



OPEN ACCESS

Original research

End of rarity: the evolving burden of interstitial lung diseases in England – trends from a national population-based cohort study

Georgie May Massen ,¹ Ann Dorothy Morgan ,² Peter M George ,^{3,4}
Jennifer K Quint ,¹ Vidya Navaratnam ^{1,5,6}

► Additional supplemental material is published online only. To view, please visit the journal online (<https://doi.org/10.1136/thorax-2026-224977>).

¹School of Public Health, Imperial College London Faculty of Medicine, London, UK

²National Heart and Lung Institute, Imperial College London, London, UK

³Interstitial Lung Disease Unit, Royal Brompton and Harefield NHS Foundation Trust, London, UK

⁴National Heart and Lung Institute, London, UK

⁵Faculty of Health, Medicine and Behavioural Sciences, The University of Queensland, Saint Lucia, Queensland, Australia

⁶Institute for Respiratory Health, Nedlands, Western Australia, Australia

Correspondence to

Professor Jennifer K Quint;
j.quint@imperial.ac.uk

GMM and ADM are joint first authors.

JKQ and VN are joint senior authors.

Received 24 February 2026
Accepted 21 August 2026

ABSTRACT

Background The disease burden from interstitial lung diseases (ILDs) is thought to be increasing globally, but there is a paucity of data on epidemiological trends in non-idiopathic pulmonary fibrosis (IPF) ILD. We calculated incidence, prevalence and mortality rates of all ILD, IPF and non-IPF ILD in England from 2005 to 2022 and investigated trends over time.

Methods Using linked primary and secondary electronic healthcare records, we used three code-based algorithms to estimate incidence, prevalence and mortality rates of all ILD, IPF and non-IPF ILD subtypes. We additionally stratified our results by sex, age-group and region. Poisson regression was used to investigate changes to trends over time. We used Kaplan-Meier methods to calculate median survival.

Results There were 153 975 incident cases of ILD in either primary or secondary care records. Incidence rates of all ILD, IPF and non-IPF ILD were 30% higher in 2022 compared with 2005, with rates plateauing from 2018 onwards. The prevalence of IPF and non-IPF ILD increased by 6% and 7%, respectively, over the study period. Mortality rates adjusted for age and sex in people with all ILD and non-IPF ILD were similar in 2022 compared with 2005 but in people with IPF it was 15% lower. Median survival of our overall ILD cohort was 5.4 years (IQR 1.88–13.03).

Conclusions Our population-based study highlights that ILD are no longer rare. Currently, there are 200 000 people living with ILD and 60 000 new cases are diagnosed each year. Healthcare and research funding needs to reflect this growing burden.

INTRODUCTION

Interstitial lung diseases (ILDs) are characterised by inflammation and fibrosis of the lung parenchyma and are associated with high morbidity and mortality. They are classified based on aetiology, with idiopathic pulmonary fibrosis (IPF) being the most common subtype. ILDs can develop in association with systemic autoimmune rheumatic diseases (SARDs) such as rheumatoid arthritis or scleroderma, or because of environmental or occupational exposures (eg, hypersensitivity pneumonitis (HP)).

ILD disease burden is reported to be on the rise, due to increased disease awareness and availability of CT scanning. Of the ILD subtypes, the incidence, prevalence and mortality of IPF is the

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Epidemiological data for interstitial lung disease (ILD), especially non-idiopathic pulmonary fibrosis (IPF) ILD remains sparse and inconsistent, partly due to heterogeneous study designs and outcomes.

WHAT THIS STUDY ADDS

⇒ Updated incidence, prevalence and mortality of ILD in England, including non-IPF ILD subtypes.
⇒ Demonstrates increasing ILD incidence and prevalence in England over almost two decades.
⇒ Highlights that all-cause mortality rates in people with IPF were 15% lower in 2022 compared with 2005.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ We estimate that 60 000 new ILD cases are diagnosed annually.
⇒ Choosing appropriate data sources is crucial for estimating ILD disease burden in England and allows for effective health service planning and availability of holistic care for these patients.

most extensively reported, with incidence rates across Europe varying between 0.2 and 27.2 per 100 000 person-years.^{1 2} However, data on non-IPF ILD incidence and prevalence are still limited, partly due to its heterogeneous nature. Recently, a study using UK primary care electronic healthcare records reported that the incidence of all ILDs was stable between 2010 and 2019, with an overall rate of 28.71 per 100 000 person-years.³ Using administrative healthcare data from insurance claims in Canada, Johansson *et al* also found stable overall ILD incidence rates over a similar period of time, with increasing prevalence and decreasing 1-year mortality.⁴

Given the growing interest in non-IPF ILD, recognition that fibrotic phenotypes have similar prognoses to IPF and new emerging therapies for patients with non-IPF ILD, we aimed to provide up to date estimates of ILD (IPF and non-IPF ILD subtypes) incidence, prevalence and mortality in England over almost two decades, using linked electronic healthcare records.



© Author(s) (or their employer(s)) 2026. Re-use permitted under CC BY. Published by BMJ Group.

To cite: Massen GM, Morgan AD, George PM, *et al*. *Thorax* Epub ahead of print: [please include Day Month Year]. doi:10.1136/thorax-2026-224977

METHODS

Data sources

We used data from the Clinical Practice Research Datalink (CPRD) Aurum database, March 2025 build, linked to National Health Service England's Hospital Episode Statistics Admitted Patient Care (HES-APC) database and Index of Multiple Deprivation data (IMD). CPRD Aurum contains pseudonymised primary care data covering 24% of general practice registered patients in England and is broadly representative of the wider UK population in terms of age, sex, ethnicity, region and deprivation.⁵ CPRD Aurum data contains information on patient demographics, clinical diagnoses, specialist referrals, laboratory testing and prescriptions from primary care, captured from participating general practices. Further information on data sources used in this study can be found in the online supplemental methods section. Ethical approval was obtained to conduct this research study (protocol number: 25_004889), full details can be found in the ethics section at the end of this manuscript.

ILD definition

People with a record of ILD were identified in either primary or secondary care records. Three code-based algorithms were applied: (1) algorithm one captured people with a record of ILD present in primary care (CPRD Aurum) only, (2) algorithm two captured people with a record of ILD present in secondary care (HES-APC) only, while (3) algorithm three captured people with a record of ILD in either primary (CPRD Aurum) or secondary care (HES-APC). All ILD codes used in this study, including individual subtypes, were reviewed independently by two ILD experts. The code lists and files used to define ILD in our population are available at: <https://github.com/NHLI-Respiratory-Epi/Incidence-prevalence-and-mortality-of-ILD>.

Our initial study population consisted of individuals with any ILD code, including those with only a broad ILD code and codes for multiple ILD subtypes. This was the study population for our all ILD cohort. We later stratified our all ILD study population into IPF and non-IPF ILD cohorts. At this point, people with only a broad ILD code or multiple ILD subtype codes were excluded. Our third analyses consisted of individual ILD subtype cohorts. When performing these analyses, we re-included people with only a broad ILD code as well as those with multiple ILD subtype codes as separate and distinct groups.

We additionally stratified people into the following ILD subtypes: exposure-related ILD (including HP), connective tissue disease (CTD) (or SARD) associated ILD, drug or radiation induced ILD, pulmonary sarcoidosis and other ILD (eg, smoking-related ILD, organising pneumonia). We chose to use the term exposure-related ILD as the diagnostics codes in this subtype were not specific enough for HP. Further information on the code-based algorithms used to construct our ILD cohort is found in the online supplemental methods section.

Study population

Our study period was from 1 January 2005 to 31 December 2022. The final study population included people who were aged 18 years or older at entry to the study, enrolled at a general practice eligible for linkage to HES and IMD, had non-missing sex information and with at least 1 day of follow-up during our study period after they met all study inclusion criteria. The start of follow-up was the latest of study start date (1 January 2005), the patient's 18th birthday, the date when the patient joined the general practice or the date of first record of ILD. The end of follow-up was the earliest of study end date (31 December

2022), date the patient left the practice, date of last data collection to CPRD Aurum or date of death.

STATISTICAL ANALYSIS

Incidence and prevalence analyses

People were included in the incidence analyses if their first record of ILD was during the study period and had at least 1 year of registration with their current general practitioner (GP) prior to first record of ILD. We calculated crude and age-standardised annual incidence rates as well as point prevalence of all ILD, IPF and non-IPF ILD between 2005 and 2022 for each code-based algorithm (primary care only, secondary care only, primary or secondary care). We then repeated these analyses, stratifying by ILD subtype. For our incidence and prevalence analyses, the denominator consisted of adults over the age of 18 years derived from CPRD Aurum. Incidence rates were calculated as number of new cases divided by person-time at risk for each patient, for each respective year. Annual point prevalence was estimated on the 1st of July of each year. Estimates of prevalence were further stratified by age, sex and region. To investigate changes to ILD incidence over time, we used Poisson regression to estimate annual incidence rate ratios, relative to 2005, controlling for age group and sex.

Mortality and survival analyses

Annual crude all-cause mortality rates were calculated for people with all ILD, IPF and non-IPF ILD for each code-based algorithm, dividing the number of all ILD, IPF and non-IPF ILD deaths respectively by the total time at risk contributed by eligible patients. We used Poisson regression to estimate annual mortality rate ratios, with 2005 as our baseline year, controlling for the effects of age and sex. We then calculated annual crude all-cause mortality rates for individual ILD subtypes. Finally, we estimated median survival for our ILD cohort overall and each subtype, using Kaplan-Meier methods. The proportional hazards assumption was confirmed using graphical methods.

RESULTS

Study population characteristics

Between 2005 and 2022, there were 64 434 incident cases of ILD in primary care only, 121 040 incident cases of ILD in secondary care only and 153 975 incident cases of ILD in either primary or secondary care (table 1). All incident cases of ILD had similar demographics across all three code-based algorithms (table 1). Incident cases of IPF were older at time of diagnosis, with a higher proportion being male and people who have quit smoking compared with non-IPF ILD incident cases (tables 1 and 2).

9454 (14.7%) of our ILD cases only had a broad ILD code and 962 (1.5%) had codes for multiple ILD subtypes. The most common non-IPF ILD subtype was exposure-related ILD, followed by CTD-ILD and pulmonary sarcoidosis (table 2). People with exposure-related ILD were more likely to be male compared with those with CTD-ILD or pulmonary sarcoidosis and were older at diagnosis (table 2). There were no other significant differences in demographics between the ILD subtypes (table 2).

Incidence and prevalence analyses

Across all three code-based algorithms, the incidence of all ILD, IPF and non-IPF ILD increased between 2005 and 2022 (figure 1, online supplemental table 1). The incidence rate of all ILD in primary care increased from 2.68 (95% CI 2.56 to

Table 1 Demographics of people with ILD, non-IPF ILD and IPF with records in primary care only (Aurum), secondary care only (HES) or combined (Aurum or HES)

	ILD			Non-IPF-ILD			IPF		
	Aurum	HES	Aurum or HES	Aurum	HES	Aurum or HES	Aurum	HES	Aurum or HES
Incident cases (N)	64 434	121 040	153 975	23 090	40 739	55 130	30 928	45 639	57 987
Sex (% male)	38 433 (59.65)	62 900 (51.97)	83 167 (54.01)	12 922 (55.96)	17 806 (43.71)	26 406 (47.90)	19 477 (62.98)	27 206 (59.61)	35 200 (60.70)
Age at diagnosis (Median, IQR)	74 (64–81)	70 (55–79)	70 (57–79)	68 (55–76)	61 (45–74)	63 (48–75)	77 (70–83)	76 (67–84)	76 (67–83)
Smoking status (%)									
Person who smokes currently	2408 (3.74)	22 108 (18.27)	21 868 (14.20)	854 (3.70)	6124 (15.03)	6205 (11.26)	1092 (3.53)	6107 (27.15)	5828 (10.05)
Person who has quit smoking	18 566 (28.81)	58 855 (48.62)	57 046 (37.05)	5637 (24.41)	17 687 (43.42)	18 253 (33.11)	10 183 (32.92)	26 046 (57.07)	23 654 (40.79)
Person who does not smoke	9084 (14.10)	37 127 (30.67)	21 868 (14.20)	3669 (15.89)	15 994 (39.26)	6205 (11.26)	4066 (13.15)	12 389 (27.15)	11 196 (19.31)
Unknown	34 376 (53.35)	2950 (2.44)	38 923 (25.28)	12 930 (56.00)	933 (2.29)	14 576 (26.44)	15 587 (50.40)	1097 (2.40)	17 309 (29.85)
Ethnicity (%)									
Caucasian	56 473 (87.64)	103 838 (85.79)	132 160 (85.83)	19 549 (84.66)	33 067 (81.17)	45 156 (81.91)	27 803 (89.90)	40 653 (89.08)	51 539 (88.88)
Non-Caucasian	5572 (8.65)	14 956 (12.36)	17 652 (11.46)	2638 (11.42)	7120 (17.48)	8560 (15.53)	2021 (6.53)	3961 (8.68)	4675 (8.06)
Unknown	2389 (3.71)	2245 (1.85)	4163 (2.70)	903 (3.91)	551 (1.35)	1414 (2.56)	1104 (3.57)	1025 (3.35)	1773 (3.06)
Deprivation (%)									
1 (Least deprived)	13 441 (20.86)	21 932 (18.12)	28 840 (18.73)	4600 (19.92)	7332 (18.00)	10 243 (18.58)	6538 (21.14)	8754 (19.18)	11 372 (19.61)
2	13 486 (20.93)	23 651 (19.53)	30 472 (19.79)	4730 (20.49)	7687 (19.31)	10 794 (19.58)	6551 (21.18)	9246 (20.26)	11 829 (20.40)
3	12 416 (19.27)	23 410 (19.34)	29 671 (19.27)	4526 (19.60)	8136 (19.97)	10 949 (19.86)	5879 (19.01)	8831 (19.35)	11 093 (19.13)
4	12 565 (19.50)	25 503 (21.07)	31 934 (20.74)	4678 (20.26)	8971 (22.02)	11 841 (21.48)	5935 (19.19)	9169 (20.09)	11 545 (19.91)
5 (Most deprived)	12 526 (19.44)	26 544 (21.93)	33 058 (21.47)	4556 (19.73)	8433 (20.70)	11 301 (20.50)	6025 (19.48)	9639 (21.12)	12 148 (20.96)
Region (%)									
North East	2249 (3.49)	3922 (3.24)	4989 (3.24)	826 (3.58)	1247 (3.06)	1701 (3.09)	1162 (3.76)	1527 (3.35)	1983 (3.42)
North West	14 698 (22.81)	24 330 (20.10)	31 869 (20.70)	4893 (21.19)	7365 (18.08)	10 524 (19.09)	7850 (25.38)	10 427 (22.85)	13 676 (23.58)
Yorkshire and Humber	2333 (3.62)	3701 (3.06)	4901 (3.18)	720 (3.12)	1218 (2.99)	1669 (3.03)	1332 (4.31)	1536 (3.37)	2125 (3.66)
East Midlands	1502 (2.33)	2859 (2.36)	3 583 (2.33)	562 (2.43)	946 (2.32)	1287 (2.33)	705 (2.28)	1069 (2.34)	1308 (2.26)
West Midlands	11 007 (17.08)	19 567 (16.17)	25 150 (16.33)	3606 (15.62)	6144 (15.08)	8405 (15.25)	5248 (16.97)	7948 (17.41)	9997 (17.24)
East of England	2859 (4.44)	5460 (4.51)	6804 (4.42)	1051 (4.55)	1867 (4.58)	2490 (4.52)	1127 (3.64)	2203 (4.83)	2560 (4.41)
London	8434 (13.09)	22 092 (18.25)	26 529 (17.23)	3576 (15.49)	8736 (21.44)	10 938 (19.84)	3501 (11.32)	6665 (14.60)	8064 (13.91)
South East	13 111 (20.35)	23 581 (19.48)	30 634 (19.90)	4747 (20.56)	7800 (19.15)	10 866 (19.71)	6227 (20.13)	8670 (19.00)	11 297 (19.48)
South West	8241 (12.79)	15 527 (12.83)	19 516 (12.67)	3109 (13.46)	5415 (13.29)	7250 (13.15)	3776 (12.21)	5594 (12.26)	6977 (12.03)
HES, Hospital Episode Statistics; ILD, interstitial lung disease; IPF, idiopathic pulmonary fibrosis.									

Table 2 Demographics of people with ILD, IPF, non-IPF ILD and individual subtypes with records in primary care only (CPRD Aurum)

	ILD	Non-IPF-ILD	IPF	Exposure related ILD	CTD-ILD	Drug/radiation ILD	Pulmonary sarcoid	Other ILD	Broad ILD codes	Multiple ILD subtype codes
Incident cases (N)	64 434	23 090	30 928	8895	7035	701	4844	1615	9454	962
Sex (% male)	38 433 (59.65)	12 922 (55.96)	19 477 (62.98)	6452 (72.54)	2726 (38.75)	296 (42.23)	2747 (56.71)	701 (43.41)	5515 (58.34)	519 (53.95)
Age at diagnosis (median, IQR)	74 (64–81)	68 (55–76)	77 (70–83)	72 (62–79)	71 (62–78)	70 (60–78)	51 (42–61)	65 (53–75)	75 (66–82)	65 (54–74)
Smoking status (%)										
Person who smokes currently	2408 (3.74)	854 (3.70)	1092 (3.53)	245 (2.75)	348 (4.95)	39 (5.56)	160 (3.30)	62 (3.84)	404 (4.27)	58 (6.03)
Person who has quit smoking	18 566 (28.81)	5637 (24.41)	10 183 (32.92)	2107 (23.69)	2281 (32.42)	165 (23.54)	806 (16.64)	278 (17.21)	2418 (25.58)	328 (34.10)
Person who does not smoke	9084 (14.10)	3669 (15.89)	4066 (13.15)	1139 (12.80)	1291 (18.35)	48 (6.85)	1022 (21.10)	169 (10.46)	1129 (11.94)	220 (22.87)
Unknown	34 376 (53.35)	12 930 (56.00)	15 587 (50.40)	5404 (60.75)	3115 (44.28)	449 (64.05)	2856 (58.96)	1106 (68.48)	5503 (58.21)	356 (37.01)
Ethnicity (%)										
Caucasian	56 473 (87.64)	19 549 (84.66)	27 803 (89.90)	8159 (91.73)	5963 (84.76)	656 (93.58)	3390 (69.98)	1381 (85.51)	8291 (87.70)	830 (86.28)
Non-Caucasian	5572 (8.65)	2638 (11.42)	2021 (6.53)	413 (4.64)	912 (12.96)	35 (4.99)	1124 (23.20)	154 (9.54)	804 (8.50)	109 (11.33)
Unknown	2389 (3.71)	903 (3.91)	1104 (3.57)	323 (3.63)	160 (2.27)	10 (1.43)	330 (6.81)	80 (4.95)	359 (3.80)	23 (2.39)
Deprivation (%)										
1 (Least deprived)	13 441 (20.86)	4600 (19.92)	6538 (21.14)	1601 (17.99)	1476 (20.99)	160 (22.89)	992 (20.48)	371 (22.99)	1060 (21.88)	236 (24.48)
2	13 486 (20.93)	4730 (20.49)	6551 (21.18)	1820 (20.47)	1489 (21.16)	141 (20.03)	930 (19.19)	353 (21.81)	1044 (21.56)	163 (16.98)
3	12 416 (19.27)	4526 (19.60)	5879 (19.01)	1760 (19.79)	1375 (19.54)	144 (20.60)	940 (19.40)	306 (18.96)	939 (19.38)	182 (18.96)
4	12 565 (19.50)	4678 (20.26)	5935 (19.19)	1777 (19.98)	1383 (19.66)	126 (18.03)	1113 (22.98)	277 (17.16)	905 (18.69)	186 (19.27)
5 (Most deprived)	12 526 (19.44)	4556 (19.73)	6025 (19.48)	1937 (21.78)	1312 (18.65)	130 (18.45)	869 (17.95)	308 (19.08)	896 (18.49)	195 (20.31)
Region (%)										
North East	2249 (3.49)	826 (3.58)	1162 (3.76)	421 (4.73)	245 (3.48)	25 (3.57)	93 (1.92)	42 (2.60)	231 (2.44)	30 (3.12)
North West	14 698 (22.81)	4893 (21.19)	7850 (25.38)	2101 (23.62)	1499 (21.31)	177 (25.25)	765 (15.79)	351 (21.73)	1765 (18.67)	190 (19.75)
Yorkshire and Humber	2333 (3.62)	720 (3.12)	1332 (4.31)	276 (3.10)	253 (3.60)	19 (2.71)	115 (2.37)	57 (3.53)	250 (2.64)	31 (3.22)
East Midlands	1502 (2.33)	562 (2.43)	705 (2.28)	274 (3.08)	145 (2.06)	9 (1.28)	95 (1.96)	39 (2.41)	219 (2.32)	16 (1.66)
West Midlands	11 007 (17.08)	3606 (15.62)	5248 (16.97)	1298 (14.59)	1279 (18.18)	106 (15.12)	733 (15.13)	190 (11.76)	1989 (21.04)	164 (17.05)
East of England	2859 (4.44)	1051 (4.55)	1127 (3.64)	340 (3.82)	358 (5.09)	39 (5.56)	238 (4.91)	76 (4.71)	642 (6.79)	39 (4.05)
London	8434 (13.09)	3576 (15.49)	3501 (11.32)	1046 (11.76)	1045 (14.85)	84 (11.98)	1180 (24.36)	221 (13.68)	1226 (12.97)	131 (13.62)
South East	13 111 (20.35)	4747 (20.56)	6227 (20.13)	1830 (20.57)	1356 (19.28)	139 (19.83)	1015 (20.95)	407 (25.20)	1914 (20.25)	223 (23.18)
South West	8241 (12.79)	3109 (13.46)	3776 (12.21)	1309 (14.72)	855 (12.15)	103 (14.69)	610 (12.59)	232 (14.37)	1218 (12.88)	138 (14.35)
CPRD, Clinical Practice Research Datalink; CTD, connective tissue disease; ILD, interstitial lung disease; IPF, idiopathic pulmonary fibrosis.										

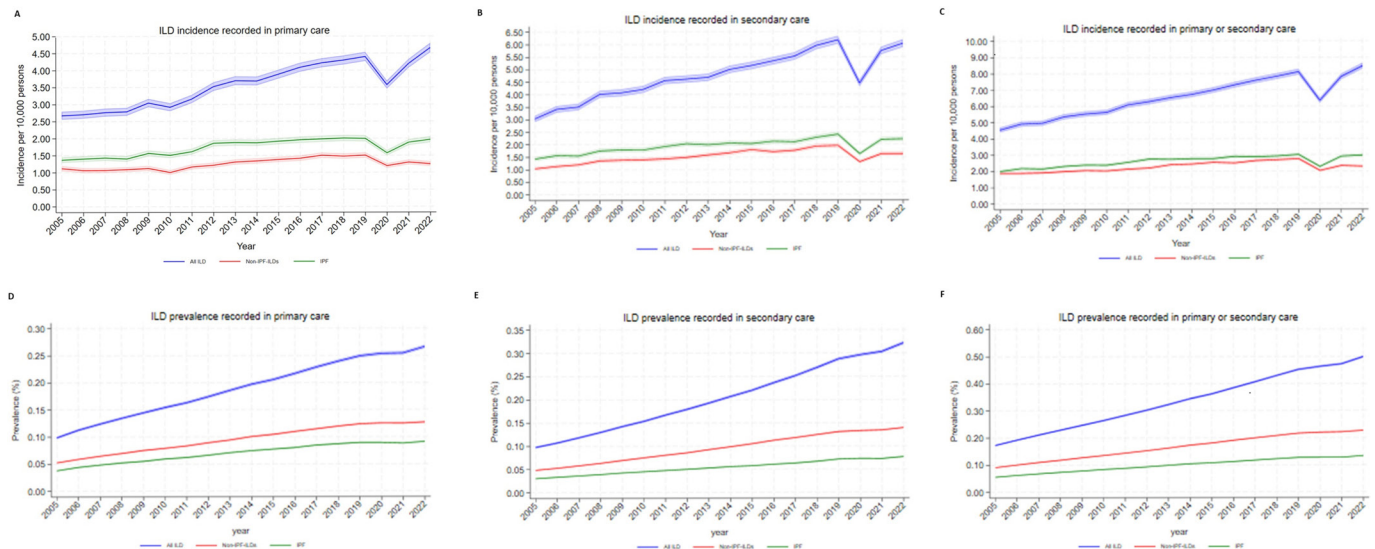


Figure 1 Incidence and prevalence of all ILD, IPF and non-IPF ILD for all three algorithms. ILD, interstitial lung disease; IPF, idiopathic pulmonary fibrosis.

2.79) per 10000 person-years in 2005 to 4.66 (95% CI 4.52 to 4.80) per 10000 person-years in 2022 (figure 1A, table 3). Using secondary care records only, the incidence rate of all ILD increased over the study period from 3.04 (95% CI 2.92 to 3.17) per 10000 person-years to 6.04 (95% CI 6.19 to 2.22) per 10000 person-years (figure 1B, online supplemental table 1). Our third code-based algorithm using a combination of either primary or secondary care records demonstrated a similar trend, with the highest all ILD incidence rate in 2022, which was 8.47 (95% CI 8.29 to 8.66) per 10000 person-years (figure 1C, online supplemental table 1). Using data from primary care only, we found that the incidence rates of ILD, IPF and non-IPF ILD were all 30% higher in 2022 compared with 2005, after controlling for age and sex. We observed a plateau in incidence rates from 2018 onwards (table 3).

We observed similar trends in all ILD, IPF and non-IPF ILD prevalence with all three algorithms (figure 1D–F, online supplemental table 2). Using primary care records only, the prevalence of all ILD increased from 0.10% (95% CI 0.10% to 0.10%) in 2005 to 0.27% (95% CI 0.26% to 0.27%) in 2022 (figure 1D, online supplemental table 2). Prevalence of all ILD was higher in secondary care records only; the 2022 point prevalence was 0.32% (95% CI 0.32% to 0.33%) (figure 1E, online supplemental table 2). Prevalence was highest when combining primary and secondary care records—in 2022, point prevalence for all ILD, non-IPF ILD and IPF was 0.50% (95% CI 0.49% to 0.50%), 0.23% (95% CI 0.22% to 0.23%) and 0.13% (95% CI 0.13% to 0.14%), respectively (figure 1F, online supplemental table 2).

Using primary care records only, we found the incidence of exposure-related ILD decreased from 0.58 (95% CI 0.53 to 0.64) per 10000 person-years in 2005 to 0.37 (95% CI 0.33 to 0.41) per 10000 person-years in 2022. In contrast, the incidence of CTD-ILD increased over the study period from 0.28 (95% CI 0.25 to 0.32) per 10000 person-years to 0.44 (95% CI 0.39 to 0.48) per 10000 person-years (figure 2A). We observed a rapid rise in the use of broad ILD codes during our study period (figure 2A,B). Prevalence of exposure-related ILD, CTD-ILD and pulmonary sarcoidosis increased, but prevalence of drug/radiation induced ILD and other ILD remained stable between 2005 and 2022 (figure 2B). Incidence and prevalence of the individual ILD subtypes using secondary care records only as

well as a combination of primary and secondary care records were higher than in primary care records only but demonstrated similar trends (online supplemental figures 1–4).

Prevalence of all ILD was higher in men and increased in both sexes (online supplemental figure 5). This trend was seen across ILD subtypes, except for CTD-ILD, where prevalence was higher in women (online supplemental table 3). Prevalence of all ILD also increased in every age group over the study period, with the most rapid increases seen in people aged 70 and above (online supplemental figure 6). This trend was present in most ILD subtypes, except pulmonary sarcoidosis and other ILD where the highest increase in prevalence was in people aged 18–59 years (online supplemental table 4A–I). We found marked regional variation in the prevalence of all ILD, IPF and non-IPF ILD in 2022 (online supplemental figure 7A–C). In 2022, ILD prevalence was highest in Northeast and Northwest England, whereas IPF prevalence was highest in Northwest, Southeast England and non-IPF ILD prevalence was highest in Northwest, Southeast England as well as the West Midlands.

Mortality rates and survival analyses

Higher all-cause mortality rates were seen in people with all ILD, IPF and non-IPF ILD with records in secondary care only and combined primary and secondary records compared with those with records in primary care only (figure 3A–C). All-cause crude mortality rates in people with all ILD, IPF and non-IPF ILD recorded in primary care increased between 2005 and 2022 (figure 3A), but for those with records in secondary care only or combined primary and secondary care, we observed decreasing crude mortality rates over the same period of time (figure 3BC).

Using data from people with records in primary care only, we found that all-cause mortality rates in people with all ILD and non-IPF ILD were not significantly different in 2022 compared with 2005 after controlling for age and sex (table 4). In contrast, all-cause mortality rates in people with IPF were 15% lower in 2022 compared with 2008 (mortality rate ratio 0.85, 95% CI 0.77 to 0.94).

Within the different ILD subtypes, the highest mortality rates throughout the study period were observed in people with IPF while the lowest were in people with pulmonary sarcoidosis

Table 3 Crude incidence rates and IRRs of ILD, IPF and non-IPF ILD between 2005 and 2022 in primary care (CPRD Aurum)

Year	ILD				IPF				Non-IPF-ILD						
	Number of incident cases	Person years	Crude incidence rate per 10000 person-years (95% CI)	Age-standardised incidence rate per 10000 person-years (95% CI)	Adjusted IRR* (95% CI)	Number of incident cases	Person years	Crude incidence rate per 10000 person-years (95% CI)	Age-standardised incidence rate per 10000 person-years (95% CI)	Adjusted IRR* (95% CI)	Number of incident cases	Person years	Crude incidence rate per 10000 person-years (95% CI)	Age-standardised incidence rate per 10000 person-years (95% CI)	Adjusted IRR* (95% CI)
2005	2061	7 702 707	2.68 (2.56–2.79)	2.77 (2.65–2.89)	Ref	1059	7 727 391	1.37 (1.29 to 1.46)	1.43 (1.33–1.51)	Ref	867	7 732 973	1.12 (1.05–1.20)	1.14 (1.06–1.21)	Ref
2006	2123	7 847 518	2.71 (2.69–2.82)	2.79 (2.67–2.91)	1.02 (0.96 to 1.09)	1104	7 872 621	1.40 (1.32 to 1.49)	1.45 (1.36–1.54)	1.02 (0.94 to 1.11)	837	7 877 950	1.06 (0.99–1.14)	1.08 (1.00–1.15)	1.03 (0.94 to 1.13)
2007	2211	7 997 633	2.76 (2.65–2.88)	2.84 (2.72–2.96)	1.05 (0.99 to 1.11)	1149	8 023 084	1.43 (1.35 to 1.52)	1.48 (1.39–1.56)	1.04 (0.96 to 1.14)	856	8 028 059	1.07 (1.00–1.14)	1.08 (1.01–1.15)	1.05 (0.95 to 1.16)
2008	2273	8 153 336	2.79 (2.67–2.90)	2.86 (2.74–2.98)	1.06 (1.00 to 1.13)	1147	8 179 015	1.40 (1.32 to 1.49)	1.45 (1.37–1.54)	1.06 (0.98 to 1.15)	891	8 283 667	1.09 (1.02–1.16)	1.09 (1.02–1.17)	1.07 (0.97 to 1.17)
2009	2515	8 260 388	3.04 (2.93–3.17)	3.11 (2.99–3.23)	1.08 (1.02 to 1.15)	1297	8 286 084	1.56 (1.48 to 1.65)	1.61 (1.52–1.70)	1.08 (1.00 to 1.17)	931	8 290 178	1.12 (1.05–1.20)	1.14 (1.05–1.20)	1.09 (0.99 to 1.19)
2010	2499	8 386 831	2.92 (2.81–3.04)	3.00 (2.89–3.12)	1.10 (1.04 to 1.17)	1269	8 412 559	1.51 (1.43 to 1.59)	1.55 (1.47–1.64)	1.10 (1.01 to 1.19)	845	8 416 251	1.00 (0.94–1.07)	1.01 (0.94–1.08)	1.11 (1.01 to 1.22)
2011	2690	8 511 375	3.16 (3.04–3.28)	3.23 (3.11–3.36)	1.12 (1.05 to 1.18)	1377	8 537 011	1.61 (1.53 to 1.70)	1.66 (1.57–1.75)	1.12 (1.03 to 1.21)	995	8 540 207	1.17 (1.09–1.24)	1.18 (1.09–1.24)	1.12 (1.02 to 1.23)
2012	3055	8 673 536	3.52 (3.40–3.65)	3.62 (3.49–3.75)	1.15 (1.09 to 1.21)	1618	8 699 151	1.86 (1.77 to 1.95)	1.91 (1.82–2.01)	1.15 (1.06 to 1.24)	1057	8 701 689	1.21 (1.14–1.29)	1.22 (1.15–1.30)	1.15 (1.06 to 1.26)
2013	3225	8 727 837	3.70 (3.57–3.82)	3.75 (3.62–3.89)	1.15 (1.09 to 1.21)	1645	8 753 092	1.88 (1.79 to 1.97)	1.91 (1.82–2.01)	1.15 (1.06 to 1.24)	1145	8 755 099	1.31 (1.23–1.39)	1.31 (1.23–1.39)	1.16 (1.06 to 1.26)
2014	3267	8 863 773	3.69 (3.56–3.81)	3.73 (3.60–3.86)	1.18 (1.12 to 1.25)	1663	8 888 854	1.87 (1.78 to 1.96)	1.89 (1.80–1.99)	1.19 (1.10 to 1.28)	1190	8 890 254	1.34 (1.26–1.42)	1.32 (1.24–1.40)	1.19 (1.09 to 1.30)
2015	3514	9 044 091	3.89 (3.75–4.01)	3.94 (3.81–4.07)	1.21 (1.14 to 1.27)	1738	9 068 793	1.92 (1.83 to 2.01)	1.93 (1.84–2.03)	1.21 (1.12 to 1.30)	1252	9 069 611	1.38 (1.31–1.46)	1.37 (1.30–1.45)	1.22 (1.11 to 1.33)
2016	3786	9 261 714	4.09 (3.96–4.22)	4.12 (3.99–4.25)	1.24 (1.18 to 1.31)	1818	9 286 067	1.96 (1.87 to 2.05)	1.97 (1.88–2.06)	1.24 (1.15 to 1.34)	1315	9 286 184	1.42 (1.34–1.49)	1.39 (1.32–1.47)	1.25 (1.14 to 1.36)
2017	3978	9 433 871	4.22 (4.09–4.35)	4.21 (4.08–4.34)	1.28 (1.21 to 1.35)	1875	9 457 548	1.98 (1.89 to 2.07)	1.98 (1.89–2.07)	1.28 (1.19 to 1.38)	1420	9 457 097	1.50 (1.42–1.58)	1.47 (1.40–1.55)	1.28 (1.17 to 1.39)
2018	4127	9 604 229	4.30 (4.17–4.43)	4.28 (4.14–4.41)	1.30 (1.23 to 1.37)	1932	9 627 191	2.01 (1.92 to 2.10)	1.99 (1.90–2.08)	1.31 (1.21 to 1.41)	1419	9 635 993	1.47 (1.40–1.55)	1.44 (1.37–1.52)	1.29 (1.19 to 1.41)
2019	4256	9 671 344	4.40 (4.27–4.53)	4.33 (4.20–4.46)	1.31 (1.24 to 1.38)	1935	9 693 177	2.00 (1.91 to 2.09)	1.95 (1.86–2.04)	1.32 (1.22 to 1.42)	1457	9 691 366	1.50 (1.43–1.58)	1.46 (1.38–1.53)	1.31 (1.20 to 1.42)
2020	3494	9 780 749	3.57 (3.45–3.69)	3.58 (3.46–3.70)	1.32 (1.25 to 1.40)	1540	9 801 548	1.57 (1.49 to 1.65)	1.56 (1.49–1.64)	1.33 (1.23 to 1.44)	1170	9 699 210	1.19 (1.13–1.26)	1.17 (1.11–1.24)	1.32 (1.21 to 1.44)
2021	4165	9 931 348	4.19 (4.07–4.32)	4.18 (4.05–4.31)	1.34 (1.27 to 1.41)	1873	9 950 503	1.88 (1.80 to 1.97)	1.86 (1.78–1.95)	1.34 (1.24 to 1.44)	1291	9 947 476	1.30 (1.23–1.37)	1.26 (1.19–1.33)	1.34 (1.23 to 1.46)
2022	4444	9 536 538	4.66 (4.52–4.80)	4.61 (4.48–4.75)	1.30 (1.24 to 1.37)	1877	9 553 387	1.96 (1.88 to 2.06)	1.93 (1.84–2.01)	1.30 (1.21 to 1.40)	1193	9 549 927	1.25 (1.18–1.32)	1.22 (1.15–1.29)	1.30 (1.19 to 1.42)

*IRRs adjusted for age and sex.

ILD, Interstitial lung disease; IPF, idiopathic pulmonary fibrosis; IRR, incidence rate ratio.

*IRRs adjusted for age and sex.

CPRD, Clinical Practice Research Datalink; ILD, interstitial lung disease; IPF, idiopathic pulmonary fibrosis; IRR, incidence rate ratio.

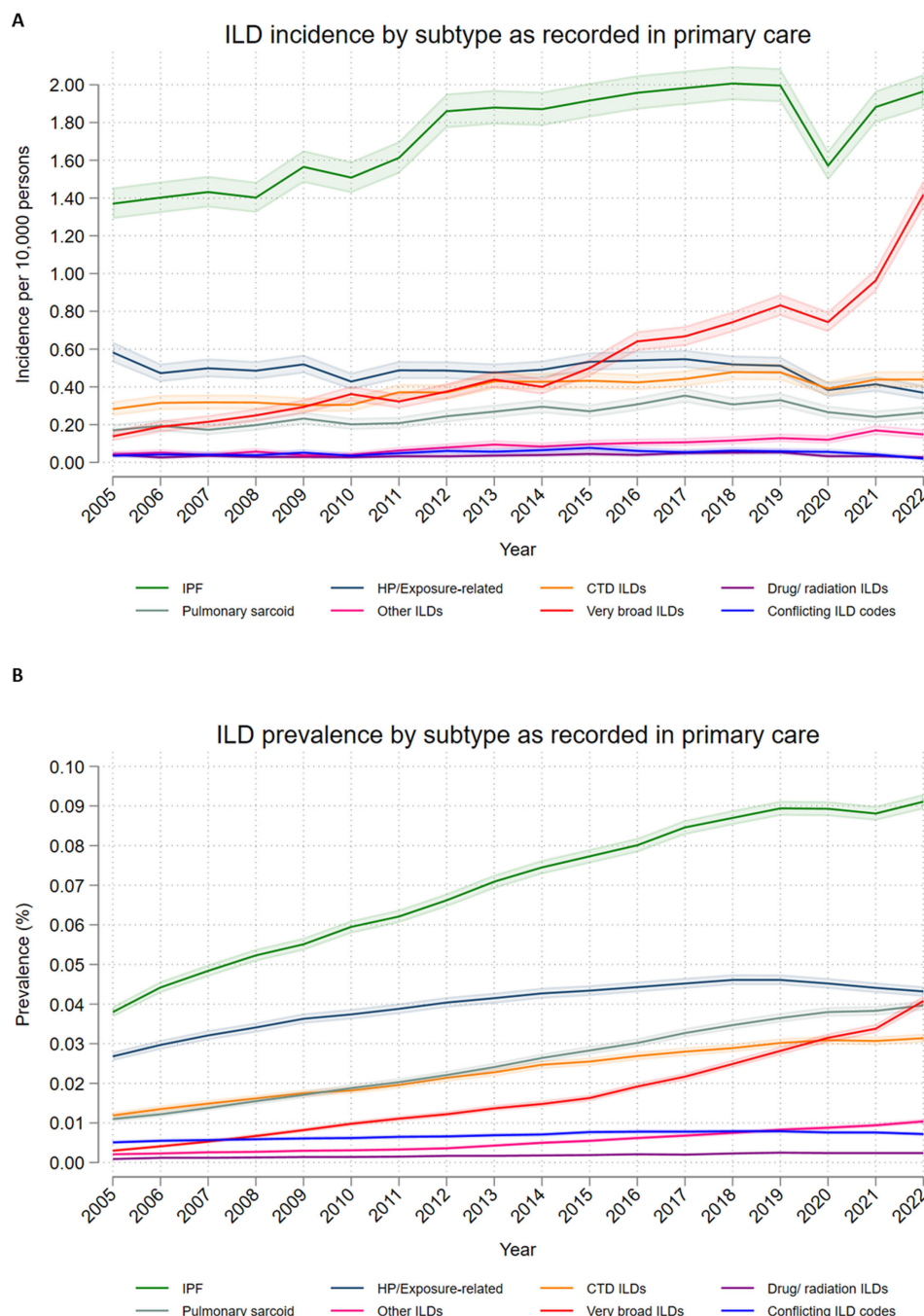


Figure 2 Incidence and prevalence of ILD subtypes using primary care data. CTD, connective tissue disease; HP, hypersensitivity pneumonitis; ILD, interstitial lung disease; IPF, idiopathic pulmonary fibrosis.

(figure 4A). Median survival of our overall ILD cohort was 5.42 years (IQR 1.88–13.03), with people with IPF having a shorter median survival (3.61 years, IQR 1.28–6.39) compared with those with non-IPF ILD (4.88 years, IQR 1.47–8.21) (figure 4B). Across the ILD subtypes, people with drug/radiation induced ILD had the worst prognosis, with a median survival of 3.44 years (IQR 0.92–11.33).

DISCUSSION

Using three algorithms across linked UK primary and secondary care electronic health records, we found that the incidence and prevalence of all ILD, IPF and non-IPF ILD increased between 2005 and 2022. The lowest estimates of incidence and prevalence

occur when using primary care records only. The incidence of all ILD, IPF and non-IPF ILD in primary care records were all 30% higher in 2022 compared with 2005, with a plateau in incidence from 2018 onwards. Among the non-IPF ILD subtypes, exposure-related ILD, CTD-ILD and pulmonary sarcoidosis were most common, with an increasing number of people with a record of a broad ILD code only. Prevalence of all ILD, non-IPF ILD and IPF were highest in men, older people and the North of England.

Lower mortality rates were observed when using all ILD, IPF and non-IPF ILD cases recorded only in primary care. While mortality rates from our all ILD and non-IPF ILD cohorts with only records in primary care were stable over the study period,

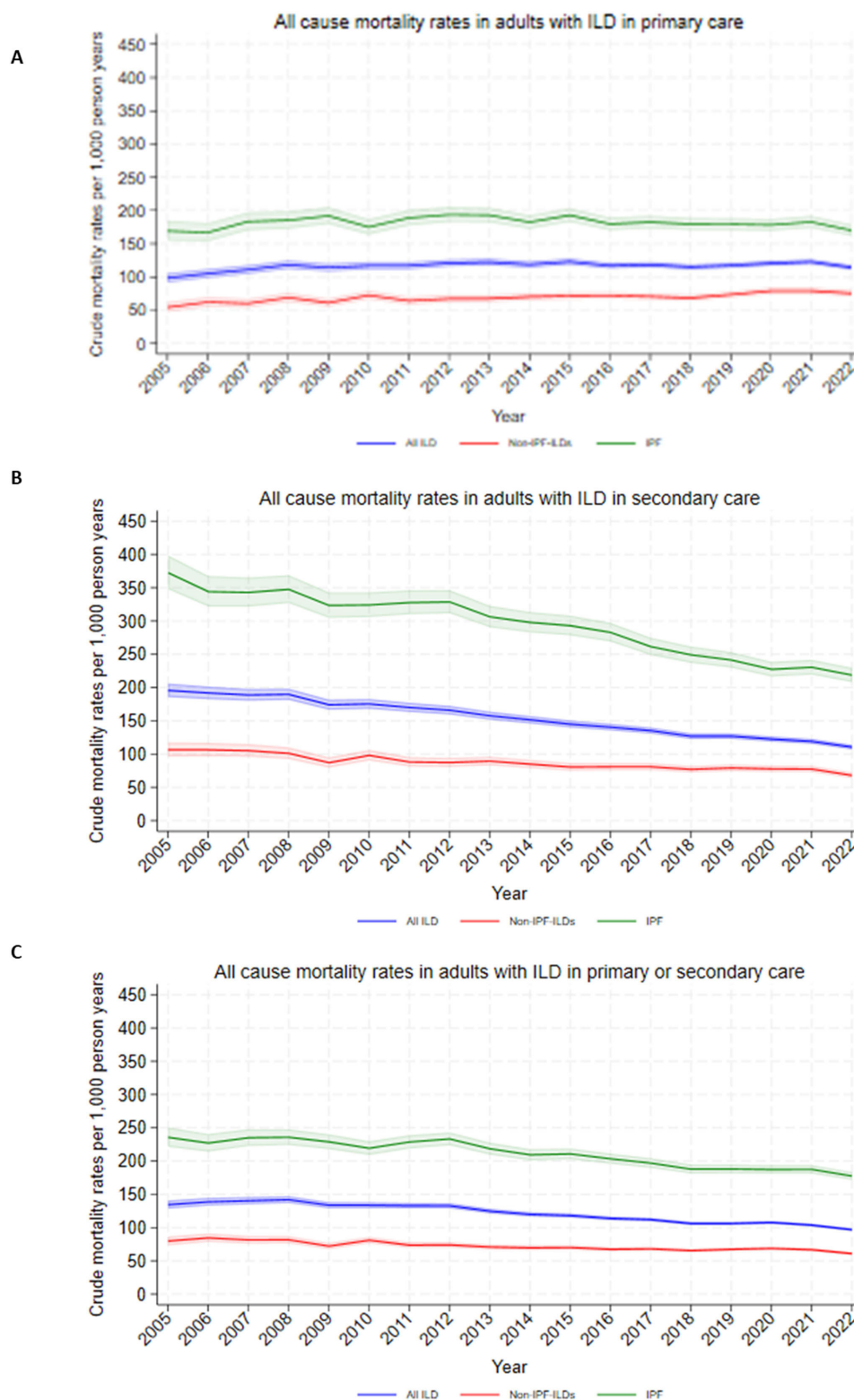


Figure 3 All-cause mortality rates for all ILD, IPF and non-IPF ILD for all three algorithms. ILD, interstitial lung disease; IPF, idiopathic pulmonary fibrosis.

all ILD and non-IPF ILD cohorts with records in secondary care only or both primary and secondary care demonstrated falling mortality rates. Overall median survival for all ILD was poor, with those with IPF and drug/radiation induced ILD having the worst prognoses.

There are still limited data on incidence and prevalence of ILD globally, as demonstrated by our recent systematic review.⁶

Published estimates of ILD incidence vary widely with the lowest estimates from Hong Kong (2.57 per 100 000 person-years in 2020)⁷ and the highest from Canada (98.7 per 100 000 person-years in 2015).⁴ A UK study using electronic primary care records only reported lower ILD incidence rates compared with ours, but only included people with IPF, CTD-ILD and HP.³ We also found that the overall ILD prevalence increased by 6–7% over

Table 4 Crude and adjusted MRRs in people with ILD, IPF and non-IPF ILD between 2005 and 2022 using primary care data (CPRD Aurum)

Year	ILD			IPF			Non-IPF-ILD		
	Number of deaths	Crude MRR (95% CI)	Adjusted MRR* (95% CI)	Number of deaths	Crude MRR (95% CI)	Adjusted MRR* (95% CI)	Number of deaths	Crude MRR (95% CI)	Adjusted MRR* (95% CI)
2005	753	Ref	Ref	499	Ref	Ref	212	Ref	Ref
2006	936	1.07 (0.97 to 1.18)	1.03 (0.94 to 1.13)	579	0.99 (0.88 to 1.11)	0.98 (0.86 to 1.09)	279	1.15 (0.96 to 1.37)	1.09 (0.92 to 1.30)
2007	1112	1.13 (1.03 to 1.24)	1.06 (0.97 to 1.17)	715	1.08 (0.97 to 1.21)	1.05 (0.94 to 1.18)	302	1.11 (0.93 to 1.32)	1.03 (0.86 to 1.22)
2008	1307	1.20 (1.10 to 1.31)	1.10 (1.01 to 1.21)	794	1.09 (0.98 to 1.22)	1.03 (0.92 to 1.16)	380	1.27 (1.07 to 1.50)	1.16 (0.98 to 1.37)
2009	1377	1.16 (1.06 to 1.27)	1.05 (0.96 to 1.15)	886	1.13 (1.02 to 1.27)	1.06 (0.95 to 1.18)	368	1.13 (0.95 to 1.33)	1.02 (0.86 to 1.20)
2010	1528	1.19 (1.09 to 1.29)	1.06 (0.97 to 1.15)	879	1.03 (0.93 to 1.15)	0.95 (0.85 to 1.06)	467	1.34 (1.14 to 1.57)	1.19 (1.02 to 1.41)
2011	1637	1.19 (1.09 to 1.29)	1.05 (0.96 to 1.14)	1005	1.12 (1.00 to 1.24)	1.01 (0.91 to 1.12)	442	1.18 (1.00 to 1.39)	1.04 (0.89 to 1.23)
2012	1841	1.23 (1.13 to 1.34)	1.07 (0.98 to 1.16)	1120	1.14 (1.03 to 1.27)	1.02 (0.92 to 1.14)	503	1.24 (1.06 to 1.46)	1.09 (0.93 to 1.28)
2013	1999	1.24 (1.14 to 1.35)	1.08 (0.99 to 1.17)	1199	1.14 (1.03 to 1.27)	1.01 (0.91 to 1.12)	536	1.24 (1.06 to 1.46)	1.09 (0.93 to 1.28)
2014	2089	1.21 (1.11 to 1.31)	1.04 (0.96 to 1.13)	1214	1.08 (0.97 to 1.20)	0.95 (0.86 to 1.06)	600	1.29 (1.11 to 1.51)	1.12 (0.96 to 1.31)
2015	2308	1.25 (1.15 to 1.36)	1.07 (0.99 to 1.16)	1357	1.14 (1.03 to 1.26)	1.00 (0.90 to 1.11)	647	1.32 (1.13 to 1.54)	1.14 (0.97 to 1.33)
2016	2369	1.19 (1.09 to 1.29)	1.00 (0.92 to 1.08)	1348	1.06 (0.96 to 1.18)	0.91 (0.82 to 1.01)	695	1.33 (1.14 to 1.55)	1.12 (0.96 to 1.31)
2017	2572	1.20 (1.11 to 1.30)	1.00 (0.92 to 1.09)	1460	1.08 (0.97 to 1.19)	0.92 (0.83 to 1.02)	723	1.30 (1.12 to 1.52)	1.09 (0.94 to 1.28)
2018	2666	1.17 (1.08 to 1.27)	0.96 (0.89 to 1.04)	1511	1.06 (0.96 to 1.17)	0.90 (0.81 to 0.99)	743	1.26 (1.08 to 1.47)	1.04 (0.90 to 1.22)
2019	2849	1.19 (1.10 to 1.29)	0.97 (0.89 to 1.05)	1561	1.06 (0.96 to 1.17)	0.89 (0.80 to 0.98)	826	1.35 (1.16 to 1.57)	1.10 (0.95 to 1.28)
2020	3021	1.23 (1.13 to 1.33)	1.00 (0.92 to 1.08)	1572	1.05 (0.95 to 1.16)	0.88 (0.80 to 0.98)	908	1.46 (1.25 to 1.69)	1.19 (1.02 to 1.38)
2021	3128	1.25 (1.15 to 1.35)	1.02 (0.94 to 1.10)	1609	1.08 (0.98 to 1.19)	0.91 (0.82 to 1.01)	914	1.46 (1.26 to 1.69)	1.18 (1.02 to 1.37)
2022	2929	1.16 (1.07 to 1.30)	0.94 (0.87 to 1.02)	1491	1.00 (0.91 to 1.11)	0.85 (0.77 to 0.94)	845	1.39 (1.20 to 1.62)	1.11 (0.91 to 1.29)

*MRRs adjusted for age and sex.

CPRD, Clinical Practice Research Datalink; ILD, interstitial lung disease; IPF, idiopathic pulmonary fibrosis; MRR, mortality rate ratio.

almost 20 years, driven by increases in both IPF and non-IPF ILD prevalence which is consistent with that of other studies.^{4 8} These estimates are smaller compared with increases in prevalence of ILD from Canada⁴ and Hong Kong.⁷

We found an overall increase in the incidence of all ILD, IPF and non-IPF ILD between 2008 and 2022 but rates plateaued towards the end of our study period. Increased case ascertainment through greater access to high resolution CT scanning or greater disease awareness resulted in some of the increase in incidence rates. Other possible explanations for the overall observed trends include standardisation of ILD diagnostic practices through commissioning of ILD specialist centres or that the level of case ascertainment of ILD at the end of our study period is close to being complete, resulting in the plateau in incidence rates. We also observed a sharp decline in incidence rates during the COVID pandemic, which can be attributed to disruption in access to healthcare and diagnostics during the pandemic period. Similar trends have been observed in other chronic conditions, including malignancies.⁹

Previous studies on the incidence of HP report stable incidence rates,^{3 10 11} which are in contrast with our findings. It is possible that the decrease in incidence seen over our study period is related to changes in HP diagnostic guidelines^{12 13} or increased MDT involvement at diagnosis,¹⁴ resulting in a shift of diagnoses to other non-IPF ILD subtypes.

Ours is the first study to report regional variations in ILD, IPF and non-IPF ILD prevalence. The regional pattern in IPF prevalence which we observed is similar to those previously reported for IPF incidence.¹⁵ It is not possible to determine if the variations observed are due to aetiological importance or arise from differences in healthcare practice and/or access and further research is required before any firm conclusions can be drawn.

We had mixed findings in mortality trends across our three algorithms—with increasing mortality rates in all ILD recorded in primary care only but falling rates in all ILD recorded in secondary care only or combined records. One explanation for this is earlier diagnosis, and fewer cases diagnosed at advanced stage as an inpatient. Within our cohorts with primary care records only, we demonstrated that all-cause mortality rates in 2022 were similar compared with 2005 for people with overall ILD and non-IPF ILD. However, we found that all-cause mortality rates in people with IPF were lower in 2022 compared with 2005. The reduction in all-cause mortality rates in people with IPF in this study occurs from 2018 onwards. This is in contrast to our previous study that showed increasing IPF-specific mortality rates between 2008 and 2018.² Possible explanations in the differences in mortality estimates include the different definitions of mortality and the different starting study populations between the studies. It is also possible that improved care for people with IPF following the introduction of ILD specialist centres and greater access to antifibrotics have contributed to the falling mortality rates. Pirfenidone and Nintedanib were approved by the National Institute for Health and Care Excellence in 2013 and 2016 respectively which may partly explain the improved survival seen in the latter part of our study period.

Due to differences in methodology and follow-up time, there are few comparable studies on mortality rates or median survival within ILD subtypes. People with IPF had a median survival of 3.6 years, which is marginally improved from previous studies using UK primary care data.^{15 16} A recent study from the USA using Medicare beneficiaries' data estimated median survival of 3.9 years for people with IPF, which is similar to ours.¹⁷ Johansson *et al* also found improved 1 year mortality in people

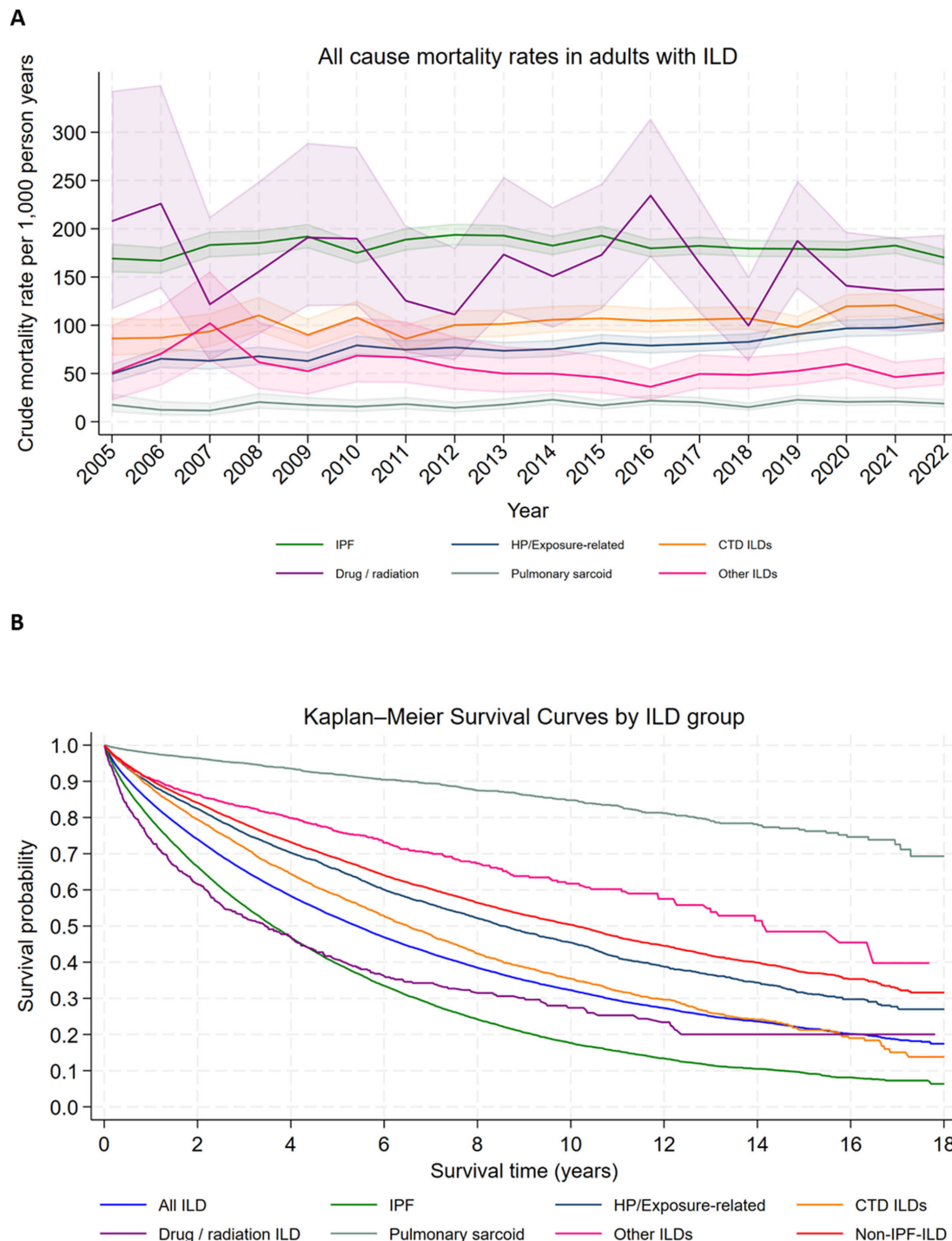


Figure 4 All-cause mortality rates and Kaplan-Meier survival graphs of ILD subtypes using primary care data. CTD, connective tissue disease; HP, hypersensitivity pneumonitis; ILD, interstitial lung disease; IPF, idiopathic pulmonary fibrosis.

with IPF in 2018 compared with 2011, suggesting a global trend of improved survival in people with IPF.⁴ Data on estimated median survival in non-IPF ILD subtypes are sparse, outside of tertiary centre cohorts. Two previous studies using UK primary care data reported a similar median survival in people with CTD-ILD to ours.^{3,18} A study from the USA found that median survival in people with fibrotic HP was 7.2 years, which is similar to our estimates.¹⁹ Our drug and radiation induced ILD cohort had the worst median survival. It is possible that this is related to concurrent comorbidities, particularly that of malignancy. A

similar poor survival in these people was observed by a study by Amar *et al.*²⁰

The main strength of our study is the use of linked primary and secondary care health records and the use of different case-finding algorithms to define a large population-based cohort of all ILD, including non-IPF ILD subtypes. To our knowledge, this is the first study to report UK population-level estimates of incidence, prevalence and mortality rates of overall ILD, non-IPF ILD and individual ILD subtypes over almost 20 years. The large number of ILD cases included in our study as well as the

long study period means we have enabled precise estimates on incidence, prevalence and mortality rates stratified by age, sex, region as well as investigate trends over time. Studies of this nature and magnitude are not possible using cohorts from secondary or tertiary care centres or from disease registries.

A potential weakness of our study is the validity of the diagnosis of ILD, especially within the non-IPF ILD subtypes. We used routinely collected clinical information for this study; it does not include standardised diagnostic criteria for diagnosing ILD. Multidisciplinary (MDT) discussion for the diagnosis and management for ILD was recommended standard practice in the UK early in our study period.¹⁴ Although there are no published national data on the proportion of ILD cases discussed at MDT, a global study on characterisation of ILD MDT meetings found that a median of 75% of new ILD cases were discussed at formal MDT meetings.²¹ Therefore, it is likely that most people with a record of ILD in primary or secondary care would have been discussed at MDT. We have previously demonstrated the validity of IPF diagnoses in UK routinely collected healthcare records using a similar code-based algorithm approach.²² Although we have not validated the non-IPF ILD codes used for this study, they were reviewed independently by two ILD experts and a similar case-finding approach to that of our validated IPF codes²² was used. Given the challenges of obtaining non-IPF ILD diagnoses and changes to diagnostic criteria,^{13 23} misclassification of individual ILD subtypes is still possible. However, demographic and mortality data in our cohorts defined using both primary and secondary care records are similar to other ILD cohorts,^{24 25} providing further validation of our case-finding algorithms.

To accurately estimate the burden of all ILD, we initially included people with a broad ILD code in our initial study population. It is unlikely that patients with a broad ILD would be coded as having ILD by their GP without confirmation from a respiratory specialist or secondary care, as demonstrated by a recent qualitative study.²⁶ Therefore, excluding people with a broad ILD code only will lead to an underestimate in ILD disease burden. When we stratified our cohort into IPF and non-IPF ILD, we then excluded people with a broad ILD code as well as those with multiple ILD subtype codes. This approach may have increased the specificity but would have reduced the sensitivity of the individual ILD subtype estimates.

We found changes in ILD coding practice in primary care, with increasing use of broad ILD codes, particularly from 2016 onwards. The cause for this shift in coding practice remains unclear but has also been seen in IPF studies using UK primary care data.^{2 3} One hypothesis is that changes to code use frequencies may be related to changes in ILD nomenclature and diagnostic criteria over time, resulting in GPs opting to use a more generic or broad ILD code.

A further limitation of our study is the use of all-cause mortality rather than ILD-specific mortality when estimating ILD mortality burden. Although up to 80% of people with ILD die from their underlying disease,²⁷ it is possible that we have overestimated the mortality burden of ILD by using all-cause mortality rates.

In summary, our population-based study demonstrates the growing burden of ILD. Our findings suggest that at present, over 200 000 people in the UK are living with ILD and between 30 000 and 60 000 new cases of ILD are diagnosed each year. Despite changes in how care is delivered to people with ILD and the introduction of new therapies, survival is worse than most malignancies.²⁸ Our estimates suggest ILD incidence and prevalence no longer meets criteria for classification as a rare disease and should be recognised as an important public health problem.

Contributors GMM, ADM, PMG, JKQ and VN developed the research questions. GMM and ADM contributed to conceptualisation, study design, data management, statistical analyses and manuscript review. VN and JKQ contributed to conceptualisation, review of study design and manuscript writing. PMG supported with the understanding of ILD and manuscript review. All authors interpreted and reviewed the results. VN drafted the manuscript. All authors read, commented on and approved the manuscript. JKQ is the guarantor of the paper.

Funding This work was funded by an IIS from Boehringer Ingelheim, grant number 1199-0566. VN is funded by Future Health Research Innovation Fund WA Emerging Leaders Near Miss Award (WANMAEL2025-26/5).

Map disclaimer The depiction of boundaries on this map does not imply the expression of any opinion whatsoever on the part of BMJ (or any member of its group) concerning the legal status of any country, territory, jurisdiction or area or of its authorities. This map is provided without any warranty of any kind, either express or implied.

Competing interests GMM and ADM have no competing interests. PMG has received investigator-led grants from Boehringer Ingelheim; consulting fees from Boehringer Ingelheim, AstraZeneca, Avalyn Pharma, Daiichi-Sankyo, Brainomix; honoraria from Boehringer Ingelheim, AstraZeneca, GSK; has stock options in Brainomix. JKQ received funding for this work via an IIS from Boehringer Ingelheim (grant number 1199-0566); has received grants from Medical Research Council, NIHR, Health Data Research, GSK, BI, AZ, Insmed, Sanofi; honoraria from GlaxoSmithKline, BI, Sanofi, Chiesi, AstraZeneca; is joint EIC of Thorax. VN received funding for this work from the WA Emerging Leaders Near Miss Awards; has received consulting fees, support for attending meetings and honoraria from Boehringer Ingelheim and Johnson & Johnson.

Patient consent for publication Not applicable.

Ethics approval CPRD has NHS Health Research Authority (HRA) Research Ethics Committee (REC) approval to allow the collection and release of anonymised primary care data for observational research (NHS HRA REC reference number: 05/MRE04/87), this approval was granted on the 11/02/2025. Each year CPRD obtains Section 251 regulatory support through the HRA Confidentiality Advisory Group (CAG), to enable patient identifiers, without accompanying clinical data, to flow from CPRD contributing GP practices in England to NHS Digital, for the purposes of data linkage (CAG reference number: 21/CAG/0008). The protocol for this research was approved by CPRD's Research Data Governance (RDG) Process (protocol number: 25_004889) and the approved protocol is available upon request. Linked pseudonymised data was provided for this study by CPRD. Data is linked by NHS Digital, the statutory trusted third party for linking data, using identifiable data held only by NHS Digital. Select general practices consent to this process at a practice level with individual patients having the right to opt-out. This study is based in part on data from the Clinical Practice Research Datalink obtained under licence from the UK Medicines and Healthcare products Regulatory Agency. The data is provided by patients and collected by the NHS as part of their care and support. The Office for National Statistics (ONS) was the provider of the ONS Data contained within the CPRD Data and maintains a Copyright (2024). The Hospital Episode Statistics (HES) was the provider of HES-Admitted Patient Care database contained within the CPRD Data and maintains a Copyright (2024). Linked data were reused with the permission of The Health and Social Care Information Centre, all rights reserved. The interpretation and conclusions contained in this study are those of the author/s alone.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data may be obtained from a third party and are not publicly available. Data are available on request from the CPRD. Their provision requires the purchase of a licence, and this licence does not permit the authors to make them publicly available to all. This work used data from the version collected in December 2023 and has clearly specified the data selected within each Methods section. To allow identical data to be obtained by others, via the purchase of a licence, the code lists will be provided upon request. Licences are available from the CPRD (<http://www.cprd.com>): The Clinical Practice Research Datalink Group, The Medicines and Healthcare products Regulatory Agency, 10 South Colonnade, Canary Wharf, London E14 4PU.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution 4.0 Unported (CC BY 4.0) license, which permits others to copy, redistribute, remix, transform and build upon this work for any

purpose, provided the original work is properly cited, a link to the licence is given, and indication of whether changes were made. See: <https://creativecommons.org/licenses/by/4.0/>.

ORCID iDs

Georgie May Massen <https://orcid.org/0000-0003-4355-6546>

Ann Dorothy Morgan <https://orcid.org/0000-0003-4794-0860>

Peter M George <https://orcid.org/0000-0003-1316-4891>

Jennifer K Quint <https://orcid.org/0000-0003-0149-4869>

Vidya Navaratnam <https://orcid.org/0000-0002-2914-8012>

REFERENCES

- Kornum JB, Christensen S, Grijota M, *et al.* The incidence of interstitial lung disease 1995-2005: a Danish nationwide population-based study. *BMC Pulm Med* 2008;8:24.
- Gupta R, Morgan AD, George PM, *et al.* Incidence, prevalence and mortality of idiopathic pulmonary fibrosis in England from 2008 to 2018: a cohort study. *Thorax* 2024;79:624-31.
- Gonnelli F, Eleangovan N, Smith U, *et al.* Incidence and survival of interstitial lung diseases in the UK in 2010-2019. *ERJ Open Res* 2025;11.
- Johannson KA, Liu Z, Southern DA, *et al.* Incidence, Prevalence, and Mortality of Interstitial Lung Diseases in Alberta, Canada: A Population-based Study. *Ann Am Thorac Soc* 2025;22:1009-17.
- Wolf A, Dedman D, Campbell J, *et al.* Data resource profile: Clinical Practice Research Datalink (CPRD) Aurum. *Int J Epidemiol* 2019;48:1740-1740g.
- Shah Gupta R, Koteci A, Morgan A, *et al.* Incidence and prevalence of interstitial lung diseases worldwide: a systematic literature review. *BMJ Open Respir Res* 2023;10:e001291.
- Ye Y, Sing C-W, Hubbard R, *et al.* Prevalence, incidence, and survival analysis of interstitial lung diseases in Hong Kong: a 16-year population-based cohort study. *Lancet Reg Health West Pac* 2024;42:100871.
- Maher TM, Bendstrup E, Dron L, *et al.* Global incidence and prevalence of idiopathic pulmonary fibrosis. *Respir Res* 2021;22:197.
- Howlader N, Bhattacharya M, Scoppa S, *et al.* Cancer and COVID-19: US cancer incidence rates during the first year of the pandemic. *J Natl Cancer Inst* 2024;116:208-15.
- Fernández Pérez ER, Kong AM, Raimundo K, *et al.* Epidemiology of Hypersensitivity Pneumonitis among an Insured Population in the United States: A Claims-based Cohort Analysis. *Ann Am Thorac Soc* 2018;15:460-9.
- Solaymani-Dodaran M, West J, Smith C, *et al.* Extrinsic allergic alveolitis: incidence and mortality in the general population. *QJM* 2007;100:233-7.
- Raghu G, Remy-Jardin M, Myers JL, *et al.* Diagnosis of Idiopathic Pulmonary Fibrosis. An Official ATS/ERS/JRS/ALAT Clinical Practice Guideline. *Am J Respir Crit Care Med* 2018;198:e44-68.
- Fernández Pérez ER, Travis WD, Lynch DA, *et al.* Diagnosis and Evaluation of Hypersensitivity Pneumonitis: CHEST Guideline and Expert Panel Report. *Chest* 2021;160:e97-156.
- Bradley B, Branley HM, Egan JJ, *et al.* Interstitial lung disease guideline: the British Thoracic Society in collaboration with the Thoracic Society of Australia and New Zealand and the Irish Thoracic Society. *Thorax* 2008;63 Suppl 5:v1-58.
- Navaratnam V, Fleming KM, West J, *et al.* The rising incidence of idiopathic pulmonary fibrosis in the U.K. *Thorax* 2011;66:462-7.
- Strongman H, Kausar I, Maher TM. Incidence, Prevalence, and Survival of Patients with Idiopathic Pulmonary Fibrosis in the UK. *Adv Ther* 2018;35:724-36.
- Raghu G, Chen S-Y, Yeh W-S, *et al.* Idiopathic pulmonary fibrosis in US Medicare beneficiaries aged 65 years and older: incidence, prevalence, and survival, 2001-11. *Lancet Respir Med* 2014;2:566-72.
- Navaratnam V, Ali N, Smith CJP, *et al.* Does the presence of connective tissue disease modify survival in patients with pulmonary fibrosis? *Respir Med* 2011;105:1925-30.
- Vourlekis JS, Schwarz MI, Cherniack RM, *et al.* The effect of pulmonary fibrosis on survival in patients with hypersensitivity pneumonitis. *Am J Med* 2004;116:662-8.
- Amar RK, Jick SS, Rosenberg D, *et al.* Drug-/radiation-induced interstitial lung disease in the United Kingdom general population: incidence, all-cause mortality and characteristics at diagnosis. *Respirology* 2012;17:861-8.
- Richeldi L, Launders N, Martinez F, *et al.* The characterisation of interstitial lung disease multidisciplinary team meetings: a global study. *ERJ Open Res* 2019;5:00209-2018.
- Morgan A, Gupta RS, George PM, *et al.* Validation of the recording of idiopathic pulmonary fibrosis in routinely collected electronic healthcare records in England. *BMC Pulm Med* 2023;23:256.
- Raghu G, Remy-Jardin M, Ryerson CJ, *et al.* Diagnosis of Hypersensitivity Pneumonitis in Adults. An Official ATS/JRS/ALAT Clinical Practice Guideline. *Am J Respir Crit Care Med* 2020;202:e36-69.
- Fisher JH, Kolb M, Algamdi M, *et al.* Baseline characteristics and comorbidities in the Canadian Registry for Pulmonary Fibrosis. *BMC Pulm Med* 2019;19:223.
- Khan F, Stewart I, Howard L, *et al.* Comprehensive characterisation of individuals with fibrotic interstitial lung disease: baseline insights from the INJUSTIS study. *BMJ Open Respir Res* 2026;13:e003112.
- Tulloch JS, Beadsworth MB, Vivancos R, *et al.* GP coding behaviour for non-specific clinical presentations: a pilot study. *BJGP Open* 2020;4.
- Salonen J, Purokivi M, Bloigu R, *et al.* Prognosis and causes of death of patients with acute exacerbation of fibrosing interstitial lung diseases. *BMJ Open Respir Res* 2020;7:e000563.
- UK NT. How does five-year cancer survival in England vary by cancer type?, 2024. Available: <https://www.nuffieldtrust.org.uk/resource/cancer-survival-rates> [Accessed 25 Oct 2025].