Journal Pre-proof

Incidence and Prevalence of Progressive Pulmonary Fibrosis (PPF): A Systematic Literature Review and Meta-analysis

Negar Golchin MPH PharmD , Tamara Lesperance MPH , Julia Scheuring MSc , Victoria Wan MSc , Kimberly Hofer BSc , Jean-Paul Collet MD MSc PhD , Aditya Patel MD MBA

PII: S0002-9629(26)00099-6
DOI: https://doi.org/10.1016/j.amjms.2026.02.015
Reference: AMJMS 2495

To appear in: *The American Journal of the Medical Sciences*

Received date: 5 November 2025
Accepted date: 23 February 2026

Please cite this article as: Negar Golchin MPH PharmD , Tamara Lesperance MPH , Julia Scheuring MSc , Victoria Wan MSc , Kimberly Hofer BSc , Jean-Paul Collet MD MSc PhD , Aditya Patel MD MBA , Incidence and Prevalence of Progressive Pulmonary Fibrosis (PPF): A Systematic Literature Review and Meta-analysis, *The American Journal of the Medical Sciences* (2026), doi: https://doi.org/10.1016/j.amjms.2026.02.015

Incidence and Prevalence of Progressive Pulmonary Fibrosis (PPF): A Systematic Literature Review and Meta-analysis

Short Title: Incidence and Prevalence of Progressive Pulmonary Fibrosis

Negar Golchin MPH PharmD[1], Tamara Lesperance MPH[1], Julia Scheuring MSc[1], Victoria Wan MSc[2], Kimberly Hofer BSc[2], Jean-Paul Collet MD MSc PhD[2], Aditya Patel MD MBA[1]*

[1] Bristol Myers Squibb, Lawrenceville, NJ, USA

[2] Evidinno Outcomes Research Inc., Vancouver, BC, Canada

***Corresponding Author:**

Negar Golchin, PharmD MPH

3401 Princeton Pike, Lawrence Township, NJ 08648, United States

negar.golchin@bms.com

+1 (425) 301-6931

Competing Interests

Negar Golchin, Julia Scheuring, Tamara Lesperance, and Aditya Patel are employees and/or shareholders of Bristol Myers Squibb; Victoria Wan, Kimberly Hofer, and Jean-Paul Collet are employees of Evidinno Outcomes Research Inc. (Vancouver, BC, Canada), which was contracted

by Bristol Myers Squibb to conduct this study.

Funding

This study was funded by Bristol Myers Squibb.

Abstract

Background: Progressive pulmonary fibrosis (PPF) has only recently been defined. This study aimed to estimate the global incidence and prevalence of PPF and assess regional variations.

Methods: MEDLINE®, Embase, and Cochrane Database of Systematic Reviews were searched from 01/01/2000 to 11/07/2023 for English-language studies reporting incidence or prevalence of PPF. A DerSimonian-and-Laird random-effects model was used to calculate pooled weighted-mean incidence/prevalence estimates.

Results: Of 3,823 abstracts, five studies were included for meta-analysis. Pooled global incidence across four studies was 10.9/100,000 (95% confidence interval: 1.3–20.5). In Europe, pooled incidence across two studies was 6.7/100,000 (2.2–11.3). Reported incidence rates were 2.4/100,000 in South Korea and 27.6/100,000 in the United States (US). Pooled global prevalence across all studies was 37.0/100,000 (23.6–50.5). Within Europe, pooled prevalence from two studies was 30.2/100,000 (0.0–61.9); prevalence estimates were 6.4 in South Korea, and 57.8 and 60.7 in two US studies.

Conclusions: PPF appears rare globally, with higher incidence and prevalence in the US.

Standardized definitions are needed for more accurate estimates.

Keywords: Pulmonary fibrosis; lung diseases; Progressive pulmonary fibrosis; non-idiopathic pulmonary fibrosis; interstitial lung disease; Epidemiology; Incidence; Prevalence

Introduction

Interstitial lung diseases (ILDs), also known as diffuse parenchymal lung diseases, represent a heterogenous and potentially fatal group of restrictive lung conditions with overlapping clinical presentations and lung injury patterns.[1] A prototypical and archetypal form of ILD is idiopathic pulmonary fibrosis (IPF), characterized by worsening dyspnea, frequent exacerbations and reduced lung function, with a poor prognosis and early mortality.[2,3] Of patients with an ILD other than IPF, it has been estimated that over half will exhibit a benign course of stable chronic disease, showing improvement following immunomodulatory therapy.[1] However, other patients with non-IPF ILD may experience progressive fibrosis, referred to as PPF, which presents similarly to untreated IPF despite appropriate treatment. Notably, PPF has been historically described as non-IPF progressive fibrosing ILD (non-IPF PF-ILD).[4] This older terminology, however, is associated with less standardization in the literature, as studies on non-IPF PF-ILD employed inconsistent methods for evaluating progression such as considering only pulmonary function versus a blend of both clinical symptoms and imaging.[5-7]

PPF has only recently been classified in guidelines as a manifestation of ILD, defined in patients with a non-IPF ILD as having had at least two of the following three criteria occurring within the past year with no alternative explanation: 1) worsening respiratory symptoms, 2) physiological evidence of disease progression based on objective findings of lung function decline, and 3) radiological evidence of disease.[3] Thus, PPF represents a subset of fibrotic ILDs that is distinct from the more commonly occurring IPF, although they share common features with overlapping

symptoms.[3,8] Progressive clinical burden due to decline in respiratory function, dyspnea, and fatigue associated with PPF have a profound impact on quality of life, with substantial humanistic and economic burden, and a high risk of early death.[1,6,9,10] Patients with PPF have a median survival time of three to five years following diagnosis and a five-year survival rate of approximately 40% to 50%.[11,12]

Given that the criteria for PPF have only recently been agreed upon and published in a consensus clinical practice guideline,[3] clear epidemiological information is still rare. In addition, there is a notable lack of specific ICD coding for identifying patients with PPF in databases outside of the United States (US).[11,13,14] This systematic literature review and meta-analysis sought to provide an up-to-date quantitative estimate of the incidence and prevalence of PPF globally and across different regions, while exploring factors contributing to variability across studies.

Methods

This systematic review and meta-analysis followed standard accepted methodologies for conducting and reporting systematic reviews as recommended by the Cochrane Handbook for Systematic Reviews of Interventions,[15] the Joanna Briggs Institute (JBI) methodological guidance for incidence and prevalence data from systematic reviews,[16] and the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA).[17]

Study Eligibility

Study eligibility criteria were predefined using the CoCoPop framework (Condition, Context, Population).[16] Eligible studies reported on the incidence and/or prevalence of PPF among adults (typically aged ≥18 years) from the general population in any country. Only epidemiological studies reporting population-level incidence or prevalence data were included. Searches were restricted to studies published in English.

Studies were excluded if they focused on pediatric populations, evaluated conditions other than PPF, or reported outcomes not relevant to incidence or prevalence. In addition, studies were excluded when the study design or data source was not appropriate for estimating population-level epidemiology, such as non-original research (e.g., reviews, editorials, case reports), or when insufficient data were available for extraction or quantitative synthesis.

Data Sources

Relevant studies were identified by searching MEDLINE®, Embase, and the Cochrane Database of Systematic Reviews via the Ovid platform using pre-defined search strategies. Databases were searched from January 1, 2000, to November 7, 2023, to focus on recent data. Gray literature searches were also conducted to capture studies from sources that were not included in the main literature databases. Abstracts from the American Thoracic Society, British Thoracic Society, European Respiratory Society, Canadian Society of Respiratory Therapists Annual Conference, American College of Rheumatology, and European Congress of Rheumatology were searched for the years 2021 – 2023. Bibliographies of identified literature reviews were also searched manually. Search strategies for MEDLINE® and Embase are provided in **Supplementary Table S1** and **Supplementary Table S2**, respectively.

Study Selection, Data Extraction, and Study Quality Appraisal

All identified abstracts and conference proceedings were reviewed by a senior reviewer according to eligibility criteria. In a subsequent screening round, selected studies were submitted to full-text screening for eligibility by two independent reviewers. Any unresolved discrepancies in study selection between the two reviewers were resolved by a third reviewer. A senior reviewer extracted all pre-specified study and population characteristics from the final list of included studies, and two independent reviewers extracted all relevant outcomes data from the final list of included studies. Data on study characteristics (e.g., study design, study region, study period), population characteristics (e.g., age, sex, race, accompanying disease context), and outcomes (e.g., incidence, prevalence, numerator and/or denominator, i.e., total population or

sample size) were extracted.

Investigators assessed the quality of the included studies using the Johanna Briggs Institute checklist for prevalence studies.[18]

Statistical Analysis

Where necessary, data provided by included studies were used to calculate denominators, numerators, as well as 95% confidence intervals (CI). When studies offered both crude and adjusted rates, the adjusted rate (e.g., adjusted for sex and age) was selected. Reported rates were converted to rates per 100,000 persons as needed. For studies that reported incidence and prevalence by year or gender, these values were averaged according to the population size for each year or the sample size for each gender to produce a single rate per study.

The meta-analysis was performed using the DerSimonian-and-Laird random-effects model, a version of the inverse variance method endorsed by the Cochrane Handbook.[15,19] The analysis was stratified by regions or globally, with pooled rates categorized by region where possible.

To calculate random-effects pooled estimates, variance in the reported rates was required. For studies that did not provide a 95% confidence interval (CI), standard errors were calculated using the number of cases of PPF and population size, if available. In studies reporting annual incidence

or prevalence, a weighted average over the study period was calculated for the meta-analysis, while reported point estimates for a specific study period were used as is for quantitative analysis

The degree of heterogeneity among the studies in the meta-analysis was assessed using the Cochrane I^2 statistic. Where possible, potential sources of heterogeneity were explored. Pooled weighted random-effects incidence and prevalence estimates were generated using the *metafor* R-package.[20]

Results

Study Selection

In total, 3,823 non-duplicate records were identified from the database searches (**Figure 1**). Following full-text screening, six studies were included in the review. Of these, one study by Alaws et al. (2023)[21] was only available as a conference abstract at the time of this review; authors reported the incidence rate of acute interstitial pneumonitis, an acute and progressive type of ILD, stratified by age and sex. This study also did not report the total sample size; for these reasons, this study was excluded from the meta-analysis. Five studies were available for meta-analysis.[11,13,14,22,23]

Figure 1: PRISMA flow diagram

Study Characteristics and Study Quality Appraisal

The characteristics of the five included studies are presented in **Table 1**. They were all retrospective studies conducted between 2010 and 2019, generally well designed to select relevant patients and detect fibrosis progression. The length of the study period ranged from 5 to 8 years with an overall median of 5 years. The source population size ranged from just under 3 million[23] to approximately 65 million people.[11] One study in France[11] and one in South Korea[14] used the country's general population as denominator (>97% of the population). The two US studies[22,23] used different claims databases: Olson et al. used an insurance database (IBM MarketScan)[22] while Singer et al. used data from both a private insurance (Optum) and Medicare Advantage with Part D[23] (**Table 1**).

Definitions and diagnostics criteria for progressive pulmonary fibrosis are shown in **Table 2**. All included studies used ICD codes to identify ILD and related conditions combined with proxy criteria likely associated with progression of pulmonary fibrosis (for instance increased prescription of respiratory function tests or lung CT scanner). Finally, Hilberg's study (2022),[13] was carried out across six European countries (Belgium, Denmark, Finland, Greece, Norway and Portugal) using different sub-populations from hospital databases, specialized centers, regional and national registries. The investigators also used ICD codes to identify candidate patients with PPF in a first phase, followed by a second phase with manual review by clinicians to validate cases of progressive fibrosis using a combination of a relative decline in lung function measured by

forced vital capacity (FVC) and/or radiological progression,[13] referred to as clinical diagnosis.

Regarding quality, studies were generally well designed, with noted transparency on decisions regarding the algorithms to select patients and detect fibrosis progression. However, most of these algorithms were not validated and may have identified patients that did not have PPF, with a possible inflation of the incidence or prevalence rates by inclusion of false positive cases. Otherwise, studies used appropriate sampling of participants and conducted appropriate statistical analyses. Two studies[14,21] did not include a measure of dispersion when reporting incidence or prevalence. A summary of the assessment using the JBI quality of assessment tool for studies included in the meta-analysis is presented in (**Supplementary Table S3**).

Incidence of PPF

Four studies reported the incidence of patients with PPF.[11,13,14,22] Pooled weighted global annual incidence of PPF per 100,000 across the four studies was 10.9 ([95% CI: 1.3, 20.5]; I^2: 100%, **Figure 2**). Regionally, the one study from Asia[14] reported a lower annual incidence rate (2.4 [95% CI: 2.3, 2.6]) than the pooled rate across the two European studies (6.7 [2.2, 11.3]; I^2: 98.2%).[11,13] The annual incidence rate from the one US study was 27.6 (95% CI: 27.1, 28.0),[22] much higher than the pooled incidence rate (5.2 [3.2, 7.2]) across the three studies from other countries[11,13,14] (data not shown).

Figure 2: Forest plot of incidence of PPF by region (n = 4)

Prevalence of PPF

Five studies reported the average annual prevalence rate of PPF per 100,000 persons using two types of source populations (denominators): general population[11,14] and more narrowly defined or identified sub-populations.[22,23] Joung et al. (2023) from South Korea reported the lowest prevalence of 6.4 ([95% CI: 6.2, 6.6]).[14] Hilberg's study used different source populations from the participating countries.[13] Pooled global annual prevalence of PPF across the five studies was 37.0 ([95% CI: 23.6, 50.5]; I^2: 99.7%; **Figure 3**). The pooled prevalence across the two US studies (59.3 [95% CI: 56.4, 62.1]; I^2: 90.1%; data not shown) was higher than that for studies conducted in either Europe (14.1 [13.8, 14.4], I^2: 0.0%) or pooled globally.

The wide 95% CI of the European pooled prevalence estimate of 30.2 ([95% CI: 0.0, 61.9]; I^2: 99.7%) is related to the large difference of reported PPF prevalence between the two European studies (**Figure 3**).[11,13]

Figure 3: Forest plot of prevalence of PPF by region (n = 5)

Discussion

The current study is based on systematically identifying related literature published between January 2000 and November 2023, with the included studies conducted between 2010 and 2019 from US, Europe and South Korea. The pooled global incidence (10.9 per 100,000) and prevalence (37.0 per 100,000) estimates of PPF suggest that it is a rare condition. Another important aspect of our results is the important heterogeneity across countries with US estimates for PPF incidence and prevalence much higher than those from Europe and South Korea. Several factors may explain the general heterogeneity observed between studies and more specifically the difference between US estimates and estimates from other countries.

Although the two US studies are methodologically sound, their incidence and prevalence estimates are likely overestimated due to the limited representativeness of their denominators. Olson et al.[22] used the IBM MarketScan database, which includes employees and their families covered by employer-sponsored insurance. This population does not fully reflect the US demographic distribution in terms of urban-rural residence, race and age.[24] Similarly, Singer et al.[23] analyzed data from Optum's private insurance database and Medicare Advantage with Part D, which covers individuals aged 65 and older – representing only about 20% of the US population. In contrast, studies by Joung et al. (2023)[14] and Nasser et al. (2021)[11] used data from near-complete national population, covering 98.8% of France and 97% of South Korea, respectively. The broader population coverage in these studies may explain why their estimates are lower than those reported in the US studies.

All included studies in the meta-analysis identified patients with a high probability of PPF using ICD codes from large administrative claims databases.[11,13,14,22,23] During the study periods reported here, non-specific ICD-9,[22] ICD-10,[11,13,14] and ICD-10 CM[23] codes were incorporated into the algorithms. A major challenge in each study was the absence of a specific ICD code for PPF. Furthermore, to identify the progressive behavior of fibrosis, the algorithms relied on available databases information, including the number and frequency of prescribed tests, imaging or specific procedure such as home oxygen. However, assessing progression based on prescriptions is potentially flawed, as claims databases do not provide test results, which may lead to misclassifications. Finally, the inclusion of ICD codes from various systemic conditions that are risk factors for PPF, such as sarcoidosis, rheumatoid arthritis, or unclassifiable interstitial lung diseases, can only help in narrowing down to the likely diagnosis, without providing certainty.

Furthermore, another source of uncertainty in identifying PPF cases is the potential for incorrect entry of ICD codes or later updates to a different diagnosis. Different approaches have been used to mitigate this challenge. The study by Hilberg et al. (2022) included a manual review by clinicians of records identified through ICD-9/10 codes.[13] Nasser et al. (2021)[11] described a preliminary validation of their third algorithm (hospital-related) to identify PPF cases from the database, although the results showed only 46% sensitivity and 63% specificity. Finally, Singer et al. (2022) attempted to minimize this type of error by requiring that two diagnostic codes were entered within 365 days of each other.[23] Another possible source of heterogeneity in the study results may stem from the publication of a new international clinical practice guideline

introducing new diagnostic criteria for PPF.[3] This could influence clinical practice and increase disparity between studies published before and after its release. However, using ICD codes to assess PPF does not allow for evaluating the specific criteria clinicians used for diagnosis. In particular, the requirement that patients meet two out of three criteria outlined in the guideline cannot be verified in large administrative databases - a limitation affecting all studies except Hilberg's study that used hospital databases.[13] Stratifying studies by pre- and post-guideline publication periods to evaluate whether newer diagnostic frameworks enhance detection was not feasible due to the limited evidence base. Whether these guidelines have improved case identification and led to more consistent epidemiological trends remains an open question.

Among studies conducted in Europe, Hilberg et al.[13] stands out for implementing a standardized protocol across six European countries, using hospital databases to identify PPF. Clinical validation by clinicians was conducted on a sample of cases including indicators of progressive fibrosis such as a decline in FVC of 10% or greater over two years, a change in high-resolution computed tomography (HRCT) showing increased fibrosis, starting the use of oxygen, or death due to a respiratory event. However, the clinical validation showed that the algorithms over-selected PPF cases and exhibited considerable variation across countries; for instance, the positive predictive value ranged from 49% in Finland to 75% in Belgium and Norway. This variability likely reflects differences in databases quality and source populations across countries. For instance, data coming from Denmark, Finland, and Norway benefited from a national reference center in each country, with the denominator being the entire adult population aged

18 years or older. In contrast, data from Portugal and Greece were collected from multiple hospitals with undefined catchment areas, potentially affecting the validity of their figures. These differences likely contributed to heterogeneity in the results reported by Hilberg.

To our knowledge, this systematic review and meta-analysis is the first to provide a quantitative assessment of global and regional incidence and prevalence of PPF in adult populations. A recent review by Cottin et al. (2022) presented a qualitative review of published incidence of PPF ranging from 2.1 to 32.6 per 100,000 person-years (n=4) and prevalence ranging from 6.9 to 70.3 per 100,000 persons (n=3).[25] The present study identified PPF incidence estimates similarly ranging from 2.4 to 27.6 per 100,000 person-years (n=4), and prevalence similarly ranging from 6.4 to 60.7 per 100,000 persons (n=5). Notably, our more recent review captured three additional studies[13,14,23] that were published after the literature search reported by Cottin. Finally, the review by Cottin et al. included a study from Australia by Sweeney et al. (2020)[26] that did not meet our inclusion criteria for appropriate study design because it was a commentary.

This study provides the most comprehensive and up-to-date assessment of the literature on incidence and prevalence of PPF. We provide a quantitative synthesis of the available data through meta-analysis, revealing certain regional trends that bear further study. An important limitation to consider is that for both incidence and prevalence estimates, the Cochrane I^2 statistic of the pooled studies showed considerable heterogeneity, as per the guideline from the Cochrane Handbook for Systematic Reviews of Interventions.[15] The use of non-clinically validated

algorithms is prone to introduce biases and misclassifications contributing to the large heterogeneity in point estimates of incidence and prevalence. In addition, estimates from databases that capture specific sub-populations (e.g., elderly, Medicare in the US) introduce a systematic selection bias that likely overestimates incidence and prevalence. The study by Olson et al.[22] acknowledged limitations in their use of ICD codes that may have included some patients with IPF. In addition, only records published in English were included; data on PPF epidemiology from non-English speaking countries may have been missed.

Conclusions

Overall, this review shows that developing validated algorithms using specific ICD codes for the identification of ILD, and specifically PPF, remains a challenge that affects the validity of all epidemiological studies. This in turn makes it difficult to compare incidence and prevalence estimates across the current literature.

Global estimates of PPF suggest that it is a rare condition. Furthermore, this study, albeit with a limited evidence base, showed higher incidence and prevalence of PPF in the US versus other countries at the population level. A contributing factor to this finding may be related to the overestimation of incidence and prevalence caused by lower denominators from claim database data that did not represent the general US populations. PPF is a complex form of ILD, with various risk factors affecting progression. Further standardization in the definition and diagnostic criteria

for PPF is required for more accurate estimates of incidence and prevalence.

Declarations

Ethics Approval and Consent to Participate

Not applicable

Consent for Publication

Not applicable

Availability of Data and Materials

The datasets used and analyzed during the current study are available from the corresponding author upon reasonable request.

Authors' Contributions

NG, TL, JS, JPC, and AP were involved in the conception and design of the work. VW, KH, and JPC acquired and analyzed the prevalence and incidence data and were major contributors in the writing of the original manuscript. All authors were involved in interpretation of the prevalence and incidence data and editing of the draft manuscripts. All authors read and approved the final manuscript.

Acknowledgements

The authors would like to thank Christopher Crotty, professional Medical Writer of Evidinno Outcomes Research Inc. (Vancouver, BC, Canada) for medical writing assistance.

References

1. **Cottin V, Wollin L, Fischer A, Quaresma M, Stowasser S, Harari S**. Fibrosing interstitial lung diseases: knowns and unknowns. *European respiratory review : an official journal of the European Respiratory Society.* 2019;28(151).

2. **Lederer DJ, Martinez FJ**. Idiopathic Pulmonary Fibrosis. *N Engl J Med.* 2018;378(19):1811-1823.

3. **Raghu G, Remy-Jardin M, Richeldi L**, et al. Idiopathic pulmonary fibrosis (an update) and progressive pulmonary fibrosis in adults: an official ATS/ERS/JRS/ALAT clinical practice guideline. *American Journal of Respiratory and Critical Care Medicine.* 2022;205(9):e18-e47.

4. **Kang HK, Song JW**. Progressive Pulmonary Fibrosis: Where Are We Now? *Tuberculosis and respiratory diseases.* 2024;87(2):123-133.

5. **Behr J, Prasse A, Kreuter M**, et al. Pirfenidone in patients with progressive fibrotic interstitial lung diseases other than idiopathic pulmonary fibrosis (RELIEF): a double-blind, randomised, placebo-controlled, phase 2b trial. *The Lancet Respiratory medicine.* 2021;9(5):476-486.

6. **Flaherty KR, Wells AU, Cottin V**, et al. Nintedanib in Progressive Fibrosing Interstitial Lung Diseases. *N Engl J Med.* 2019;381(18):1718-1727.

7. **Maher TM, Corte TJ, Fischer A**, et al. Pirfenidone in patients with unclassifiable

progressive fibrosing interstitial lung disease: a double-blind, randomised, placebo-controlled, phase 2 trial. *The Lancet Respiratory medicine.* 2020;8(2):147-157.

8. **Wijsenbeek M, Cottin V**. Spectrum of Fibrotic Lung Diseases. *N Engl J Med.* 2020;383(10):958-968.

9. **Cottin V**. Criteria for Progressive Pulmonary Fibrosis: Getting the Horse Ready for the Cart. *Am J Respir Crit Care Med.* 2023;207(1):11-13.

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

11. **Nasser M, Larrieu S, Boussel L**, et al. Estimates of epidemiology, mortality and disease burden associated with progressive fibrosing interstitial lung disease in France (the PROGRESS study). *Respir Res.* 2021;22(1):162.

12. **Platenburg M, van der Vis JJ, Grutters JC, van Moorsel CHM**. Decreased Survival and Lung Function in Progressive Pulmonary Fibrosis. *Medicina (Kaunas).* 2023;59(2).

13. **Hilberg O, Hoffmann-Vold AM, Smith V**, et al. Epidemiology of interstitial lung diseases and their progressive-fibrosing behaviour in six European countries. *ERJ Open Res.* 2022;8(1).

14. **Joung KI, Park H, Park S, Shin JY, Kim YH**. Nationwide epidemiologic study for fibrosing interstitial lung disease (F-ILD) in South Korea: a population-based study. *BMC Pulm Med.* 2023;23(1):98.

15. **Higgins JPT, Thomas J, Chandler J**, et al. (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.4 (updated August 2023). Cochrane, 2023.

16. **Munn Z, Moola S, Lisy K, Riitano D, Tufanaru C**. Methodological guidance for systematic reviews of observational epidemiological studies reporting prevalence and cumulative incidence data. *Int J Evid Based Healthc.* 2015;13(3):147-153.

17. **Page MJ, McKenzie JE, Bossuyt PM**, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ.* 2021;372:n71.

18. **Munn Z, Moola S, Riitano D, Lisy K**. The development of a critical appraisal tool for use in systematic reviews addressing questions of prevalence. *International journal of health policy and management.* 2014;3(3):123.

19. **DerSimonian R, Laird N**. Meta-analysis in clinical trials. *Control Clin Trials.* 1986;7(3):177-188.

20. **Viechtbauer W**. Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software.* 2010;36(3):1-48.

21. **Alaws H, Ramalingam V, Kumar G**. Epidemiology of Patients Admitted with Acute Interstitial Pneumonitis. *Chest.* 2023;164(4).

22. **Olson AL, Patnaik P, Hartmann N, Bohn RL, Garry EM, Wallace L**. Prevalence and Incidence of Chronic Fibrosing Interstitial Lung Diseases with a Progressive Phenotype in the United States Estimated in a Large Claims Database Analysis. *Adv Ther.* 2021;38(7):4100-4114.

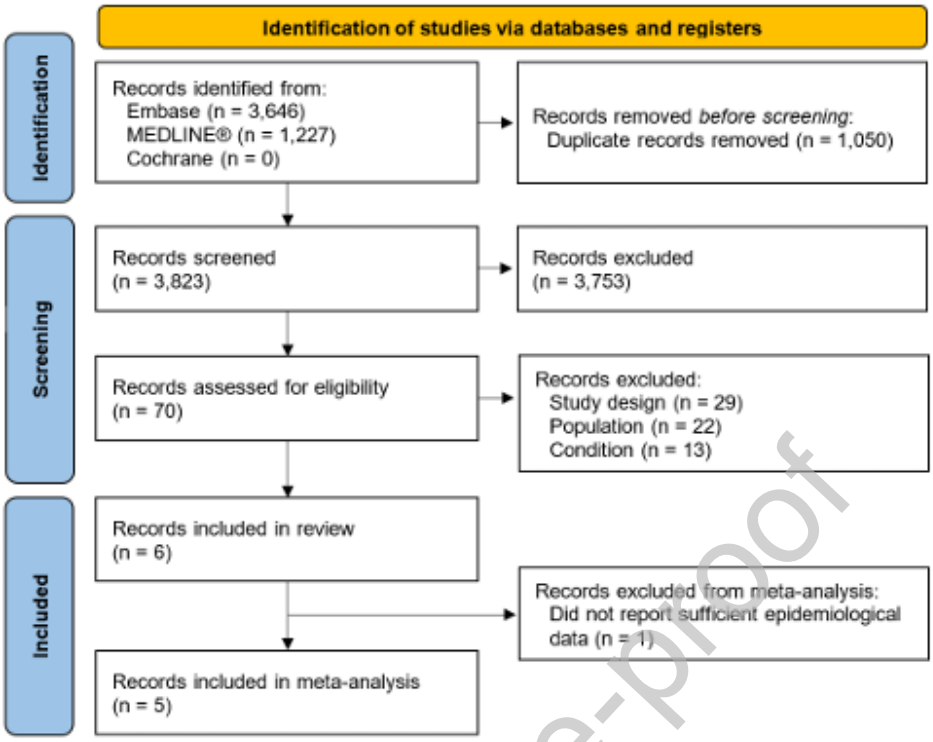
23. **Singer D, Bengtson LGS, Conoscenti CS**, et al. Claims-based Prevalence of Disease Progression among Patients with Fibrosing Interstitial Lung Disease Other than Idiopathic Pulmonary Fibrosis in the United States. *Ann Am Thorac Soc.* 2022;19(7):1112-1121.

24. **Marker DA, Hilton C, Zelko J**, et al. Real World Data Versus Probability Surveys for Estimating Health Conditions at the State Level. *Journal of survey statistics and methodology.* 2024;12(5):1515-1530.

25. **Cottin V, Brown KK, Flaherty KR, Wells AU**. Progressive Pulmonary Fibrosis: Should the Timelines Be Taken Out of the Definition? *Am J Respir Crit Care Med.* 2022;206(10):1293-1294.

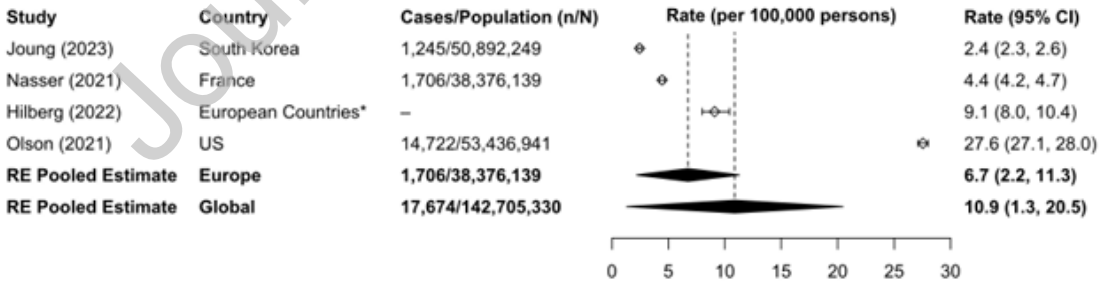
26. **Sweeney DJ, Khor YH, Goh NSL**. The unmet care needs of progressive fibrosing interstitial lung disease. *Respirology.* 2020;25(12):1231-1232.

Figure Legends



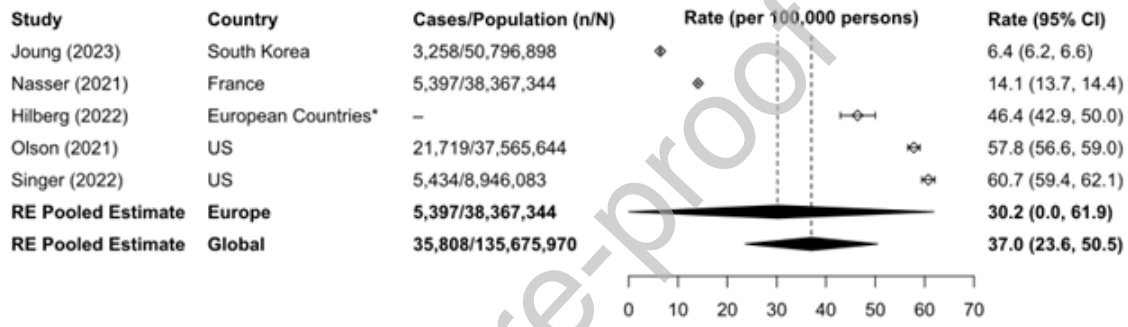
n: number of records.

Figure 1: PRISMA flow diagram



Belgium, Denmark, Finland, Greece, Norway, and Portugal. Europe: European Countries and France. –: Not reported; CI: Confidence interval; PPF: Progressive pulmonary fibrosis; RE: Random effects; US: United States

Figure 2: Forest plot of incidence of PPF by region (n = 4)



Belgium, Denmark, Finland, Greece, Norway, and Portugal. Europe: European Countries and France. –: Not reported; CI: Confidence interval; PPF: Progressive pulmonary fibrosis; RE: Random effects; US: United States.

Figure 3: Forest plot of prevalence of PPF by region (n = 5)

Table 1: Characteristics of the studies included in the meta-analysis (n = 5)

Author (Year); Country	Study Design	Data Source	Time Period	Source Population Size	Source Population Type	Age, Years	Males, n (%)	Comorbidities, n (%)	Disease Context, n (%)
Hilberg (2022); Belgium, Denmark, Finland, Greece, Norway, Portugal	Retrospective	Belgium: Leuven University Hospital, G-hent University Hospital, Liege University Hospital Centre; Denmark: Lilleb aelt	2014-2018	Variable sampling from each country	Different populations from hospital databases, specialized centers, regional and national registries	NR	18+: – (100 %)	NR	NR

		Hospital; Finland: Turku Hospital; Greece: Heraklion University Hospital, University Hospital of Larissa, General Hospital of Thessaloniki, Athens Medical Centre;							

		Norway: Oslo University Hospital; Portugal: Coimbra Hospital and University Centre, São João University Hospital Centre, Vila Nova de Gaia/ Espinho Hospital Centr							

		e, Beatriz Ângelo Hospital							
Joung (2023); South Korea	Retrospective	Korean Health Insurance Review and Assessment (HIRA) database	2011-2018	50,796,898	General population	Overall F-ILD (Mean): 70.6 (SD: 10.7) 18-29: 49 (0.2%) 30-39: 192 (1.0%) 40-49: 544 (2.7%) 50-59: 2,060 (10.3	Overall F-ILD: 12,830 (64.2) PPF: 2784 (52.3)	Overall F-ILD: Charlson comorbidity index: 3.73 PPF: Charlson comorbidity index: 3.99	Overall F-ILD: NR PPF: HP: 57 (1.1) Autoimmune-ILD: 2,484 (46.6) RA-ILD: 1,672 (31.4) SSc-ILD: 296 (5.6) Other CTD-ILD: 420 (7.9) SLE or

						%) 60-69: 5,432 (27.2 %) 70-79: 7,754 (38.8 %) 80+: 3,963 (19.8 %) PPF (Mean): 69.2 (SD: 12) 18-29: 19 (0.4%) 30-39: 89 (1.7%) 40-49:			dermatomyositis ILD: 245 (4.6) Sarcoidosis ILD: 28 (0.5) External ILD: 119 (2.2) Unclassifiable ILD: 2,669 (50.1)

						248 (4.7%) 50-59: 712 (13.4%) 60-69: 1,386 (26.0%) 70-79: 1,814 (34.1%) 80+: 1,057 (19.8%)			
Nasser (2021); France	Retrospective	Système National des Données de Santé	2010-2016	98.8% of the French population of over 66 million people	General population	NR	NR	Arterial hypertension: 9,193 (63.8) Gastroesophageal reflux disease: 7,991 (55.4)	Exposure-related ILD other than HP: 3486 (24.2) IIP: 3113 (21.6)

								Cardiac arrythmias: 3,155 (21.9) Depression: 2,953 (20.5) Congestive heart failure: 2,886 (20.0) Chronic coronary disease: 2,227 (15.5) Lung cancer: 940 (6.5) Anemia: 826 (5.7) Pulmonary hypertension: 765 (5.3)	HP: 728 (5.1) RA-ILD: 2521 (17.5) SSc-ILD: 907 (6.3) Mixed CTD-ILD: 655 (4.5) Other autoimmune ILD: 1,503 (10.4) Sarcoidosis ILD: 1,500 (10.4)

								Diarrhea: 486 (3.4) Digital ulcer: 475 (3.3) Osteoporosis: 399 (2.8) Acute coronary syndrome: 284 (2.0) Cirrhosis: 172 (1.2)	
Olson (2021); US	Retrospective	IBM MarketScan Research Databases	2012-2015	37,565,644	Subpopulation (Selective)	F-ILD: 18-39: 1,040 (2.9%) 40-49: 2,470 (6.9%) 50-	F-ILD: 16,463 (46) PPF: 10,269 (47.3)	NR	F-ILD: NR PPF: Unclassifiable IIP: 15,515 (71.4) Sarcoidosis ILD: 1,160

						59: 6,455 (18.0 %) 60-69: 9,178 (25.6 %) 70-79: 8,955 (24.9 %) 80+: 8,703 (24.3 %)			(5.3) RA-ILD: 1,777 (8.2) SSc-ILD: 618 (2.8) Mixed CTD-ILD: 62 (0.3) Exposure-related ILD other than HP: 481 (2.2) HP: 346 (1.6) Non-specific IIP: 150 (0.7)
Singer (2022); US	Retrospective	Medicare Optum	2016-2019	Medicare: 2,936,729 Commercial: 6,009,363	Sub-population (Sele	Medicare: 18-39:	Medicare: 1,239,487	NR	NR

		Research Database			ctive)	18,218 (0.6%) 40-49: 40,724 (1.4%) 50-59: 134,391 (4.6%) 60-69: 775,141 (26.4%) 70-79: 1,323,328 (45.1%) 80+: 644,918 (22.0%)	(42.2) Commercial: 3,068,327 (51.1)		

						Commercial: 18-39: 2,574,993 (42.8%) 40-49: 1,281,403 (21.3%) 50-59: 1,308,552 (21.8%) 60-69: 729,766 (12.1%) 70-79: 82494 (1.4%) 80+:			

						3215 5 (0.5%)			

sub-population (selective) represents a patient population selected from within a database. ctd: connective tissue disease; f-ild: fibrosing interstitial lung disease; hp: hypersensitivity pneumonitis; iip: idiopathic interstitial pneumonia; ppf: progressive pulmonary fibrosis; ra: rheumatoid arthritis; sd: standard deviation; sle: systemic lupus erythematosus; ssc: systemic sclerosis; us: united states; nr: not reported.

Table 2: Inclusion criteria for studied patients and definitions and diagnostics criteria for PPF across the selected studies (n = 5)

Author (Year); Country	**Inclusion Criteria for Studied Patients**	**Definition or Diagnostic Criteria for Progression (PPF)**
Hilberg (2022); Belgium, Denmark, Finland, Greece, Norway, Portugal	Patients ⩾18 years old listed in the database/registry at each year of the study period identified with ILD by either 9th/10th revision of International Classification of Diseases (ICD-9/ICD-10) codes, local codes and/or keywords	One of the following progression criteria: • Relative decline ≥10% in FVC over 2 years • FVC between 5 and <10% but with any of the following: ≥1 ILD-related hospitalization (excluding emergency visits), increasing extent of fibrosis on HRCT, starting or increasing oxygen use, or death due to respiratory event
Joung (2023); South Korea	Patients with f-ILD were defined as those aged ⩾18 years on the index date and had at least one lung disease diagnosis and one f-ILD diagnosis (ICD-10: J84.1), or at least two f-ILD diagnosis (ICD-10: J84.1) for the year based on ICD-10 diagnostic codes	One of the following progression criteria: • Receiving more than one oxygen therapy • Being hospitalized in internal respiratory medicine, rheumatology, or internal medicine, or visiting the emergency department with the address of ILD • History of lung transplantation • Satisfying the following (a) and (b) and [(c) or (d)] for medical record base on the claim data were

Author (Year); Country	Inclusion Criteria for Studied Patients	Definition or Diagnostic Criteria for Progression (PPF)
	Index date was defined as the first date that satisfies the case definition by the disease spectrum	included in the PFILD: (a) at least three respiratory or rheumatology visits, (b) receiving prescriptions for corticosteroids or immunosuppressants, (c) at least three x-rays and at least three pulmonary function test, (d) at least two HRCT or chest CT exams
Nasser (2021); France	Patients who met the following criteria were included: (1) aged⩾20 years; (2) met the criteria for progression defined; (3) ⩾2-year history in the database prior to index date; and (4) affiliated with the general reimbursement scheme Patients were excluded if they had IPF based on case definitions	One of the following progression criteria: • ≥3 claims each for pulmonologist consultations and pulmonary function tests within 12 months; and glucocorticoid or immunosuppressive therapy; plus palliative care • ≥3 HRCT or chest CT scans • ≥1 claim for oxygen therapy, respiratory hospitalization in an intensive care unit following an emergency visit or lung transplant
Olson (2021); US	Patients aged ⩾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 F-ILD and/or progressive fibrosing ILD. Gaps of up to 30 days in coverage were permitted F-ILD diagnosis was identified using ICD-9 codes	Any of the following proxy progression criteria • ≥2 pulmonary function tests or ≥2 oxygen titration tests within 90 days • ≥2 HRCT or ≥3 chest 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
Singer (2022); US	The study included commercial enrollees and Medicare Advantage with Part D beneficiaries 18 years of age or older with known sex, geographic region, and insurance type, identified between October 1, 2016, and September 30, 2019 (identification period).	At least one or two of the following proxy criteria of progression: • At least 2 pulmonary function tests on different dates of service within 90 days of each other

Author (Year); Country	Inclusion Criteria for Studied Patients	Definition or Diagnostic Criteria for Progression (PPF)
	Continuous health plan enrollment on and during the 12 months preceding September 30, 2019, was required The first step in the algorithm was the identification of patients with non-IPF fibrosing ILD diagnoses using ICD-10-CM diagnosis codes selected by clinical experts (practicing pulmonologists) as being likely to indicate relevant non-IPF ILDs. Patients were required to have at least two non-IPF fibrosing ILD diagnoses on different dates within 365 days during the identification period The date of the first non-IPF fibrosing ILD diagnosis was designated as the index date	• At least 2 oxygen titrations tests on different dates of service within 90 days of each day • At least 2 inpatient or outpatient HRCT scans on different dates of service within 360 days of each other • At least 1 pharmacy claim for an oral corticosteroid with a prednisone-equivalent dose greater than 20mg/d • At least 1 pharmacy claim for a new Immunosuppressive medication • At least one claim for lung transplant • At least one claim for oxygen therapy • At least one claim for palliative care • At least one respiratory hospitalization (inpatient stay with a respiratory-related diagnosis code)

ct: computed tomography; f-ild: fibrosing interstitial lung disease; hrct: high-resolution computer tomograph; icd: international classification of disease; ild: interstitial lung disease; ipf: idiopathic pulmonary fibrosis; ppf: progressive pulmonary fibrosis; us: united states.

Additional Files

File name: Additional File 1.docx. **Title of data**: Supplementary Table S1: Search strategy for MEDLINE® (via OvidSP). **Description**: A supplementary table of the search strategy for MEDLINE database

File name: Additional File 2.docx. **Title of data**: Supplementary Table S2: Search strategy for Embase (via OvidSP). **Description**: A supplementary table of the search strategy for Embase database

File name: Additional File 3.docx. **Title of data**: Supplementary Table S3: Summary of the Joanna Briggs Institute quality assessment for reviewed studies (n = 6). **Description**: A supplementary table of a summary of the Johanna Briggs Institute quality assessment took for the included studies

Competing Interests

Negar Golchin, Julia Scheuring, Tamara Lesperance, and Aditya Patel are employees and/or shareholders of Bristol Myers Squibb; Victoria Wan, Kimberly Hofer, and Jean-Paul Collet are employees of Evidinno Outcomes Research Inc. (Vancouver, BC, Canada), which was contracted by Bristol Myers Squibb to conduct this study.