



Epidemiology of Pulmonary Fibrosis: A Cohort Study Using Healthcare Data in Sweden

Giovanni Ferrara · Lisen Arnheim-Dahlström · Karen Bartley ·
Christer Janson · Klaus-Uwe Kirchgässler · Aaron Levine ·
C. Magnus Sköld

Received: December 27, 2018 / Published online: March 5, 2019
© The Author(s) 2019

ABSTRACT

Introduction: Data on the epidemiology of idiopathic pulmonary fibrosis (IPF) in Sweden are lacking. This study estimates the incidence and prevalence of IPF in Sweden, and describes the demographic and clinical characteristics and the overall survival of patients with IPF.

Methods: Two cohorts were studied: a national cohort of 17,247 patients with pulmonary fibrosis (ICD-10 code J84.1 with no competing diagnosis) from the Swedish National Patient

Register (cohort 1 [C1]); and an electronic medical record-based regional subset of C1 comprising 1755 patients having pulmonary fibrosis and a radiology procedure (C2).

Results: The incidence of pulmonary fibrosis in C1 ranged from 10.4 to 15.4 cases per 100,000 population per year between 2001 and 2015. The prevalence increased from 15.4 to 68.0 cases per 100,000 population per year. Patients ≥ 70 years and men had a higher incidence and prevalence of pulmonary fibrosis. Common comorbidities included respiratory infections and cardiovascular disorders. Approximately one-third of patients in each cohort were hospitalised with pulmonary fibrosis within a year of diagnosis. The median survival time from disease diagnosis was 2.6 years in C1 and 5.2 years in C2. Older patients had a higher risk of hospitalisation and

Enhanced digital features To view enhanced digital features for this article go to <https://doi.org/10.6084/m9.figshare.7725917>.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s41030-019-0087-9>) contains supplementary material, which is available to authorized users.

G. Ferrara · C. M. Sköld
Department of Respiratory Medicine and Allergy,
Karolinska University Hospital, Stockholm, Sweden

G. Ferrara · C. M. Sköld (✉)
Department of Medicine Solna, Karolinska
Institutet, Stockholm, Sweden
e-mail: Magnus.Skold@ki.se

L. Arnheim-Dahlström · A. Levine
IQVIA, Solna, Sweden

L. Arnheim-Dahlström
Department of Medical Epidemiology and
Biostatistics, Karolinska Institutet, Stockholm,
Sweden

K. Bartley
Genentech, Inc., South San Francisco, CA, USA

C. Janson
Department of Medical Sciences: Respiratory,
Allergy and Sleep Research, Uppsala University,
Uppsala, Sweden

K.-U. Kirchgässler
F. Hoffmann-La Roche, Ltd., Basel, Switzerland

mortality. Women had a better prognosis than men.

Conclusion: This study underscores the importance of pulmonary fibrosis as a cause of respiratory-related morbidity and mortality in Sweden. The stable incidence and increasing prevalence over time suggests longer survival. The higher morbidity and mortality in older patients highlights the importance of early case detection, diagnosis and management for better prognosis.

Funding: F. Hoffmann-La Roche, Ltd./Genentech, Inc.

Keywords: Epidemiology; Idiopathic pulmonary fibrosis; Real-world evidence; Retrospective cohort; Sweden

INTRODUCTION

Idiopathic pulmonary fibrosis (IPF) is a chronic, progressive, and fatal lung disease characterised by the irreversible loss of lung function [1]. In a systematic review of studies from year 2000 onwards, a conservative annual incidence range of 3–9 cases per 100,000 population was estimated for Europe and North America, with a majority observing an increase in disease incidence over time [2].

The prognosis for untreated patients with IPF is poor, with a reported median survival of 3–5 years from the time of diagnosis [3]. Risk factors for mortality include male sex, older age (> 70 years), tobacco use, reduced lung function, pronounced fibrosis at time of diagnosis and pulmonary hypertension [4]. IPF is furthermore associated with comorbidities such as gastroesophageal reflux disease [5], obstructive sleep apnoea [6] and vascular disease [7, 8].

Available data on incidence and prevalence of IPF in Sweden are limited. A recently launched web-based registry of IPF patients in Sweden recorded 71 patients from 11 respiratory clinics within the first year of its launch [9]. In a nationwide study in Finland that used hospital records combined with familial screening for case detection, the prevalence of IPF was estimated to be 16–18 per 100,000 population [10].

In this study, the data from the Swedish national healthcare registries, maintained by the National Board of Health and Welfare (NBHW), and electronic medical records (EMR) were used to estimate the incidence and prevalence of pulmonary fibrosis in Sweden, and to describe the demographic and clinical characteristics and the overall survival of patients in Sweden with IPF.

METHODS

Study Design and Setting

The study was a retrospective cohort study comprising two cohorts of patients in Sweden diagnosed with pulmonary fibrosis: (1) a national cohort based on linked patient-level data from national population-based health registers (cohort 1); and (2) a subnational cohort based on patients from two healthcare regions, Stockholm and Uppsala county councils, including patient-level information from EMRs in primary care and pulmonary clinics linked to the national health register data (cohort 2). Cohort 2 was an in-depth EMR-based regional subset of cohort 1.

Data Sources

The following data sources were used in this study:

1. *The Swedish National Patient Register*

The Swedish National Patient Register, which is held by NBHW, was established in 1964 [11]. Information on all completed in-patient admissions in publicly operated hospitals were added to the database in 1987, and collection of outpatient data commenced a decade later, in 1997. The key variables available in the database include age, gender, diagnosis, surgery, referral, region, hospital, specialty, and hospital admissions and discharges.

The National Patient Register was used primarily to identify the study population and to obtain information on patient demographic and clinical characteristics (age, gender,

comorbidities) before and after disease diagnosis.

2. *The Swedish Cause of Death Register*

The Swedish Cause of Death Register contains data from 1952 onwards, updated annually by NBHW. The main variables included in the register are date and country of birth, nationality, residency, sex, marital status, date and place of death, underlying cause(s) of death based on ICD codes, autopsy type (if autopsied) and whether operated on within 4 weeks before death [12]. Although the quality of the death statistics may vary, in part due to the declining rates of autopsies over the past few decades and changes in the disease classification system and methods for statistical processing, the recorded date of death is accurate. The Cause of Death Register was used to ascertain information on mortality.

3. *EMR Data*

EMR data from primary care and from four respiratory clinics in two Swedish County Councils, Stockholm and Uppsala, were extracted to obtain more detailed clinical information to include only patients with IPF. Patient identification and EMR data extraction was carried out using the Pygargus Customised Extraction Platform (CXP), which since 2005 has been used in many research projects on the Swedish healthcare system [13, 14]. Anonymous data on age, gender, prescriptions, diagnoses, measures, laboratory tests, healthcare professional visits, referrals, and other information can be collected from the EMRs into an encrypted database [15].

Data Linkage

Patients identified in the Swedish National Patient Register were linked to the National Prescribed Drug Register and the Cause of Death Register using their personal identity number (PIN). The linked data were used to create the database for cohort 1. Linkages across the national health registers were handled by NBHW and were made possible by PINs carried by each Swedish citizen from birth to death. After extraction, the EMR data were linked to the same Swedish national healthcare registers

(Swedish National Patient Register, Cause of Death Register) to create the database for cohort 2.

The linkages were performed in accordance with national and EU data privacy legislation. The original key codes, containing the unique PIN, were separately delivered to the NBHW. All data were pseudonymised following linkage, and a new individual key code was produced and stored at NBHW. The merged and anonymised database was delivered to the principal investigator (PI) and no identification of patients was possible from then on.

Case Definition

The International Classification of Diseases, Tenth Revision (ICD-10) was used to identify patients with pulmonary fibrosis. ICD-10 code J84.1, i.e., patients with “Other interstitial pulmonary diseases with fibrosis”, was used as an inclusion criterion. For cohort 1, all patients with a registration of ICD-10 diagnosis code J84.1 in the Swedish National Patient Register between 1 January 2001 and 31 December 2015, and aged 40 years or older at the time of diagnosis were eligible for inclusion. Patients with pulmonary fibrosis resulting from inhalation of chemicals, gases, fumes or vapours (J68.4) or following radiation (J70.1, radiation fibrosis) were excluded. Patients were also excluded if they had a record of the following diagnoses (ICD-10 code) on or after the ICD-10 J84.1 diagnosis code: (1) pneumoconiosis (J60); (2) asbestosis (J61); (3) silicosis or talcosis (J62.8); (4) berylliosis and other inorganic dusts (J63.0–J63.6); (5) unspecified pneumoconiosis (J64); (6) hypersensitivity pneumonitis (J67); (7) rheumatoid arthritis and other inflammatory polyarthropathies (M06.9); and (8) diffuse connective tissue diseases (M32.10, M33.20, M33.90, M34.0, M34.1, M34.9, M35.0, M35.01, M35.5, M35.8, M35.9).

For cohort 2, in addition to the eligibility criteria mentioned above, patients with a registration of ICD-10 diagnosis code J84.1 in EMRs between 1 November 2004 and 12 March 2017 were also eligible. To obtain a population of IPF patients, the cohort 2 subpopulation was

further refined to include only those having one or more radiology procedure (CT or chest X-ray) records in their EMRs following the diagnosis of pulmonary fibrosis.

Key Variables and Outcomes

1. *Index Date*

The date of first diagnosis of pulmonary fibrosis (cohort 1) or IPF (cohort 2) in a patient was considered the index date. For cohort 1 patients, this was the date of initial registration of ICD-10 J84.1 diagnosis code in the National Patient Register. For cohort 2 patients, this was the date of initial registration of ICD-10 J84.1 diagnosis code in the EMR and/or the National Patient Register. In the case of discrepancies between these two sources, EMRs were considered as the primary source. From the index date, patients were followed for the entire disease duration, i.e., until death or until the end of the study period (31 December 2015 and March 2017 for cohorts 1 and 2, respectively), whichever occurred first.

2. *Disease Severity*

Forced vital capacity (FVC), expressed as a percentage of the predicted value, was used as a measure of disease severity. It was defined as follows: normal—FVC \geq 100%; mild—FVC \geq 80%, < 100%; moderate—FVC \geq 50%, < 80%; and severe—FVC < 50% [16]. Data on spirometry test results were collected from the EMRs.

3. *Comorbidities*

Comorbidities were extracted from the National Patient Register and EMRs using ICD-10 diagnosis codes.

4. *Hospitalisation due to Pulmonary Fibrosis*

Number and duration of hospitalisations due to pulmonary fibrosis (i.e., having ICD-10 J84.1 as the main diagnosis code) were collected from the National Patient Register and from EMRs (as free text search).

5. *Mortality*

Information on all-cause and pulmonary fibrosis-related mortality was collected from the Cause of Death Register. Deaths with an ICD-10 J84.1 diagnosis code as the underlying cause of death were regarded as having been caused by

pulmonary fibrosis. Non-pulmonary fibrosis-related mortality was divided into two categories: deaths caused by another respiratory disease (records with a J diagnosis code as the underlying cause of death) and deaths due to non-respiratory causes (records with any other diagnosis code as the underlying cause of death).

Statistical Analysis

Continuous variables were described using summary statistics (number of observations, mean, median, standard deviation [SD] and IQR [25th and 75th percentiles]). Categorical variables were described using frequencies or percentages. The annual incidence and prevalence were calculated for each calendar year from 2001 to 2015, and presented as the number of incident or prevalent cases per 100,000 population per year. Incidence and prevalence information was presented for the overall study population as well as stratified by age and sex.

The prevalence of comorbidities occurring in > 10% of patients was reported for the period before the index diagnosis (pre-index), as well as for the entire study period. Associations of gender, age at index date, and disease severity with annual number of hospitalisations were all assessed using generalised linear modelling with negative binomial distribution. Mortality was described using number of events, median time to event and 1-year survival. Cox proportional hazards models were used to ascertain the effect of age, gender and disease severity on mortality and corresponding hazard ratios (HR) with 95% confidence intervals (95% CI) presented.

Statistical comparisons were not conducted for differences between the two cohorts, as cohort 2 was a regional subset of cohort 1, and thus most patients in cohort 2 were also part of the larger national cohort.

Ethics Statement

This is a retrospective observational study that used anonymised patient data from the Swedish national health registries and primary care EMRs. The study was approved by the

Stockholm County Ethical Committee (Regionala etikprövningsnämnden i Stockholm) with reference number 2016/1364-31/2 dated 17 August 2016.

RESULTS

The incidence and prevalence of pulmonary fibrosis in Sweden, along with demographic and clinical characteristics, and the morbidity and mortality profile of pulmonary fibrosis/IPF patients are presented here. Results are presented separately for cohorts 1 and 2.

Cohort 1

A total of 18,750 patients with pulmonary fibrosis (ICD-10 code of J84.1) were recorded in the Swedish National Patient Register between 2001 and 2015, of which 17,247 (92%) met the eligibility criteria and were included in the final analysis (Fig. 1). Many patients (44%) were from the three most populous counties: Västra

Götaland (16.5%), Stockholm (15.8%) or Skåne (11.7%).

The mean (SD) age of patients at the time of disease diagnosis was 74.6 (10.5) years. The majority of patients (10,771, 62.5%) were men. Disease duration, as measured by the median (IQR) duration of follow-up, was 1.8 (0.5, 4.3) years. A total of 5056 (29.3%) patients were diagnosed at a pulmonary clinic.

Incidence and Prevalence of Pulmonary Fibrosis

The incidence of pulmonary fibrosis in the Swedish population ranged from 10.4 to 15.4 cases per 100,000 population per year between 2001 and 2015; it remained stable over time. In 2015 (the last year of the study period), the incidence was 13.9 cases per 100,000 population. On the other hand, the prevalence of pulmonary fibrosis increased steadily from 15.4 cases per 100,000 population in 2001 to 68.0 cases per 100,000 population in 2015 (Fig. 2).

The age-specific incidence and prevalence of pulmonary fibrosis in Sweden is presented as

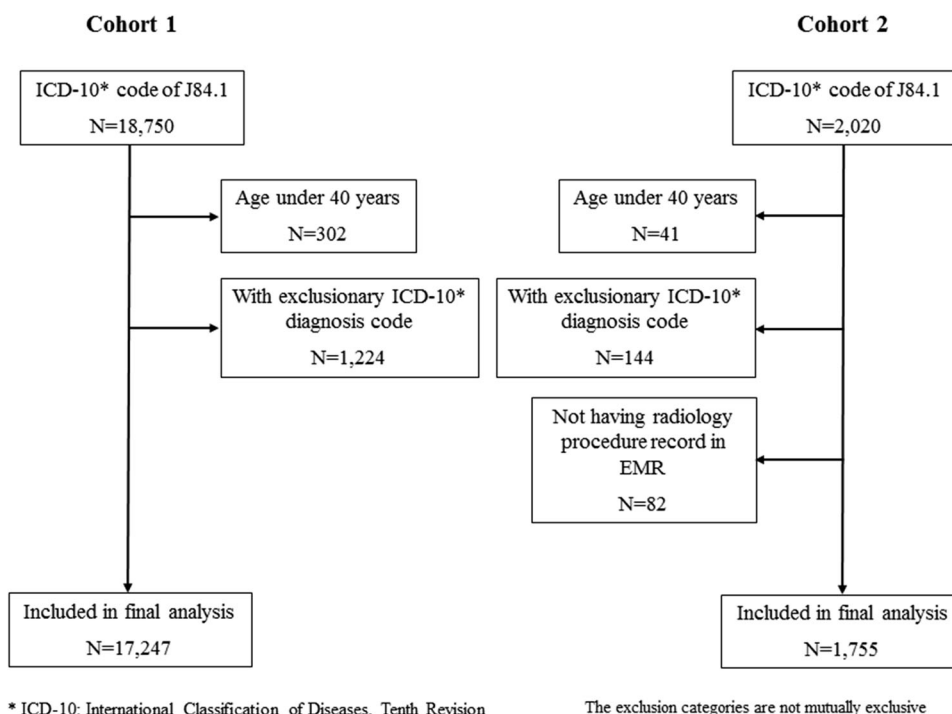


Fig. 1 Flowchart of the number of patients in each cohort

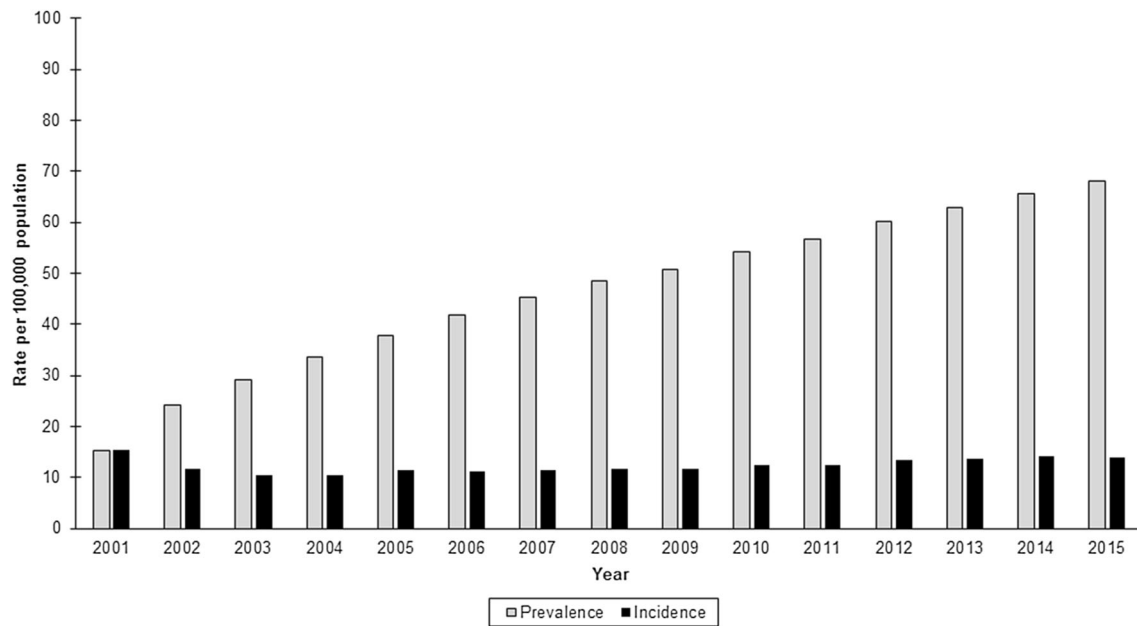


Fig. 2 Incidence and prevalence of pulmonary fibrosis in Sweden between 2001 and 2015

Figure S1 in the electronic supplementary material. The incidence and the prevalence of pulmonary fibrosis increased with increasing age, and both were higher in men than women (Figure S2).

Comorbidities

The prevalence of comorbidities occurring in > 10% of patients during the pre-index period (i.e., any time before the index diagnosis) and during the entire study period are reported in Supplementary Table S1. Early potential symptoms of pulmonary fibrosis such as dyspnoea, cough and fatigue occurred in 12.2, 3.7 and 1.5% of patients, respectively, before their index diagnosis. Respiratory infections and most cardiovascular disorders affected considerable proportions of patients before and after the index diagnosis (Table S1).

Hospitalisations

Within the first year of diagnosis, a total of 6073 (35.2%) patients were hospitalised with the main diagnosis of pulmonary fibrosis (ICD-10 code of J84.1). Most of these hospitalisations occurred in non-pulmonary wards (4857, 80%)

as opposed to pulmonary wards (1216, 20%). During the fifth year after diagnosis, 403 (8.8%) of the 4604 surviving patients were hospitalised with the main diagnosis of pulmonary fibrosis, of which 104 (25.8%) were admitted to the pulmonary ward. A total of 8159 (47.3%) patients were hospitalised with pulmonary fibrosis as the main diagnosis during all follow-up after the index date; 1837 (22.5%) of them were admitted in the pulmonary wards.

Potential predictors of the annual number of pulmonary fibrosis-related hospitalisations are presented in Table 1. Age at index diagnosis was positively associated with an increased number of hospitalisations, with a 1-year increase in age resulting in a 2% increase in fibrosis-related hospitalisation ($p < 0.001$). Additionally, the expected number of hospitalisations per year was 25% lower in women compared to men, which was statistically significant ($p < 0.001$).

Mortality

A total of 11,663 (67.6%) of the 17,247 patients died during the study period, of whom 3981 (34.1%) had pulmonary fibrosis as the primary cause of death (ICD code of J84.1). A total of 1441 (12.4%) patients died from other

Table 1 Generalised linear regression analysis to ascertain the factors associated with pulmonary fibrosis-related hospitalisations

Cohort	Parameter	Level	N	Estimate ^b	95% CI ^a		p value
					Lower	Upper	
Cohort 1							
	Sex	Female	6476	− 0.29	− 0.41	− 0.16	< 0.001
		Male ^c	10,771	0	–	–	–
	Age at index		17,247	0.02	0.01	0.03	< 0.001
Cohort 2							
	Sex	Female	651	− 0.30	− 0.61	0.02	0.06
		Male ^c	1104	0	–	–	–
	Age at index		1755	0.03	0.01	0.04	< 0.001
	Disease severity at index ^d	Mild (FVC ^a ≥ 80%, < 100%)	98	0.38	− 1.15	1.9	0.63
		Moderate (FVC ^a ≥ 50%, < 80%)	242	1.23	− 0.19	2.65	0.09
		Severe (FVC ^a < 50%)	63	2.01	0.44	3.59	0.01
		Normal (FVC ^a ≥ 100%) ^c	37	0	–	–	–

^a CI confidence interval, FVC forced vital capacity

^b Parameter estimates are log-linear

^c Reference category

^d Parameter estimates of 1315 patients with ‘unknown’ FVC status are not presented

respiratory diseases, and 6241 (53.5%) died from non-respiratory causes. The mean age at the time of death was 76.9 years.

The Kaplan–Meier plot of survival probability with respect to death from any cause is shown in Fig. 3A. The median survival time was 2.6 years; the 5-year survival was approximately 30%. The median survival time among women was 3.1 years, which was better than that for men, who had a median survival time of 2.3 years (Table 2).

Factors associated with all-cause mortality are presented in Table 2. Older age at index date was associated with a significantly higher risk of mortality within a year of disease diagnosis: each year, the increase in age led to an increased risk of mortality by 6% (Table 2). Female patients had a lower risk of mortality than male patients (HR 0.82, 95% CI 0.79, 0.85).

Cohort 2

A total of 2020 patients were identified with the J84.1 ICD-10 code in EMRs between November 2004 and March 2017. Of these, 1755 (86.9%) met the eligibility criteria and were included in the final analysis (Fig. 1).

The average (SD) age of patients at the time of disease diagnosis was 72.0 (10.4) years; 1104 (62.9%) patients were male. The median (IQR) duration of follow-up after the index date was 2.8 (1.0, 5.4) years. Information on the specialty of diagnosis was available for 1512 (86.2%) patients, the majority of whom (862, 57.0%) were diagnosed in a pulmonary clinic.

Comorbidities

Among the early potential signs of IPF, the prevalence of dyspnoea was highest during the pre-index period (22.0%), followed by cough

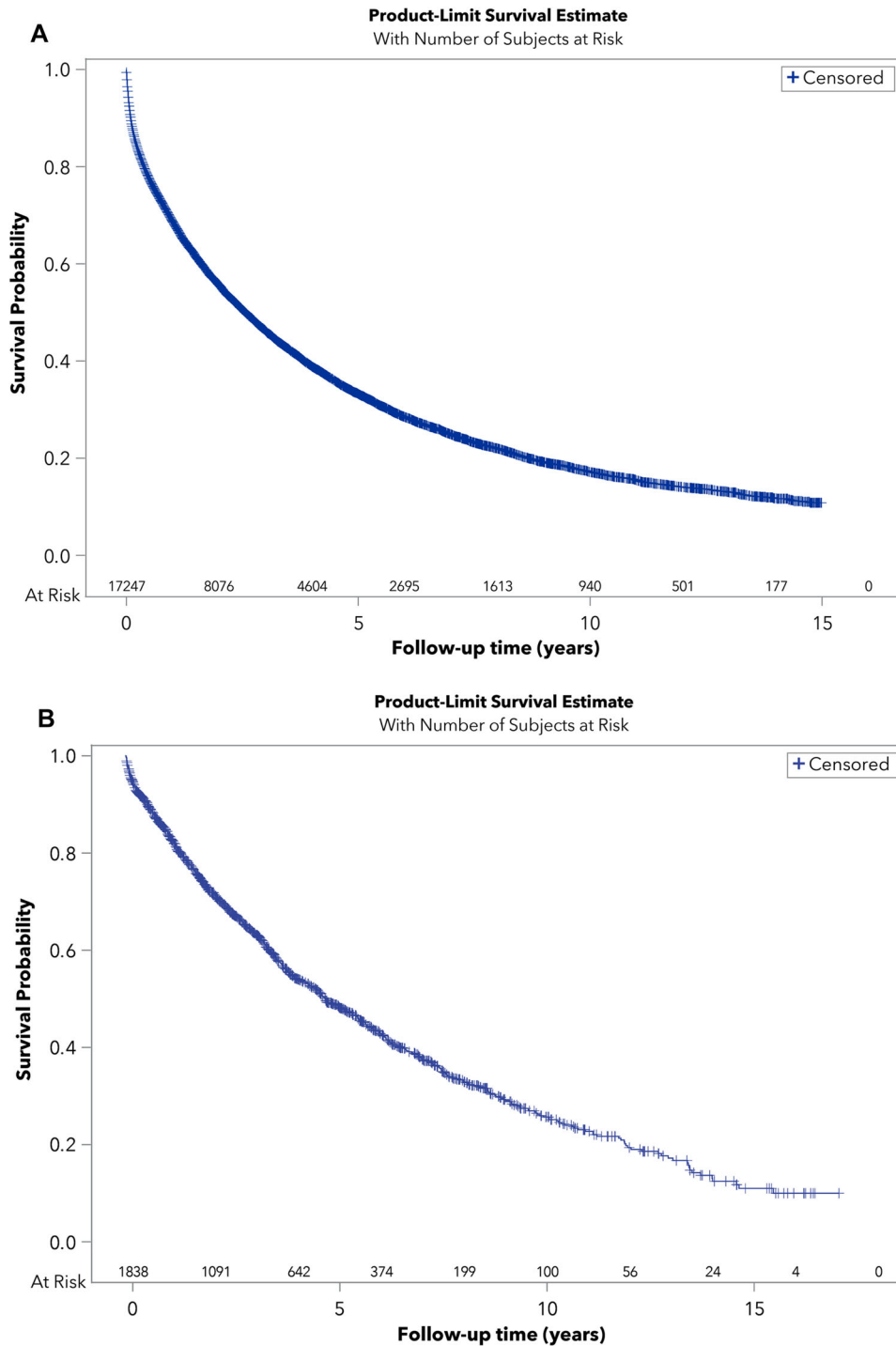


Fig. 3 Kaplan-Meier estimates of patient survival time with respect to all-cause deaths in **a** cohort 1 and **b** cohort 2

(17.3%) and fatigue (7.6%). In general, most patients suffered from respiratory and cardiovascular disorders, both before and after the

index diagnosis (Table S1). Information on smoking status was available for 60.3% of the patients: 74.1% of the classifiable patients were

Table 2 Cox proportionate hazards regression analysis to assess factors associated with death from any cause within 1 year of the index pulmonary fibrosis/IPF diagnosis

Cohort	Variable	<i>N</i>	Total no. of events	Patients alive 1 year after index diagnosis, % (95% CI ^a)	Median survival, years	Unadjusted HR (95% CI ^a) ^b
Cohort 1						
	Sex					
	Female	6476	4179	72.0 (70.9–73.1)	3.1	0.82 (0.79, 0.85)
	Male ^c	10,771	7484	66.6 (65.7–67.5)	2.3	–
	Age at index	17,247	11,663	68.6 (67.9–69.3)	–	1.06 (1.06, 1.06)
Cohort 2						
	Sex					
	Female	651	302	86.7 (84.0%, 89.4%)	6.2	0.74 (0.64, 0.85)
	Male ^c	1104	583	82.8% (80.5%, 85.1%)	4.6	–
	Age at index	1755	885	84.3% (82.5%, 86.0%)	–	1.06 (1.06, 1.07)
	Disease severity at index ^d					
	Mild (FVC ^a ≥ 80%, < 100%) ^c	98	36	94.8% (90.3%, 99.2%)	6.2	1
	Moderate (FVC ^a ≥ 50%, < 80%)	242	96	90.8% (87.0%, 94.5%)	6.2	1.13 (0.77, 1.65)
	Severe (FVC ^a < 50%)	63	33	81.1% (71.1%, 91.2%)	4.4	1.56 (0.97, 2.51)
	Normal (FVC ^a ≥ 100%)	37	12	81.7% (68.3%, 95.0%)	5.6	1.14 (0.59, 2.18)

^a CI confidence interval, FVC forced vital capacity

^b Statistically significant if 95% CI does not include 1

^c Reference category

^d Parameter estimates of 1315 patients with ‘unknown’ FVC status are not presented

current or former smokers, whereas 25.9% had never smoked.

Hospitalisations

Approximately one-third of the 1755 patients (582, 33.2%) were hospitalised with the main diagnosis of pulmonary fibrosis (ICD Code J84.1 as the main diagnosis) within the first year of disease diagnosis, of whom 186 (32%) were admitted to pulmonary wards. In the fifth year following disease diagnosis, 65 (10.4%) of the surviving 626 patients were hospitalised with pulmonary fibrosis as the main diagnosis, of which 21 (32.3%) were admitted to pulmonary wards. During the entire follow-up period after the index date, a total of 950 (54.1%) of the

1755 IPF patients were hospitalised for pulmonary fibrosis-related treatment; 328 (34.5%) of them were admitted to a pulmonary ward.

Potential predictors of the annual number of pulmonary fibrosis-related hospitalisations are presented in Table 1. The number of hospitalisations increased with age: 1-year increase in age resulted in 1.03% increase in pulmonary fibrosis-related hospitalisation rate. Female patients had 26% lower hospitalisation rates than male patients, although this difference was not statistically significant. Data on disease severity at index diagnosis were known for 440 (25.1%) patients. Patients with severe disease (FVC < 50%) had significantly higher

pulmonary fibrosis-related hospitalisation rates than those with normal FVC (Table 1).

Mortality

Of the 1755 patients included in the final analysis, 885 (50.4%) died during the study period. Pulmonary fibrosis was attributed as the primary cause of death in 342 (38.6%) of the 885 patients, whereas in 79 (8.9%) patients, respiratory disease other than pulmonary fibrosis was the primary cause of death. A total of 464 (52.4%) of the 885 patients died from non-respiratory causes. The mean age at the time of death was 74.5 years.

The Kaplan–Meier plot of survival probability for all-cause deaths is presented in Fig. 3B. The median survival time for all-cause deaths was 5.2 years from the time of disease diagnosis, and the 5-year survival probability was 51%. Female patients survived for a median duration of 6.2 years from the time of their disease diagnosis. This was longer than for male patients, who survived for a median duration of 4.6 years (Table 2).

Factors related to all-cause mortality were ascertained (Table 2). Age was associated with an increased risk of mortality: 1-year increase in age resulted in 6% increase in the risk of mortality within a year of disease diagnosis. Compared to male patients, female patients had a significantly lower risk of mortality within a year of disease diagnosis (HR 0.74, 95% CI 0.64, 0.85). Disease severity was not associated with an increased risk of mortality (Table 2).

DISCUSSION

This study retrospectively identified two cohorts of patients in Sweden with an incident diagnosis of pulmonary fibrosis/IPF between 2001 and 2015: a national cohort (cohort 1) of pulmonary fibrosis patients (ICD-10 code of J84.1) and a subnational cohort (cohort 2) of IPF patients from Stockholm and Uppsala county councils, which was an in-depth EMR-based regional subset of cohort 1. Analysis of data from the national cohort depicted an annual incidence of between 10 and 15 cases of pulmonary fibrosis per 100,000 population and an

annual prevalence of between 15 and 68 cases per 100,000 population. The disease incidence and prevalence were higher in older adults and men, who also had a poorer disease prognosis (higher hospitalisation and mortality rates).

Following a review of the literature, this is the first study linking the national registers with EMR data to describe the epidemiology of pulmonary fibrosis/IPF in the Swedish population. The method of combining data from EMRs and national registers has been used in numerous previous studies [17–21]. This allows patients to be followed between primary and secondary care settings, and to capture disease epidemiology and healthcare resource utilisation. Variability in the accuracy and degree of completeness of patient records due to differences in the standards of physician practice can, however, pose methodological challenges [22, 23].

The incidence and prevalence of pulmonary fibrosis in the national cohort was higher than what has been previously estimated for the European population by studies investigating the epidemiology of IPF [2, 24, 25]. This finding may reflect a true under-recognised burden of IPF in the Swedish population. Alternatively, it may be an artefact of the case definition used in this study, as the exclusionary criteria may not have excluded all patients with non-IPF interstitial lung disease, leading to a diagnosis misclassification [26, 27]. It is, however, important to note that similar diagnosis algorithms have been used by previous real-world studies on IPF [28–30].

The incidence of pulmonary fibrosis was higher in the older age groups and in the male population, which is generally consistent with the findings from other population-based studies in Europe and elsewhere [29, 31–34]. While the higher incidence and prevalence of pulmonary fibrosis in older adults is probably due to age-related degenerative changes, it is possible that differences by gender may be partly due to differences in environmental and occupational exposures, which needs to be explored further. People engaged in occupations related to dust exposure, such as farmers, stone cutters and carpenters, have been found to have an increased risk of IPF [35–37].

The incidence of pulmonary fibrosis in this study remained stable over the 15 years of observation, whereas the prevalence increased steadily over time (Fig. 2). Similar patterns in the incidence and prevalence of IPF in patients aged 45 years and older have been reported from studies using US national healthcare claims data [33, 38]. While the stable incidence of pulmonary fibrosis represents a constant influx of new cases, the increasing prevalence can be attributed in part to the longer survival of patients, which may be linked to early case detection and referral due to increased awareness of IPF among primary care physicians. Early access to specialist care has been associated with a lower risk of mortality in IPF patients, irrespective of disease severity [39]. Availability of newer treatment options, such as anti-fibrotic medications, can also affect patient survival. Treatment with anti-fibrotic medications has been shown to reduce disease progression and may have the potential to prolong survival in patients with IPF [40, 41].

Age at diagnosis was positively associated with hospitalisation and with mortality from any cause within 1 year of the index diagnosis in both cohorts. Similarly, compared to men, women had a lower risk of hospitalisation or death from any cause within 1 year of the index diagnosis in the national cohort. The higher risk of mortality in older patients and in men has been reported previously [42–44]. Comorbid conditions in this population were also generally consistent with previously studied populations of patients with IPF [45, 46].

This study has several limitations that are inherent to real-world data. Despite the use of a diagnosis algorithm to refine the disease definition, it is possible that a proportion of non-IPF patients were misclassified as IPF patients in the national registers and EMRs. Earlier studies have observed that the use of ICD codes to identify IPF patients may overestimate the incidence and prevalence of IPF [26, 27]. The probability of disease misclassification is higher in cohort 1 than in cohort 2, as patients in cohort 2 were additionally refined by the presence of at least one radiology procedure record (CT or X-ray) after their ICD-10 J84.1 diagnosis. Nonetheless, the effect of potential

misclassification is unlikely to have significantly altered the major findings of this study, as the disease patterns and risk factors (except for the mortality experience) showed similar trends in both cohorts. The temporal trends in the national incidence and prevalence, and the trends across age and gender are not affected, as the misclassification, if any, is likely to be similar across the categories.

Other limitations of the data set include possible differences in the standards of physician practice, and the fact that clinical assessments may have been performed but not recorded in the EMR, and hence not reported to the National Patient Register. Similarly, diagnoses reflecting potential early symptoms of pulmonary fibrosis/IPF (e.g., dyspnoea, cough and fatigue) or comorbidities (e.g., depression, anxiety and diabetes) may have been underreported to the National Patient Register, as they are in most similar databases, resulting in underestimation of their prevalence. Furthermore, as the ICD-10 codes stay in the National Patient Register once entered, some of the comorbidities (e.g., chronic obstructive pulmonary disease [COPD]) may have been wrongly attributed to patients with pulmonary fibrosis/IPF.

The major strength of this study is the ability to link patient-level data from multiple sources, including nationwide administrative health registers and in-depth clinical information from EMRs. This allowed for a long and virtually complete follow-up of the patients in the study. In addition, combining EMR data from primary and secondary care allowed a more complete view of the patient journey than if only one of the EMR data sources would have been used. Lastly, the Pygargus methodology used to extract data from EMRs enabled the assessment of detailed clinical data on a relatively large scale, i.e., among a total of 1755 patients included in cohort 2.

CONCLUSIONS

This study underscores the importance of pulmonary fibrosis as a cause of respiratory-related morbidity and mortality in the Swedish

population. The stable incidence and increasing prevalence over time could suggest longer survival of patients with pulmonary fibrosis. The higher morbidity and mortality in older patients highlights the importance of early case detection, diagnosis and management for better prognosis.

ACKNOWLEDGEMENTS

Funding. This study and article processing charges were sponsored by F. Hoffmann-La Roche, Ltd./Genentech, Inc. All authors had full access to all of the data in this study and take complete responsibility for the integrity of the data and accuracy of the data analysis.

Authorship. All named authors meet the International Committee of Medical Journal Editors (ICMJE) criteria for authorship for this article, take responsibility for the integrity of the work as a whole and have given their approval for this version to be published.

Medical Writing, Editorial, and other Assistance. Medical writing support was provided by Rajiv Sarkar on behalf of IQVIA, India, funded by F. Hoffmann-La Roche, Ltd./Genentech, Inc.

Disclosures. Giovanni Ferrara has received fees for lectures from Boehringer Ingelheim and Roche. Lisen Arnheim-Dahlström was an employee of IQVIA, now an employee of Celgene. She has received research funding for Karolinska Institutet from GlaxoSmithKline, MSD and SPMSD. Karen Bartley is an employee of Roche-Genentech and holds Roche-Genentech shares. Christer Janson has received consultancy or speaker fees from AstraZeneca, Boehringer Ingelheim, GlaxoSmithKline, Meda, Novartis and TEVA. Klaus-Uwe Kirchgässler is an employee of F. Hoffmann-La Roche Ltd. and holds Roche shares. Aaron Levine is an employee of IQVIA. C. Magnus Sköld has received research grants from Boehringer Ingelheim, F. Hoffmann-La Roche, Ltd. and Sandoz, and consultancy or speaker fees from

AstraZeneca, Boehringer Ingelheim, GlaxoSmithKline, Meda, Novartis, Mundipharma, Sandoz, Chiesi, Almirall, InterMune and F. Hoffmann-La Roche, Ltd.

Compliance with Ethics Guidelines. This is a retrospective observational study that used anonymised patient data from the Swedish national health registries and primary care EMRs. The study was approved by the Stockholm County Ethical Committee (Regionala etikprövningsnämnden i Stockholm) with reference number 2016/1364-31/2 dated 17 August 2016.

Data Availability. The data sets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Open Access. This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any non-commercial use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

REFERENCES

1. Raghu G, Collard HR, Egan JJ, Martinez FJ, Behr J, Brown KK, et al. An official ATS/ERS/JRS/ALAT statement: idiopathic pulmonary fibrosis: evidence-based guidelines for diagnosis and management. *Am J Respir Crit Care Med.* 2011;183(6):788–824.
2. Hutchinson J, Fogarty A, Hubbard R, McKeever T. Global incidence and mortality of idiopathic pulmonary fibrosis: a systematic review. *Eur Respir J.* 2015;46(3):795–806.
3. Puglisi S, Torrisi SE, Vindigni V, Giuliano R, Palmucci S, Mule M, et al. New perspectives on management of idiopathic pulmonary fibrosis. *Ther Adv Chronic Dis.* 2016;7(2):108–20.
4. Ley B, Collard HR, King TE Jr. Clinical course and prediction of survival in idiopathic pulmonary

- fibrosis. *Am J Respir Crit Care Med*. 2011;183(4):431–40.
5. Raghu G, Freudenberger TD, Yang S, Curtis JR, Spada C, Hayes J, et al. High prevalence of abnormal acid gastro-oesophageal reflux in idiopathic pulmonary fibrosis. *Eur Respir J*. 2006;27(1):136–42.
 6. Lancaster LH, Mason WR, Parnell JA, Rice TW, Loyd JE, Milstone AP, et al. Obstructive sleep apnea is common in idiopathic pulmonary fibrosis. *Chest*. 2009;136(3):772–8.
 7. Collard HR, Ward AJ, Lanes S, Courtney Hayflinger D, Rosenberg DM, Hunsche E. Burden of illness in idiopathic pulmonary fibrosis. *J Med Econ*. 2012;15(5):829–35.
 8. Hubbard RB, Smith C, Le Jeune I, Gribbin J, Fogarty AW. The association between idiopathic pulmonary fibrosis and vascular disease: a population-based study. *Am J Respir Crit Care Med*. 2008;178(12):1257–61.
 9. Ferrara G, Carlson L, Palm A, Einarsson J, Olivesten C, Skold M. Idiopathic pulmonary fibrosis in Sweden: report from the first year of activity of the Swedish IPF-Registry. *Eur Clin Respir J*. 2016;3:31090.
 10. Hodgson U, Laitinen T, Tukiainen P. Nationwide prevalence of sporadic and familial idiopathic pulmonary fibrosis: evidence of founder effect among multiplex families in Finland. *Thorax*. 2002;57(4):338–42.
 11. National Board of Health and Welfare (NBHW). Hälsodataregister Stockholm: NBHW. <http://www.socialstyrelsen.se/register/halsodataregister>. Cited 2018 Mar 22.
 12. Brooke HL, Talback M, Hornblad J, Johansson LA, Ludvigsson JF, Druid H, et al. The Swedish cause of death register. *Eur J Epidemiol*. 2017;32(9):765–73.
 13. Larsson K, Janson C, Lisspers K, Jorgensen L, Stratalis G, Telg G, et al. Combination of budesonide/formoterol more effective than fluticasone/salmeterol in preventing exacerbations in chronic obstructive pulmonary disease: the PATHOS study. *J Intern Med*. 2013;273(6):584–94.
 14. Martinell M, Stalhammar J, Hallqvist J. Automated data extraction—a feasible way to construct patient registers of primary care utilization. *Ups J Med Sci*. 2012;117(1):52–6.
 15. Ringborg A, Yin DD, Martinell M, Stalhammar J, Lindgren P. The impact of acute myocardial infarction and stroke on health care costs in patients with type 2 diabetes in Sweden. *Eur J Cardiovasc Prev Rehabil*. 2009;16(5):576–82.
 16. Hedenstrom H, Malmberg P, Fridriksson HV. Reference values for lung function tests in men: regression equations with smoking variables. *Ups J Med Sci*. 1986;91(3):299–310.
 17. Franzen S, Janson C, Larsson K, Petzold M, Olsson U, Magnusson G, et al. Evaluation of the use of Swedish integrated electronic health records and register health care data as support clinical trials in severe asthma: the PACEHR study. *Respir Res*. 2016;17(1):152.
 18. Lisspers K, Larsson K, Johansson G, Janson C, Costa-Scharplatz M, Gruenberger JB, et al. Economic burden of COPD in a Swedish cohort: the ARCTIC study. *Int J Chron Obstruct Pulmon Dis*. 2018;13:275–85. <https://doi.org/10.2147/COPD.S149633>.
 19. Rockberg J, Jorgensen L, Taylor B, Sobocki P, Johansson G. Risk of mortality and recurrent cardiovascular events in patients with acute coronary syndromes on high intensity statin treatment. *Prev Med Rep*. 2017;6:203–9.
 20. Janson C, Larsson K, Lisspers KH, Stallberg B, Stratalis G, Goike H, et al. Pneumonia and pneumonia related mortality in patients with COPD treated with fixed combinations of inhaled corticosteroid and long acting beta2 agonist: observational matched cohort study (PATHOS). *BMJ*. 2013;346:f3306.
 21. Kjeldsen SE, Stalhammar J, Hasvold P, Bodegard J, Olsson U, Russell D. Effects of losartan vs candesartan in reducing cardiovascular events in the primary treatment of hypertension. *J Hum Hypertens*. 2010;24(4):263–73.
 22. Sorensen HT, Sabroe S, Olsen J. A framework for evaluation of secondary data sources for epidemiological research. *Int J Epidemiol*. 1996;25(2):435–42.
 23. Wasserman RC. Electronic medical records (EMRs), epidemiology, and epistemology: reflections on EMRs and future pediatric clinical research. *Acad Pediatr*. 2011;11(4):280–7.
 24. Nalysnyk L, Cid-Ruzafa J, Rotella P, Esser D. Incidence and prevalence of idiopathic pulmonary fibrosis: review of the literature. *Eur Respir Rev*. 2012;21(126):355–61.
 25. Ley B, Collard HR. Epidemiology of idiopathic pulmonary fibrosis. *Clin Epidemiol*. 2013;5:483–92.
 26. Esposito DB, Lanes S, Donneyong M, Holick CN, Lasky JA, Lederer D, et al. Idiopathic pulmonary fibrosis in United States automated claims. Incidence, prevalence, and algorithm validation. *Am J Respir Crit Care Med*. 2015;192(10):1200–7.

27. Ley B, Urbania T, Husson G, Vittinghoff E, Brush DR, Eisner MD, et al. Code-based diagnostic algorithms for idiopathic pulmonary fibrosis case validation and improvement. *Ann Am Thorac Soc*. 2017;14(6):880–7.
28. Raghu G, Weycker D, Edelsberg J, Bradford WZ, Oster G. Incidence and prevalence of idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med*. 2006;174(7):810–6.
29. Lai CC, Wang CY, Lu HM, Chen L, Teng NC, Yan YH, et al. Idiopathic pulmonary fibrosis in Taiwan—a population-based study. *Respir Med*. 2012;106(11):1566–74.
30. Olson AL, Swigris JJ, Lezotte DC, Norris JM, Wilson CG, Brown KK. Mortality from pulmonary fibrosis increased in the United States from 1992 to 2003. *Am J Respir Crit Care Med*. 2007;176(3):277–84.
31. Navaratnam V, Fleming KM, West J, Smith CJ, Jenkins RG, Fogarty A, et al. The rising incidence of idiopathic pulmonary fibrosis in the U.K. *Thorax*. 2011;66(6):462–7.
32. Fernandez Perez ER, Daniels CE, Schroeder DR, St Sauver J, Hartman TE, Bartholmai BJ, et al. Incidence, prevalence, and clinical course of idiopathic pulmonary fibrosis: a population-based study. *Chest*. 2010;137(1):129–37.
33. Raghu G, Chen SY, Hou Q, Yeh WS, Collard HR. Incidence and prevalence of idiopathic pulmonary fibrosis in US adults 18–64 years old. *Eur Respir J*. 2016;48(1):179–86.
34. von Plessen C, Grinde O, Gulsvik A. Incidence and prevalence of cryptogenic fibrosing alveolitis in a Norwegian community. *Respir Med*. 2003;97(4):428–35.
35. Baumgartner KB, Samet JM, Coultas DB, Stidley CA, Hunt WC, Colby TV, et al. Occupational and environmental risk factors for idiopathic pulmonary fibrosis: a multicenter case-control study. Collaborating centers. *Am J Epidemiol*. 2000;152(4):307–15.
36. Awadalla NJ, Hegazy A, Elmetwally RA, Wahby I. Occupational and environmental risk factors for idiopathic pulmonary fibrosis in Egypt: a multicenter case-control study. *Int J Occup Environ Med*. 2012;3(3):107–16.
37. Miyake Y, Sasaki S, Yokoyama T, Chida K, Azuma A, Suda T, et al. Occupational and environmental factors and idiopathic pulmonary fibrosis in Japan. *Ann Occup Hyg*. 2005;49(3):259–65.
38. Raghu G, Chen SY, Yeh WS, Maroni B, Li Q, Lee YC, 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(7):566–72.
39. Lamas DJ, Kawut SM, Bagiella E, Philip N, Arcasoy SM, Lederer DJ. Delayed access and survival in idiopathic pulmonary fibrosis: a cohort study. *Am J Respir Crit Care Med*. 2011;184(7):842–7.
40. Nathan SD, Albera C, Bradford WZ, Costabel U, Glaspole I, Glassberg MK, et al. Effect of pirfenidone on mortality: pooled analyses and meta-analyses of clinical trials in idiopathic pulmonary fibrosis. *Lancet Respir Med*. 2017;5(1):33–41.
41. Richeldi L, Cottin V, du Bois RM, Selman M, Kimura T, Bailes Z, et al. Nintedanib in patients with idiopathic pulmonary fibrosis: combined evidence from the TOMORROW and INPULSIS(R) trials. *Respir Med*. 2016;113:74–9.
42. Gribbin J, Hubbard RB, Le Jeune I, Smith CJ, West J, Tata LJ. Incidence and mortality of idiopathic pulmonary fibrosis and sarcoidosis in the UK. *Thorax*. 2006;61(11):980–5.
43. Mannino DM, Etzel RA, Parrish RG. Pulmonary fibrosis deaths in the United States, 1979–1991. An analysis of multiple-cause mortality data. *Am J Respir Crit Care Med*. 1996;153(5):1548–52.
44. Hubbard R, Johnston I, Coultas DB, Britton J. Mortality rates from cryptogenic fibrosing alveolitis in seven countries. *Thorax*. 1996;51(7):711–6.
45. Behr J, Kreuter M, Hoepfer MM, Wirtz H, Klotsche J, Koschel D, et al. Management of patients with idiopathic pulmonary fibrosis in clinical practice: the INSIGHTS-IPF registry. *Eur Respir J*. 2015;46(1):186–96.
46. Raghu G, Amatto VC, Behr J, Stowasser S. Comorbidities in idiopathic pulmonary fibrosis patients: a systematic literature review. *Eur Respir J*. 2015;46(4):1113–30.