

Supplementary materials

Supplementary Text 1. Prevalence estimation

Supplementary Table 1. List of terminology used.

Supplementary Table 2. Performance of EHRead® identifying records that contain key study variables.

Supplementary Table 3. Population's EHRs ascertainment and data pattern assessment

Supplementary Table 4. Crude prevalence by sex and age of RA and RA-ILD

Supplementary Table 5. Age-adjusted prevalence by gender and age of RA and RA-ILD

Supplementary Table 6. Chest High Resolution Computerized Tomography pattern of ILD at the time of inclusion

Supplementary Table 7. Frequency of comorbidities at the time of inclusion

Supplementary Table 8. Blood tests at the time of inclusion

Supplementary Table 9. Sensitivity analysis with RA-support population

Supplementary Table 10. Supplementary analysis in patients treated with DMARDs or glucocorticoids

Supplementary Table 11. Healthcare Resource Utilization during follow-up

Supplementary Text 1. Prevalence estimation

Prevalence was calculated as follows:

- **Crude prevalence.** Calculated according to the formula:

$$\text{Prevalence} = \frac{\text{Total number of patients diagnosed with RA in the study period} *}{\text{Total number of patients at risk in the areas covered by the hospital sites} **}$$

- **Age/sex-adjusted prevalence.** Calculated by estimating the crude prevalence by sex and age (age groups spanning 10 years). These estimations were weighted according to the distribution of the population in Spain. Final adjusted estimates were obtained by adding all weighted estimations in the interest age groups (i.e., 18-29, 30-44, 45-59, 60-74, and 75+ years of age).

*The total number of patients diagnosed with RAILD in the study period was defined as the sum of all patients with at least one mention of RA and ILD in the EHRs available within the study period (FAS).

**The total number of people at risk was obtained using the databases detailed below:

- *Spatial databases*, including geographical areas covered by the participating hospital sites and Spanish census databases; the latter contain information regarding the total population in each area in 2011. These databases make possible to identify the population belonging to the care area of each hospital.
- *Temporal databases*, which contain the distribution of the population by sex and age between 1970 and 2021 for each Spanish province, which make possible to extrapolate the population to the specific time of our study.

Supplementary Table 1. List of terminology used.

Term name	SNOMED ID*
Abatacept	421412005
Substance abuse	415658005
ACE inhibitors	678506343435815000
Acute interstitial pneumonia	236302005
Acute myocardial infarction	57054005
Adalimumab	398728003
Alanine Aminotransferase	33
Albumin	27
Alcohol consumption	219006
Anemia	271737000
Anti-carbamylated protein antibodies	50127
Anti-citrullinated protein antibodies	257
Antidepressants	663127017671042000
Anti-inflammatories	687455536609505000
Antiplatelet agents	895808641729358000
Anxiolytics	866579482144794000
Aspartate Aminotransferase	34
Asthma	195967001
Atrial fibrillation	49436004
Azathioprine	111165009
Baricitinib	70034
Beta-blockers	904597696138226000
Body mass index	226
Bone erosion	50129
Bone fracture	125605004
Calcium antagonists	883947310140411000
Cardiovascular disease	49601007
Certolizumab pegol	430307008
Chest X-ray	399208008
Cholesterol	84698008
Chronic Obstructive Pulmonary Disease	13645005
Clinical Disease Activity Index	50132
Combined emphysema-fibrosis syndrome	94894
Corticosteroids	79440004
C-reactive protein	48
Crohn's disease	34000006
Cryptogenic organizing pneumonia	68409003, 4047
Cyclophosphamide	74470007
Depression	35489007

Desquamative interstitial pneumonia	8549006
Diabetes mellitus	73211009
Diffusing Capacity of the Lung for Carbon Monoxide	3609, 124
Disease Activity Score 28	50130
Diuretics	543324962549680000
Dry crackles	48409008
Dyslipidemia	370992007
Erythrocyte sedimentation rate	8
Etanercept	118259007
Ex-drinker	82581004
Ex-smoker	8392000
Ex-substance abuse	105550004,1073,3295,416479009, 44870007
Factor Xa inhibitors	978329880435347000
Fatty liver	197321007
Filgotinib	10014
Flare-up	201791009
Forced vital capacity	366151006, 122
Gamma-glutamyl transferase	35
Glomerular filtration rate (GFR)	81
Glucose	67079006
Golimumab	442318002
Gout	90560007
HbA1c (Glycated hemoglobin)	61
HDL cholesterol	23
Heart failure	84114007
Height	246
Hemoglobin	1
High-resolution computed tomography of the lung	241541005
Hypertension	38341003
Hypertransaminasemia	99
Hyperuricemia	35885006
Hypolipidemic	856752805180429000
Idiopathic inflammatory myopathies	1473
Infections	40733004
Inflammatory bowel disease	24526004
Infliximab	108675009
In-hospital mortality	99999000563101, 599
Interstitial lung disease	233703007
Interstitial pneumonia	64667001
Interstitial pneumonia with autoimmune features	870573008
Kidney failure	236423003, 314983004

Latent tuberculosis infection	11999007
LDL cholesterol	24
Leflunomide	109143003
Lung biopsy	78603008
Lung ultrasound	88628008
Lymphocytes	11
Lymphocytic interstitial pneumonia	44274007
Malignancy	363346000
Metabolic syndrome	237602007
Methotrexate	68887009
Mixed connective tissue disease	3933
MMRC (Modified Medical Research Council dyspnea scale)	391120009,391123006,391124000, 391125004,391126003
Mycophenolate	109131004
Neutrophils	10
Nintedanib	712494002
No substance abuse	1
Non-drinker	105542008
Non-smoker	8517006
Nonspecific interstitial pneumonia	129452008
Non-Steroidal Anti-Inflammatory Drugs	109506597124226000
Obesity	414916001
Obliterative bronchiolitis	40100001
Oral antidiabetics	791432907517847000
Oral contraceptives	59261009
Osteoporosis	64859006
Other lung diseases	19829001
Pirfenidone	439581000
Platelets	15
Psoriasis	9014002
Psychiatric pathology	74732009
Pulmonary fibrosis	51615001
RAILD	10713006
RAPID3 (Routine Assessment of Patient Index Data 3)	50133
Reactive arthritis	239783001
Respiratory bronchiolitis	4048
Respiratory bronchiolitis associated with interstitial lung disease	4050
Respiratory bronchiolitis associated with rheumatoid arthritis	4049
Rheumatoid arthritis	69896004
Rheumatoid factor	247

Rheumatoid factor negative	165840002
Rheumatoid factor positive	165839004
Rheumatoid nodule	50782009
Rituximab	108809004
Sarilumab	70024
SDAI (Simplified Disease Activity Index)	50131
Serum creatinine	19
Severe infections	91302008
Shared epitope	50128
Sjogren's syndrome	83901003
Smoker	77176002
Sulfasalazine	45844004
Systemic lupus erythematosus	55464009
Systemic sclerosis	299276009
Tocilizumab	444649004
Tofacitinib	70054
Total bilirubin	25
Transient ischemic attack	266257000
Triglycerides	21
Ulcerative colitis	64766004
Upadacitinib	10015
Usual interstitial pneumonia	33172
Uveitis	128473001
Vitamin K antagonists	780908461421528000
Weight	259
White blood cell count	9

**SNOMED IDs Included terminology related terms of an inferior hierarchy in SNOMED classification. New IDs were created for laboratory and other measurable parameters.*

Supplementary Table 2. Performance of EHRead® identifying records that contain key study variables.

Variable	Precision
Bronchiolitis obliterans	0.917
Cryptogenic organizing pneumonia	0.784
Abatacept	0.771
Usual interstitial pneumonia	0.856
Rheumatoid arthritis	0.764
Anti-citrullinated protein antibodies	0.493
Rheumatoid factor	0.594
Nonspecific interstitial pneumonia	0.717
Desquamative interstitial pneumonia	0.750
Interstitial lung disease	0.794
Nintedanib	0.778

Precision = True Positives / (True Positives + False Positives).

Supplementary Table 3. Population's EHRs ascertainment and data pattern assessment.

	RA (13958)	RAILD (712)	RAnonILD (13246)
RA (support subset) n (%)	12538 (89.8)	696 (97.8)	11842 (89.4)
Multiple RA mentions n (%)	10909 (78.2)	595 (83.6)	10314 (77.9)
RA treatments* n (%)	10315 (73.9)	651 (91.4)	9664 (73)
Positive RF, ACPA or DAI** n (%)	3979 (28.5)	247 (34.7)	3732 (28.2)
Multiple ILD mentions n (%)	546 (3.9)	546 (76.7)	0 (0)
Follow-up time (months)			
N (%)	13958 (100)	712 (100)	13246 (100)
Mean (SD)	19.1 (18.1)	25.2 (18.8)#	18.8 (18)
Median (Q1, Q3)	13.3 (3.8, 31)	22.7 (7.8, 41.4)	12.9 (3.5, 30.3)
Look-back time (months)			
N (%)	13958 (100)	712 (100)	13246 (100)
Mean (SD)	26.6 (30.3)	21.8 (27.5)#	26.8 (30.4)
Median (Q1, Q3)	17.6 (0, 46.2)	14.1 (0, 37.3)	17.9 (0, 46.6)
Number of EHRs/year			
N (%)	13207 (94.6)	702 (98.6)	12505 (94.4)
Mean (SD)	34.1 (154.5)	42.9 (122.2)#	33.6 (156.1)
Median (Q1, Q3)	7.8 (3.5, 17.5)	15.1 (6.8, 36.4)	7.5 (3.5, 16.6)
Number of follow-up EHRs/year			
N (%)	12290 (88)	691 (97.1)	11599 (87.6)
Mean (SD)	42.6 (172.6)	62 (179.8)#	41.5 (172.1)
Median (Q1, Q3)	8.2 (3.8, 19.2)	15.1 (7.1, 41.1)	8 (3.7, 18.3)
Number of look-back EHRs/year			
N (%)	10051 (72)	507 (71.2)	9544 (72.1)
Mean (SD)	46.1 (270.9)	65.4 (269.7)#	45.1 (270.9)
Median (Q1, Q3)	7.7 (3.4, 17.9)	13.4 (6, 36.3)	7.5 (3.3, 17.3)

DAI: Disease activity index; ILD: Interstitial lung disease; EHR: electronic health record; RA: Rheumatoid arthritis. The occurrence count of each feature was analyzed and divided by follow-up years. (#) Median (Q1, Q3) is preferred over mean (SD) for interpretation as the feature is non-normal. *RA treatments refer to disease modifying anti-rheumatic drugs and glucocorticoids. **Refer to DAS28, SDAI or CDAI.

Supplementary Table 4. Crude prevalence by sex and age of RA and RAILD

Age	Prevalence of RA, % (CI)			Prevalence of RAILD, % (CI)		
	Overall	Male	Female	Overall	Male	Female
All	0.60 (0.27; 0.92)	0.33 (0.13; 0.53)	0.83 (0.40; 1.26)	0.03 (0.02; 0.03)	0.02 (0.02; 0.03)	0.03 (0.03; 0.04)
18-29	0.18 (0.07; 0.29)	0.11 (0.02; 0.20)	0.24 (0.11; 0.37)	0.00 (0.00; 0.00)	0.00 (0.00; 0.00)	0.00 (0.00; 0.00)
30-44	0.31 (0.09; 0.52)	0.12 (0.02; 0.22)	0.48 (0.16; 0.80)	0.01 (0.01; 0.01)	0.00 (0.00; 0.00)	0.01 (0.01; 0.01)
45-59	0.53 (0.26; 0.79)	0.26 (0.11; 0.41)	0.78 (0.40; 1.16)	0.02 (0.02; 0.02)	0.02 (0.01; 0.02)	0.02 (0.02; 0.03)
60-74	0.90 (0.47; 1.33)	0.57 (0.26; 0.88)	1.19 (0.67; 1.7)	0.06 (0.05; 0.07)	0.05 (0.04; 0.06)	0.07 (0.06; 0.07)
>75	1.27 (0.62; 1.91)	0.89 (0.43; 1.35)	1.50 (0.74; 2.27)	0.07 (0.05; 0.10)	0.07 (0.0; 0.09)	0.07 (0.05; 0.10)

CI: Confident interval; ILD: Interstitial lung disease; RA: Rheumatoid arthritis.

Supplementary Table 5. Age-adjusted prevalence by sex and age of RA and RAILD

Age	Prevalence of RA, % (CI)			Prevalence of RAILD, % (CI)		
	Overall	Male	Female	Overall	Male	Female
All	0.53 (0.24; 0.81)	0.28 (0.11; 0.45)	0.75 (0.36; 1.14)	0.02 (0.02; 0.03)	0.02 (0.01; 0.02)	0.03 (0.03; 0.03)
18-29	0.03 (0.01; 0.05)	0.02 (0.01; 0.04)	0.04 (0.02; 0.07)	0.00 (0.00; 0.00)	0.00 (0.00; 0.00)	0.00 (0.00; 0.00)
30-44	0.10 (0.03; 0.17)	0.04 (0.01; 0.07)	0.15 (0.05; 0.24)	0.00 (0.00; 0.00)	0.00 (0.00; 0.00)	0.00 (0.00; 0.00)
45-59	0.12 (0.06; 0.18)	0.06 (0.02; 0.10)	0.18 (0.09; 0.26)	0.00 (0.00; 0.01)	0.00 (0.00; 0.01)	0.01 (0.01; 0.01)
60-74	0.15 (0.08; 0.22)	0.09 (0.04; 0.14)	0.20 (0.11; 0.29)	0.01 (0.01; 0.01)	0.01 (0.01; 0.01)	0.01 (0.01; 0.01)
>75	0.12 (0.06; 0.19)	0.07 (0.03; 0.11)	0.18 (0.09; 0.28)	0.01 (0.01; 0.01)	0.01 (0.00; 0.01)	0.01 (0.01; 0.01)

CI: Confident interval; ILD: Interstitial lung disease; RA: Rheumatoid arthritis.

Supplementary Table 6. Chest High Resolution Computerized Tomography pattern of ILD at the time of inclusion

	RAILD (712)
With specified ILD subtype n (%)	396 (55.6)
Usual interstitial pneumonia, n (%)	116 (29.3)
Non-specific interstitial pneumonia, n (%)	90 (22.7)
Interstitial lymphocytic pneumonia, n (%)	12 (3)
Cryptogenic organizing pneumonia, n (%)	80 (20.2)
Desquamative interstitial pneumonia, n (%)	6 (1.5)
Acute interstitial pneumonia, n (%)	25 (6.3)
Obliterans bronchiolitis, n (%)	61 (15.4)
Emphysema-fibrosis syndrome, n (%)	6 (1.5)
Time from RA to ILD diagnosis, years	
N (%)	492 (69.1)
Mean (SD)	5.8 (9.4)#
Median (Q1, Q3)	1.9 (0.1, 6.9)

ILD: Interstitial lung disease; RA: Rheumatoid arthritis. Data extracted and analyzed considering from the first EHR available to 2 months after the inclusion date.

Supplementary Table 7. Frequency of comorbidities at the time of inclusion

	RA (13958)	RAILD (712)	RAnonILD (13246)	p value	OR (95% CI)
Hypertension, n (%)	5194 (37.2)	342 (48)	4852 (36.6)	<0.001*	1.60 (1.37, 1.87)
Diabetes mellitus, n (%)	2111 (15.1)	160 (22.5)	1951 (14.7)	<0.001*	1.68 (1.39, 2.02)
Dyslipidemia, n (%)	6580 (47.1)	404 (56.7)	6176 (46.6)	<0.001*	1.50 (1.29, 1.75)
Cardiovascular disease, n (%)	2023 (14.5)	184 (25.8)	1839 (13.9)	<0.001*	2.16 (1.80, 2.58)
Atrial Fibrillation, n (%)	952 (6.8)	79 (11.1)	873 (6.6)	<0.001*	1.77 (1.37, 2.26)
Heart failure, n (%)	1198 (8.6)	128 (18)	1070 (8.1)	<0.001*	2.49 (2.02, 3.06)
Acute myocardial infarction, n (%)	290 (2.1)	19 (2.7)	271 (2)	0.278	1.31 (0.77, 2.11)
TIA, n (%)	184 (1.3)	10 (1.4)	174 (1.3)	0.737	1.07 (0.50, 2.03)
Obesity, n (%)	877 (6.3)	65 (9.1)	812 (6.1)	0.002*	1.54 (1.16, 2.01)
Fatty Liver, n (%)	455 (3.3)	35 (4.9)	420 (3.2)	0.016**	1.58 (1.08, 2.25)
Malignancy, n (%)	1243 (8.9)	113 (15.9)	1130 (8.5)	<0.001*	2.02 (1.62, 2.50)
Infections, n (%)	2431 (17.4)	239 (33.6)	2192 (16.5)	<0.001*	2.55 (2.16, 3.00)
Latent tuberculosis infection, n (%)	51 (0.4)	9 (1.3)	42 (0.3)	<0.001*	4.02 (1.72, 8.43)
COPD, n (%)	718 (5.1)	93 (13.1)	625 (4.7)	<0.001*	3.03 (2.38, 3.84)
Asthma, n (%)	844 (6)	69 (9.7)	775 (5.9)	<0.001*	1.73 (1.31, 2.24)
Depression, n (%)	1036 (7.4)	68 (9.6)	968 (7.3)	0.033*+	1.34 (1.02, 1.74)
Kidney failure, n (%)	1260 (9)	109 (15.3)	1151 (8.7)	<0.001*	1.90 (1.52, 2.36)
Inflammatory Bowel Disease, n (%)	417 (3)	27 (3.8)	390 (2.9)	0.212	1.30 (0.84, 1.94)
Crohn's Disease, n (%)	147 (1.1)	14 (2)	133 (1)	0.022*	1.98 (1.05, 3.46)
Ulcerative colitis, n (%)	84 (0.6)	9 (1.3)	75 (0.6)	0.039*+	2.25 (0.99, 4.52)
Psoriasis, n (%)	629 (4.5)	26 (3.7)	603 (4.6)	0.307	0.79 (0.51, 1.19)
Reactive arthritis, n (%)	40 (0.3)	1 (0.1)	39 (0.3)	0.722	0.48 (0.01, 2.82)
Uveitis, n (%)	237 (1.7)	4 (0.6)	233 (1.8)	0.011*	0.32 (0.09, 0.82)
Gout, n (%)	628 (4.5)	27 (3.8)	601 (4.5)	0.403	0.83 (0.54, 1.23)
Hyperuricemia, n (%)	599 (4.3)	31 (4.4)	568 (4.3)	0.924	1.02 (0.68, 1.47)
Hypertransaminasemia, n (%)	202 (1.4)	13 (1.8)	189 (1.4)	0.336	1.28 (0.67, 2.27)
Osteoporosis, n (%)	1711 (12.3)	134 (18.8)	1577 (11.9)	<0.001*	1.72 (1.40, 2.09)
Bone fracture, n (%)	2319 (16.6)	150 (21.1)	2169 (16.4)	<0.002*	1.36 (1.12, 1.65)

CI: Confidence interval; COPD: Chronic obstructive pulmonary disease; ILD: Interstitial lung disease; OR: Odds ratio (Fisher test); RA: Rheumatoid arthritis; SD: Standard deviation; TIA: Transient Ischemic Attack. Data extracted and analyzed considering a window of [-6, +2 months] around the inclusion date. *Statistical differences were considered significant when $p < 0.05$ in two-tailed test. +When correcting for multiple testing (adjusted p values), $p > 0.05$.

Supplementary Table 8. Blood tests at the time of inclusion

	RA (13958)	RAILD (712)	RAnonILD (13246)	p value	Estimate ¹ (95% CI)
Hemoglobin, N (%)	7178 (51.4)	422 (59.3)	6756 (51)		
Mean (SD), g/dl	13.8 (5.1)	13 (4.1)#	13.8 (5.1)	<0.001*	-0.80 (-1.21, -0.39)
Median (Q1, Q3), g/dl	13 (11.6, 14.3)	12.6 (11.1, 14)	13 (11.6, 14.3)		
Anemia [†] , n (%)	2780 (19.9)	205 (28.8)	2575 (19.4)	<0.001*	1.68 (1.41, 1.99) ²
Platelets, N (%)	4458 (31.9)	271 (38.1)	4187 (31.6)		
Mean (SD), 1000/mm ³	251.3 (118.1)#	247.5 (120)#	251.5 (117.9)#	0.248	-7.00 (-20.00, 5.00) ³
Median (Q1, Q3), 1000/mm ³	243 (187, 308)	238 (174.5, 314.5)	243 (188, 307)		
Thrombocytopenia, n (%)	482 (3.5)	39 (5.5)	443 (3.3)	0.004*	1.67 (1.16, 2.35)
Thrombocytosis, n (%)	541 (3.9)	41 (5.8)	500 (3.8)	0.012*	1.56 (1.09, 2.17)
HbA1c, N (%)	1057 (7.6)	97 (13.6)	960 (7.2)		
Mean (SD), %	7.7 (6.2)#	7.6 (5.7)#	7.7 (6.2)#	0.296	0.10 (-0.10, 0.40)
Median (Q1, Q3), %	6.1 (5.6, 7.2)	6.4 (5.7, 7.2)	6.1 (5.6, 7.2)		
HDL, N (%)	1525 (10.9)	112 (15.7)	1413 (10.7)		
Mean (SD), mg/dl	52.3 (20.9)#	47.4 (20)#	52.7 (21)#	0.004*	-0.05 (-9.00, -2.00)
Median (Q1, Q3), mg/dl	51 (38, 64)	45 (35, 57.2)	51 (39, 64)		
LDL, N (%)	2123 (15.2)	144 (20.2)	1979 (14.9)		
Mean (SD), mg/dl	113.5 (42.3)#	109.6 (41.5)#	113.7 (42.4)#	0.204	-4.00 (-11.00, 2.00)
Median (Q1, Q3), mg/dl	112 (84, 140)	107.5 (80, 132)	112 (84, 140)		
Triglycerides, N (%)	2973 (21.3)	223 (31.3)	2750 (20.8)		
Mean (SD), mg/dl	135.7 (186.5)#	148.7 (114.5)#	134.7 (191.1)#	<0.001*	14.00 (6.00, 22.00)
Median (Q1, Q3), mg/dl	110 (77, 163)	124 (93, 169)	109 (77, 163)		
Albumin, N (%)	1477 (10.6)	156 (21.9)	1321 (10)		
Mean (SD), mg/dl	5.8 (8.1)#	7.9 (10.8)#	5.6 (7.6)#	0.681	2.28 (0.52, 4.03)
Median (Q1, Q3), mg/dl	3.8 (3.2, 4.2)	3.7 (3.2, 4.3)	3.8 (3.1, 4.2)		
Total bilirubin, N (%)	2778 (19.9)	185 (26)	2593 (19.6)		
Mean (SD), mg/dl	1.2 (4.1)#	1 (3.4)#	1.2 (4.1)#	0.605	0.01 (-0.06, 0.03)
Median (Q1, Q3), mg/dl	0.5 (0.3, 0.8)	0.4 (0.3, 0.7)	0.5 (0.3, 0.8)		
Creatinine, N (%)	2648 (19)	209 (29.4)	2439 (18.4)		
Mean (SD), mg/dl	1 (0.8)#	1 (0.9)#	1 (0.8)#	0.851	0.01 (-0.05, 0.03)
Median (Q1, Q3), mg/dl	0.9 (0.7, 1.1)	0.9 (0.7, 1.1)	0.9 (0.7, 1)		
GFR, N (%)	1373 (9.8)	94 (13.2)	1279 (9.7)		
Mean (SD), ml/min/1.73m ²	64.8 (25.3)#	64.9 (25.3)#	64.8 (25.3)#	0.672	0.01 (-5.00, 3.00)
Median (Q1, Q3), ml/min/1.73m ²	63 (47, 88)	60 (49.2, 87.2)	63.7 (47, 88)		
ALT, N (%)	5761 (41.3)	335 (47.1)	5426 (41)		
Mean (SD), u/l	84.9 (562.5)	75.4 (379.1)#	85.5 (571.9)	0.646	-10.18 (-53.66, 33.29) ²
Median (Q1, Q3), u/l	20 (13, 36)	19 (12, 34)	20 (13, 36)		
AST, N (%)	4350 (31.2)	276 (38.8)	4074 (30.8)		
Mean (SD), u/l	53 (388.3)#	56.3 (362.1)#	52.8 (390.1)#	0.219	1.00 (-1.00, 2.00)

	RA (13958)	RAILD (712)	RAnonILD (13246)	p value	Estimate ¹ (95% CI)
Median (Q1, Q3), u/l	22 (17, 32)	23.5 (17, 35)	22 (17, 32)		
GGT, N (%)	3365 (24.1)	260 (36.5)	3105 (23.4)		
Mean (SD), u/l	72.9 (154.4)#	104.1 (277.3)#	70.3 (139.1)#	0.004	5.00 (2.00, 9.00)
Median (Q1, Q3), u/l	36 (19, 71)	42 (23, 86)	36 (19, 70)		

ACPA: Anticitrullinated protein antibody; ALT: Alanine aminotransferase; AST: Aspartate aminotransferase; CI: Confidence interval; CRP: C-reactive protein; ESR: Erythrocyte sedimentation rate; GFR: Glomerular filtration rate; GGT: gamma glutamyl transferase; ILD: Interstitial lung disease; HbA1c: Hemoglobin A1c; HDL: High density lipoprotein; LDL: Low density lipoprotein; RA: Rheumatoid arthritis; RF: Rheumatoid factor; SD: Standard deviation. Data extracted and analyzed considering a window of [-6, +2 months] around the inclusion date. If multiple values were detected in the time window, only the closest to index was retained. #Median (Q1, Q3) is preferred over mean (SD) for interpretation as the feature is non-normal. ¹ Wilcoxon test (difference of location) was performed for comparing RAILD vs RAnonILD patients. ² Odds ratio (Fisher test) ³ Welch t-test was performed for statistical analysis of ALT. *Statistical differences were considered significant when $p < 0.05$ in two-tailed tests. +When correcting for multiple testing (adjusted p values), $p > 0.05$.

Supplementary Table 9. Sensitivity analysis with RA-support population at the time of inclusion

	RA (support) (12538)	RAILD (support) (696)	RAnonILD (support) (11842)
General characteristics			
Sex n (%)	12538 (100)	696 (100)	11842 (100)
Female n (%)	9280 (74)	432 (62.1)	8848 (74.7)
Male n (%)	3258 (26)	264 (37.9)	2994 (25.3)
Age, years			
N (%)	12538 (100)	696 (100)	11842 (100)
Mean (SD)	60.8 (16.6)	64.9 (13.4)#	60.6 (16.7)
Median (Q1, Q3)	62 (49, 74)	65 (57, 75)	61 (49, 74)
Tobacco n (%)	4171 (33.3)	332 (47.7)	3839 (32.4)
Tobacco yes/ex* n (%)	2457 (58.9)	218 (65.7)	2239 (58.3)
Tobacco no* n (%)	1714 (41.1)	114 (34.3)	1600 (41.7)
Alcohol n (%)	927 (7.4)	46 (6.6)	881 (7.4)
Alcohol yes/ex* n (%)	285 (30.7)	17 (37)	268 (30.4)
Alcohol no* n (%)	642 (69.3)	29 (63)	613 (69.6)
Comorbidities			
Hypertension n (%)	4720 (37.6)	335 (48.1)	4385 (37)
Diabetes mellitus n (%)	1916 (15.3)	156 (22.4)	1760 (14.9)
Dyslipidemia n (%)	5944 (47.4)	395 (56.8)	5549 (46.9)
Cardiovascular disease n (%)	1905 (15.2)	181 (26)	1724 (14.6)
Atrial Fibrillation n (%)	892 (7.1)	77 (11.1)	815 (6.9)
Heart failure n (%)	1140 (9.1)	128 (18.4)	1012 (8.5)
Acute myocardial infarction n (%)	276 (2.2)	18 (2.6)	258 (2.2)
TIA n (%)	173 (1.4)	10 (1.4)	163 (1.4)
Obesity n (%)	806 (6.4)	65 (9.3)	741 (6.3)
Fatty Liver n (%)	409 (3.3)	33 (4.7)	376 (3.2)
Malignancy n (%)	1163 (9.3)	110 (15.8)	1053 (8.9)
Infections n (%)	2289 (18.3)	237 (34.1)	2052 (17.3)
Latent tuberculosis infection n (%)	49 (0.4)	9 (1.3)	40 (0.3)
COPD n (%)	668 (5.3)	93 (13.4)	575 (4.9)
Asthma n (%)	775 (6.2)	66 (9.5)	709 (6)
Depression n (%)	940 (7.5)	65 (9.3)	875 (7.4)
Kidney failure n (%)	1184 (9.4)	108 (15.5)	1076 (9.1)
Inflammatory Bowel Disease n (%)	396 (3.2)	26 (3.7)	370 (3.1)
Crohn's Disease n (%)	141 (1.1)	14 (2)	127 (1.1)
Ulcerative Colitis n (%)	78 (0.6)	9 (1.3)	69 (0.6)
Psoriasis n (%)	542 (4.3)	26 (3.7)	516 (4.4)
Reactive arthritis n (%)	35 (0.3)	1 (0.1)	34 (0.3)
Uveitis n (%)	206 (1.6)	3 (0.4)	203 (1.7)

	RA (support (12538)	RAILD (support (696)	RAnonILD (support (11842)
Gout n (%)	567 (4.5)	27 (3.9)	540 (4.6)
Hyperuricemia n (%)	545 (4.3)	29 (4.2)	516 (4.4)
Hypertransaminasemia n (%)	183 (1.5)	13 (1.9)	170 (1.4)
Osteoporosis n (%)	1592 (12.7)	132 (19)	1460 (12.3)
Bone fracture n (%)	2155 (17.2)	148 (21.3)	2007 (16.9)

ILD: Interstitial lung disease; RA: Rheumatoid arthritis. *The presence of each feature was analyzed at the time of inclusion with a window of [-6, 2] months. *Percentages calculated considering patients with available data. (#) Median (Q1, Q3) is preferred over mean (SD) for interpretation as the feature is non-normal.*

Supplementary Table 10. Supplementary analysis in patients treated with DMARDs or glucocorticoids.

	RA (treated) (10315)	RAILD (treated) (651)	RAnonILD (treated) (9664)
Conventional synthetic (cs) DMARDs n (%)	5381 (52.2)	318 (48.8)	5063 (52.4)
Methotrexate n (%)	4994 (48.4)	271 (41.6)	4723 (48.9)
Leflunomide n (%)	1014 (9.8)	117 (18)	897 (9.3)
Sulfasalazine n (%)	117 (1.1)	13 (2)	104 (1.1)
Biological (b) DMARDs n (%)	2257 (21.9)	194 (29.8)	2063 (21.3)
TNF inhibitors n (%)	1853 (18)	108 (16.6)	1745 (18.1)
Adalimumab n (%)	766 (7.4)	49 (7.5)	717 (7.4)
Certolizumab pegol n (%)	179 (1.7)	5 (0.8)	174 (1.8)
Etanercept n (%)	890 (8.6)	41 (6.3)	849 (8.8)
Golimumab n (%)	178 (1.7)	6 (0.9)	172 (1.8)
Infliximab n (%)	251 (2.4)	28 (4.3)	223 (2.3)
Other MoA n (%)	647 (6.3)	124 (19)	523 (5.4)
Abatacept n (%)	347 (3.4)	62 (9.5)	285 (2.9)
Rituximab n (%)	332 (3.2)	70 (10.8)	262 (2.7)
Tocilizumab n (%)	386 (3.7)	34 (5.2)	352 (3.6)
Targeted synthetic (ts) DMARDs, JAK inhibitors n (%)	76 (0.7)	2 (0.3)	74 (0.8)
Tofacitinib n (%)	56 (0.5)	2 (0.3)	54 (0.6)
Baricitinib n (%)	24 (0.2)	0 (0)	24 (0.2)
Filgotinib n (%)	1 (0)	0 (0)	1 (0)
Upadacitinib n (%)	1 (0)	0 (0)	1 (0)
Others			
Cyclophosphamide n (%)	86 (0.8)	28 (4.3)	58 (0.6)
Mycophenolate n (%)	113 (1.1)	64 (9.8)	49 (0.5)
Azathioprine n (%)	184 (1.8)	49 (7.5)	135 (1.4)
Glucocorticoids n (%)	2878 (27.9)	377 (57.9)	2501 (25.9)

DMARD: Disease-modifying antirheumatic drug; ILD: Interstitial lung disease; JAK: Janus kinase; MoA: Mechanism of action; RA: Rheumatoid arthritis; TNF: Tumor necrosis factor. *The presence of each feature was analyzed between the time of inclusion and two years of follow-up.*

Supplementary Table 11. Healthcare resource utilization during follow-up

	RA (13958)	RAILD (712)	RAnonILD (13246)	p value	OR (95% CI)
Number of hospitalizations					
N (%)	4473 (32)	442 (62.1)	4031 (30.4)	<0.001**	3.74 (3.19, 4.39)
Mean (SD)	25.6 (70.9)#	22.3 (58.3)#	25.9 (72.1)#		
Median (Q1, Q3)	2.8 (0.9, 10.7)	4.3 (1.3, 15.7)	2.7 (0.9, 10.2)	<0.001**	0.65 (0.32, 1.05)
Number of outpatient visits					
N (%)	10731 (76.9)	617 (86.7)	10114 (76.4)	<0.001**	2.01 (1.61, 2.53)
Mean (SD)	7.4 (18.6)	9.1 (12.9)#	7.3 (18.9)	0.001**	1.76 (0.68, 2.84)
Median (Q1, Q3)	3.9 (1.7, 8)	6.3 (2.8, 11.4)	3.8 (1.6, 7.8)		
Number of emergency room visits					
N (%)	6102 (43.7)	435 (61.1)	5667 (42.8)	<0.001**	2.10 (1.79, 2.46)
Mean (SD)	3.1 (16.8)	2.1 (3.2)#	3.2 (17.4)	<0.001**	-1.09 (-1.64, -0.55)
Median (Q1, Q3)	1 (0.5, 2)	1.1 (0.5, 2.3)	1 (0.5, 2)		

ILD: Interstitial lung disease; OR: Odds ratio (Fisher test); RA: Rheumatoid arthritis; SD: Standard deviation. *The presence of each feature was analyzed between the time of inclusion and the last follow-up. For numerical features, the occurrence count of each feature was analyzed and divided by follow-up years. (#) Median (Q1, Q3) is preferred over mean (SD) for interpretation as the feature is non-normal.*