identified as a major RA-ILD risk factor, associated with 3-fold higher odds of presence of RA-ILD compared with RA without ILD (23). In a recent study comparing patients with and those without reported clinical RA-ILD from the FinnGen study—a collection of data from prospective epidemiologic and disease-based cohorts and hospital biobank samples in Finland—the presence of the *MUC5B* rs35705950 variant was associated with an increased lifetime risk of clinical RA-ILD, highlighting the importance of genetic predisposition in the occurrence of RA-ILD (19). However, the FinnGen study design included patients with symptomatic RA-ILD (i.e., ILD identified in health care registries), which did not allow for the identification of specific risk factors for subclinical ILD.

The identification of high-risk RA patients who may benefit from HRCT screening at an early and subclinical stage of the disease is an important unmet need in RA-ILD (26). Consequently, we aimed to develop and validate a risk score for subclinical ILD in patients with RA.

PATIENTS AND METHODS

Study populations and study design. This cross-sectional study included a discovery and a replication population of patients with RA from 2 prospective cohorts.

All patients included in the study met either the ACR 1987 classification criteria for RA (27) or the 2010 ACR/EULAR classification criteria for RA (28). To meet the World Health Organization guidelines for screening of a high-risk patient from an asymptomatic population (29), patients with RA were considered to be asymptomatic at the time of chest HRCT scan screening according to a definition that reflected a real-life situation in rheumatology. According to this definition, patients should have no history of ILD and no pulmonary signs or respiratory symptoms (i.e., dyspnea, cough, clubbing, and crackles at lung auscultation), as demonstrated from systematic questioning and physical examination by a senior rheumatologist. Patients with RA-ILD or with symptoms suggestive of ILD were not included in the study. The 6-minute walk test was not performed.

The discovery population consisted of patients from the prospective French ESPOIR cohort ([ClinicalTrials.gov](https://clinicaltrials.gov) identifier: NCT03666091). The ESPOIR cohort included patients with early RA who were included as participants from January 2003 to April 2005 (30); patients included in the ESPOIR cohort were evaluated every 6 months in the first 2 years and then yearly. For our discovery population, we included patients who had agreed to undergo a chest HRCT for research purposes between year 9 and year 12 of their follow-up.

The replication population consisted of patients from the independent prospective TRANSLATE2 cohort ([ClinicalTrials.gov](https://clinicaltrials.gov) identifier NCT04227535). For our replication population, we included patients who were being investigated for RA-ILD in the TRANSLATE2 study who had received consecutive and

Toulouse III–Paul Sabatier, INSERM UMR 1043, F-31024, and Service de Rhumatologie, Hôpital Purpan, F-31024, Toulouse, France; ⁹Bernard Combe, MD, PhD: Université de Montpellier and Département de Rhumatologie, Hôpital Lapeyronie, F-34000, Montpellier, France; ¹⁰René-Marc Flipo, MD, PhD: Université de Lille, and Service de Rhumatologie, Hôpital Salengro, F-59000, Lille, France; ¹¹Xavier Mariette, MD, PhD: Université Paris-Saclay, INSERM UMR 1184, CEA, F-94270, and Service de Rhumatologie, Hôpital Bicêtre, AP-HP, F-94270, Le Kremlin Bicêtre, France; ¹²Olivier Vittecoq, MD, PhD: Rouen University Hospital, Service de Rhumatologie, CIC-CRB 1404, F-76000, and Normandy University, UNIROUEN, INSERM, U1234, FR-76000, Rouen, France; ¹³Alain Saraux, MD, PhD, Guillermo Carvajal-Alegriax, MD, PhD: Université de Bretagne Occidentale, INSERM UMR 1227, F-29200, and Service de Rhumatologie, Hôpital de la Cavale Blanche, F-2900, Brest, France; ¹⁴Jean Sibilia, MD, PhD: Université de Strasbourg, INSERM UMR 51109, F-67000, and Service de Rhumatologie, RESO: Centre de Référence des Maladies Autoimmunes Systémiques Rares Est Sud-Ouest, Hôpital De Haute-pierre, F-67000, Strasbourg, France; ¹⁵Francis Berenbaum, MD, PhD: Sorbonne

Université, CRSA, INSERM UMR 938, F-75012, Service de Rhumatologie, Hôpital Saint-Antoine, AP-HP, F-75012, Paris, France; ¹⁶Caroline Kannengiesser, PharmD, PhD: Université de Paris, INSERM UMR 1152, F-75018, Département de Génétique Moléculaire, Hôpital Bichat-Claude Bernard, AP-HP, FR-75018, Paris, France; ¹⁷Catherine Boileaux, PharmD, PhD: Département de Génétique Moléculaire, Hôpital Bichat-Claude Bernard, AP-HP, FR-75018, Université de Paris, INSERM UMR 1148, Paris, F-75018, France; ¹⁸Jeffrey A. Sparks, MD, MMSc: Division of Rheumatology, Inflammation, and Immunity, Brigham and Women's Hospital, and Harvard Medical School, Boston, Massachusetts.

Author disclosures are available at <https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1002%2Fart.42162&file=art42162-sup-0001-Disclosureform.pdf>.

Address correspondence via email to Philippe Dieudé, MD, PhD, at philippe.dieude@aphp.fr.

Submitted for publication November 19, 2021; accepted in revised form May 10, 2022.

systematic chest HRCT from February 2020 to February 2021 at Bichat Hospital (Paris, France).

For patients in both the discovery and replication populations, chest HRCT scans were centrally read by an experienced radiologist and pulmonologist (MPD and RB), who remained blinded with regard to patient phenotype and genotype data. Results from the chest HRCT scans were classified as ILD, no ILD, or not interpretable, and ILD extension and pattern were evaluated according to previously reported criteria (31). Inconsistencies between the individual reviewers were resolved by consensus. Only patients with interpretable chest HRCT scans were included in the analyses.

The institutional review boards (ethics committee of Montpellier, France, no. 020307, Northern and Western French Ethic Committee III no. 2019-31) approved all protocols, and all patients provided written informed consent. This study was performed without direct patient and public involvement.

Genotyping. Included patients underwent genotyping for the *MUC5B* rs35705950 variant and subtyping to identify the presence of the shared epitope of HLA-DRB1, as previously described (32,33).

Data collection. In the discovery step, we prospectively collected potential predictors for subclinical RA-ILD (a list of all collected data is provided in the Supplementary Appendix, available on the *Arthritis & Rheumatology* website at <https://onlinelibrary.wiley.com/doi/10.1002/art.42162>). Baseline clinical and biologic data were collected at time of patient inclusion. All longitudinal variables were systematically collected at every follow-up visit (every 6 months during the first 2 years of RA and then yearly) (33). For variables that were found to be independently associated with RA-ILD in the discovery step, we systematically collected these retrospective variables at inclusion of the replication population from their medical records. Detailed information is provided in the Supplementary Appendix.

Statistical analysis. We used R program version 4.1.1 for all statistical analyses. We used GraphPad Prism 9.0 to create graphics in Figure 1 and in the supplementary figures. Further details on all methods are provided in the Supplementary Appendix.

Identification of independent risk factors for subclinical RA-ILD. For the discovery population (ESPOIR cohort), we tested the association of each collected variable with subclinical RA-ILD occurrence in bivariate and then multivariate analysis by logistic regression. Odds ratios (ORs) or effect sizes and 95% confidence intervals (95% CIs) were estimated. *P* values less than 0.05 were considered statistically significant. Because of the low number of missing data (<1%) (Supplementary Table 1, available on the *Arthritis & Rheumatology* website at <https://onlinelibrary.wiley.com/doi/10.1002/art.42162>), no imputation methods were used.

Construction of a risk score for subclinical RA-ILD. According to the categorized multivariate model in the discovery population (in which which continuous variables are transformed into categorical variables), we generated an aggregate-weighting score for each independent risk factor for subclinical RA-ILD (no missing data). For each patient, a risk score for subclinical RA-ILD was calculated by summing the weighted scores of each independent risk factor. We calculated performance of the risk score, which included calculating the area under the receiver operating characteristic (ROC) curve (AUC), sensitivity, specificity, and likelihood ratio for a proposed total cutoff value providing a sensitivity of $\geq 70\%$. The risk score was developed according to the transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD) guidelines (Supplementary Table 2, available on the *Arthritis & Rheumatology* website at <https://onlinelibrary.wiley.com/doi/10.1002/art.42162>) (34).

Validation in an independent cohort. We tested the risk score for validation in the replication population. We calculated the performance of all corresponding risk scores, which included calculations of sensitivity, specificity, likelihood ratio, and ROC curve analysis with AUC calculation based on the proposed corresponding cutoff values for each total risk score.

RESULTS

Study populations. The discovery population (ESPOIR cohort) included 163 patients (see flow chart in Supplementary Figure 1, available on the *Arthritis & Rheumatology* website at <https://onlinelibrary.wiley.com/doi/10.1002/art.42162>). Patients from the discovery population received chest HRCT between year 9 and year 12 of follow-up. Among the 163 patients, 35 (21.5%) were men, 150 (92.6%) were White, the median age at RA onset was 49.4 years (interquartile range [IQR] 41.2–55.1), 96 (58.9%) were positive for ACPAs, 128 (78.5%) were positive for RF, 77 (47.2%) were ever smokers, and the *MUC5B* rs35705950 T risk allele frequency was 11.3% (Table 1). Chest HRCT scan was performed after a median RA disease duration of 13.9 years (IQR 13–14.1). At the time of chest HRCT, the median Disease Activity Score in 28 joints using the erythrocyte sedimentation rate (DAS28-ESR) over the follow-up was 2.9 (IQR 2.4–3.7). Among the 163 patients, 138 (84.7%) had received methotrexate (MTX), 66 (40.5%) had received a biologic disease-modifying antirheumatic drug (bDMARD), and 32 (19.6%) had moderate to high tobacco smoking exposure (Table 1). Subclinical RA-ILD was detected in 31 (19.0%) of 163 patients. Missing data are provided in Supplementary Table 1. Characteristics of the patients who were not included in the analysis (i.e., refusal to participate or uninterpretable HRCT scan) are summarized in Supplementary Table 3 (available on the *Arthritis & Rheumatology* website at <https://onlinelibrary.wiley.com/doi/10.1002/art.42162>).

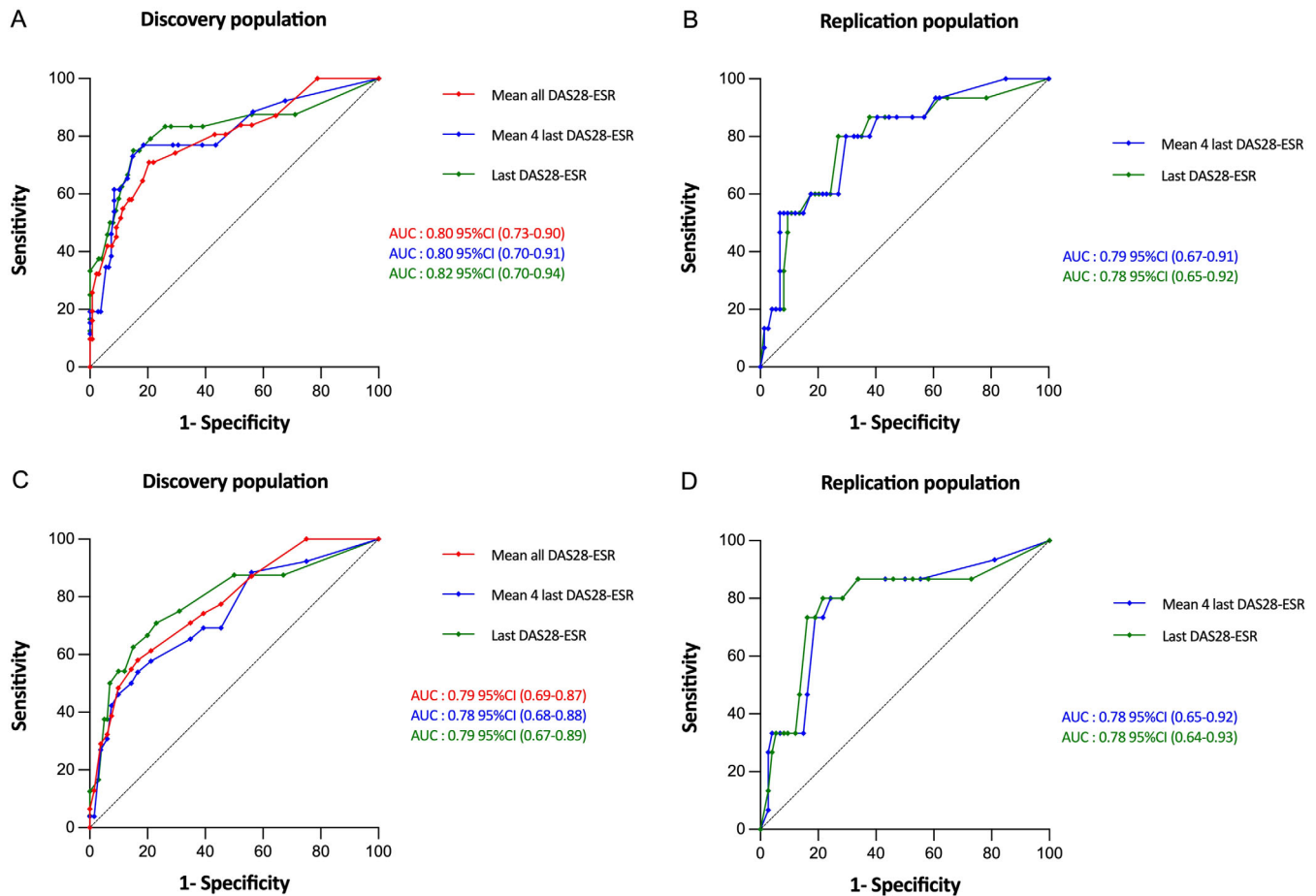


Figure 1. Performance of the proposed risk scores for detection of subclinical rheumatoid arthritis-associated interstitial lung disease in the discovery and replication populations. Results are shown as area under the receiver operating characteristic curves (AUCs) (with 95% confidence intervals [95% CIs]) in the full model (including *MUC5B* rs35705950) (A and B) and in the simplified model (without *MUC5B* rs35705950) (C and D). Curves labeled “mean all DAS28-ESR” represent models that used the mean of all Disease Activity Score in 28 joints using the erythrocyte sedimentation rate (DAS28-ESR) values over the follow-up until chest high-resolution computed tomography (HRCT) was performed to calculate the risk score. Curves labeled “mean 4 last DAS28-ESR” represent models that used the mean of the last 4 DAS28-ESR values available before chest HRCT to calculate the risk score. Curves labeled “mean last DAS28-ESR” represent models that used mean of the last DAS28-ESR value before chest HRCT to calculate the risk score.

The replication population (TRANSLATE2 cohort) included 89 patients, and their characteristics are summarized in Table 2. Subclinical RA-ILD was detected in 15 (16.9%) of 89 patients.

Identification of independent risk factors for subclinical RA-ILD. *Bivariate analysis.* Compared with patients with RA without ILD, those with subclinical RA-ILD more frequently carried the *MUC5B* rs35705950 T risk allele (minor allele frequency of 19% versus 9.5%, $P = 0.02$), were more frequently men (38.7% versus 17.4%, $P = 0.01$), were older at RA onset (median age 56.2 years [IQR 50.6–61.4] versus 47.6 years [IQR 38.5–53.7], $P < 0.0001$), had higher DAS28-ESR (median 3.4 [IQR 2.7–4.1] versus 2.9 [IQR 2.3–3.5], $P = 0.03$), and had higher scores for the Health Assessment Questionnaire (median 0.6 [IQR 0.3–1.1] versus 0.4 [IQR 0.2–0.6], $P = 0.007$) over the follow-up (Table 1). Patients with subclinical RA-ILD had numerically higher

body mass index (BMI), longer tobacco smoking exposure (pack-years), and higher mean C-reactive protein level during the follow-up compared with results shown in RA patients without ILD; however, differences between RA patients with and those without ILD were not statistically significant (Table 1). We detected no differences in the ACPA or RF positivity rates or ACPA or RF titers according to RA-ILD status versus RA without ILD status, presence versus absence of the shared epitope for HLA-DRB1 status, or level of exposure to MTX or bDMARD (Table 1). The relatively small number of patients with RA-ILD did not allow subanalyses according to the HRCT patterns.

Multivariate analysis. Logistic regression analysis identified 4 variables independently associated with subclinical RA-ILD: the *MUC5B* rs35705950 T risk allele (OR 3.74 [95% CI 1.37, 10.39]) ($P = 0.01$), male sex (OR 3.93 [95% CI 1.40, 11.39]) ($P = 0.01$), older age at RA onset (for each year, OR 1.1 [95% CI 1.04, 1.16])

Table 1. Characteristics of patients with RA in the discovery sample (ESPOIR)*

Characteristic	Patient group			OR or ES or HR (95% CI)	P†
	Overall RA (n = 163)	RA-ILD (n = 31)	RA without ILD (n = 132)		
Characteristic at RA onset					
Male sex	35 (21.5)	12 (38.7)	23 (17.4)	2.99 (1.28, 7.01)	0.01
White race/ethnicity	150 (92.6)	27 (87.1)	123 (93.9)	2.28 (0.64, 8.11)	0.25
Age, median (IQR) years	49.4 (41.2–55.1)	56.2 (50.6–61.4)	47.6 (38.5–53.7)	−0.84 (−1.24, −0.43)‡	<0.0001
BMI median (IQR) kg/m ²	24.0 (21.5–27.4)	24.9 (23.0–28.9)	23.7 (21.4–27.0)	−0.23 (−0.62, −0.17)‡	0.11
Ever smoker	77 (47.2)	17 (54.8)	60 (45.5)	1.46 (0.66, 3.20)	0.42
Smoking exposure, median (IQR) pack-years	0 (0–15)	8 (0–20)	0 (0–13)	−0.37 (−0.77, 0.03)‡	0.11
DAS28-ESR, median (IQR)	5.2 (4.5–6.1)	5.5 (4.7–6.2)	5.1 (4.4–6.0)	−0.04 (−0.43, 0.36)‡	0.48
CRP, median (IQR) mg/dl	10.0 (4.0–22.8)	9.0 (4.5–20.5)	10.0 (4.0–22.5)	0.7 (−0.33, 0.43)‡	0.69
ACPA positive	96 (58.9)	20 (64.5)	76 (57.6)	1.34 (0.59, 3.02)	0.55
ACPA titer, median (IQR) units/ml	394 (0.0–500)	429 (0–500)	325 (0–500)	−0.10 (−0.50, 0.29)‡	0.68
RF positive	128 (78.5)	24 (77.4)	104 (78.8)	0.92 (0.36, 2.36)	0.81
RF titer, median (IQR) units/ml	30 (5–95)	50 (5–161)	16 (6–43)	−0.16 (−0.55, 0.24)‡	0.33
ANA positive	61 (37.4)	10 (32.3)	51 (38.6)	0.76 (0.33, 1.74)	0.54
Sicca syndrome	111 (68.1)	19 (61.3)	92 (69.7)	0.69 (0.31, 1.55)	0.40
Erosive status	29 (17.8)	7 (11.1)	22 (17.7)	1.46 (0.56–3.80)	0.44
Total SHS, median (IQR)	2.0 (0.0–4.0)	2.0 (0.0–4.0)	2.0 (0.0–4.0)	−0.19 (−0.61, 0.23)‡	0.59
HAQ score, median (IQR)	1.0 (0.5–1.5)	1.1 (0.5–1.4)	1.0 (0.5–1.5)	−0.07 (−0.47, 0.32)‡	0.83
HLA-DRB1*SE allele presence					0.93
0	68 (42.8)	14 (45.2)	54 (42.2)	1 (referent)	
1	65 (40.9)	13 (42.0)	52 (40.6)	0.96 (0.41, 2.25)	
2	26 (16.4)	4 (12.9)	22 (17.2)	0.70 (0.21, 2.37)	
MUC5B rs35705950 GT/TT genotype, %	11.3	19	9.5	2.91 (1.22, 6.95)	0.02
Longitudinal variables at time of chest HRCT scan					
RA duration, median (IQR) years	13.9 (13–14.1)	13.9 (13.3–14)	13.9 (13–14.1)	0.82 (0.55, 1.22)§	0.30
Smoking status trajectory					
Never-stop/low	131 (80.4)	22 (71.0)	109 (82.6)	1 (referent)	0.21
Maintained/moderate-high	32 (19.6)	9 (29.0)	23 (17.4)	1.94 (0.79, 4.75)	
DAS28-ESR, median (IQR)	2.9 (2.4–3.7)	3.4 (2.7–4.1)	2.9 (2.3–3.5)	−0.47 (−0.87, −0.07)‡	0.03
CRP, median (IQR) mg/dl	5.8 (3.9–9.2)	6.6 (3.9–9.2)	5.8 (4.1–8.7)	−0.38 (−0.78, 0.01)‡	0.42
CRP trajectory					
None-low	147 (90.2)	25 (80.6)	122 (92.4)	1 (referent)	0.09
Moderate-high	16 (9.8)	6 (19.4)	10 (7.6)	2.93 (0.91, 8.79)	
HAQ score, median (IQR)	0.4 (0.2–0.8)	0.6 (0.3–1.1)	0.4 (0.2–0.6)	−0.01 (−0.4, 0.39)‡	0.007
MTX exposure	138 (84.7)	27 (87.1)	111 (84.1)	1.28 (0.40, 4.03)	0.79
MTX exposure trajectory					
None-low	51 (31.3)	12 (38.7)	39 (29.6)	1 (referent)	0.64
Moderate	76 (46.6)	13 (41.9)	63 (47.7)	0.67 (0.28, 1.62)	
High	36 (22.1)	6 (19.4)	30 (22.7)	0.65 (0.22, 1.93)	
bDMARD exposure	66 (40.5)	14 (45.2)	52 (39.4)	1.27 (0.58, 2.79)	0.68
TNF inhibitor exposure	54 (33.1)	14 (45.2)	40 (30.3)	1.89 (0.85, 4.21)	0.14
Glucocorticoid exposure trajectory					0.84
None-low	60 (36.8)	13 (41.9)	47 (35.6)	1 (referent)	
Moderate	30 (18.4)	5 (16.2)	25 (18.9)	0.72 (0.23, 2.26)	
High	73 (44.8)	13 (41.9)	60 (45.5)	0.78 (0.33, 1.85)	
ILD pattern on chest HRCT					
UIP and probable UIP	–	4 (12.9)	–	–	–
NSIP	–	4 (12.9)	–	–	–
Indeterminate	–	23 (74.2)	–	–	–
Extent of ILD on chest HRCT					
<5%	–	7 (22.6)	–	–	–
5–10%	–	15 (48.4)	–	–	–
>10%	–	9 (29.0)	–	–	–
Pulmonary function test results, median (IQR)					
FVC, % predicted	–	110 (91.8–117)	–	–	–
FEV ₁ , % predicted	–	96 (90.3–113)	–	–	–

(Continued)

Table 1. (Cont'd)

Characteristic	Patient group			OR or ES or HR (95% CI)	P†
	Overall RA (n = 163)	RA-ILD (n = 31)	RA without ILD (n = 132)		
TLC, % predicted	-	104 (92.3–108.3)	-	-	
DLco, % predicted	-	77 (62–80)	-	-	

* Except where indicated otherwise, results (qualitative variables) are the number (%); quantitative variables are shown as median and interquartile range (IQR). The odds ratio (OR) with 95% confidence interval (95% CI) represents the likelihood of a qualitative variable being higher in the rheumatoid arthritis-associated interstitial lung disease (RA-ILD) group than in the RA without ILD group. Erosive status indicates presence of articular erosions due to RA. The effect size (ES) with 95% CI represents the effect size of a quantitative variable being higher in the RA-ILD group than in the RA without ILD group. BMI = body mass index; DAS28-ESR = Disease Activity Score in 28 joints using the erythrocyte sedimentation rate; CRP = C-reactive protein; ACPA = anti-citrullinated peptide antibody; RF = rheumatoid factor; ANA = antinuclear antibody; SHS = modified Sharp/van der Heijde score of radiographic progression; HAQ = Health Assessment Questionnaire; HLA-DRB1*SE = HLA-DRB1 shared epitope; HRCT = high-resolution computed tomography; MTX = methotrexate; bDMARD = biologic disease-modifying antirheumatic drug; TNF = tumor necrosis factor; UIP = usual interstitial pneumonia; NSIP = nonspecific interstitial pneumonia; FVC = forced vital capacity, FEV₁ = forced expiratory volume in 1 second; TLC = total lung capacity; DLco = diffusing capacity for carbon monoxide.

† Results from the bivariate analysis.

‡ ES with 95% CI.

§ Hazard ratio (HR) with 95% CI.

($P < 0.001$), and increased mean DAS28-ESR over the follow-up (for each unit increase, OR 2.03 [95% CI 1.24, 3.42]) ($P = 0.006$) (Table 3).

Development of a risk score for subclinical RA-ILD.

We represented the performance of the multivariate model from the discovery population by ROC curve analysis, in which the AUC was calculated as 0.81 (95% CI 0.73, 0.90) (Supplementary Figure 2, available on the *Arthritis & Rheumatology* website at

<https://onlinelibrary.wiley.com/doi/10.1002/art.42162>). We had similar findings when we used categorized variables to determine performance (Supplementary Appendix), with the AUC calculated as 0.80 (95% CI 0.73, 0.9) (Figure 1A). The corresponding risk matrix is provided in Table 4. The probability (OR) for subclinical RA-ILD ranged from 2 (0.3–5.7) for female patients not carrying the *MUC5B* rs35705950 T risk allele and ≤ 49 years of age at RA onset and with a mean DAS2-ESR ≤ 2.9 over follow-up to 94.9 (72.1–99.3) for male patients carrying the

Table 2. Characteristics of patients with RA in the replication population (TRANSLATE2)*

Characteristic	Patient group		
	Overall RA (n = 89)	RA-ILD (n = 15)	RA without ILD (n = 74)
Male sex	23 (25.8)	7 (46.7)	16 (21.6)
RA duration, median (IQR) years	11 (5–18)	9 (4.5–11.5)	12 (6–18)
Age at RA onset			
Median (IQR) years	46.0 (34.0–55.0)	50.0 (43.5–57.5)	42.5 (30.5–53.8)
≤ 49 years	53 (59.6)	5 (33.3)	48 (64.9)
50–58 years	11 (12.4)	4 (26.7)	7 (9.5)
> 58 years	25 (28.1)	6 (40.0)	19 (25.7)
ACPA positive	84 (94.4)	14 (93.3)	70 (94.6)
RF positive	66 (75.9)	11 (73.3)	55 (76.4)
Ever smoker	48 (59.3)	8 (61.5)	40 (58.8)
<i>MUC5B</i> rs35705950 GT/TT genotype, %	17.4	26.7	15.5
All DAS28-ESR values during the 4 years before HRCT			
Median (IQR)	3.1 (2.2–4.6)	4.8 (4.0–5.1)	2.9 (2.0–3.9)
< 2.9	29 (32.6)	1 (6.6)	28 (37.8)
2.9–4.3	37 (41.6)	4 (26.7)	33 (44.6)
> 4.3	23 (25.8)	10 (66.7)	13 (17.6)
Last 4 DAS28-ESR values available before HRCT			
Median (IQR)	3.5 (2.4–4.4)	4.7 (4.2–5.0)	3.2 (2.3–4.1)
< 2.9	31 (34.8)	1 (6.6)	30 (40.5)
2.9–4.3	35 (39.4)	4 (26.7)	31 (41.9)
> 4.3	23 (25.8)	10 (66.7)	13 (17.6)
Last DAS28-ESR value available before HRCT			
Median (IQR)	3.1 (2.2–4.6)	4.8 (4.0–5.1)	2.9 (2.0–3.9)
< 2.9	40 (45.0)	2 (13.3)	38 (51.4)
2.9–4.3	22 (24.7)	2 (13.3)	20 (27.0)
> 4.3	27 (30.3)	11 (73.4)	16 (21.6)

* Except where indicated otherwise, values are the number (%). See Table 1 for definitions.

Table 3. Independent risk factors associated with subclinical RA-ILD in the discovery population (ESPOIR)*

	Model with MUC5B rs35705950		Categorized model with MUC5B rs35705950		Weighted coefficient with MUC5B rs35705950		Model without MUC5B rs35705950		Categorized model without MUC5B rs35705950		Weighted coefficient without MUC5B rs35705950	
	OR (95% CI)	P†	OR (95% CI)	P†	OR (95% CI)	P†	OR (95% CI)	P†	OR (95% CI)	P†	OR (95% CI)	P†
MUC5B rs35705950												
GG	1 (referent)		1 (referent)		0		NA		NA		NA	
GT/TT	3.74 (1.37–10.39)	0.01	3.52 (1.28–9.77)	0.01	27		NA		NA		NA	
Sex												
Female	1 (referent)		1 (referent)		0		1 (referent)		NA		NA	
Male	3.93 (1.40–11.39)	0.01	3.59 (1.26–10.42)	0.02	28		3.8 (1.4–10.6)	<0.01	NA		NA	
Age at RA onset												
Per year change	1.10 (1.04–1.16)	<0.001	–		0		1.1 (1.0–1.2)	0.001	–		–	
≤49 years	–		1 (referent)		0		–		1 (referent)		–	
50–58 years	–		3.10 (1.08–9.66)	0.04	25		–		2.68 (0.97–7.89)	0.06	–	
>58 years	–		9.98 (3.04–36.36)	<0.001	50		–		8.81 (2.8–30.09)	>0.001	–	
Mean DAS28-ESR												
Per unit change	2.03 (1.24–3.42)	0.006	–		0		1.9 (1.2–3.2)	0.01	–		–	
<2.9	–		1 (referent)		0		–		1 (referent)		–	
2.9–4.3	–		2.83 (1.02–8.46)	0.05	23		–		2.54 (0.94–7.36)	0.07	–	
>4.3	–		5.93 (1.45–24.78)	0.01	39		–		5.93 (1.4–23.47)	0.01	–	

* A categorized model is a model in which continuous variables are transformed into categorical variables. NA = not applicable. See Table 1 for other definitions.
 † Results from the multivariate analysis.

Table 4. Association matrices for risk scores associated with the risk of subclinical RA-ILD in the discovery population*

Risk matrix variable	DAS28-ESR <2.9		DAS28-ESR 2.9–4.3		DAS28-ESR >4.3	
	Female	Male	Female	Male	Female	Male
Model with <i>MUC5B</i> rs35705950						
≤49 years (GG)	2.0 (0.3, 5.7)	7.1 (1.0, 18.6)	6.7 (1.2, 16.5)	21.3 (3.2, 50.3)	12.5 (2.0, 29.7)	34.9 (5.9, 71.4)†
≤49 years (GT/TT)	6.7 (1.4, 17.6)	21.3 (5.4, 48.0)	20.3 (5.2, 38.2)	48.9 (16.0, 80.1)†	33.5 (6.3, 59.4)†	65.4 (16.3, 90.8)‡
50–58 years (GG)	6.2 (1.5, 15.6)	19.9 (4.5, 41.6)	18.9 (5.9, 30.6)	46.8 (15.0, 71.7)†	31.6 (8.3, 57.1)†	63.4 (18.1, 87.5)‡
50–58 years (GT/TT)	18.9 (3.8, 50.1)	46.8 (13.0, 80.8)†	45.2 (15.6, 71.7)†	75.6 (39, 94.6)§	62 (17.9, 87.4)‡	86.0 (40.7, 97.7)§
>58 years (GG)	16.7 (5.0, 39.1)	42.9 (20.6, 72.0)†	41.4 (15.0, 69.5)†	72.7 (35.6, 92.3)‡	58.3 (25.1, 84.8)‡	84.0 (48.3, 97.2)§
>58 years (GT/TT)	41.4 (17.4, 76.6)†	72.7 (46.9, 92.7)‡	71.4 (40.8, 91.7)‡	90.4 (66.9, 98.3)§	83.1 (47.3, 96.8)§	94.9 (72.1, 99.3)§
Model without <i>MUC5B</i> rs35705950						
≤49 years	3.6 (0.8, 7.1)	11.9 (3.3, 24.7)	8.6 (2.3, 17.6)	25.6 (7.2, 53.7)†	17.5 (4.0, 40.4)	43.6 (9.7, 81.8)†
50–58 years	9.1 (2.2, 20.3)	26.6 (8.8, 53.5)†	20.2 (7.9, 35.8)	48.0 (21.7, 73.8)†	36.3 (10.1, 66.2)†	67.4 (22.0, 91.5)‡
>58 years	24.6 (8.0, 51.2)	54.4 (26.8, 79.3)‡	45.4 (20.2, 71.9)†	75.2 (45.2, 93.7)§	65.1 (35.3, 91.1)‡	87.2 (59.1, 98.3)§

* Risk matrix models were stratified by the presence or absence of each independent risk factor for subclinical RA-ILD (age at RA onset, *MUC5B* rs35705950 genotype [GG or GT/TT], DAS28-ESR disease activity scores, and sex). See Table 1 for definitions.

† High risk level for RA-ILD.

‡ Higher level of risk for RA-ILD.

§ Highest level of risk for RA-ILD.

MUC5B rs35705950 T risk allele and >58 years of age at RA onset and with a mean DAS28-ESR >4.3 over follow-up (Table 4).

To generate a risk score for subclinical RA-ILD, we attributed a weighted coefficient to each independent risk factor, which led to a total risk score ranging from 0 to 144 (Table 3). In a model in which we used a total risk score cutoff of 51 for defining subclinical RA-ILD, the sensitivity was 71.0% and the specificity was 79.6% (Supplementary Table 4, available on the *Arthritis & Rheumatology* website at <https://onlinelibrary.wiley.com/doi/10.1002/art.42162>).

To simplify the score for use in daily clinical practice, we aimed to estimate the best approximation of the mean DAS28-ESR over follow-up. The last 4 DAS28-ESR values obtained from each patient before they underwent chest HRCT scan (i.e., 4 years before the chest HRCT scan) provided a good estimate of the mean DAS28-ESR for all values available over the entire follow-up, with an AUC of 0.80 (95% CI 0.70, 0.91) (not significantly different from the model with all available DAS28-ESR values, bootstrap $P = 0.08$). For the model with the last 4 DAS28-ESR values and with total risk score cutoff of 51 for defining subclinical RA-ILD, the sensitivity was 73.1% and the specificity was 85.2% (Figure 1A, Supplementary Table 4). We then validated the score based on the last 4 DAS28-ESR values obtained from each patient before chest HRCT in the replication cohort. In this validation model, the AUC was 0.79 (95% CI 0.67, 0.91) (Figure 1B); the sensitivity and specificity values based on a total risk score cutoff of 51 are listed in Supplementary Table 4.

When we evaluated the last DAS28-ESR value available within the year before chest HRCT scan in the discovery population, we found that the performance was similar and not significantly different from the model with all available DAS28-ESR values (bootstrap $P = 0.23$). In this model for detection of

subclinical RA-ILD based on a total risk score cutoff of 51, the AUC was 0.82 (95% CI 0.70, 0.94) (Figure 1A), the sensitivity was 75.0%, and the specificity was 85.0% (Supplementary Table 4). We then validated the score derived from the last DAS28-ESR value in the replication population; in the validated model, the AUC was 0.78 (95% CI 0.65, 0.92) (Figure 1B). The sensitivity, specificity, and likelihood ratios based on a total risk score cutoff of 51 are listed in the Supplementary Table 4.

Because *MUC5B* rs35705950 genotyping is not available yet in daily practice, we created a simplified model that excluded *MUC5B* rs35705950. Male sex (OR 3.8 [95% CI 1.4, 10.6]) ($P < 0.01$), older age at RA onset (OR 1.1 [95% CI 1.0, 1.2]) ($P = 0.001$), and increased mean DAS28-ESR value over the follow-up (OR 1.9 [95% CI 1.2, 3.2]) ($P = 0.01$) were independently associated with subclinical RA-ILD (Table 3). A new weighted coefficient was attributed to each independent risk factor, which led to a total risk score that ranged from 0 to 113 (Table 3). The corresponding risk matrix is provided in Table 4.

The performance of the ROC curve for the simplified model was comparable to the full model (bootstrap $P = 0.25$), with an AUC of 0.79 (95% CI 0.69, 0.87) (Figure 1C). When we included the mean of the last 4 DAS28-ESR values available in the 4 years before the chest HRCT, performance of the risk score for the simplified model was comparable between the discovery population (AUC 0.78 [95% CI 0.68, 0.88]) and after validation in the replication population (AUC 0.78 [95% CI 0.65, 0.92]) (Figure 1C; see also Supplementary Table 5, available on the *Arthritis & Rheumatology* website at <https://onlinelibrary.wiley.com/doi/10.1002/art.42162>) (not significantly different from the model with all available DAS28-ESR values, bootstrap $P = 0.21$). The simplified risk score derived from the last DAS28-ESR value available within the year before the chest HRCT scan was similar to the score derived from

the last 4 DAS28-ESR values (not significantly different from the model with all available DAS28-ESR values, bootstrap $P = 0.42$) in both the discovery and replication populations, with an AUC of 0.79 (95% CI 0.67, 0.89) and 0.78 (95% CI 0.64, 0.93), respectively (Figures 1C and D). Corresponding sensitivity and specificity values in the model based on a total risk score cutoff value of 25 for defining subclinical RA-ILD were 75.0% and 69.0%, respectively, in the discovery population and 86.7% and 47.3%, respectively, in the replication population (Supplementary Table 5).

Of note, if the ROC AUCs of the models with and without *MUC5B* rs35705950 were found comparable (bootstrap $P = 0.25$) (Figure 1), the model that included *MUC5B* rs35705950 had better goodness of fit than the model without, with an Akaike's information criterion value of 133 and 138, respectively (likelihood ratio test, $P = 0.01$).

DISCUSSION

A scoring system that allows stratification of patients at high risk for RA-ILD before the onset of their pulmonary symptoms (i.e., subclinical RA-ILD) may help clinicians identify patients who would most benefit from chest HRCT screening.

In this study, we proposed and validated a risk score for subclinical RA-ILD that included 4 variables (sex, age at RA onset, RA disease activity using DAS28-ESR, and the *MUC5B* rs35705950 genetic variant). Although the risk score without *MUC5B* rs35705950 was found to be appropriate to discriminate patients with subclinical RA-ILD, the model with *MUC5B* rs35705950 had better performance, suggesting an important contribution of the genetic variant to the overall risk of subclinical RA-ILD. In our study, the contribution of *MUC5B* rs35705950, a common variant with a relatively high magnitude of association, to risk of subclinical RA-ILD was similar to that previously reported for clinical RA-ILD (23). Indeed, in both our present study and our earlier study, the odds of RA-ILD developing in patients carrying the *MUC5B* rs35705950 T risk allele was at least 3 times greater than in those carrying the GG genotype (23). Of note, a similar magnitude of association was reported in individuals with interstitial lung abnormalities and without RA (35). Our findings are in good agreement with the results of the FinnGen epidemiologic cohort (19) and support a pivotal role of the *MUC5B* rs35705950 variant for both clinical and subclinical RA-ILD risk stratification and add to the possible interest for genotyping the risk variant in future clinical practice (36).

The performance of the risk score when the last DAS28-ESR value available in the year before the chest HRCT scan was used was similar to the performance when the mean DAS28-ESR over follow-up was used. This finding is concordant with the previously reported effect of the increase in annual DAS28 on the risk of incident RA-ILD within the year before RA-ILD onset (21). Estimation of the risk for subclinical RA-ILD using only the last DAS28-ESR

available makes our risk score easy to use for daily practice. However, the performance of our risk score cannot be directly compared with the performance of other scores because of the different study design (i.e., a systematic exploration of asymptomatic patients by chest HRCT) and the integration of different variables in our model (i.e., the *MUC5B* rs35705950 variant and DAS28-ESR). Of interest, a risk score for clinical RA-ILD defined in a recent case-control study identified both male sex and disease activity (i.e., Clinical Disease Activity Index score >28 and ESR >80 mm/hour) as independent risk factors, which reinforces their contribution to the excess risk for ILD in patients with RA (17). In their case-control study, Paulin et al also identified smoking and the presence of extraarticular manifestations as predictors of ILD among patients with RA. However, in a nested case-control study that matched incident RA-ILD cases to RA non-ILD controls on age, sex, RA duration, RF, and time from exposure assessment to RA-ILD, the investigators identified obesity, CRP level, functional status, and heavy smoking as potential risk factors for RA-ILD (37). Even if our study was not designed to assess the impact of RA treatments, MTX use was not found to contribute to the risk of subclinical RA-ILD, which is consistent with previous studies that concluded that MTX was not a risk factor for RA-ILD (22,25,38).

Our study has some limitations. The relatively low occurrence of RA-ILD in the discovery population and the relatively small sample size may have decreased the power to detect other RA-ILD risk factors (17,37). Conversely, if such factors were not identified in our study because of lack of power, the magnitude of their likely association would be small, with a limited contribution to the excess risk for subclinical RA-ILD. In addition, the limited sample size did not allow us to perform subanalyses to identify risk factors for specific HRCT patterns. Indeed, the potential of including *MUC5B* rs35705950 in a future risk score for patients at high risk of subclinical usual interstitial pneumonia-type RA should be considered according to the restricted association demonstrated between the risk variant and the usual interstitial pneumonia pattern shown on HRCT scan of the chest of patients with RA-ILD (23). In our study, the ACPA positivity rate was different between the discovery and replication populations (58.9% versus 94.4%, respectively), which could be the consequence of sampling bias. However, this difference should not affect the performance of the risk score, as the discovery stage did not identify ACPA status as an independent risk factor for subclinical RA-ILD. Lastly, several patients included in the ESPOIR cohort did not agree to participate in this cross-sectional study, which may have implied a selection bias. However, the characteristics of patients included in the study and those not included were not different. In addition, the prevalence of subclinical RA-ILD in our discovery population was comparable to that previously reported in the literature (1-6). Our risk score was developed and validated in patients having established RA with a mean disease duration of 10 years. The predictive value of our risk score will need validation for early or

longstanding RA. Therefore, future large prospective studies are needed to investigate 1) the effects of other potential risk factors, including smoking, BMI, RF-positive and ACPA-positive status, and serum biomarkers (39), on the risk score for subclinical RA-ILD, 2) the performance of our risk score for other RA durations, notably at RA onset, and 3) the identification of risk score for a specific HRCT pattern and progression to clinical lung fibrosis.

In conclusion, this is the first study that identified and validated a risk score (with and without inclusion of *MUC5B* rs35705950) that would allow the identification of patients at high risk for subclinical RA-ILD who are eligible for chest HRCT screening. The fact that the highest-performance model was the one that included *MUC5B* rs35705950 in the risk score illustrates the significant contribution of this genetic variant to the risk of subclinical RA-ILD. These findings could help clinicians in their daily practice and could affect future recommendations of RA-ILD screening.

ACKNOWLEDGMENTS

We thank the participating patients for their contributions to this work, Nathalie Rincheval for expert monitoring and data management, and the investigators who recruited and followed the patients.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Dr. Juge had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Juge, Granger, Debray, Kedra, Borie, Crestani, Fautrel, Dieudé.

Acquisition of data. Juge, Debray, Borie, Kannengiesser, Boileau, Crestani, Fautrel, Dieudé.

Analysis and interpretation of data. Juge, Granger, Ebstein, Louis-Sidney, Kedra, Doyle, Borie, Constantin, Combe, Flipo, Mariette, Vittecoq, Sarau, Carvajal-Alegria, Sibilia, Berenbaum, Kannengiesser, Boileau, Sparks, Crestani, Fautrel, Dieudé.

REFERENCES

- Bongartz T, Nannini C, Medina-Velasquez YF, et al. Incidence and mortality of interstitial lung disease in rheumatoid arthritis: a population-based study. *Arthritis Rheum* 2010;62:1583–91.
- Koduri G, Norton S, Young A, et al. Interstitial lung disease has a poor prognosis in rheumatoid arthritis: results from an inception cohort. *Rheumatology* 2010;49:1483–9.
- Gabbay E, Tarala R, Will R, et al. Interstitial lung disease in recent onset rheumatoid arthritis. *Am J Respir Crit Care Med* 1997;156:528–35.
- Reynisdottir G, Karimi R, Joshua V, et al. Structural changes and antibody enrichment in the lungs are early features of anti-citrullinated protein antibody-positive rheumatoid arthritis. *Arthritis Rheumatol* 2014;66:31–9.
- Salaffi F, Carotti M, Di Carlo M, et al. High-resolution computed tomography of the lung in patients with rheumatoid arthritis: prevalence of interstitial lung disease involvement and determinants of abnormalities. *Medicine (Baltimore)* 2019;98:e17088.
- Chen J, Shi Y, Wang X, et al. Asymptomatic preclinical rheumatoid arthritis-associated interstitial lung disease. *Clin Dev Immunol* 2013;2013:406927.
- Robles-Perez A, Luburich P, Rodriguez-Sanchon B, et al. Preclinical lung disease in early rheumatoid arthritis. *Chron Respir Dis* 2016;13:75–81.
- Gochuico BR, Avila NA, Chow CK, et al. Progressive preclinical interstitial lung disease in rheumatoid arthritis. *Arch Intern Med* 2008;168:159–66.
- Hyltdgaard C, Ellingsen T, Hilberg O, et al. Rheumatoid arthritis-associated interstitial lung disease: clinical characteristics and predictors of mortality. *Respiration* 2019;98:455–60.
- Raimundo K, Solomon JJ, Olson AL, et al. Rheumatoid arthritis-associated interstitial lung disease in the United States: prevalence, incidence, and healthcare costs and mortality. *J Rheumatol* 2019;46:360–9.
- Spagnolo P, Ryerson CJ, Putman R, et al. Early diagnosis of fibrotic interstitial lung disease: challenges and opportunities. *Lancet Respir Med* 2021;9:1065–76.
- Vij R, Strek ME. Diagnosis and treatment of connective tissue disease-associated interstitial lung disease. *Chest* 2013;143:814–24.
- Kanne JP. Interstitial lung disease (ILD): imaging finding, and the role of imaging in evaluating the patient with known or suspected ILD. *Semin Roentgenol* 2010;45:3.
- Bradley B, Branley HM, Egan JJ, et al. Interstitial lung disease guideline: the British Thoracic Society in collaboration with the Thoracic Society of Australia and New Zealand and the Irish Thoracic Society. *Thorax* 2008;63 Suppl:v1–58.
- Hansell DM, Goldin JG, King TE Jr, et al. CT staging and monitoring of fibrotic interstitial lung diseases in clinical practice and treatment trials: a position paper from the Fleischner Society. *Lancet Respir Med* 2015;3:483–96.
- Flaherty KR, Wells AU, Cottin V, et al. Nintedanib in progressive fibrosing interstitial lung diseases. *N Engl J Med* 2019;381:1718–27.
- Paulin F, Doyle TJ, Mercado JF, et al. Development of a risk indicator score for the identification of interstitial lung disease in patients with rheumatoid arthritis. *Reumatol Clin (Engl Ed)* 2021;17:207–11.
- Mena-Vazquez N, Perez Albaladejo L, Manrique-Ariza S, et al. Analysis of clinical-analytical characteristics in patients with rheumatoid arthritis and interstitial lung disease: case-control study. *Reumatol Clin (Engl Ed)* 2021;17:197–202.
- Palomaki A, FinnGen Rheumatology Clinical Expert Group, Palotie A, et al. Lifetime risk of rheumatoid arthritis-associated interstitial lung disease in *MUC5B* mutation carriers. *Ann Rheum Dis* 2021;80:1530–6.
- Kelly CA, Saravanan V, Nisar M, et al. Rheumatoid arthritis-related interstitial lung disease: associations, prognostic factors and physiological and radiological characteristics—a large multicentre UK study. *Rheumatology* 2014;53:1676–82.
- Sparks JA, He X, Huang J, et al. Rheumatoid arthritis disease activity predicting incident clinically apparent rheumatoid arthritis-associated interstitial lung disease: a prospective cohort study. *Arthritis Rheumatol* 2019;71:1472–82.
- Kiely P, Busby AD, Nikiphorou E, et al. Is incident rheumatoid arthritis interstitial lung disease associated with methotrexate treatment? Results from a multivariate analysis in the ERAS and ERAN inception cohorts. *BMJ Open* 2019;9:e028466.
- Juge PA, Lee JS, Ebstein E, et al. *MUC5B* Promoter variant and rheumatoid arthritis with interstitial lung disease. *N Engl J Med* 2018;379:2209–19.
- Hyltdgaard C, Hilberg O, Pedersen AB, et al. A population-based cohort study of rheumatoid arthritis-associated interstitial lung disease: comorbidity and mortality. *Ann Rheum Dis* 2017;76:1700–6.

25. Juge PA, Lee JS, Lau J, et al. Methotrexate and rheumatoid arthritis associated interstitial lung disease. *Eur Respir J* 2020;57:2000337.
26. Kelly C, Emery P, Dieude P. Current issues in rheumatoid arthritis-associated interstitial lung disease. *Lancet Rheumatol* 2021;3: E798–807.
27. Arnett FC, Edworthy SM, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum* 1988;31:315–24.
28. Aletaha D, Neogi T, Silman AJ, et al. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Arthritis Rheum* 2010;62:2569–81.
29. World Health Organization. Screening programmes: a short guide. Increase effectiveness, maximize benefits and minimize harm. Geneva: World Health Organization; 2020.
30. Combe B, Rincheval N, Berenbaum F, et al. Current favourable 10-year outcome of patients with early rheumatoid arthritis: data from the ESPOIR cohort. *Rheumatology (Oxford)* 2021;60:5073–9.
31. Raghu G, Remy-Jardin M, Myers JL, et al. Diagnosis of idiopathic pulmonary fibrosis. An official ATS/ERS/JRS/ALAT clinical practice guideline. *Am J Respir Crit Care Med* 2018;198:e44–68.
32. Seibold MA, Wise AL, Speer MC, et al. A common MUC5B promoter polymorphism and pulmonary fibrosis. *N Engl J Med* 2011;364: 1503–12.
33. Combe B, Benessiano J, Berenbaum F, et al. The ESPOIR cohort: a ten-year follow-up of early arthritis in France: methodology and baseline characteristics of the 813 included patients. *Joint Bone Spine* 2007;74:440–5.
34. Collins GS, Reitsma JB, Altman DG, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD Statement. *Eur J Clin Invest*. 2015;45:204–14.
35. Hunninghake GM, Hatabu H, Okajima Y, et al. MUC5B promoter polymorphism and interstitial lung abnormalities. *N Engl J Med* 2013;368: 2192–200.
36. Sparks JA. Towards clinical significance of the MUC5B promoter variant and risk of rheumatoid arthritis-associated interstitial lung disease. *Ann Rheum Dis* 2021;80:1503–4.
37. Kronzer VL, Huang W, Dellaripa PF, et al. Lifestyle and clinical risk factors for incident rheumatoid arthritis-associated interstitial lung disease. *J Rheumatol* 2021;48:656–63.
38. Ibfelt EH, Jacobsen RK, Kopp TI, et al. Methotrexate and risk of interstitial lung disease and respiratory failure in rheumatoid arthritis: a nationwide population-based study. *Rheumatology (Oxford)* 2021; 60:346–52.
39. Doyle TJ, Patel AS, Hatabu H, et al. Detection of rheumatoid arthritis-interstitial lung disease is enhanced by serum biomarkers. *Am J Respir Crit Care Med* 2015;191:1403–12.