Rheumatology, 2024, **00**, 1–8
https://doi.org/10.1093/rheumatology/keae615
Advance access publication 6 November 2024
Original Article

Clinical science

MUC5B promoter variant and survival in rheumatoid arthritis-associated interstitial lung disease

Jacob Klein[1], Austin M. Wheeler[1], Joshua F. Baker[2], Yangyuna Yang[1], Punyasha Roul[1], Halie Frideres[1], Katherine D. Wysham[3], Gail S. Kerr[4], Andreas Reimold[5], Dana P. Ascherman[6], Gary A. Kunkel[7], Grant W. Cannon[7], Paul A. Monach[8], Jill A. Poole[9], Geoffrey M. Thiele[1], Ted R. Mikuls[1], Bryant R. England[1,*]

[1]Department of Internal Medicine, Division of Rheumatology & Immunology, VA Nebraska Western Iowa Health Care System & University of Nebraska Medical Center, Omaha, NE, USA
[2]Department of Medicine, Division of Rheumatology, Corporal Michael J. Crescenz VA Medical Center and University of Pennsylvania, Philadelphia, PA, USA
[3]Department of Internal Medicine, Division of Rheumatology, VA Puget Sound Health Care System and University of Washington, Seattle, WA, USA
[4]Department of Medicine, Division of Rheumatology, Washington D.C. VA, Howard University, & Georgetown University, Washington, DC, USA
[5]Department of Internal Medicine, Rheumatic Diseases Division, Dallas VA & University of Texas Southwestern, Dallas, TX, USA
[6]Department of Medicine, Division of Rheumatology & Clinical Immunology, Pittsburgh VA & University of Pittsburgh, Pittsburgh, PA, USA
[7]Department of Internal Medicine, Division of Rheumatology, VA Salt Lake City Health Care System and University of Utah, Salt Lake City, UT, USA
[8]Boston VA, Boston, MA, USA
[9]Department of Internal Medicine, Division of Allergy & Immunology, University of Nebraska Medical Center, Omaha, NE, USA

*Correspondence to: Bryant England, Department of Internal Medicine, Division of Rheumatology & Immunology, VA Nebraska Western Iowa Health Care System & University of Nebraska Medical Center, 986270 Nebraska Medicine, Omaha, NE 68198-6270, USA. E-mail: Bryant.england@unmc.edu

Abstract

Objective: The objective of this study was to investigate the association between the *MUC5B* rs35705950 promoter variant and survival in RA-associated interstitial lung disease (RA-ILD).

Methods: We studied participants in the Veteran Affairs Rheumatoid Arthritis (VARA) registry with validated ILD diagnoses. Participants were followed until death or till the end of the study period. The *MUC5B* rs35705950 promoter variant was measured using an Infinium genotyping array, assuming autosomal dominant inheritance. Survival and cause of death were determined from VA death records and the National Death Index. Associations of the *MUC5B* promoter variant with survival were tested in Cox regression models, adjusting for potential confounders.

Results: Among 263 participants with RA-ILD (mean age 69 years, 95% male, 73% White, 85% smoking history), the *MUC5B* promoter variant was present in 33.5%. The mortality rate was similar between those with [12.2/100 PY (95% CI: 9.4, 15.8)] and without [11.1/100 PY (95% CI: 9.1, 13.5)] the variant. *MUC5B* status was not significantly associated with survival overall [aHR 0.97 (95% CI: 0.68, 1.37)] or when stratified by ILD pattern [clinical usual interstitial pneumonia (UIP) aHR 0.86 (95% CI: 0.55, 1.35); clinical non-UIP aHR 1.15 (95% CI: 0.63, 2.09)]. Further, *MUC5B* status was not significantly associated with respiratory-related [aHR 0.83 (95% CI: 0.42, 1.66)] or non-respiratory causes of death [aHR 1.08 (95% CI: 0.72, 1.62)].

Conclusion: While associated with RA-ILD risk, the *MUC5B* promoter variant was not predictive of survival among RA-ILD patients in this multicentre cohort. Further studies are needed to identify other genetic and non-genetic prognostic factors in RA-ILD to inform disease management.

Keywords: rheumatoid arthritis, interstitial lung disease, genetics, survival.

Rheumatology key messages

- The *MUC5B* promoter variant is the strongest known genetic risk factor for rheumatoid arthritis-associated interstitial lung disease (RA-ILD).
- In this multicenter study, *MUC5B* status was not significantly associated with survival in RA-ILD patients.
- *MUC5B* appears to be more indicative of RA-ILD susceptibility than RA-ILD survival prognosis.

Received: 5 July 2024. **Accepted:** 26 October 2024

Introduction

RA is a systemic inflammatory disease that affects ∼0.5–1% of the population, with lung involvement being a common extra-articular manifestation [1]. RA-associated interstitial lung disease (RA-ILD) is clinically diagnosed in ∼10–15% of patients with RA [2–4], and 30–40% of RA patients may have subclinical interstitial lung abnormalities detectable with advanced imaging [5–7]. RA-ILD causes substantial morbidity and premature mortality, with a median survival of 3–8 years after diagnosis [8, 9]. Risk factors for RA-ILD that have been identified include older age, male sex, cigarette smoking, severity of RA disease activity, and the *MUC5B* rs35705950 promoter variant [9–12]. Despite the significant impact of ILD on quality of life and survival, a standard screening approach for diagnosing and risk stratifying these patients is lacking.

RA-ILD and idiopathic pulmonary fibrosis (IPF) have significant histo-radiologic overlap and share several disease characteristics, including environmental risk factors, predilection for a usual interstitial pneumonia (UIP) pattern, and a high mortality rate [1, 13, 14]. The gain of function *MUC5B* rs35705950 promoter variant is the strongest known genetic risk factor for IPF [10, 12, 15, 16]. Given the numerous similarities between IPF and RA-ILD, the *MUC5B* promoter variant has been evaluated as a genetic risk factor for RA-ILD. An international study by Juge *et al.* demonstrated the *MUC5B* promoter variant to be strongly associated with RA-ILD, but only RA-ILD in a UIP pattern [15]. Additional studies confirmed a 2- to 3-fold higher odds of RA-ILD in a UIP pattern when the *MUC5B* promoter variant was present [10]. Importantly, the *MUC5B* promoter variant has not been associated with ILD related to other autoimmune diseases, including SSc or myositis [15, 17, 18].

The *MUC5B* gene encodes the MUC5B protein found in bronchial epithelium, which contributes to mucus viscosity and mucociliary clearance of the respiratory system [15]. Mechanisms underpinning the potential pathogenic role of the *MUC5B* promoter variant on ILD development have not been fully elucidated. However, it is hypothesized that overexpression of MUC5B disrupts repair mechanisms and interferes with mucociliary clearance [19]. While studies have consistently replicated an association of the *MUC5B* promoter variant with RA-ILD risk, it is unclear whether this mutation is also prognostic of a more severe disease course. In IPF, survival outcomes stratified by the presence of the *MUC5B* promoter variant have been conflicting. Some studies have demonstrated no impact on survival, while others have shown improved survival when the variant was present [20–23]. There is limited evidence on survival outcomes related to the *MUC5B* promoter variant in RA-ILD. An international, retrospective cohort study of RA-ILD patients found no influence of *MUC5B* on survival or decline in pulmonary function in RA-ILD patients [24]. However, the median follow-up duration was only 3.5 years in that study, and patient selection methods from specialized centres may have introduced referral bias.

The objective of this study was to determine whether the *MUC5B* rs35705950 promoter variant was prognostic of survival in a multicentre cohort of US veterans with RA-ILD. Based on the existing literature in IPF and RA-ILD, we hypothesized that the presence of the *MUC5B* rs35705950 promoter variant would not significantly predict survival.

Methods

Study design

We conducted a cohort study of participants in the Veterans Affairs Rheumatoid Arthritis (VARA) registry with RA-ILD. The VARA registry is a multicentre, prospective cohort study of US veterans with rheumatologist-diagnosed RA fulfilling the 1987 or 2010 ACR criteria, and the registry initiated enrolment in 2003 [25, 26]. All participants provided written informed consent prior to enrolment. Each participating site in this study received Institutional Review Board approval.

RA-ILD validation

ILD was identified through systematic screening and validation of all participants in the VARA registry, as reported in prior work [9, 10, 27–29]. Participants who screened positive for ILD had diagnoses validated through a standardized medical record review performed by a board-certified rheumatologist. Requirements for validated ILD diagnoses included documentation of a provider diagnosis of RA-ILD and either imaging (predominately chest CT) or lung biopsy findings consistent with RA-ILD. Participants who did not meet the RA-ILD validation requirements were excluded. Participants with RA-ILD were followed from the latest of either VARA enrolment or ILD diagnosis (i.e. index date) until death or till the end of the study period (31 December 2021).

The clinical ILD pattern was obtained from chest CT clinical reports and categorized as UIP or non-UIP, as previously reported [30]. When an ILD pattern was not documented, the presence of honeycombing on chest CT reports was also considered evidence of clinical UIP pattern. Those with clinical non-specific interstitial pneumonia, other, and unknown radiologic pattern were classified as clinical non-UIP. Pulmonary function test (PFT) results were obtained from medical records. Baseline values of forced vital capacity (FVC) and diffusion capacity of carbon monoxide (DLCO) were the values most proximate to the index date [median (IQR) 91 (21, 368) days].

Genotyping

Genotyping was performed using the Infinium Global Screening Array-24 v2.0 (Illumina, Inc; San Diego, CA), and DNA was collected at registry enrolment, as previously described [10]. *MUC5B* rs35705950 was directly genotyped and was analysed assuming autosomal dominant inheritance, based on prior literature [10, 15]. The wild-type allele was defined by the guanine (G) nucleotide, while the variant allele was defined by the thymine (T) nucleotide.

Study outcomes

The primary outcome of this study was survival, determined from VA death records and the National Death Index (NDI). Cause of death was determined by International Classification of Diseases (ICD)-10 codes in the NDI and categorized as respiratory or non-respiratory according to the Centers for Disease Control and Prevention (CDC) chapters [31]. One exception to this categorization was made for 'rheumatoid lung' codes (ICD-10: M05.1X), which were considered respiratory causes of death.

Demographic and clinical variables

Demographic variables and health behaviour data, including age, sex, race, cigarette smoking status (categorized as ever *vs*

never used), and education level, were collected at registry enrolment. RA disease activity measures [28-joint Disease Activity Score (DAS28)] and functional status [multidimensional HAQ (MDHAQ)] were collected longitudinally during routine care. DMARD use at the index date was obtained from VA pharmacy dispensing data in the VA Corporate Data Warehouse (CDW). RA disease duration was obtained from the registry, while anti-CCP and RF were measured from serum collected at registry enrolment with standardized assays. Comorbid chronic obstructive pulmonary diseases (COPDs) or asthma were identified by the presence of at least two diagnostic codes from inpatient or outpatient encounters within linked administrative VA data, as described previously [27].

Statistical analysis

Crude mortality rates and Kaplan–Meier curves were generated, stratified by *MUC5B* status. We utilized unadjusted and adjusted Cox regression models to assess the association of *MUC5B* status with survival. Covariables in all adjusted models included age, sex, race, smoking history, baseline DAS28, and ILD duration. We further adjusted for baseline FVC% predicted in separate models, given missingness. A subgroup analysis of incident RA-ILD cases that developed after registry enrolment was conducted to assess for the potential of survival bias in the primary analyses (i.e. patients with more severe RA-ILD dying before enrolling in the registry). These analyses used similar Cox regression models as in the primary analysis.

In the secondary analyses, we evaluated the association of *MUC5B* status with survival stratified by clinical UIP or non-UIP ILD pattern. We additionally evaluated cause-specific mortality risk related to *MUC5B* status, separately evaluating respiratory-related deaths and non-respiratory–related deaths, while censoring for other causes of death. Secondary analyses used similar Cox regression models as the primary analyses. The proportional hazards assumption was tested by Schoenfeld residuals and negative log-log survival plots, which did not indicate violations. Complete case analysis was utilized in regression models. All analyses were completed using Stata MP 18.0 within the VA Informatics and Computing Infrastructure environment.

Results

Baseline patient characteristics

We studied 263 participants within the VARA registry with validated RA-ILD and *MUC5B* genotyping. The cohort had a mean age of 69.0 years and was predominantly male (94.7%) and reported White race (73.0%) (Table 1). A history of former or current cigarette smoking was present in 84.8% of participants, and 61.1% had concomitant COPD or asthma at baseline. The mean RA disease duration at index date was 12.4 years, and the majority of participants were seropositive for anti-CCP antibody (84.0%). The mean (S.D.) ILD duration at enrolment was 1.4 years (2.9), with a baseline FVC % predicted of 78.3% (17.5) and DLCO % predicted of 61.8% (21.4). Clinical UIP was the most common (56.7%) radiologic pattern. MTX (29.7%), biologic DMARDs (33.8%) and prednisone (38.0%) were the most commonly used medications.

The *MUC5B* variant (GT/TT genotype) was present in 33.5% of those with the variant allele, 31.9% were heterozygous (GT), and 1.5% were homozygous (TT). White race (85.2% *vs* 66.9%), older age (mean 71.3 *vs* 67.9 years), male sex (98.9% *vs* 92.6%) and COPD/asthma (69.3 *vs* 57.1%) were more frequent among RA-ILD participants with the GT/TT *vs* GG genotype. Other patient characteristics were similar between GT/TT and GG genotypes.

Table 1. Baseline patient characteristics

Characteristic	All RA-ILD ($n = 263$)	*MUC5B* GT/TT ($n = 88$)	*MUC5B* GG ($n = 175$)	*P*-value
Demographics and comorbidities				
Age, mean (S.D.), years	69.0 (9.0)	71.3 (8.1)	67.9 (9.2)	0.004
Male sex, %	94.7	98.9	92.6	0.03
White race, %	73.0	85.2	66.9	0.002
High school education, %	79.5	79.5	79.4	0.75
Ever smoker, %	84.8	84.1	85.1	0.65
BMI, mean (S.D.), kg/m^2	28.2 (5.6)	28.4 (5.4)	28.0 (5.8)	0.60
COPD/asthma, %	61.2	69.3	57.1	0.03
RA disease characteristics				
RA duration, mean (S.D.), years	12.4 (12.1)	11.3 (10.9)	12.9 (12.6)	0.34
Anti-CCP positive, %	84.0	87.5	82.3	0.28
RF positivity, %	82.9	80.7	84.0	0.67
DAS28, mean (S.D.)	3.9 (1.3)	3.9 (1.3)	4.0 (1.4)	0.50
MDHAQ, mean (S.D.)	1.0 (0.6)	0.9 (0.6)	1.0 (0.6)	0.35
MTX use, %	29.7	32.0	25.0	0.29
Biologic use, %	33.8	33.7	34.1	0.85
Prednisone use, %	38.0	38.3	37.5	0.99
ILD-related measures				
ILD duration, mean (S.D.), years	1.4 (2.9)	1.7 (3.0)	1.2 (2.9)	0.21
Clinical UIP pattern, %	56.7	61.4	54.3	0.43
FVC % predicted, mean (S.D.)	78.3 (17.5)	76.0 (15.9)	79.3 (18.2)	0.21
DLCO % predicted, mean (S.D.)	61.8 (21.4)	61.5 (18.4)	61.9 (22.8)	0.92

Values are mean (S.D.) or % of non-missing. Missing data: high school education, $n = 21$; smoking history, $n = 5$; BMI, $n = 8$; COPD, $n = 8$; RA disease duration, $n = 5$; DAS28, $n = 15$; MDHAQ, $n = 14$; MTX use, $n = 8$; biologic use, $n = 8$; prednisone use, $n = 8$; FVC % predicted, $n = 64$; DLCO % predicted, $n = 128$. Anti-CCP: anti-CCP antibody; COPD: chronic obstructive pulmonary disease; DAS28: 28-joint Disease Activity Score; MDHAQ: multi-dimensional HAQ; ILD: interstitial lung disease; UIP: usual interstitial pneumonia; FVC: forced vital capacity; DLCO: diffusing capacity of carbon monoxide.

Mortality rates and causes of death

Over 1363 patient-years (PY) of follow-up, 156 deaths occurred in the primary analysis, with a median follow-up of 4.3 years. The mortality rates based on *MUC5B* status were similar, with the GG genotype experiencing a mortality rate of 11.1/100 PY (95% CI: 9.1, 13.5) and the GT/TT genotype a rate of 12.2/100 PY (95% CI: 9.4, 15.8). The median survival did not differ by *MUC5B* status (GG 6.3 *vs* GT/TT 6.1 years). Circulatory and respiratory-related deaths were the two leading causes of death, accounting for 28.9% and 26.3%, respectively (Table 2). Other common causes of death included neoplasms (17.8%), diseases of the musculoskeletal system (8.6%), and infection (3.2%). When comparing cause of death by *MUC5B* status, diseases of the circulatory system (GG: 28.9% and GT/TT: 29.1%) and the respiratory system (GG 27.8% and GT/TT: 23.6%) remained the most common causes.

Survival based on *MUC5B* status

Kaplan–Meier curves stratified by *MUC5B* status are shown in Fig. 1, with no differences in crude survival based on *MUC5B* status ($P = 0.52$ by log-rank test). In unadjusted Cox regression models, *MUC5B* status was not associated with survival [hazard ratio (HR) 1.12 (95% CI: 0.80, 1.55)] (Fig. 2). Findings were similar in adjusted models [adjusted HR (aHR) 0.97 (95% CI: 0.68, 1.37)], including those adjusting for baseline FVC % predicted [aHR 0.80 (95% CI: 0.52, 1.22)].

To address potential survival bias, we evaluated survival by *MUC5B* status among those with incident RA-ILD ($n = 156$). Participants with the GG genotype had a mortality rate of 11.7/100 PY (95% CI: 9.1, 15.0), and those with the GT/TT genotype had a rate of 12.2/100 PY (95% CI: 8.3, 17.9). There were no associations between *MUC5B* status and survival in unadjusted [HR 1.06 (95% CI: 0.67, 1.68)] or adjusted [aHR 0.95 (95% CI: 0.58, 1.56)] regression models (Fig. 2).

Secondary analysis of ILD pattern and cause-specific mortality

Among RA-ILD participants with clinical UIP ($n = 149$), the findings were similar, demonstrating no significant association of *MUC5B* status with survival [aHR of 0.86 (95% CI: 0.55, 1.35)] (Fig. 3). The findings were also similar among 114 RA-ILD participants with a clinical non-UIP pattern [aHR of 1.15 (95% CI: 0.63, 2.09)]. In cause-specific mortality analyses, *MUC5B* status was not significantly associated with either respiratory-related death [aHR 0.83 (95% CI: 0.42, 1.66)] or non-respiratory causes of death [aHR 1.08 (95% CI: 0.72, 1.62)].

Discussion

RA-ILD is a common extra-articular manifestation of RA that causes significant morbidity and premature mortality among affected individuals. Despite recognizing its impact on patient outcomes, we lack effective approaches to screen for RA-ILD and risk-stratify RA-ILD patients early in the disease process. The *MUC5B* rs35705950 promoter variant has been established as the strongest genetic risk factor for RA-ILD [10, 15, 16], yet few studies have explored the prognostic value of this variant in RA-ILD. Utilizing a multicentre prospective RA cohort and robust data linkages, we systematically identified participants with validated RA-ILD, to evaluate the association of *MUC5B* genotype with survival. Our study did not observe any influence of the *MUC5B* promoter variant on overall survival, among participants with clinical UIP RA-ILD, or for specific causes of death. Thus, while a strong risk factor for RA-ILD onset, it does not appear to inform the likelihood of disease progression to death.

Our findings are in agreement with those of Juge *et al.*, who showed that the presence of the *MUC5B* rs35705950 promoter variant did not influence transplant-free survival or decline in pulmonary function tests in a large, international, retrospective cohort study of 261 RA-ILD patients [24]. However, the participants of that study were predominantly referrals from large tertiary centres, which has the potential to introduce referral or selection bias into the results. Our study was uniquely able to assess RA-ILD prognosis by *MUC5B* status by systematically capturing ILD cases within a large, multicentre, prospective RA cohort study. Further, we were able to confirm our findings among a subgroup of patients with incident RA-ILD. Our findings further establish that the *MUC5B* rs35705950 promoter variant aids in identifying individuals at risk for developing RA-ILD, but does not effectively predict disease outcomes in RA-ILD.

Among CTD-ILD, RA-ILD patients are uniquely predisposed to developing ILD in a UIP pattern [1, 13]. In RA-ILD,

Table 2. Causes of death in RA-ILD participants

ICD-10 Chapter	Overall *N* (%) of deaths	*MUC5B* GT/TT *N* (%) of deaths	*MUC5B* GG *N* (%) of deaths
Total available recorded causes of death[a]	$N = 152$	$N = 55$	$N = 97$
Diseases of circulatory system	44 (28.9)	16 (29.1)	28 (28.9)
Diseases of respiratory system[b]	40 (26.3)	13 (23.6)	27 (27.8)
Neoplasms[c]	27 (17.8)	10 (18.2)	17 (17.5)
Diseases of musculoskeletal system	13 (8.6)	7 (12.7)	6 (6.2)
Certain infection and parasitic disease	5 (3.2)	0 (0.0)	5 (5.2)
Diseases of digestive system	4 (2.6)	2 (3.6)	2 (2.1)
Diseases of nervous system	3 (2.0)	2 (3.6)	1 (1.0)
Other[d]	16 (10.5)	5 (10.0)	11 (11.3)

[a] Missing data for cause of death by ICD-10 code ($n = 4$).
[b] Deaths due to interstitial lung disease ($n = 21$; 52.5% of respiratory).
[c] Deaths due to lung cancer ($n = 19$, 70.4% of neoplasms).
[d] Other includes: symptoms, signs and abnormal clinical and laboratory findings not classified elsewhere ($n = 3$); codes for special purposes ($n = 3$); external causes of morbidity and mortality ($n = 3$); diseases of genitourinary system ($n = 2$); endocrine, nutritional, and metabolic diseases ($n = 2$); mental and behavioural disorders ($n = 2$); diseases of skin and s.c.tissue ($n = 1$).

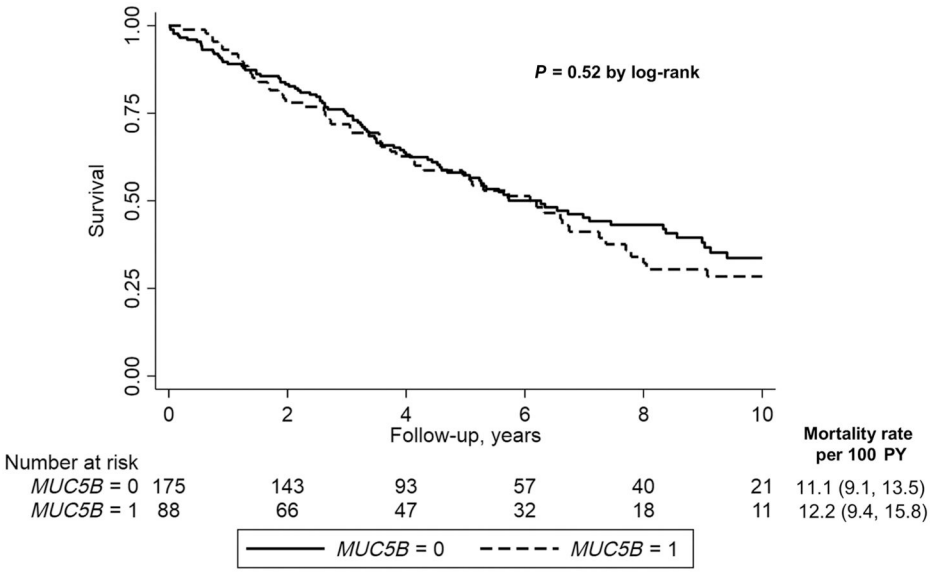
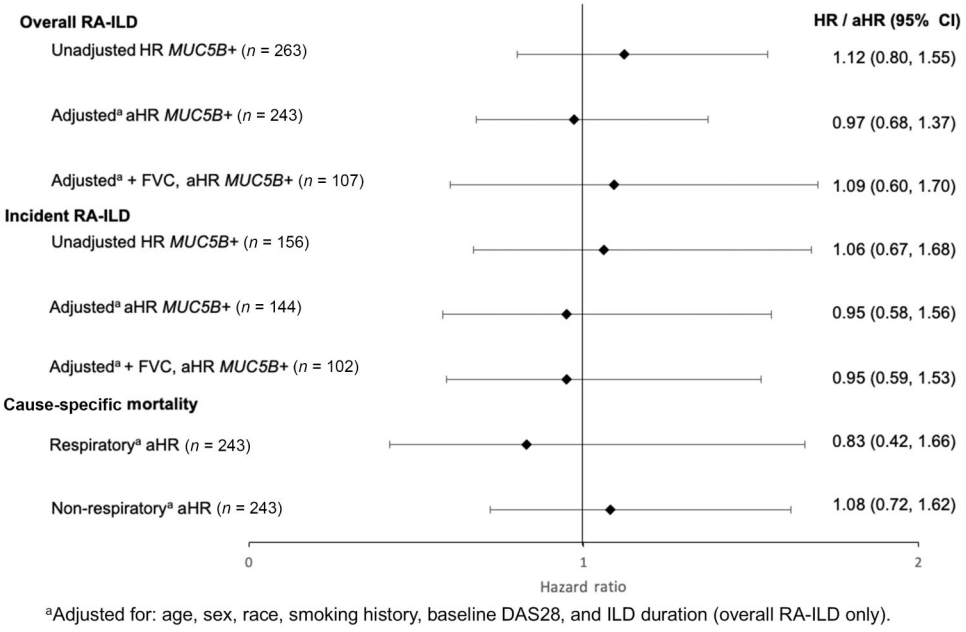


Figure 1. Survival in RA-associated interstitial lung disease (RA-ILD) stratified by *MUC5B* status. Kaplan–Meier curve depicting survival in patients with RA-ILD based on *MUC5B* rs35705950 promoter variant genotype (GG *vs* GT/TT). PY: patient years; *MUC5B* = 0: absence of *MUC5B* rs35705950 promoter variant (GG genotype); *MUC5B* = 1: presence of *MUC5B* rs35705950 promoter variant (GT/TT genotype)



[a]Adjusted for: age, sex, race, smoking history, baseline DAS28, and ILD duration (overall RA-ILD only).

Figure 2. Association between *MUC5B* status and survival in unadjusted and adjusted Cox regression models. Forest plot depicting unadjusted and adjusted associations of *MUC5B* genotypes (GT/TT *vs* GG) with survival in the overall RA-ILD and incident RA-ILD cohorts. Adjusted models account for age, sex, race, smoking status, baseline DAS28, and ILD duration (overall RA-ILD only). Additional analyses adjusted for baseline FVC % predicted. aHR: adjusted hazard ratio; HR: hazard ratio; DAS28: 28-joint DAS; FVC: forced vital capacity; ILD: interstitial lung disease

the UIP pattern has generally been associated with worse survival outcomes compared with non-UIP [32], and the *MUC5B* promoter variants confers risk only for RA-ILD in a UIP pattern [10, 15]. Therefore, we investigated whether the prognostic value of *MUC5B* was dependent upon RA-ILD pattern. Among both clinical UIP and non-UIP patterns, we found no significant impact of *MUC5B* status on survival. Similarly, we evaluated whether *MUC5B* may be differentially prognostic of cause-specific mortality. As expected, the main causes of death among the cohort, regardless of *MUC5B* status, were deaths due to cardiovascular- and respiratory-related diseases, with ILD being the most frequent cause of respiratory-related deaths. The *MUC5B* promoter variant was not significantly associated with either respiratory or non-respiratory mortality risk. These null findings by clinical ILD pattern and for specific causes of death strengthen the conclusion that the *MUC5B* rs35705950 promoter variant does not have a meaningful impact on RA-ILD survival.

Although our study did not demonstrate any significant association between survival and the presence of the *MUC5B* rs35705950 promoter variant, there is the potential for index event or collider bias. Index event bias is a type of selection

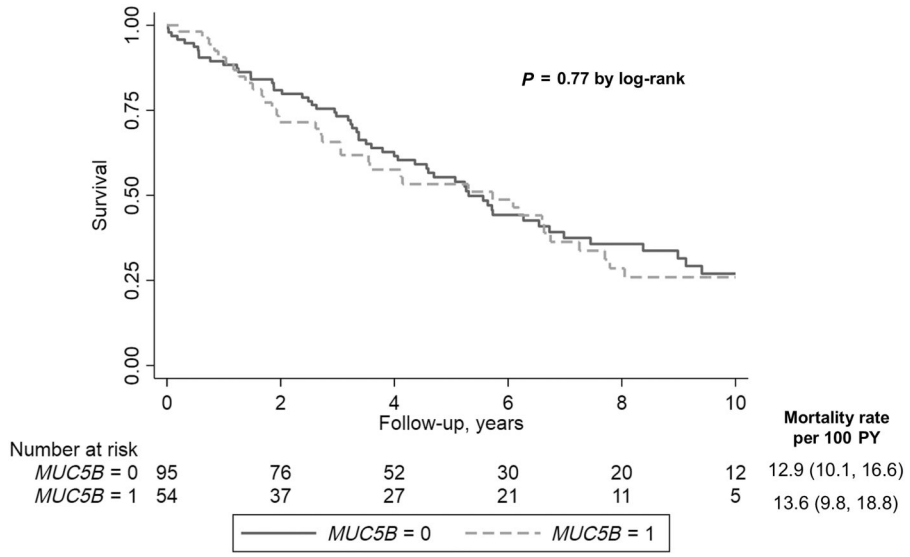


Figure 3. Survival in RA-associated clinical usual interstitial pneumonia (UIP) stratified by *MUC5B* status. Kaplan–Meier survival curve depicting survival outcomes over a 10-year period in clinical UIP RA-associated interstitial lung disease (RA-ILD). The comparison is based on the presence or absence of the *MUC5B* rs35705950 promoter variant (GT/TT *vs* GG genotype). PY: patient years; *MUC5B* = 0: absence of *MUC5B* rs35705950 promoter variant (GG genotype); *MUC5B* = 1: presence of *MUC5B* rs35705950 promoter variant (GT/TT genotype)

bias recognized as contributing to the 'risk factor paradox' observed in some rheumatic disease studies [33]. This bias arises when conditioning on an outcome leads to spurious associations between risk factors and a subsequent outcome. In our study, restricting the population to patients with RA-ILD could lead to such a bias, since the *MUC5B* promoter variant is a risk factor for the development of RA-ILD. This could alter associations of risk factors with survival if patients with the *MUC5B* rs35705950 promoter variant do not require the same burden of other risk factors to initially develop RA-ILD. The presence of the *MUC5B* variant may falsely protect RA-ILD patients from other known and unknown risk factors, as epidemiologic models may estimate the *MUC5B* variant to falsely protect RA-ILD patients from other known and unknown risk factors. The anticipated result in this situation would be a bias towards the null or even protection from death when estimating associations between the *MUC5B* promoter variant and survival. As genome-wide association studies have increased, the potential for collider bias (also termed index event bias) to impact prognostic studies has become a greater concern. Dudbridge *et al.* performed a simulation study exploring the potential for index event bias in relation to the *MUC5B* rs35705950 promoter and survival in IPF [34]. The findings reversed the paradoxical association and demonstrated worse survival when the variant was present. However, Dudbridge *et al.* revisited this concept in a later study using alternative methods and found insufficient evidence to suggest index event bias significantly influenced the association between the *MUC5B* promoter variant and survival and IPF [23]. Thus, the degree to which index event bias may impact genetic prognostic studies remains unsettled.

Additional limitations to this study include the potential for limited generalizability, given that the participants were from the U.S. Veteran predominately male population, with high smoking rates. While these characteristics are risk factors for RA-ILD, particularly RA-UIP, our findings may not be generalizable to RA-ILD cohorts with more females, younger ages, and less RA-UIP. Reflecting variation in real-world testing frequency, baseline FVC values were not available for all participants. Other longitudinal RA-ILD severity measures were also not available for these analyses, including the degree of involvement on chest CT and dyspnea measures. Smoking status was included as a covariate, but we did not have data on smoking intensity or duration. Similarly, information on other inhalant-related exposures were not available. RA-ILD characterization was retrospective, which could have resulted in misclassification of ILD and limited the availability of ILD pattern data. Moreover, ILD pattern was determined by clinically available imaging reports, pathology reports, and the treating provider's determination. While this reflects typical clinical practice situations, there may be misclassification of ILD pattern compared with other research method determinations, such as independent review by expert chest radiologists.

In conclusion, in this large, multicentre study of US veterans with RA-ILD, the *MUC5B* rs35705950 promoter variant was not associated with all-cause or cause-specific mortality. Though the *MUC5B* promoter variant is a well-established genetic risk factor for RA-ILD susceptibility, this genetic marker does not serve as an effective prognostic marker to risk stratify patients. Future studies should be conducted exploring how other genetic and non-genetic prognostic factors might influence or predict the natural course of RA-ILD. With the advancement in therapeutic modalities for RA-ILD, there is a significant need to risk stratify patients to provide individualized, patient-centred treatment plans.

This study complies with the Declaration of Helsinki. All participants provided written informed consent, and each site received Institutional Review Board Approval. This study was approved by the VA Central Institutional Review Board.

Data availability

Data are available upon reasonable request to the corresponding author (Bryant.england@unmc.edu) after the obtainment of necessary regulatory approvals, including data use agreements and institutional review board approvals.

Funding

B.R.E. received support from the U.S. Department of Veterans Affairs Clinical Sciences Research & Development (CX002203) and the Rheumatology Research Foundation. J. F.B. is supported by the U.S. Department of Veterans Affairs Clinical Sciences Research & Development (CX001703) and the U.S. Department of Veterans Affairs Rehabilitation Research & Development (RX003644). T.R.M. is supported by the U.S. Department of Veterans Affairs Biomedical Laboratory Research & Development (BX004660 and BX003635), the U.S. Department of Defense (PR200793), the Rheumatology Research Foundation, and the National Institute of General Medical Sciences (U54 GM115458). K. D.W. is supported by the U.S. Department of Veterans Affairs Clinicl Sciences Research & Development (CX002351). J.A.P. is supported through grants from the National Institute for Occupational Safety (R01OH012045 (JAP), the Department of Defence (PR200793), and the Central States Center of Agricultural Safety and Health (CS-CASH)).

Disclosure statement: B.R.E. has received consulting and research support from Boehringer-Ingelheim. J.F.B. has received consulting fees from Bristol-Myers Squibb, Pfizer, CorEvitas, and Burns-White, L.L.C. has received research funding from Horizon Therapeutics. T.R.M. has consulted with Pfizer, UCB, Olatech Therapeutics and Horizon Therapeutics and received research funding from Horizon Therapeutics. G.S.K. has consulted with BMS, Sanofi, Janssen, Horizon, Pfizer and UCB. J.A.P. has received research reagents from AstraZeneca (no monies) and has been a site investigator for allergy and asthma clinical studies for Takeda, GlaxoSmithKline, Regeneron, Areteia, and AstraZeneca (no monies).

Acknowledgements

This work was supported by the Joint VA-DoD Mortality Data Repository and the Centers for Disease Control and Prevention, National Death Index. The views expressed in this article are those of the authors and do not necessarily reflect the position or policy of the Department of Veterans Affairs or the United States government.

References

1. Kim EJ, Collard HR, King TE. Jr. Rheumatoid arthritis-associated interstitial lung disease: the relevance of histopathologic and radiographic pattern. Chest 2009;136:1397–405.
2. Spagnolo P, Lee JS, Sverzellati N, Rossi G, Cottin V. The lung in rheumatoid arthritis: focus on interstitial lung disease. Arthritis Rheumatol 2018;70:1544–54.
3. Bendstrup E, Møller J, Kronborg-White S, Prior TS, Hyldgaard C. Interstitial lung disease in rheumatoid arthritis remains a challenge for clinicians. J Clin Med 2019;8:2038.
4. Olson AL, Swigris JJ, Sprunger DB *et al.* Rheumatoid arthritis-interstitial lung disease-associated mortality. Am J Respir Crit Care Med 2011;183:372–8.
5. Habib HM, Eisa AA, Arafat WR, Marie MA. Pulmonary involvement in early rheumatoid arthritis patients. Clin Rheumatol 2011; 30:217–21.
6. Juge P-A, Granger B, Debray M-P *et al.* A risk score to detect subclinical rheumatoid arthritis-associated interstitial lung disease. Arthritis Rheumatol 2022;74:1755–65.
7. 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.
8. 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.
9. Brooks R, Baker JF, Yang Y *et al.* The impact of disease severity measures on survival in U.S. veterans with rheumatoid arthritis-associated interstitial lung disease. Rheumatology (Oxford) 2022; 61:4667–77.
10. Wheeler AM, Baker JF, Poole JA *et al.* Genetic, social, and environmental risk factors in rheumatoid arthritis-associated interstitial lung disease. Semin Arthritis Rheum 2022;57:152098.
11. Samhouri BF, Vassallo R, Achenbach SJ *et al.* Incidence, Risk Factors, and Mortality of Clinical and Subclinical Rheumatoid Arthritis–Associated Interstitial Lung Disease: a Population-Based Cohort. Arthritis Care &amp. Research 2022;74:2042–9.
12. McDermott G, Gill R, Gagne S *et al.* Associations of the MUC5B promoter variant with timing of interstitial lung disease and rheumatoid arthritis onset. Rheumatology (Oxford) 2022; 61:4915–23.
13. Lee H-K, Kim DS, Yoo B *et al.* Histopathologic pattern and clinical features of rheumatoid arthritis-associated interstitial lung disease. Chest 2005;127:2019–27.
14. Kelly CA, Saravanan V, Nisar M *et al.*; British Rheumatoid Interstitial Lung (BRILL) Network. Rheumatoid arthritis-related interstitial lung disease: associations, prognostic factors and physiological and radiological characteristics—a large multicentre UK study. Rheumatology (Oxford) 2014;53:1676–82.
15. Juge P-A, Lee JS, Ebstein E *et al.* MUC5B promoter variant and rheumatoid arthritis with interstitial lung disease. N Engl J Med 2018;379:2209–19.
16. Palomäki A, Palotie A, Koskela J *et al.*; FinnGen. Lifetime risk of rheumatoid arthritis-associated interstitial lung disease in MUC5B mutation carriers. Ann Rheum Dis 2021;80:1530–6.
17. Johnson C, Rosen P, Lloyd T *et al.* Exploration of the MUC5B promoter variant and ILD risk in patients with autoimmune myositis. Respir Med 2017;130:52–4.
18. Lee MG, Lee YH. A meta-analysis examining the association between the MUC5B rs35705950 T/G polymorphism and susceptibility to idiopathic pulmonary fibrosis. Inflamm Res 2015; 64:463–70.
19. Zhang Q, Wang Y, Qu D, Yu J, Yang J. The possible pathogenesis of idiopathic pulmonary fibrosis considering MUC5B. Biomed Res Int 2019;2019:9712464.
20. Peljto AL, Zhang Y, Fingerlin TE *et al.* Association between the MUC5B promoter polymorphism and survival in patients with idiopathic pulmonary fibrosis. JAMA 2013;309:2232–9.
21. van der Vis JJ, Snetselaar R, Kazemier KM *et al.* Effect of Muc5b promoter polymorphism on disease predisposition and survival in idiopathic interstitial pneumonias. Respirology 2016;21:712–7.
22. Dhooria S, Bal A, Sehgal IS *et al.* MUC5B promoter polymorphism and survival in indian patients with idiopathic pulmonary fibrosis. Chest 2022;162:824–7.
23. Cai S, Allen RJ, Wain LV, Dudbridge F. Reassessing the association of MUC5B with survival in idiopathic pulmonary fibrosis. Ann Hum Genet 2023;87:248–53.
24. Juge P-A, Solomon JJ, van Moorsel CHM *et al.* MUC5B promoter variant rs35705950 and rheumatoid arthritis associated interstitial lung disease survival and progression. Semin Arthritis Rheum 2021;51:996–1004.
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;31:315–24.
26. Kay J, Upchurch KS. ACR/EULAR 2010 rheumatoid arthritis classification criteria. Rheumatology 2012;51(Suppl 6):vi5–9.

27. England BR, Sayles H, Michaud K *et al.* Chronic lung disease in U.S. Veterans with rheumatoid arthritis and the impact on survival. Clin Rheumatol 2018;37:2907–15.
28. England BR, Roul P, Mahajan TD *et al.* Performance of administrative algorithms to identify interstitial lung disease in rheumatoid arthritis. Arthritis Care Res (Hoboken) 2020;72:1392–403.
29. England BR, Duryee MJ, Roul P *et al.* Malondialdehyde-acetaldehyde adducts and antibody responses in rheumatoid arthritis-associated interstitial lung disease. Arthritis Rheumatol 2019;71:1483–93.
30. Luedders BA, Wheeler AM, Ascherman DP *et al.* Plasma matrix metalloproteinase concentrations and risk of interstitial lung disease in a prospective rheumatoid arthritis cohort. Arthritis Rheumatol 2024; 76:1013–22.
31. England BR, Sokolove J, Robinson WH *et al.* Associations of circulating cytokines and chemokines with cancer mortality in men with rheumatoid arthritis. Arthritis Rheumatol 2016;68:2394–402.
32. Singh N, Varghese J, England BR *et al.* Impact of the pattern of interstitial lung disease on mortality in rheumatoid arthritis: a systematic literature review and meta-analysis. Semin Arthritis Rheum 2019;49:358–65.
33. Choi HK, Nguyen US, Niu J, Danaei G, Zhang Y. Selection bias in rheumatic disease research. Nat Rev Rheumatol 2014; 10:403–12.
34. Dudbridge F, Allen RJ, Sheehan NA *et al.* Adjustment for index event bias in genome-wide association studies of subsequent events. Nat Commun 2019;10:1561.

Rheumatology, 2024, 00, 1–8
https://doi.org/10.1093/rheumatology/keae615
Original Article