Prevalence and Incidence of Chronic Fibrosing Interstitial Lung Diseases with a Progressive Phenotype in the United States Estimated in a Large Claims Database Analysis

Amy L. Olson · Padmaja Patnaik · Nadine Hartmann ·

Rhonda L. Bohn · Elizabeth M. Garry · Laura Wallace

ABSTRACT

Introduction: Many fibrosing interstitial lung diseases (ILDs) develop a chronic progressive phenotype. While idiopathic pulmonary fibrosis, which is always progressive, is well characterized with established treatment options, the epidemiology of other chronic fibrosing ILDs with a progressive phenotype has not been widely investigated. Treatment options are limited, with a high unmet need. This claims database study estimates the incidence and prevalence of these diseases in the USA.

Methods: Diagnosis, procedure and resource utilization codes from insurance claims were used to identify patients with fibrosing ILD and those with a chronic progressive phenotype among 37,565,644 adult patients in the IBM® MarketScan® Research Database 2012–2015. Two eligible ILD claims were required for a fibrosing ILD diagnosis. Progression was defined using a novel algorithm constituted by criteria considered proxies for progression. Patients were defined as having incident (new) or existing diagnoses based on claims during a 365-day period before study entry.

Results: The estimated age- and sex-adjusted prevalence per 100,000 persons of fibrosing ILD (95% confidence interval) was 117.82 (116.56, 119.08) and of chronic fibrosing ILDs with a progressive phenotype was 70.30 (69.32, 71.27). The estimated adjusted incidence per 100,000 patient-years of fibrosing ILD was 51.56 (50.88, 52.24) and of chronic fibrosing ILDs with a progressive phenotype was 32.55 (32.01, 33.09). Among incident fibrosing ILD patients, 57.3% experienced progression over a median of 117 days (interquartile range 63–224), with largely comparable rates of progression among different diseases.

Conclusions: Chronic fibrosing ILDs with a progressive phenotype comprise a relatively new disease construct requiring varied approaches to obtain reliable estimates of prevalence

A. L. Olson (✉)
Interstitial Lung Disease Program, Division of Pulmonary and Critical Care Medicine,
Autoimmune Lung Center, National Jewish Health,
Denver, CO, USA
e-mail: amy.olson@boehringer-ingelheim.com

P. Patnaik
Boehringer Ingelheim Pharmaceuticals, Inc.,
Ridgefield, CT, USA

N. Hartmann · L. Wallace
Boehringer Ingelheim International GmbH,
Ingelheim am Rhein, Germany

R. L. Bohn
Bohn Epidemiology, LLC, Boston, MA, USA

E. M. Garry
Aetion Inc, Boston, MA, USA

and incidence. This is the first large claims database study using real-world data to provide estimates of the prevalence and incidence of these diseases among a very large segment of the US population and could form the groundwork for future studies.

PLAIN LANGUAGE SUMMARY

Progressive lung fibrosis occurs in idiopathic pulmonary fibrosis; however, interstitial lung fibrosis may occur in other diseases such as rheumatoid arthritis, chronic hypersensitivity pneumonitis and sarcoidosis, and may or may not be progressive in these diseases. Little is known about the frequency of lung fibrosis among patients with these diseases or how often such fibrosis is progressive. This study used information from a large insurance claims database (IBM® MarketScan®) to estimate the frequency and progression of lung fibrosis associated with different diseases.

Keywords: Epidemiology; Lung disease; Pulmonary fibrosis; Progression

Key Summary Points
Why carry out this study?
Treatment options for chronic fibrosing ILDs with a progressive phenotype other than idiopathic pulmonary fibrosis are limited, with a high unmet need
The epidemiology of these diseases has not been widely investigated
This claims database study estimates their incidence and prevalence in the USA
What was learned from the study?
The estimated age- and sex-adjusted prevalence per 100,000 persons of fibrosing ILD (95% confidence interval) was 117.82 (116.56, 119.08) and of chronic fibrosing ILDs with a progressive phenotype was 70.30 (69.32, 71.27)
The estimated incidence per 100,000 patient-years of fibrosing ILD was 51.56 (50.88, 52.24) and of chronic fibrosing ILDs with a progressive phenotype was 32.55 (32.01, 33.09)
This is the first such study to provide estimates of prevalence and incidence of these diseases and could form the groundwork for future studies

DIGITAL FEATURES

This article is published with digital features, including a plain language summary, to facilitate understanding of the article. To view digital features for this article go to https://doi.org/10.6084/m9.figshare.14560515.

INTRODUCTION

Interstitial lung disease (ILD) encompasses a large group of disorders that can be characterized by fibrosis of the lungs. Idiopathic pulmonary fibrosis (IPF), which is always progressive, is one of the most common and most studied of the ILDs [1–3]. Other ILDs may or may not have a progressive phenotype and include idiopathic interstitial pneumonias (IIPs), hypersensitivity pneumonitis, exposure-related ILD such as pneumoconiosis, sarcoidosis, and ILDs associated with connective tissue disease and autoimmune disease [1–3].

Chronic fibrosing ILDs other than IPF that have a progressive phenotype are increasingly referred to as 'progressive fibrosing ILD' [2, 4–7]. They appear to share a common pathology, characterized by increasing extent of fibrosis on high-resolution computed tomography (HRCT), declining lung function, worsening symptoms, and early mortality [3]. There is a high medical need for effective treatment. One treatment, nintedanib, which was previously approved to slow lung function decline in SSc-ILD [8, 9], is now approved for treatment of patients with chronic fibrosing ILDs with a progressive phenotype [10].

Given the clinical and pathophysiologic similarities between progressive fibrosing ILDs, it has been postulated that the progressive phenotype involves a common pathologic mechanism regardless of cause, and thus these diseases could respond to similar treatment. Antifibrotic treatment with nintedanib has been evaluated in progressive fibrosing ILD with convincing results [11] that contributed to the approval in the US [10] and in 45 other countries, including Japan [12], Canada [13], and the European Union [14]. However, devising effective strategies for appropriate disease management and allocation of resources for this condition is hampered by limited understanding of its epidemiology.

Rare disease frequencies are inherently difficult to quantify at a population level, and the relatively recent emergence of progressive fibrosing ILD as a disease concept that combines an underlying disease, lung fibrosis, and a progressive phenotype makes it further challenging to quantify. Various approaches to data generation are required to build a complete evidence base of disease frequencies over time, and these estimates need to be fine-tuned until a range of plausible estimates can be achieved. These approaches include smaller clinic-based studies with potentially valid estimates derived from review of complete data directly accessible from patient charts but based on small numbers of patients, and larger claims database studies, which offer the strength of including very large numbers of patients, but do not allow for direct access to all clinically relevant data available in patient charts. Not surprisingly, data on overall prevalence and incidence of progressive fibrosing ILD are limited.

Progressive fibrosing ILD is increasingly being accepted as a disease concept [3, 15–18]. However, publications have focused on aspects of pathogenesis, diagnosis, and disease management, with few data available on prevalence and incidence. Wijsenbeek et al. [15] reported that 18–32% of ILDs other than IPF were estimated to progress, according to physician surveys. The prevalence of progressive fibrosing ILDs has been estimated at up to 28 per 100,000 persons, based on the prevalence of fibrosing ILDs and survey-derived estimates of how many

patients develop a progressive phenotype [19–24].

To complement the existing evidence base, we have conducted the first large database study using medical insurance claims to identify patients with progressive fibrosing ILD using diagnostic and procedure/resource utilization codes in insurance claims and to then estimate the prevalence and incidence of progressive fibrosing ILD in this large population of patients. This approach allows estimation of the overall burden of progressive fibrosing ILD in the US population and provides valuable information on the prevalence and incidence of this group of conditions that has not been well described.

METHODS

Study Design

This study used data from the IBM® MarketScan® Research Databases of US medical and prescription commercial claims to assess the prevalence and incidence of fibrosing ILD and progressive fibrosing ILD, both overall and by different underlying diseases. Codes included on claims for diagnoses (International Classification of Diseases, Ninth Revision, Clinical Modification [ICD-9 CM]) [25] and for procedures (Current Procedural Terminology, 4th Edition [CPT-4]), prescriptions, and resource utilization were used to identify patients with fibrosing ILD and, among this patient group, those with a progressive phenotype. In the absence of existing validated algorithms or previously published data to guide this process, a combination of codes was selected in consultation with practicing pulmonologists to capture both the fibrosing ILD and the progressive phenotype through proxy criteria.

Patients

Patients aged at least 18 years were required to have a period of 365 days with continuous medical and pharmacy insurance coverage (baseline period) before study entry. After entry

to the study, patients were required to maintain coverage up to the time of diagnosis of fibrosing ILD and/or progressive fibrosing ILD. Gaps of up to 30 days in coverage were permitted.

Identification of Fibrosing ILD

Eligible ICD-9 codes for a diagnosis of fibrosing ILD are shown in the supplementary material (Table S1). The code for IPF (516.31) was not eligible, although patients with this code were not excluded if they had another eligible code. Given the lack of precedence on attempting to identify fibrosing ILD patients using codes on claims in a large insurance database, two approaches were considered. Initially, because it was thought that a highly sensitive algorithm might be required to capture the outcome of interest, a single claim with a diagnostic code for lung fibrosis as indicated above was used as a diagnosis of fibrosing ILD. Subsequently, the narrower, more specific definition that required a second claim for one of the fibrosing ILD codes listed above within 30–365 days of the first was used to establish a diagnosis. Since this more specific definition excludes patients who did not survive or remain in the database long enough to satisfy the criteria, the second claim was considered to be the index date for study entry to eliminate the potential for selection bias [26].

For both definitions, incident cases of fibrosing ILD comprised patients without an eligible code during the baseline period prior to the initial claim. All others were considered existing cases. Prevalent cases included both incident and existing cases.

Identification of Progressive Fibrosing ILD

Patients with progressive fibrosing ILD were a subgroup of the fibrosing ILD cohort. In the absence of diagnostic or procedure codes on lung function and imaging methods specific to the assessment of ILD progression in clinical practice, proxies for progression based on plausible markers of progression were selected in consultation with expert pulmonologists based on their clinical experience. Patients were considered to have progressive fibrosing ILD if they satisfied any of the following criteria for progression: ≥ 2 pulmonary function tests or ≥ 2 oxygen titration tests within 90 days, ≥ 2 HRCT or ≥ 3 chest computed tomography (CT) scans within 360 days, respiratory hospitalization, palliative care, lung transplant, any use of oxygen therapy or a corticosteroid > 20 mg, or new use of immunosuppressive therapy. A sensitivity analysis was conducted in which clinical guidance was used to identify proxy criteria that may potentially be too sensitive: first, oxygen titration tests and use of corticosteroids were omitted as criteria for progression; second, respiratory hospitalization was retained, but only if there were two instances within 360 days; and oxygen therapy was retained only if there were two recorded claims during the study period.

Incident progressive fibrosing ILD cases included all patients with proxies for progression on or after the date of diagnosis of incident fibrosing ILD. Patients with existing fibrosing ILD (i.e., with a claim for fibrosing ILD during the baseline period) were also considered to have incident progressive fibrosing ILD if they had a proxy for progression on or after the date of the existing fibrosing ILD claim, without any of the proxies for progression during the baseline period.

Identification of Underlying Disease

An underlying clinical diagnosis category was assigned based on diagnosis and lung fibrosis code(s) identified on claims during the baseline period. Patients without any of the specific diagnosis codes were categorized as having idiopathic non-specific interstitial pneumonia or unclassifiable IIPs based on ICD-9 CM algorithms or as having 'no underlying condition' if none of the algorithms applied. If patients had multiple eligible codes, that nearest to the date of diagnosis was used. Those with eligible codes for different conditions on the same day were categorized as having multiple underlying conditions. Diagnosis and lung fibrosis codes used in categorization of the underlying diagnoses are provided in the supplementary material (Table S2).

Outcomes

The primary outcomes were the crude and age- and sex-adjusted (standardized to the 2014 US Census estimates) prevalence and incidence rates of fibrosing ILD and progressive fibrosing ILD. Secondary outcomes included the age and sex distribution, the prevalence and incidence rates by age and sex and by underlying clinical diagnosis, and the proportion of patients with fibrosing ILD who subsequently progressed, overall and by underlying clinical diagnosis.

Statistical Methods

All analyses were descriptive. Crude prevalences were calculated by dividing the total number of patients with incident and existing diagnosis by the total number of eligible patients in the database. The crude incidence rate was calculated by dividing the number of patients with an incident diagnosis by the sum of follow-up time for all eligible participants. Follow-up time was calculated starting when a patient had completed 365 days of continuous eligibility (baseline period). For calculation of follow-up time, patients were censored on the date of a diagnosis of fibrosing ILD or progressive fibrosing ILD, discontinuation of insurance, death, or the end of the study period.

All analyses were conducted using the Aetion Evidence Platform® (2020), v3.19, software for real-world data analysis, which has been validated for a range of studies [27]. The use of this de-identified data source was approved for exemption by the New England Independent Review Board (NEIRB).

For age- and sex-standardized rates, crude prevalence and incidence rates were calculated for age strata among males and females and weights assigned to each stratum based on the 2014 US Census estimates. The crude estimates for each stratum were multiplied by the corresponding stratum weight and weighted rates summed to provide age- and sex-adjusted rates.

RESULTS

Prevalence and Incidence Rates of Fibrosing ILD and Progressive Fibrosing ILD

A total of 37,565,644 patients aged at least 18 years were identified from 1 October 2012 to 30 September 2015 with at least 365 days' continuous enrollment. Of these, 147,678 (0.39%) had at least one claim for lung fibrosis, yielding a crude prevalence of fibrosing ILD of 393.12 per 100,000 persons (95% confidence interval [CI] 391.12, 395.12), using the sensitive (single-claim) definition. When the specific definition, requiring two claims, was used for diagnosis, 35,825 (0.10%) patients were identified as having fibrosing ILD, yielding a crude prevalence of 95.37 per 100,000 persons (95% CI 94.38, 96.35). When adjusted for age and sex, the prevalence increased to 453.88 per 100,000 persons (95% CI 451.46, 456.29) using the sensitive definition and to 117.82 per 100,000 persons (95% CI 116.56, 119.08) using the specific definition of fibrosing ILD (Table 1). Patients with progressive fibrosing ILD were a subset of the fibrosing ILD cohort, and using two claims to define fibrosing ILD, the crude prevalence of progressive fibrosing ILD was 57.82 per 100,000 persons (95% CI 57.05, 58.58), increasing to 70.30 per 100,000 persons (95% CI 69.32, 71.27) when adjusted for age and sex. Crude and age- and sex-adjusted prevalence and incidence rates using both definitions are summarized in Table 1.

Analyses using the specific, two-claim definition of fibrosing ILD are presented below, while estimates using the less specific, more sensitive, single-claim definition are available in the supplementary material (Tables S3 and S4).

During a median follow-up of 1.25 years, a total of 21,719 patients satisfied the proxy criteria indicating progression, representing 60.1% of all fibrosing ILD patients. Of these, 14,722 (67.8%) were incident cases whereas 6977 (32.2%) were existing cases (i.e., they fulfilled criteria for progression at the time of study entry).

Table 1 Crude and adjusted prevalence and incidence of fibrosing ILD and progressive fibrosing ILD using a sensitive definition (requiring a single eligible claim) or a more specific definition requiring two claims

	Prevalence, per 100,000 persons (95% CI)		Incidence, per 100,000 PY (95% CI)	
	Crude	Adjusted*	Crude	Adjusted*
Fibrosing ILD				
Sensitive: one claim	393.12	453.88	240.60	273.15
	(391.12, 395.12)	(451.46, 456.29)	(239.28, 241.92)	(271.61, 274.69)
Specific: two claims	95.37	117.82	44.14	51.56
	(94.38, 96.35)	(116.56, 119.08)	(43.57, 44.70)	(50.88, 52.24)
Progressive fibrosing ILD				
Sensitive: one claim	158.60	190.94	92.47	108.17
	(157.33, 159.87)	(189.34, 192.54)	(91.65, 93.28)	(107.19, 109.16)
Specific: two claims	57.82	70.30	27.55	32.55
	(57.05, 58.58)	(69.32, 71.27)	(27.11, 28.00)	(32.01, 33.09)

CI confidence interval, *ILD* interstitial lung disease, *PY* patient-years

*Adjusted for age and sex

Most of the incident cases of progression (13,518/14,722, 91.8%) were derived from the incident fibrosing ILD cohort (*n* = 23,577), representing 57.3% of this cohort, with progression seen over a median of 117 days (interquartile range [IQR] 63–224 days) from the first of the two lung fibrosis claims (Table 2). Median time from the date of the second claim (index date) to the first proxy for progression was 12 days (IQR 0–55 days).

Distribution of Underlying Disease Categories in the Study Cohorts

The distribution of underlying diseases was comparable in the fibrosing ILD and progressive fibrosing ILD incident cohorts. Approximately three-quarters of patients were categorized as having unclassifiable IIP. Rheumatoid arthritis-associated ILD (RA-ILD) (7.6% of fibrosing ILD cases and 7.7% of progressive fibrosing ILD) and sarcoidosis (7.3% and 5.4%, respectively) were the next most common categories (Fig. 1). The distribution of patients progressing did not differ markedly according to underlying disease category, with the exception of sarcoidosis, which had the lowest percent of patients who progressed (42.8%) over a median (IQR) of 148 (73–255) days, and hypersensitivity pneumonitis, which had the highest percent (74.2%) over a median (IQR) of 109 (63–232) days (Table 2).

Age and Sex Distribution of Fibrosing ILD and Progressive Fibrosing ILD

The prevalence of fibrosing ILD was slightly lower in males than in females, but the prevalence of progressive fibrosing ILD was similar in males and females (Table 3). For both fibrosing ILD and progressive fibrosing ILD, the crude incidence rate was similar for males and females, and at least two-thirds of patients were aged > 60 years. Incidence increased steadily with age: the incidence rate in patients aged 50–59 years was ten times that in patients aged < 40 years, while in those aged ≥ 80 years the incidence rate was 100 times that in the < 40 years group (Table 4). There was no evidence for a higher occurrence of progression within any age group.

Table 2 Distribution of patients with fibrosing ILD (specific, two-claim definition) who progressed during follow-up, according to underlying disease

	Incident and existing fibrosing ILD			Incident fibrosing ILD only			
	Patients, *n*	Patients with progression at any time, *n**	Percent with progression at any time	Patients, *n*	Patients with incident progression, *n**	Percent with incident progression	Median time to progression, days [IQR]†
All fibrosing ILD	35,825	21,719	60.6	23,577	13,518	57.3	117 [63, 224]
Unclassifiable IIP	24,881	15,515	62.4	17,385	10,241	58.9	115 [62, 219]
iNSIPs	196	150	76.5	106	69	65.1	67 [50, 176]
Sarcoidosis	2522	1160	46.0	1740	744	42.8	148 [73, 255]
RA-ILD	3040	1777	58.5	1798	1018	56.6	128 [67, 235]
SSc-ILD	1066	618	58.0	482	248	51.5	125 [74, 246]
MCTD-ILD	119	62	52.1	38	18	47.4	72 [61, 338]
No underlying diagnosis	313	153	48.9	237	109	46.0	145 [61, 275]
Multiple underlying diagnoses	404	240	59.4	204	119	58.3	111 [58, 219]
Exposure-related ILD	692	481	69.5	419	284	67.8	133 [65, 232]
Hypersensitivity pneumonitis	455	346	76.0	190	141	74.2	109 [63, 232]

Percentages are shown for the overall cohort (incident and existing fibrosing ILD) and for the incident fibrosing ILD cohort only

IIP idiopathic interstitial pneumonia, *ILD* interstitial lung disease, *iNSIP* idiopathic non-specific interstitial pneumonia, *IQR* interquartile range, *MCTD* mixed connective tissue disease, *RA* rheumatoid arthritis, *SSc* systemic sclerosis

*Patients in the existing fibrosing ILD cohort could have existing or incident progression, while patients in the incident fibrosing ILD cohort could have incident, but not existing, progression (diagnosed by proxies). Most patients with incident progression were from the incident fibrosing ILD cohort (13,518/14,722, 91.8%)

† Time to progression was calculated from the first of the two claims that were required to establish a diagnosis of incident fibrosing ILD to the time of the claim indicating progression

Distribution of Proxy Criteria Used to Define Progression

To identify the main drivers of progressive fibrosing ILD diagnosis in this population, the distribution of proxy criteria used to define progression was analyzed. For completeness, the distribution of criteria was analyzed in all patients who had a proxy for progression, including those with only one claim for fibrosing ILD. Of 49,377 patients with a qualifying claim for incident progression, 44,672 (90.5%) qualified based on only one proxy criterion. Oxygen therapy was the most common proxy,
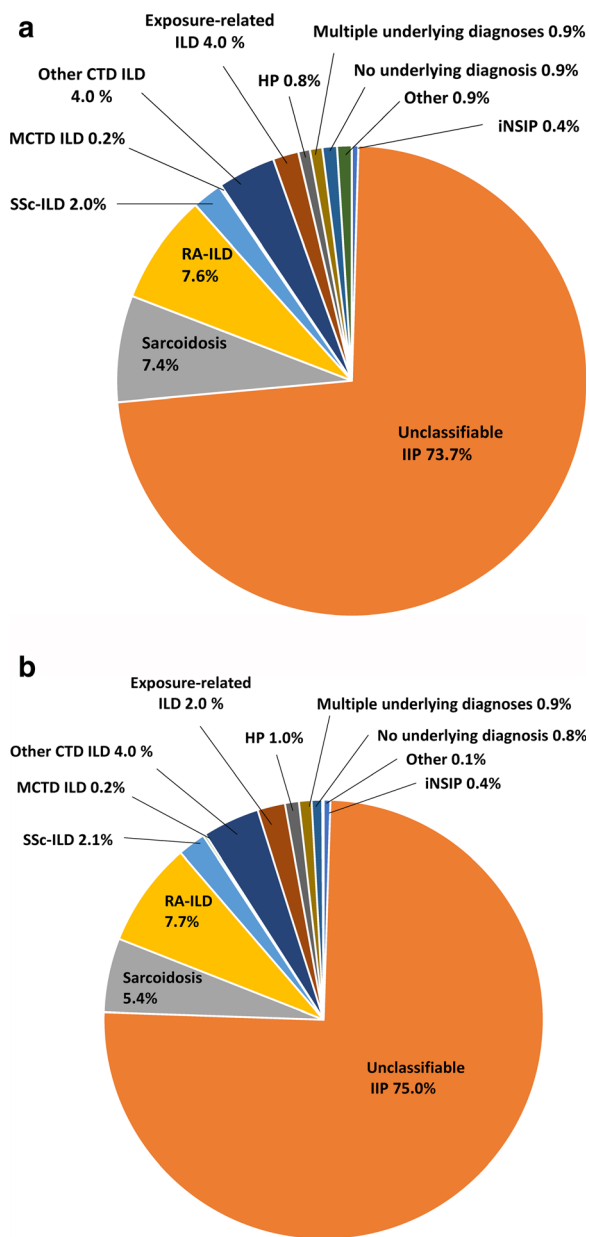
a

Exposure-related
ILD 4.0 %

Multiple underlying diagnoses 0.9%

Other CTD ILD
4.0 %

HP 0.8%

No underlying diagnosis 0.9%

Other 0.9%

MCTD ILD 0.2%

iNSIP 0.4%

SSc-ILD 2.0%

RA-ILD
7.6%

Sarcoidosis
7.4%

Unclassifiable
IIP 73.7%

b

Exposure-related
ILD 2.0 %

Multiple underlying diagnoses 0.9%

Other CTD ILD 4.0 %

HP 1.0%

No underlying diagnosis 0.8%

Other 0.1%

MCTD ILD 0.2%

iNSIP 0.4%

SSc-ILD 2.1%

RA-ILD
7.7%

Sarcoidosis
5.4%

Unclassifiable
IIP 75.0%

◀**Fig. 1** Distribution of underlying clinical diagnosis in patients assessed as having **a** incident fibrosing ILD and **b** incident progressive fibrosing ILD. The high proportion of patients classed as having unclassifiable ILD is likely to be a consequence of coding practices, where the code for unclassifiable ILD is used initially and not updated when another diagnosis is made. Patients categorized as 'other' had a claim for an underlying condition but did not satisfy any of the specific coding algorithms. *CTD* connective tissue disease, *HP* hypersensitivity pneumonitis, *IIP* idiopathic interstitial pneumonia, *ILD* interstitial lung disease, *iNSIP* idiopathic non-specific interstitial pneumonia, *MCTD* mixed connective tissue disease, *RA* rheumatoid arthritis, *SSc* systemic sclerosis

seen in 28.3% of progression claims overall and 27.6% of those with only one proxy. Other frequently used proxies were respiratory hospitalization (14.5% overall, 12.7% of those with one proxy), ≥ 2 pulmonary function tests within 90 days (15.0% overall, 14.1% of those with one proxy), ≥ 2 HRCT scans within 360 days (22.0% overall, 14.5% of those with one proxy), and three or more chest CT scans within 360 days (19.9% overall, 12.3% of those with one proxy). Other criteria for progression were found in < 3% of patients (Fig. 2). Findings were similar in the overall cohort with incident or existing progressive fibrosing ILD (Figure S1).

Sensitivity Analysis of Proxy Criteria for Progression

Removing claims for oxygen titration tests and corticosteroids as criteria and requiring two respiratory hospitalizations within 360 days or two claims for oxygen therapy for these to be proxy criteria led to very similar results, with no changes to the incidence and prevalence rates and median time to progression changed from 117 to 122 days. It was concluded that these potentially more sensitive progression proxies were not strong drivers of the observed prevalence and incidence estimates.

Table 3 Crude prevalence of fibrosing ILD and progressive fibrosing ILD by age and sex (specific, two-claim definition)

Characteristic	Number of cases (%)	Persons at risk	Crude prevalence per 100,000 persons (95% CI)
Fibrosing ILD			
Overall study persons	35,825 (100)	37,565,644	95.37 (94.38, 96.35)
Age category, years			
18–39	1040 (2.9)	15,363,409	6.77 (6.36, 7.18)
40–49	2470 (6.9)	8,313,711	29.71 (28.54, 30.88)
50–59	6455 (18)	9,012,816	71.62 (69.87, 73.37)
60–69	9178 (25.6)	5,674,094	161.75 (158.45, 165.06)
70–79	8955 (24.9)	1,736,038	515.83 (505.17, 526.49)
80 +	8703 (24.3)	1,136,978	765.45 (749.43, 781.47)
Gender			
Male	16,463 (46.0)	17,818,937	92.39 (90.98, 93.80)
Female	19,362 (54.0)	19,746,707	98.05 (96.67, 99.43)
Progressive fibrosing ILD			
Overall study persons	21,719 (100)	37,565,644	57.82 (57.05, 58.58)
Age category, years			
18–39	544 (2.5)	15,363,409	3.54 (3.24, 3.84)
40–49	1312 (6.0)	8,313,711	15.78 (14.93, 16.64)
50–59	3630 (16.7)	9,012,816	40.28 (38.97, 41.59)
60–69	5488 (25.2)	5,674,094	96.72 (94.16, 99.28)
70–79	5602 (25.8)	1,736,038	322.69 (314.25, 331.13)
80 +	5143 (23.7)	1,136,978	452.34 (440.00, 464.67)
Gender			
Male	10,269 (47.3)	17,818,937	57.63 (56.52, 58.74)
Female	11,450 (52.7)	19,746,707	57.98 (56.92, 59.05)

CI confidence interval, *ILD* interstitial lung disease

DISCUSSION

Progressive fibrosing ILDs other than IPF have been considered as a distinct group only relatively recently, and there are no published data from large studies on their overall epidemiology. Estimation of the prevalence and incidence rates has been limited to relatively small, clinic-based studies [19–21], and real-world data on larger populations are needed.

This is the first large study using a claims database to estimate the prevalence and incidence rates of fibrosing ILD and progressive fibrosing ILD. We initially used a broad, sensitive definition to define fibrosing ILD, requiring only one claim for a diagnosis. However, these

Table 4 Crude incidence rate of fibrosing ILD and progressive fibrosing ILD by age and sex

Characteristic	Number of new cases (%)	Person-years at risk	Crude incidence per 100,000 PY (95% CI)
Fibrosing ILD			
Overall study persons	23,577	53,416,575.09	44.14 (43.57, 44.70)
Age category, years			
18–39	699 (3.0)	18,697,610.47	3.74 (3.46, 4.02)
40–49	1565 (6.6)	10,576,534.84	14.80 (14.06, 15.53)
50–59	4064 (17.2)	12,116,258.50	33.54 (32.51, 34.57)
60–69	5748 (24.4)	7,517,305.06	76.46 (74.49, 78.44)
70–79	5733 (24.3)	2,681,331.63	213.81 (208.28, 219.35)
80 +	5768 (24.5)	1,827,534.58	315.62 (307.47, 323.76)
Gender			
Male	11,118 (47.2)	25,205,477.71	44.11 (43.29, 44.93)
Female	12,459 (52.8)	28,211,097.38	44.16 (43.39, 44.94)
Progressive fibrosing ILD			
Overall study persons	14,722 (100)	53,436,940.50	27.55 (27.11, 28.00)
Age category, years			
18–39	382 (2.6)	18,698,182.64	2.04 (1.84, 2.25)
40–49	852 (5.8)	10,577,953.22	8.05 (7.51, 8.60)
50–59	2353 (16.0)	12,119,899.76	19.41 (18.63, 20.20)
60–69	3629 (24.7)	7,522,234.76	48.24 (46.67, 49.81)
70–79	3884 (26.4)	2,686,127.79	144.59 (140.05, 149.14)
80 +	3622 (24.6)	1,832,542.35	197.65 (191.21, 204.09)
Gender			
Male	7142 (48.5)	25,214,279.74	28.33 (27.67, 28.98)
Female	7580 (51.5)	28,222,660.76	26.86 (26.25, 27.46)

CI confidence interval, *ILD* interstitial lung disease, *PY* patient-years

are complex conditions that may be difficult to diagnose and are frequently only diagnosed some time after the onset of symptoms. Diagnosis codes, particularly in the early stages, may be driven by reimbursement considerations rather than clinical diagnosis, and a single claim for 'lung fibrosis' may not necessarily indicate a diagnosis of fibrosing ILD. We found that a more specific definition, requiring two claims, produced estimates of the prevalence and incidence of fibrosing ILD more in line with clinic-based studies. Both estimates are provided, though we focus on the more specific definition, with which we have estimated the age-
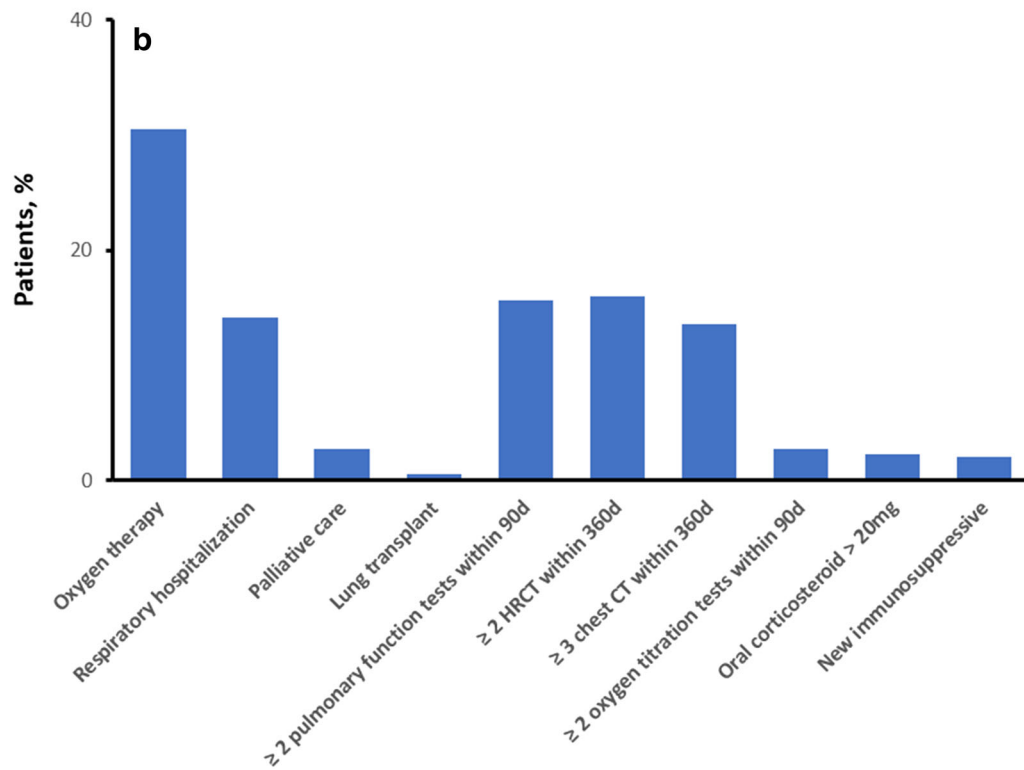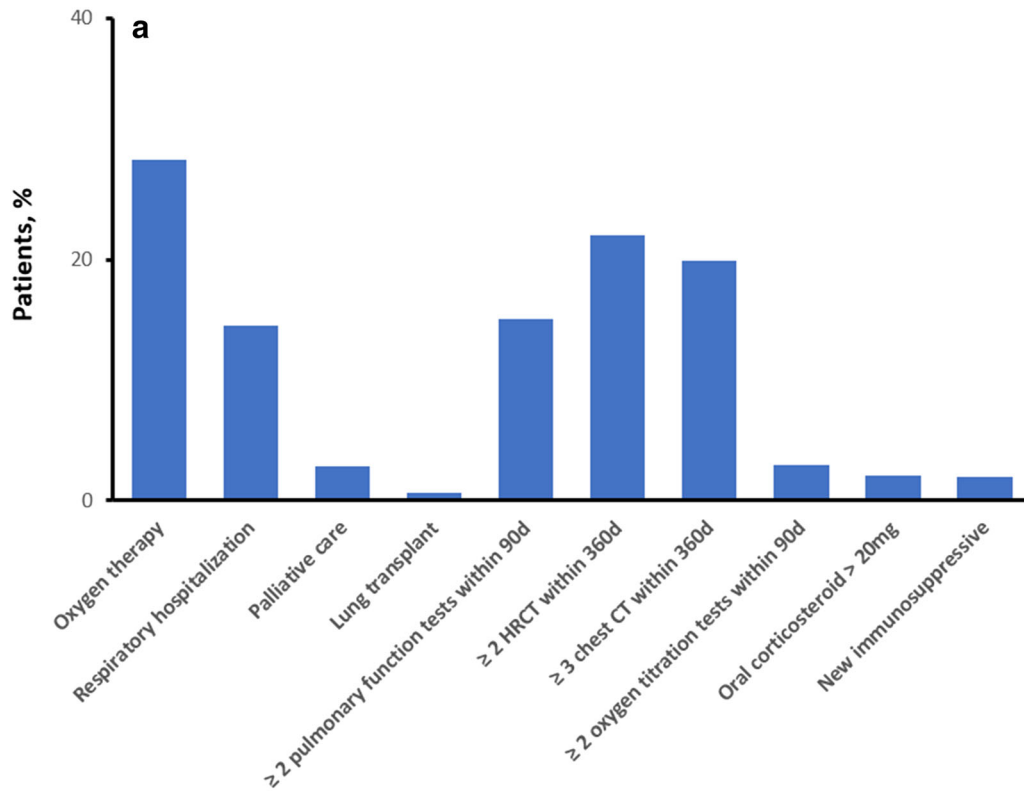
◀**Fig. 2** Proxy criteria for progression applied during the study in patients with incident progressive fibrosing ILD. Criteria were not mutually exclusive, and an individual patient could have more than one criterion applied. The majority of patients (90.5%) had only one proxy applied, and data are shown **a** for all patients and **b** for those with only one proxy applied. *CT* computed tomography, *HRCT* high-resolution computed tomography, *ILD* interstitial lung disease

and sex-adjusted prevalence of progressive fibrosing ILD to be 70.33 per 100,000 persons and the incidence to be 32.55 per 100,000 patient-years. This is higher than what has been previously reported (prevalence of progressive fibrosing ILDs up to 28 per 100,000 persons) in clinic-based studies [19–21]. Although there was a high proportion of patients in the 'unclassifiable IIP' group, the distribution of underlying conditions among other patients was comparable to previous findings [19–21], with RA-ILD and sarcoidosis most common. The prevalence of both fibrosing ILD and progressive fibrosing ILD increased markedly with age, as would be anticipated. In the absence of precise validated estimates, our study may provide potential upper limits for the true estimates of prevalence and incidence, whereas clinic-based studies may provide potential lower limits.

Given the lack of a code for progressive fibrosing ILD at the time the study was conducted and the absence of other claims database studies, progression was defined using proxy measures (as plausible markers of actual progression) that have not been validated. The most common proxies were claims for oxygen therapy and respiratory hospitalization and frequency of pulmonary function tests, chest CT scans, and HRCT scans. To ensure that single instances of oxygen therapy and respiratory hospitalization did not result in inclusion of patients that were not truly progressing, we conducted a sensitivity analysis requiring at least two claims for these two proxies that also excluded claims for oxygen titration tests and corticosteroids to create a stricter definition of progression. This led to minimal changes in results with no meaningful impact on the conclusions of the study.

Using the specific definition of fibrosing ILD, more than half of patients experienced progression within a median of 117 days (IQR 63–224 days) after presentation. There were few differences in the rate of progression by age, sex, or underlying diagnosis, except for sarcoidosis, where just under half of patients progressed according to these proxies. The proportion of progression reported in patients with fibrosing ILD in the current study is somewhat higher than has been previously suggested [15, 18–21]. This seems to be true both for the population overall and for individual diseases; in particular, the proportion of patients with sarcoidosis and progressive fibrosing ILD seems higher than has been reported previously and higher than is generally seen in clinical practice [28]. However, comparisons are difficult as methodologies used to estimate prevalence and incidence are variable. Possibly our estimates may represent the upper limit of the likely prevalence and incidence of these conditions, with patients included in the estimates that may have been excluded if clinical data were available. For instance, patients with a stable fibrosing ILD may require oxygen therapy to manage exertional dyspnea, meeting the criteria for progression in our analysis and potentially contributing to higher estimates for incidence and prevalence of the progressive phenotype than have been previously reported. Nevertheless, while the claims-based proxies for clinical progression may have been overly sensitive and led to some overestimation, this study makes an important contribution to the growing body of knowledge in this area by being the first to provide estimates of progressive fibrosing ILD from a very large population.

The current study has limitations, including relying solely on claims data without access to clinical data, which may lead to overestimates of the prevalence of fibrosing ILD/progressive fibrosing ILD, as confirmatory clinical data would increase the likelihood of identifying true cases and excluding patients without these conditions. In addition, progressive fibrosing ILD is a relatively new disease construct with multiple components (including ILD, an underlying disease and a progressive phenotype) and has a lengthy and potentially

complicated diagnostic pathway, with an algorithm that has not yet been validated. The crucial factor in identification in a claims study is the code entered by the physician. The ILD category may often be assigned early in the disease history when full diagnosis has not been made. Notably, a very high percentage of patients in the current study were categorized as 'unclassifiable IIP.' Most patients in this group were identified through the diagnostic code for post-inflammatory fibrosis (515) (n = 9716; 87.9%) at time of diagnosis, which is often used during diagnostic workup. In our study, patients were censored at date of diagnosis of progression, so it is unknown how diagnosis may have changed or been further specified during follow-up. In addition, more specific diagnoses may be made in patient documentation unavailable in the database, with no billing incentive for clinicians to update the claim. In general, it is possible that a code used initially for making a preliminary diagnosis was not updated when additional information became available, obscuring the true distribution of underlying diseases. In some cases, an IPF code may have appeared later in the follow-up as the final diagnosis, particularly for those who initially received a diagnosis of unclassifiable IIP. It was not possible to examine this because patients were censored at the date of diagnosis of progression.

The primary analysis of the current study was standardized by age and sex to the 2014 US population census estimates, but care should be taken in considering generalizability. The population under consideration excluded uninsured patients, so likely excluded many patients from lower socioeconomic backgrounds. Race is not captured in IBM® MarketScan® so we were unable to determine whether the findings of this study are applicable to all racial groups.

Progressive fibrosing ILD is a disease with a high unmet need, and well-designed clinical studies are required to optimize treatment strategies. Understanding the epidemiology of progressive fibrosing ILD is essential, but is further complicated by the grouping together of different ILDs that share a progressive phenotype. While a claims database study has the benefit of size, it remains limited by the reliance on physician coding. Because no specific code existed for progressive fibrosing ILD at the time of the study, a large database was required to estimate prevalence and incidence using proxy codes. We have categorized ILDs as incident or existing based on claims during the baseline period, but the absence of a claim does not guarantee that a patient did not have the disease during this period.

CONCLUSIONS

The limitations associated with a claims database notwithstanding, this is the first study to our knowledge providing insight into the prevalence and incidence of progressive fibrosing ILD and the likelihood of developing a progressive phenotype, using a very large population. The evidence provides a foundation for further studies to validate algorithms to identify patients and provide essential information on the epidemiology of progressive fibrosing ILD to help develop effective treatment strategies.

Further studies are needed to refine and/or replicate these findings and to maximize validity and reliability of these estimates. The findings serve as a strong beginning to the process of evidence generation for estimation of the prevalence and incidence of progressive fibrosing ILD, particularly in the absence of validated algorithms and no ICD code. Validation of the algorithm and additional data generation are needed as a next step, until a unique ICD-10 code for progressive fibrosing ILD becomes widely used and integrated into clinical practice.

ACKNOWLEDGEMENTS

Boehringer Ingelheim International GmbH, Ingelheim am Rhein, Germany.

REFERENCES

1. Wells AU, Brown KK, Flaherty KR, Kolb M, Thannickal VJ. What's in a name? That which we call IPF, by any other name would act the same. Eur Respir J. 2018;51(5):1800692.

2. Cottin V, Wollin L, Fischer A, Quaresma M, Stowasser S, Harari S. Fibrosing interstitial lung diseases: knowns and unknowns. Eur Respir Rev. 2019;28(151):180100.

3. Kolb M, Vasakova M. The natural history of progressive fibrosing interstitial lung diseases. Respir Res. 2019;20(1):57.

4. Brown KK, Martinez FJ, Walsh SLF, Thannickal VJ, Prasse A, Schlenker-Herceg R, et al. The natural history of progressive fibrosing interstitial lung diseases. Eur Respir J. 2020;55(6):2000085.

5. Gibson CD, Kugler MC, Deshwal H, Munger JS, Condos R. Advances in targeted therapy for progressive fibrosing interstitial lung disease. Lung. 2020;198(4):597–608.

6. Guler SA, Ellison K, Algamdi M, Collard HR, Ryerson CJ. Heterogeneity in unclassifiable interstitial lung disease. A systematic review and meta-analysis. Ann Am Thorc Soc. 2018;15(7):854–63.

7. George PM, Spagnolo P, Kreuter M, Altinisik G, Bonifazi M, Martinez FJ, et al. Progressive fibrosing interstitial lung disease: clinical uncertainties, consensus recommendations, and research priorities. Lancet Respir Med. 2020;8(9):925–34.

8. Distler O, Highland KB, Gahlemann M, Azuma A, Fischer A, Mayes MD, et al. Nintedanib for systemic sclerosis–associated interstitial lung disease. N Engl J Med. 2019;380(26):2518–28.

9. U.S. Food and Drug Administration. FDA approves first treatment for patients with rare type of lung disease. 2019 [cited 2020 Oct 9]. Available from: https://www.fda.gov/news-events/press-announce ments/fda-approves-first-treatment-patients-rare-type-lung-disease

10. U.S. Food & Drug Administration. FDA approves first treatment for group of progressive interstitial lung diseases. 2020 [cited 2020 Oct 9]. Available from: https://www.fda.gov/news-events/press-announcements/fda-approves-first-treatment-group-progressive-interstitial-lung-diseases

11. Flaherty KR, Wells AU, Cottin V, Devaraj A, Walsh SLF, Inoue Y, et al. Nintedanib in progressive fibrosing interstitial lung diseases. N Engl J Med. 2019;381:1718–27.

12. Nippon Boehringer Ingelheim Co. Ltd. Press release. 2020 [cited 2020 Oct 9]. Available from: https://www.boehringer-ingelheim.jp/press-release/20200529_01

13. Boehringer Ingelheim (Canada) Ltd. OFEV®: product monograph. 2020 [cited 2021 June 8]. Available from: https://www.boehringer-ingelheim.ca/sites/ca/files/documents/ofevcien.pdf

14. Boehringer Ingelheim. European Commission approves third nintedanib indication in pulmonary fibrosis. 2020 [cited 2020 Oct 9]. Available from: https://www.boehringer-ingelheim.com/press-release/ecapprovalnintedanibildpf

15. Wijsenbeek M, Kreuter M, Olson A, Fischer A, Bendstrup E, Wells CD, et al. Progressive fibrosing interstitial lung diseases: current practice in diagnosis and management. Curr Med Res Opin. 2019;35(11):2015–24.

16. Cottin V, Hirani NA, Hotchkin DL, Nambiar AM, Ogura T, Otaola M, et al. Presentation, diagnosis and clinical course of the spectrum of progressive-fibrosing interstitial lung diseases. Eur Respir Rev. 2018;27(150):180076.

17. De Sadeleer LJ, Goos T, Yserbyt J, Wuyts WA. Towards the essence of progressiveness: bringing progressive fibrosing interstitial lung disease (PF-ILD) to the next stage. J Clin Med. 2020;9(6):1722.

18. Nasser M, Larrieu S, Si-Mohamed S, Ahmad K, Boussel L, Brevet M, et al. Progressive fibrosing interstitial lung disease: a clinical cohort (the PRO-GRESS study). Eur Respir J. 2020;57(1):2002718.

19. Olson A, Hartmann N, Schlenker-Herceg R. Prevalence of progressive fibrosing interstitial lung disease. J Rheumatol. 2019. https://doi.org/10.3899/jrheum.190333.

20. Karakatsani A, Papakosta D, Rapti A, Antoniou KM, Dimadi M, Markopoulou A, et al. Epidemiology of interstitial lung diseases in Greece. Respir Med. 2009;103(8):1122–9.

21. Duchemann B, Annesi-Maesano I, de Naurois CJ, Sanyal S, Brillet PY, Brauner M, et al. Prevalence and incidence of interstitial lung diseases in a multi-ethnic county of Greater Paris. Eur Respir J. 2017;50(2):1602419.

22. Olson AL, Gifford AH, Inase N, Fernandez Perez ER, Suda T. The epidemiology of idiopathic pulmonary fibrosis and interstitial lung diseases at risk of a progressive-fibrosing phenotype. Eur Respir Rev. 2018;27(150):180077.

23. Coultas DB, Zumwalt RE, Black WC, Sobonya RE. The epidemiology of interstitial lung diseases. Am J Respir Crit Care Med. 1994;150(4):967–72.

24. Arkema EV, Grunewald J, Kullberg S, Eklund A, Askling J. Sarcoidosis incidence and prevalence: a nationwide register-based assessment in Sweden. Eur Respir J. 2016;48(6):1690–9.

25. Centers for Disease Control and Prevention. International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM). 2015 [cited 2020 Oct 9]. Available from: https://www.cdc.gov/nchs/icd/icd9cm.htm

26. Suissa S. Effectiveness of inhaled corticosteroids in chronic obstructive pulmonary disease: immortal time bias in observational studies. Am J Respir Crit Care Med. 2003;168(1):49–53.

27. Wang S, Verpillat P, Rassen J, Patrick A, Garry E, Bartels D. Transparency and reproducibility of observational cohort studies using large healthcare databases. Clin Pharmacol Ther. 2016;99(3):325–32.

28. Nardi A, Brillet P-Y, Letoumelin P, Girard F, Brauner M, Uzunhan Y, et al. Stage IV sarcoidosis: comparison of survival with the general population and causes of death. Eur Respir J. 2011;38(6):1368–73.