RHEUMATOLOGY

Rheumatology 2021;60:3689–3698
doi:10.1093/rheumatology/keaa836
Advance Access publication 18 January 2021

Original article

Prevalence, incidence and cause-specific mortality of rheumatoid arthritis–associated interstitial lung disease among older rheumatoid arthritis patients

Jeffrey A. Sparks [1,2], Yinzhu Jin [3], Soo-Kyung Cho[4], Seanna Vine[3], Rishi Desai[2,3], Tracy J. Doyle[2,5] and Seoyoung C. Kim [1,2,3]

Abstract

Objective. We aimed to investigate the prevalence, incidence and cause-specific mortality of RA-associated interstitial lung disease (RA-ILD) among older US patients with RA.

Methods. We performed a nationwide cohort study using Medicare claims data (parts A, B and D for 2008–2017). RA was identified with a validated algorithm using RA diagnosis codes and DMARD prescription. RA-ILD was identified with a validated algorithm using ILD diagnosis codes by a rheumatologist/pulmonologist. RA-ILD was categorized as prevalent or incident relative to the initial RA observation (baseline/index date). We compared the total mortality of RA-ILD to RA without ILD using multivariable Cox regression, adjusting for baseline covariates. For cause-specific mortality, Fine and Gray subdistribution hazard ratios (sdHRs) were estimated to handle competing risks of alternative mortality causes.

Results. Among 509 787 RA patients (mean age 72.6 years, 76.2% female), 10 306 (2.0%) had prevalent RA-ILD at baseline. After baseline, 13 372 (2.6%) developed RA-ILD during 1 873 127 person-years of follow-up (median 3.0 years/person). During follow-up, 38.7% of RA-ILD patients died compared with 20.7% of RA patients without ILD. After multivariable adjustment, RA-ILD had an HR of 1.66 (95% CI 1.60, 1.72) for total mortality. Accounting for competing risk of other causes of death, RA-ILD had an sdHR of 4.39 (95% CI 4.13, 4.67) for respiratory mortality and an sdHR of 1.56 (95% CI 1.43, 1.71) for cancer mortality compared with RA without ILD.

Conclusions. RA-ILD was present or developed in nearly 5% of patients in this nationwide study of older patients with RA. Compared with RA without ILD, RA-ILD was associated with excess total, respiratory and cancer mortality that was not explained by measured factors.

Key words: rheumatoid arthritis, interstitial lung disease, epidemiology, respiratory, mortality

CLINICAL SCIENCE

Rheumatology key messages

- Rheumatoid arthritis–associated interstitial lung disease (RA-ILD) affected nearly 5% of patients with RA in Medicare.
- Male sex, smoking, respiratory morbidities and RA-related medications were associated with incident RA-ILD.
- RA-ILD was associated with excess total, respiratory and cancer mortality compared with RA without ILD.

[1]Division of Rheumatology, Inflammation, and Immunity, Brigham and Women's Hospital, [2]Harvard Medical School, [3]Division of Pharmacoepidemiology and Pharmacoeconomics, Brigham and Women's Hospital, Boston, MA, USA, [4]Division of Rheumatology, Hanyang University Hospital for Rheumatic Diseases, Seoul, South Korea and [5]Division of Pulmonary and Critical Care Medicine, Brigham and Women's Hospital, Boston, MA, USA

Submitted 20 September 2020; accepted 5 November 2020

Correspondence to: Jeffrey A. Sparks, Division of Rheumatology, Inflammation, and Immunity, Brigham and Women's Hospital, 60 Fenwood Road, #6016U, Boston, MA 02115, USA.
E-mail: jsparks@bwh.harvard.edu

Introduction

Despite therapeutic advances over the last few decades, patients with RA have excess mortality, particularly due to respiratory and cardiovascular causes [1, 2]. RA-associated interstitial lung disease (RA-ILD) is one of the most serious extra-articular manifestations [3], previously associated with worsened survival compared with patients with RA without ILD as well as the general population [4–7]. RA-ILD is also associated with worsened quality of life, functional impairment and increased

healthcare utilization and costs [8, 9]. Some studies suggest that the prevalence of RA-ILD may be increasing [5, 9, 10]. Thus, understanding the epidemiology and clinical course of RA-ILD is an urgent need for researchers, clinicians and patients.

Researchers have used a variety of methods to identify RA-ILD, which has resulted in a wide range of estimates for its prevalence and incidence [11]. Registry-based studies identify RA-ILD retrospectively through a combination of medical records review and research reviews of chest imaging. Prospective measures incorporating protocolized research reads of chest imaging of RA patients without pulmonary symptoms could detect clinically insignificant abnormalities and overestimate RA-ILD prevalence [12–14]. RA-ILD risk factors include older age, male sex and smoking [15–19], but most studies have been unable to investigate cause-specific mortality due to relatively small sample sizes and short follow-ups. Some studies in administrative datasets used ILD diagnosis codes among patients with RA to identify RA-ILD [5, 9], but diagnosis codes alone may be inaccurate [positive predictive value (PPV) of a single ILD code: 43.4%] [20]. Thus, investigating the prevalence, risk factors, incidence and impact on mortality of RA-ILD in a large administrative claims database with lengthy follow-up using a validated algorithm to identify ILD would provide real-world evidence on the burden of RA-ILD among patients with RA.

We aimed to investigate the prevalence, incidence and cause-specific mortality of RA-ILD among older US patients with RA using claims data from Medicare, the national health insurance program used by most older individuals in the USA.

Methods

Study design and population

We performed a retrospective cohort study investigating the prevalence, incidence and mortality of RA-ILD among patients with RA on Medicare. We used claims data from parts A (inpatient coverage), B (outpatient coverage) and D (outpatient medications) from 1 January 2008 to 31 December 2017. Medicare is a US public health plan that primarily insures nearly all Americans $\geq$65 years of age as well as some younger individuals with certain disabilities. The Medicare database contains information on demographics, enrolment start and end dates, dispensed medications, medical diagnoses and procedures. This study protocol was approved by the Institutional Review Board of Partners HealthCare (protocol 2018P001323), which waived the requirement for obtaining informed consent.

Identification of RA

We identified patients with RA by the presence of two or more International Classification of Diseases, Ninth Revision or Tenth Revision (ICD-9 or ICD-10, respectively) codes for RA separated by 7–365 days and at least one prescription for a DMARD over the 365 days after the first RA diagnosis code, as previously reported (PPV 86.2%) [21]. The RA index date (date of entry into the cohort) was defined as the last date fulfilling this algorithm. Patients with RA who had diagnosis codes for ILD in the baseline period that did not fulfil the algorithm for RA-ILD were not included in the analysis to reduce the likelihood that patients with ambiguous RA-ILD status may have been included in the RA without ILD group. We did not require that patients have new-onset RA at the index date.

Identification of prevalent and incident RA-ILD

Among patients ever identified with RA, we identified ILD by the presence of two or more ICD-9 or ICD-10 codes for ILD in any position by a rheumatologist or pulmonologist separated by 7–365 days, as previously validated (PPV 72.4%) [20]. The RA-ILD date was defined as the second billing code for ILD that fulfilled the algorithm.

Prevalent RA-ILD was defined as the presence of RA-ILD before or at the same time as the RA index date. Incident RA-ILD was defined as occurring after the RA index date.

Total and cause-specific mortality

Vital status and date of death were obtained through linkage with the National Death Index. Data on cause of death were only available through 31 December 2016 through the National Death Index. Cause of death was categorized as cardiovascular, cancer, respiratory, infection or other (mutually exclusive) by the primary cause of death through ICD-10 codes.

Covariates

We identified demographics and baseline comorbidities, medications and healthcare utilization variables within and including 365 days of the RA index date. These factors included age, sex, US geographic region (South, West, Midwest, Northeast or other) and calendar year. The presence of individual comorbidities was identified using diagnosis codes. We also calculated a combined comorbidity score combining conditions included in both the Charlson index and the Elixhauser system, as previously validated [22]. The presence of smoking was identified through a validated algorithm using diagnosis codes, procedure codes and prescription smoking cessation aides [23]. We identified medications including DMARDs, glucocorticoids, NSAIDs, opioids, proton pump inhibitors and histamine H2-receptor antagonists using part D data. DMARDs were further classified according to the class of drugs: non-biologic (nbDMARDs) and biologic (bDMARDs) or targeted synthetic (tsDMARDs; Janus kinase inhibitors). We quantified healthcare utilization as the continuous number of clinical visits, emergency department visits, inpatients admissions, visits with rheumatologists and visits with pulmonologists within the 365 days prior to and

including the RA index date. We quantified the proportion of patients who had outpatient visits with rheumatologists or pulmonologists in the previous year.

Statistical analysis

We described both groups, prevalent RA-ILD and RA without ILD, by the baseline characteristics using descriptive statistics.

Among those with RA without ILD at baseline, we investigated incident RA-ILD throughout follow-up in Medicare. Censoring events were death, loss to follow-up and the end of the study (31 December 2017). We used Cox proportional hazards models to estimate the hazard ratios (HRs) and 95% CIs for incident RA-ILD of several factors that had previously been investigated for RA-ILD risk (sex, smoking, respiratory morbidities, DMARD class and glucocorticoid use) [11]. Base models were unadjusted and multivariable models were adjusted for age, sex and US region.

We investigated RA-ILD and mortality, comparing prevalent RA-ILD with RA without ILD at baseline. Censoring events were outcome (mortality), incident RA-ILD, loss to follow-up or the end of the study (31 December 2017 for the total mortality analyses and 31 December 2016 for the cause-specific mortality analyses). Among those with incident RA-ILD after baseline, we reported total and cause-specific mortality data for descriptive purposes, but these were not included in the comparative analyses due to possible immortal time bias (since these patients had to live long enough to develop RA-ILD after baseline). We used Cox proportional hazards models to estimate the HRs and 95% CIs for total mortality, comparing prevalent RA-ILD with RA without ILD. The base model was not adjusted for any other variable. The multivariable model for mortality included factors that were possibly associated with both RA-ILD and mortality. These included age, sex, US region, smoking, MTX use, HCQ use, TNF inhibitor use, other bDMARD or tsDMARD use, glucocorticoid use, combined comorbidity score and number of physician visits. For cause-specific mortality, Fine and Gray subdistribution HRs (sdHRs) and 95% CIs were estimated to handle competing risks of alternative causes of mortality [24]. For example, the risk of respiratory mortality for patients with RA-ILD compared with RA without ILD also accounted for the competing risk of cardiovascular, cancer, infection and other types of mortality in the analyses. Among those with cancer mortality, we listed the frequency and proportion of the most common cancer types by RA-ILD status for descriptive purposes.

We verified the proportional hazards assumption in all analyses by verifying interaction terms between the exposure variables and follow-up time was not statistically significant. We used a two-sided P-value <0.05 to define statistical significance. All analyses were performed using SAS version 9.4 (SAS Institute, Cary, NC, USA).

Results

Prevalent RA-ILD at baseline

We identified a total of 509 787 patients with RA during the study period. At baseline, 10 306 (2.0%) had prevalent RA-ILD.

Baseline characteristics

Baseline characteristics according to RA-ILD status are shown in Table 1. Patients with RA-ILD had a mean age of 72.7 years compared with 72.6 years for patients with RA without ILD. A higher proportion of men had RA-ILD at baseline (27%) compared with RA without ILD (24%). There were more smokers in the RA-ILD group than in the RA without ILD group (29% *vs* 16%). More comorbidities were present in the prevalent RA-ILD group compared with RA without ILD (mean combined comorbidity score of 3.3 *vs* 1.2). RA-ILD patients had a prevalence of chronic obstructive pulmonary disease (COPD; 56% *vs* 17%) and asthma (27% *vs* 11%). Patients with prevalent RA-ILD were more likely to be prescribed some medications (MMF, rituximab, LEF, AZA and HCQ) and less likely to be prescribed MTX. A higher proportion of patients with RA-ILD were treated with glucocorticoids compared with RA without ILD (83% *vs* 62%). Pulmonologist evaluation occurred in 94% of the prevalent RA-ILD group compared with 11% of the RA without ILD group. Compared with RA without ILD, prevalent RA-ILD had increased healthcare utilization [mean 32.6 (S.D. 29.9) *vs* 18.5 (S.D. 17.0) physician visits per year].

Incident RA-ILD

Among 499 481 patients with RA without ILD at baseline, 13 372 (2.7%) incident RA-ILD cases occurred during 1 873 127 total person-years of follow-up [median 3.0 (interquartile range 1.4–5.7), mean 3.8 (S.D. 2.7) years/patient]. When also considering prevalent RA-ILD at baseline, a total of 23 678 (4.6%) patients with RA-ILD were identified. The incidence rate (IR) of RA-ILD was 7.14 (95% CI 7.02, 7.26) per 1000 person-years among RA patients.

Associations of baseline factors with incident RA-ILD

Table 2 shows the unadjusted and multivariable HRs for incident RA-ILD according to baseline factors. Males had a multivariable HR of 1.31 (95% CI 1.26, 1.36) for RA-ILD compared with females, adjusted for age and US region. Smoking had an HR of 1.35 (95% CI 1.29, 1.41). Asthma [HR 1.57 (95% CI 1.49, 1.64)] and COPD [HR 2.00 (95% CI 1.93, 2.09)] were each associated with incident RA-ILD. Medication use *vs* non-use at baseline was also associated with incident RA-ILD: bDMARDs or tsDMARDs [HR 1.34 (95% CI 1.29, 1.40)] and glucocorticoids [HR 1.45 (95% CI 1.39, 1.50)]. nbDMARD use was inversely associated with RA-ILD [HR 0.86 (95% CI 0.81, 0.92)] compared with non-use.

Table 1 Baseline characteristics by the presence or absence of prevalent ILD among patients with RA ($N = 509\,787$)

Characteristics	Prevalent RA-ILD ($n = 10\,306$)	RA without ILD ($n = 499\,481$)
Age, years, mean (S.D.)	72.7 (6.5)	72.6 (6.8)
Male, *n* (%)	2767 (27)	118 465 (24)
Comorbidities, *n* (%)		
Hypertension	8298 (81)	381 930 (76)
Diabetes mellitus	3544 (34)	149 153 (30)
Coronary heart disease	3679 (36)	114 576 (23)
Heart failure	2613 (25)	56 821 (11)
Stroke/transient ischaemic attack	920 (9)	37 458 (7)
Peripheral vascular disease	1582 (15)	62 492 (13)
Liver dysfunction	1052 (10)	32 244 (6)
Kidney dysfunction	1747 (17)	58 239 (12)
Cancer	2418 (23)	83 254 (17)
Asthma	2787 (27)	56 494 (11)
COPD	5734 (56)	83 410 (17)
Venous thromboembolism	598 (6)	17 988 (4)
Depression	2168 (21)	86 079 (17)
Mean combined comorbidity score (S.D.)	3.3 (3.0)	1.2 (2.4)
Smoking, *n* (%)	2963 (29)	82 174 (16)
DMARDs, *n* (%)		
HCQ	4208 (41)	180 005 (36)
MTX	3799 (37)	287 421 (58)
LEF	2044 (20)	44 544 (9)
Sulfasalazine	971 (9)	40 421 (8)
AZA	881 (9)	5687 (1)
MMF	433 (4)	1124 (0)
Infliximab	856 (8)	34 997 (7)
Etanercept	616 (6)	24 628 (5)
Adalimumab	416 (4)	17 797 (4)
Certolizumab pegol	141 (1)	5392 (1)
Golimumab	94 (1)	3845 (1)
Rituximab	590 (6)	7673 (2)
Abatacept	468 (5)	14 000 (3)
Tocilizumab	125 (1)	3898 (1)
Anakinra	<10 (0)	224 (0)
Tofacitinib	63 (1)	1703 (0)
Other medications		
Glucocorticoid	8585 (83)	309 159 (62)
Opioid	6525 (63)	290 531 (58)
Proton pump inhibitor	5450 (53)	192 086 (38)
Histamine-2 receptor blocker	929 (9)	36 496 (7)
Non-steroidal anti-inflammatory drug	3690 (36)	200 819 (40)
Coxib	785 (8)	45 004 (9)
Healthcare utilization		
ED visits, *n*, mean (S.D.)	1.4 (2.8)	0.7 (1.8)
Hospital admissions, *n*, mean (S.D.)	0.8 (1.3)	0.3 (0.7)
Rheumatologist visit, *n* (%)	9080 (88)	398 637 (80)
Pulmonologist visit, *n* (%)	9687 (94)	54 376 (11)
Physician visits, *n*, mean (S.D.)	32.6 (29.9)	18.5 (17.0)
Rheumatologist visits, *n*, mean (S.D.)	4.9 (4.7)	3.7 (4.2)
Pulmonologist visits, *n*, mean (S.D.)	4.8 (4.7)	0.3 (1.3)

ED: emergency department.

Total and cause-specific mortality of RA-ILD and RA without ILD

Table 3 shows total and cause-specific mortality for prevalent RA-ILD and RA without ILD at baseline. In this analysis there were a total of 107 248 (21.0%) deaths during 1 940 404 person-years of follow-up. There were 3989 (38.7%) deaths in the prevalent RA-ILD group (IR 122.01 deaths per 1000 person-years) compared with 103 259 (20.7%) deaths in the group with RA without ILD at baseline (IR 54.13 deaths per 1000 person-years).

When considering cause-specific death (truncated as of 31 December 2016), there was also a higher proportion of total deaths in RA-ILD compared with RA without

Table 2 Incidence and risk for RA-ILD by baseline characteristics for RA patients without baseline RA-ILD ($N = 499\,481$)

Baseline characteristics	Incident RA-ILD cases	Person-years	RA-ILD incidence rate per 1000 person-years (95% CI)	Unadjusted HR (95% CI) for incident RA-ILD	Multivariable* HR (95% CI) for incident RA-ILD
Overall	13 372	1 873 127	7.14 (7.02, 7.26)	–	–
Female	9732	1 457405	6.68 (6.55, 6.81)	1.00 (Ref)	1.00 (Ref)
Male	3640	415 722	8.76 (8.47, 9.04)	1.32 (1.26, 1.36)	1.31 (1.26, 1.36)
No smoking	11 190	1 641 339	6.82 (6.69, 6.94)	1.00 (Ref)	1.00 (Ref)
Smoking	2182	231 788	9.41 (9.03, 9.82)	1.39 (1.33, 1.45)	1.35 (1.29, 1.41)
No asthma	11 356	1 679 179	6.76 (6.64, 6.89)	1.00 (Ref)	1.00 (Ref)
Asthma	2016	193 948	10.39 (9.95, 10.86)	1.54 (1.47, 1.61)	1.57 (1.49, 1.64)
No COPD	9974	1 602 565	6.22 (6.10, 6.35)	1.00 (Ref)	1.00 (Ref)
COPD	3398	270 561	12.56 (12.14, 12.99)	2.02 (1.95, 2.10)	2.00 (1.93, 2.09)
No nbDMARD use	1061	127 826	8.30 (7.81, 8.81)	1.00 (Ref)	1.00 (Ref)
nbDMARD use	12 311	1 745 301	7.05 (6.93, 7.18)	0.85 (0.89, 0.90)	0.86 (0.81, 0.92)
No bDMARD/tsDMARD use	9704	1 462 943	6.63 (6.50, 6.77)	1.00 (Ref)	1.00 (Ref)
bDMARD/tsDMARD use	3668	410 184	8.94 (8.66, 9.24)	1.45 (1.40, 1.51)	1.34 (1.29, 1.40)
No bDMARD/tsDMARD use	9704	1 462 943	6.63 (6.50, 6.77)	1.00 (Ref)	1.00 (Ref)
Rituximab use	259	24 541	10.55 (9.33, 11.90)	1.60 (1.41, 1.81)	1.59 (1.40, 1.80)
TNF inhibitor use	2935	336 076	8.73 (8.42, 9.05)	1.32 (1.26, 1.37)	1.31 (1.26, 1.36)
Other bDMARD/tsDMARD use	474	49 566	9.56 (8.73, 10.45)	1.44 (1.32, 1.58)	1.45 (1.32, 1.59)
No glucocorticoid use	4274	759 305	5.63 (5.46, 5.80)	1.00 (Ref)	1.00 (Ref)
Glucocorticoid use	9098	1 113 821	8.17 (8.00, 8.34)	1.35 (1.30, 1.40)	1.45 (1.39, 1.50)

*Adjusted for age, sex and US region.

Table 3 Total and cause-specific mortality for RA patients with and without prevalent RA-ILD at baseline ($N = 509\,787$)

Cause of death	Prevalent RA-ILD ($n = 10\,306$)			RA without ILD ($n = 499\,481$)		
	Deaths, *n* (%)	Person-years	Mortality rate per 1000 person-years (95% CI)	Deaths, *n* (%)	Total person-years	Mortality rate per 1000 person-years (95% CI)
Total	3989 (38.7)†	32 695	122.01 (118.26, 125.84)	103 259 (20.7)†	1 907 709	54.13 (53.80, 54.46)
	*n after truncation = 8876			**n after truncation = 432 155		
Total	3351 (37.8)†	27 558	121.60 (117.54, 125.77)	86 106 (19.9)†	1 592 711	54.06 (53.70, 54.42)
Cardiovascular	699 (20.9)‡	27 558	25.37 (23.54, 27.30)	29 476 (34.2)‡	1 592 711	18.51 (18.30, 18.72)
Cancer	497 (14.8)‡	27 558	18.03 (16.50, 19.67)	14 456 (16.8)‡	1 592 711	9.08 (8.93, 9.23)
Respiratory	1236 (36.9)‡	27 558	44.85 (42.40, 47.40)	11 235 (13.0)‡	1 592 711	7.05 (6.92, 7.19)
Infection	102 (3.0)‡	27 558	3.70 (3.03, 4.47)	3257 (3.8)‡	1 592 711	2.04 (1.98, 2.12)
Other	817 (20.5)‡	27 558	29.65 (27.67, 31.73)	27 682 (32.1)‡	1 592 711	17.38 (17.18, 17.59)

*Data on cause-specific death were only available as of 12/31/2016. †Denominator is all patients with prevalent RA-ILD or RA without ILD. ‡Denominator is all truncated deaths among prevalent RA-ILD or RA without ILD.

ILD (37.8% *vs* 19.9%). Among prevalent RA-ILD, the most common cause of death was respiratory (IR 44.85 per 1000 person-years). Among RA without ILD at baseline, the most common cause of death was cardiovascular (IR 18.51 per 1000 person-years).

Table 4 shows deaths among patients with incident RA-ILD occurring after baseline. There were 4410 (33.0%) deaths during 34 616 person-years of follow-up [median 2.5 (IQR 1.0–4.7), mean 3.2 (S.D. 2.6) years/patient]. The most common cause of death was respiratory (IR 40.27 per 1000 person-years).

Prevalent RA-ILD and risk for total and cause-specific mortality

Table 5 shows the unadjusted and multivariable HRs for total and cause-specific mortality comparing prevalent RA-ILD with RA without ILD at baseline. RA-ILD had an HR for mortality of 2.36 (95% CI 2.28, 2.45) compared with RA without ILD in unadjusted analyses. When considering cause-specific mortality in unadjusted analyses, but accounting for alternative causes of death, RA-ILD had an elevated sdHR for every cause of death, but most pronounced for respiratory mortality [sdHR 7.08 (95% CI 6.67, 7.51)].

Table 4 Total and cause-specific mortality for patients with incident RA-ILD after baseline ($N = 13\,372$)

Cause of death	Deaths, *n* (%)	Person-years	Mortality rate per 1000 person-years (95% CI)
Total	4410 (33.0)[†]	34 619	127.39 (123.67, 131.19)
		*n after truncation = 13 126	
Total	3537 (26.9)[†]	27 216	129.96 (125.73, 134.30)
Cardiovascular	910 (25.7)[‡]	27 216	33.44 (31.32, 35.66)
Cancer	585 (16.5)[‡]	27 216	21.49 (19.81, 23.29)
Respiratory	1096 (31.0)[‡]	27 216	40.27 (37.94, 42.71)
Infection	130 (3.7)[‡]	27 216	4.78 (4.01, 5.65)
Other	816 (23.1)[‡]	27 216	29.98 (27.98, 32.09)

*Data on cause-specific death were only available as of 31 December 2016. [†]Denominator is all patients with incident RA-ILD. [‡]Denominator is all truncated deaths among incident RA-ILD.

Table 5 Total and cause-specific mortality for prevalent RA-ILD compared with RA patients without ILD ($n = 509\,787$)

Group	Unadjusted HR (95% CI) for total mortality	Unadjusted sdHR (95% CI) for cardiovascular mortality	Unadjusted sdHR (95% CI) for cancer mortality	Unadjusted sdHR (95% CI) for respiratory mortality	Unadjusted sdHR (95% CI) for infection mortality	Unadjusted sdHR (95% CI) for other mortality
RA without ILD	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)
RA-ILD	2.36 (2.28, 2.45)	1.42 (1.32, 1.54)	2.08 (1.90, 2.27)	7.08 (6.67, 7.51)	1.89 (1.55, 2.30)	1.78 (1.66, 1.90)
Group	**Multivariable* HR (95% CI) for total mortality**	**Multivariable* sdHR (95% CI) for cardiovascular mortality**	**Multivariable* sdHR (95% CI) for cancer mortality**	**Multivariable* sdHR (95% CI) for respiratory mortality**	**Multivariable* sdHR (95% CI) for infection mortality**	**Multivariable* sdHR (95% CI) for other mortality**
RA without ILD	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)	1.00 (Ref)
RA-ILD	1.66 (1.60, 1.72)	1.01 (0.93, 1.09)	1.56 (1.43, 1.71)	4.39 (4.13, 4.67)	1.19 (0.97, 1.45)	1.30 (1.21, 1.40)

*Adjusted for age, sex, US region, smoking, MTX use, HCQ use, TNF inhibitor use, other bDMARD or tsDMARD use, glucocorticoid use, combined comorbidity score and number of physician visits.

Table 6 Type of cancer mortality by the presence or absence of prevalent RA-ILD ($N = 14\,953$)

Type of cancer, *n* (%)	Prevalent RA-ILD ($n = 497$ cancer deaths)	RA without ILD ($n = 14\,456$ cancer deaths)
Malignant neoplasms of respiratory and intrathoracic organs	264 (53.1)	4750 (32.9)
Malignant neoplasms of digestive organs	72 (14.5)	3130 (21.7)
Malignant neoplasms of lymphoid, hematopoietic and related tissue	56 (11.3)	1959 (13.6)
Malignant neoplasms of ill-defined, other secondary and unspecified sites	25 (5.0)	1034 (7.2)
Malignant neoplasms of breast	27 (5.4)	858 (5.9)
Other cancer types	53 (10.7)	2725 (18.9)

RA-ILD had a multivariable HR of 1.66 (95% CI 1.60, 1.72) for total mortality compared with RA without ILD. Considering cause-specific mortality in the multivariable analyses, RA-ILD was associated with increased respiratory mortality [sdHR 4.39 (95% CI 4.13, 4.67)], cancer mortality [sdHR 1.56 (95% CI 1.43, 1.71)] and other mortality [sdHR 1.30 (95% CI 1.21, 1.40)]. After multivariable adjustment there was no statistical difference between RA-ILD and RA without ILD for cardiovascular or infection mortality.

Among those with cancer mortality, the most common cancer types are listed in Table 6. Malignant neoplasms of the respiratory and intrathoracic organs were the most common in both the prevalent RA-ILD and RA

without ILD groups. In the RA-ILD group, 53.1% of deaths were due to this cancer type compared with 32.9% in the RA without ILD group.

Discussion

In this US nationwide study of older individuals we found that RA-ILD was present or developed in nearly 5% of patients with RA. Male sex, smoking and indicators of disease severity (bDMARD, tsDMARD and glucocorticoid use) were each associated with an increased risk of incident RA-ILD, while nbDMARD use was associated with a lower risk. RA-ILD was associated with excess total mortality that was not explained by measured factors, including demographics, comorbidities, medications and healthcare utilization. The associations of RA-related medications with incident RA-ILD risk should be interpreted with caution since they may be explained by unmeasured factors, including RA disease activity, severity, comorbidities and prior or concomitant medication use. Dedicated studies are needed to firmly establish the role of specific DMARDs and glucocorticoids in RA-ILD incidence and progression. RA-ILD was associated with a >4-fold increased risk of respiratory mortality and a 56% increased cancer mortality compared with RA without ILD. Overall, these findings emphasize that a subset of patients with RA are at risk for RA-ILD and confirm that RA-ILD is associated with excess mortality. Besides respiratory mortality, we found that patients with RA-ILD are also at increased risk for cancer mortality, a novel finding that requires further investigation.

The prevalence of RA-ILD varies widely in previous studies, ranging from as low as 2% to as high as 50%, likely due to different methodologies in detection [4, 5, 7, 9, 12–14, 25]. Patients with RA are known to have a wide range of pulmonary imaging abnormalities [7], some of which may never become symptomatic or progress to clinically significant RA-ILD and could overestimate prevalence. A previous study used administrative data to estimate the real-world prevalence of RA-ILD to be 3.2–6.0 cases per 100 000 people [9]. Our study was restricted to an older population, a known RA risk factor, and only among patients with RA so it is difficult to directly compare this with the baseline prevalence of 2.0% (~1 per 50 RA patients) that we report.

The incidence of RA-ILD has been investigated previously using several study designs. A recent Danish nationwide study using administrative claims data found that incident RA-ILD (1+ ILD diagnosis code) developed in 2.2% of patients with RA, similar to our findings (2.6% of patients) [5]. A population-based study in Olmsted County, Minnesota, USA estimated the lifetime incidence of RA-ILD to be 7.7% and the 10-year cumulative incidence to be 5.0–6.6%; RA-ILD was obtained through medical records review rather than administrative data and there was longer follow-up per patient than our study [3, 4]. Another study using a mortality database found that 6.6% of RA patients had ILD listed as a potential cause of death [10]. A pharmacoepidemiologic study using administrative data that defined RA-ILD among RA using either specific or sensitive definitions reported IRs of RA-ILD of 1.8 and 6.4 per 1000 person-years, respectively. These are both slightly lower than the IR we report, likely due to differences in ages and RA-ILD case definitions. Our study extends these literature by analysing a validated RA-ILD case definition to provide a real-world US incidence estimate.

We also investigated baseline factors associated with incident RA-ILD. Male sex and smoking were associated with an increased risk for RA-ILD, similar to several previous reports [15–19, 26]. Other respiratory morbidities, in particular asthma and COPD, were associated with increased RA-ILD. It is possible that shared risk factors (such as smoking), diagnostic uncertainty or early RA-ILD symptoms in patients with dyspnoea and other symptoms may explain this association. However, airways disease may be involved in the production of RA-related autoantibodies such as anti-citrullinated protein antibodies that could predispose patients to RA-ILD [27]. Regardless of the underlying mechanism, these patients may be at higher risk of RA-ILD and may be particularly amenable for screening/prevention strategies.

Several medications were associated with incident RA-ILD. nbDMARDs, including MTX and HCQ, were associated with lower risk. While MTX can rarely induce pneumonitis that can have features of ILD [28], recent reports also suggest that MTX is not associated with incident RA-ILD [19, 29]. While our results also offer reassurance on the possible safety of MTX in terms of RA-ILD, the study was not primarily designed to investigate medication associations that would require careful attention to confounding by indication and ideally ascribe to the new-user active-comparator framework. We also found that bDMARD, tsDMARD and glucocorticoid use was associated with incident RA-ILD. The relationship of these medications with RA-ILD is controversial [11]. Recently, articular RA disease activity was associated with an increased risk of incident RA-ILD, which may in part explain this association [25]. Future studies are needed to firmly establish the relationship of DMARDs and glucocorticoids with RA-ILD risk or whether this risk is due to other genetic or severity factors of underlying RA.

Previous studies have investigated RA-ILD and mortality [4, 6, 9, 10]. In a population-based study in the USA, RA-ILD confirmed by medical records review had a median survival of only 2.6 years and had an HR of 2.86 (95% CI 1.98, 4.12) compared with RA without ILD [4]. A recent Danish study reported a median survival after RA-ILD of 6.6 years and HRs ranging from 1.9 to 10.4 depending on RA-ILD duration [5]. Another recent US study also showed a median survival after RA-ILD of 7.8 years [9]. Unlike most previous studies, we also investigated cause-specific mortality of RA-ILD. In a large mortality database study, RA-ILD patients had lower odds than RA without ILD patients to have several

conditions listed as a potential cause of death, including cardiac diseases, infections, chronic obstructive lung diseases, pulmonary embolism and lung cancer [10]. This is in contrast to our study that found RA-ILD had a higher risk of cancer mortality in addition to the expected increased risk of respiratory mortality. Using a competing risk analysis, we found that RA-ILD had similar (not lower) cardiovascular and infection mortality than RA without ILD. These discrepancies may be due to differences in ascertainment of cause of death and study design (the previous study [10] being cross-sectional). While ours is the first to suggest that RA-ILD may be associated with an increased risk of cancer mortality, previous studies showed that ILD as well as interstitial lung abnormalities detected on chest imaging were each associated with subsequent risk of lung cancer and mortality, similar to our findings [30, 31]. Lung cancer was more likely to occur as the cause of cancer death in patients with RA-ILD than in those without RA-ILD in our study. Future studies are needed to confirm whether RA-ILD directly affects cancer risk and mortality or whether this may be due to shared risk factors, such as smoking, not completely accounted for in this analysis. While we found an association between RA-ILD and infection mortality in unadjusted analyses, this was no longer present after adjustment. It is possible that patients with RA-ILD who died with infection had the death ascribed to respiratory causes rather than infectious causes. The association of RA-ILD with serious infection in other clinical contexts would require dedicated analyses.

Our study has several strengths, including the longitudinal design within a large, nationwide claims study. Since Medicare covers nearly all Americans $\geq$65 years of age, we have real-world data on the epidemiology RA-ILD among older adults. Furthermore, because older age is a risk factor for RA-ILD, analysing Medicare also enriches RA-ILD identification. This study design allowed us to estimate both the prevalence and incidence of RA-ILD and investigate risk factors for incident RA-ILD. We had data on demographics, comorbidities, use of RA-related and other medications and healthcare utilization patterns in this population. Ours is the first to study the epidemiology and mortality outcomes of RA-ILD using a validated claims algorithm to identify RA and RA-ILD. Finally, we were able to investigate both total and cause-specific mortality to quantify the mortality burden of RA-ILD and to identify a potentially novel association of RA-ILD with cancer mortality. We used competing risk models to account for alternative causes of death, which is preferable to traditional Cox models, particularly in our study with high mortality.

However, our study also has some limitations. While we had detailed data on medication dispensement, other RA-specific data such as disease activity, duration, RA-related autoantibodies and severity were not available and are likely associated with both RA-ILD and mortality [11]. While the RA-ILD algorithm has been validated [20], the PPV was only modest, so it may have misclassified some patients because of the relatively low prevalence of the condition. The patients identified with RA-ILD in our study were nearly all evaluated by pulmonologists and were often prescribed medications specifically used to treat RA-ILD, suggesting acceptable validity. While we investigated associations of baseline factors with incident RA-ILD, these results should be interpreted with caution since the study was not designed for comparative safety for RA-ILD. The associations of medications and glucocorticoids with RA-ILD may be indicators of RA severity rather than directly affecting RA-ILD risk. As in all observational studies, residual confounding from unmeasured factors is possible. While we used a validated claims-based algorithm to identify smoking status [23], we did not have data available on smoking that may have occurred prior to entry into Medicare and we were unable to quantify smoking pack-years. We adjusted for baseline factors in the multivariable mortality analyses comparing prevalent RA-ILD with RA without ILD. Some of these factors, such as comorbidities, may have been on the causal pathway between RA-ILD and mortality and thus mediated the relationship. Since this would be expected to attenuate the effect size estimates, the true relationship between RA-ILD and mortality could be even stronger than we report. Therefore our results showing increased total and cause-specific mortality are conservative. In the mortality analyses, we adjusted for factors such as RA-related medications, so these differences are unlikely to explain the results that we report. Future dedicated studies should investigate whether RA-related medications may be associated with RA-ILD progression while appropriately accounting for confounding by indication or using randomized controlled trials. While we had nationwide data for nearly all older patients in the USA, these results may not be generalizable, particularly to a younger population.

In summary, nearly 5% of patients with RA in Medicare were identified as having RA-ILD using a validated algorithm. We identified associations with incident RA-ILD that included male sex, smoking and DMARD use. Patients with RA-ILD had increased total mortality compared with patients with RA without ILD, particularly for respiratory and cancer mortality. These results emphasize that clinically significant RA-ILD is not uncommon and has a significant impact on mortality. Further research is needed to understand which patients may be at risk for developing RA-ILD and strategies are needed to mitigate its excess mortality.

Acknowledgements

J.A.S. is supported by the National Institute of Arthritis and Musculoskeletal and Skin Diseases (grants K23 AR069688, R03 AR075886, L30 AR066953, P30 AR070253 and P30 AR072577), the Rheumatology Research Foundation (K Supplement Award and R Bridge Award), the Brigham Research Institute and the R. Bruce and Joan M. Mickey Research Scholar Fund. T.J.D. is supported by National Institutes of Health/

National Heart, Lung, and Blood Institute grants (K23 HL119558, R03 HL148484). J.A.S. and S.C.K. conceived the idea for the article. J.A.S., Y.J., S.-K.C., S.V., R.D., T.J.F. and S.C.K. designed the study and interpreted the results. J.A.S., Y.J., S.-K.C., S.V., R.D., T.J.F. and S.C.K. extracted data, planned analyses and interpreted results. J.A.S. wrote the first draft of the manuscript. S.V. performed the analyses. All authors critically reviewed the manuscript and agreed with the submission. J.A.S. accepts full responsibility for the finished article, had access to any data and controlled the decision to publish.

Funding: This work was supported by an investigator-initiated grant from Bristol-Myers Squibb (IM101-763). The funder had no role in the study design, analysis or drafting of the manuscript.

Disclosure statement: J.A.S. has performed consultancy for Bristol-Meyers Squibb, Gilead, Inova Diagnostics, Janssen, Optum and Pfizer unrelated to this work. T.J.D. reports research funding from Bristol-Myers Squibb and involvement in a clinical trial funded by Genentech and Bristol-Myers Squibb. R.D. has received research support to the Brigham and Women's Hospital from Novartis, Vertex and Bayer for other studies. S.C.K. has received research support to the Brigham and Women's Hospital from Pfizer, AbbVie, Roche and Bristol-Myers Squibb for other studies. The other authors have declared no conflicts of interest.

Data availability statement

No data are available.

References

1 England BR, Sayles H, Michaud K *et al.* Cause-specific mortality in male US veterans with rheumatoid arthritis. Arthritis Care Res 2016;68:36–45.

2 Sparks JA, Chang SC, Liao KP *et al.* Rheumatoid arthritis and mortality among women during 36 years of prospective follow-up: results from the Nurses' Health Study. Arthritis Care Res 2016;68:753–62.

3 Myasoedova E, Crowson CS, Turesson C, Gabriel SE, Matteson EL. Incidence of extraarticular rheumatoid arthritis in Olmsted County, Minnesota, in 1995–2007 versus 1985–1994: a population-based study. J Rheumatol 2011;38:983–9.

4 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.

5 Hyldgaard 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.

6 Kim D, Cho SK, Choi CB *et al.* Impact of interstitial lung disease on mortality of patients with rheumatoid arthritis. Rheumatol Int 2017;37:1735–45.

7 Duarte AC, Porter JC, Leandro MJ. The lung in a cohort of rheumatoid arthritis patients—an overview of different types of involvement and treatment. Rheumatology (Oxford) 2019;58:2031–8.

8 Doyle TJ, Dellaripa PF, Batra K *et al.* Functional impact of a spectrum of interstitial lung abnormalities in rheumatoid arthritis. Chest 2014;146:41–50.

9 Raimundo K, Solomon JJ, Olson AL *et al.* Rheumatoid arthritis-interstitial lung disease in the United States: prevalence, incidence, and healthcare costs and mortality. J Rheumatol 2019;46:360–9.

10 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.

11 Huang S, Kronzer VL, Dellaripa PF *et al.* Rheumatoid arthritis-associated interstitial lung disease: current update on prevalence, risk factors, and pharmacologic treatment. Curr Treatm Opt Rheumatol 2020;6:337–53.

12 Dong H, Julien PJ, Demoruelle MK, Deane KD, Weisman MH. Interstitial lung abnormalities in patients with early rheumatoid arthritis: a pilot study evaluating prevalence and progression. Eur J Rheumatol 2019;6:193–8.

13 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.

14 Habib HM, Eisa AA, Arafat WR, Marie MA. Pulmonary involvement in early rheumatoid arthritis patients. Clin Rheumatol 2011;30:217–21.

15 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.

16 Restrepo JF, del Rincon I, Battafarano DF *et al.* Clinical and laboratory factors associated with interstitial lung disease in rheumatoid arthritis. Clin Rheumatol 2015;34:1529–36.

17 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 (Oxford) 2014;53:1676–82.

18 Saag KG, Kolluri S, Koehnke RK, Georgou TA *et al.* Rheumatoid arthritis lung disease. Determinants of radiographic and physiologic abnormalities. Arthritis Rheum 1996;39:1711–9.

19 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.

20 Cho SK, Doyle TJ, Lee H *et al.* Validation of claims-based algorithms to identify interstitial lung disease in patients with rheumatoid arthritis. Semin Arthritis Rheum 2020;50:592–7.

21 Kim SY, Servi A, Polinski JM *et al.* Validation of rheumatoid arthritis diagnoses in health care utilization data. Arthritis Res Ther 2011;13:R32.

22 Gagne JJ, Glynn RJ, Avorn J, Levin R, Schneeweiss S. A combined comorbidity score predicted mortality in elderly patients better than existing scores. J Clin Epidemiol 2011;64:749–59.

23 Desai RJ, Solomon DH, Shadick N, Iannaccone C, Kim SC. Identification of smoking using Medicare data–a validation study of claims-based algorithms. Pharmacoepidemiol Drug Saf 2016;25:472–5.

24 Fine JP, Gray RG. A proportional hazards model for the subdistribution of a competing risk. J Am Stat Assoc 1999;94:496–509.

25 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.

26 Giles JT, Danoff SK, Sokolove J *et al.* Association of fine specificity and repertoire expansion of anticitrullinated peptide antibodies with rheumatoid arthritis associated interstitial lung disease. Ann Rheum Dis 2014;73:1487–94.

27 Ford JA, Liu X, Chu SH *et al.* Asthma, chronic obstructive pulmonary disease, and subsequent risk for incident rheumatoid arthritis among women: a prospective cohort study. Arthritis Rheumatol 2020;72:704–13.

28 Sparks JA, Dellaripa PF, Glynn RJ *et al.* Pulmonary adverse events of low-dose methotrexate in the randomized double-blind placebo-controlled Cardiovascular Inflammation Reduction Trial. Arthritis Rheumatol 2020;72:2065–71.

29 Juge PA, Lee JS, Lau J *et al.* Methotrexate and rheumatoid arthritis associated interstitial lung disease. Eur Respir J 2020;2000337. doi: 10.1183/13993003.00337-2020. [Online ahead of print.]

30 Axelsson GT, Putman RK, Aspelund T *et al.* The associations of interstitial lung abnormalities with cancer diagnoses and mortality. Eur Respir J 2020;56:1902154. doi: 10.1183/13993003.02154-2019.

31 Choi WI, Park SH, Park BJ, Lee CW. Interstitial lung disease and lung cancer development: a 5-year nationwide population-based study. Cancer Res Treat 2018;50:374–81.