Epidemiology of Idiopathic Pulmonary Fibrosis among U.S. Veterans, 2010–2019

Bhavika Kaul[1,2], Joyce S. Lee[3], Ning Zhang[2], Eric Vittinghoff[4], Kathleen Sarmiento[1,2], Harold R. Collard[1], and Mary A. Whooley[1,2,4]

[1]Department of Medicine and [4]Department of Epidemiology and Biostatistics, University of California San Francisco, San Francisco, California; [2]Measurement Science Quality Enhancement Research Initiative, San Francisco Veterans Affairs Healthcare System, San Francisco, California; and [3]Department of Medicine, University of Colorado, Aurora, Colorado

Abstract

Rationale: The development of novel therapies for idiopathic pulmonary fibrosis (IPF) has brought increased attention to the population burden of disease. However, little is known about the epidemiology of IPF among U.S. Veterans.

Objectives: This study examines temporal trends in incidence and prevalence, patient characteristics, and risk factors associated with IPF among a national cohort of U.S. Veterans.

Methods: We used data from the Veterans Health Administration (VHA) electronic health record system to describe the incidence, prevalence, and geographic distribution of IPF between January 1, 2010, and December 31, 2019. We evaluated patient characteristics associated with IPF using multivariate logistic regression.

Results: Among 10.7 million Veterans who received care from the VHA between 2010 and 2019, 139,116 (1.26%) were diagnosed with IPF. Using a narrow case definition of IPF, the prevalence increased from 276 cases per 100,000 in 2010 to 725 cases per 100,000 in 2019. The annual incidence increased from 73 cases per 100,000 person-years in 2010 to 210 cases per 100,000 person-years in 2019. Higher absolute incidence and prevalence rates were noted when a broader case definition of IPF was used. Risk factors associated with IPF among Veterans included older age, White race, tobacco use, and rural residence. After accounting for interactions, the average marginal difference in IPF prevalence between males and females was small. There was significant geographic heterogeneity of disease across the United States.

Conclusions: This study is the first comprehensive epidemiologic analysis of IPF among the U.S. Veteran population. The incidence and prevalence of IPF among Veterans has increased over the past decade. The effect of sex on risk of IPF was attenuated once accounting for other risk factors. The geographic distribution of disease is heterogeneous across the United States with rural residence associated with higher odds of IPF. The reasons for these trends deserve further study.

Keywords: epidemiology; Veterans; interstitial lung disease; idiopathic pulmonary fibrosis

(Received in original form March 10, 2021; accepted in final form July 26, 2021)

Supported by the Health Services Research and Development Quality Enhancement Research Initiative (QUE 15-283), Veterans Health Administration, Washington, D.C. Additional support provided by the Nina Ireland Grant for Lung Health and National Institutes of Health (NIH) HL127131. The views expressed in this article are those of the authors and do not necessarily reflect the position or policy of the Department of Veterans Affairs or the United States Government.

Author Contributions: B.K., J.S.L., E.V., K.S., H.R.C., and M.A.W. contributed to study conception, study design, and data interpretation. N.Z. and M.A.W. contributed to data acquisition. B.K., N.Z., E.V., and M.A.W. contributed to analysis. B.K. drafted the report, and all authors revised it critically. All authors approved the final version.

Correspondence and requests for reprints should be addressed to Bhavika Kaul, M.D., Division of Pulmonary and Critical Care Medicine, University of California San Francisco, 513 Parnassus Avenue, HSE 1314 Box 0111, San Francisco, CA 94143. E-mail: Bhavika.Kaul@ucsf.edu.

This article has a related editorial.

This article has an online supplement, which is accessible from this issue's table of contents at www.atsjournals.org.

Ann Am Thorac Soc Vol 19, No 2, pp 196–203, Feb 2022

DOI: 10.1513/AnnalsATS.202103-295OC
Internet address: www.atsjournals.org

Idiopathic pulmonary fibrosis (IPF) is a progressive, fibrosing interstitial lung disease that affects older adults (1). Without therapy, median survival time ranges from 2 to 5 years after diagnosis (2, 3). However, with the Food and Drug Administration (FDA) approval of two new medications that slow disease progression, pirfenidone and nintedanib, this trajectory may be changing, and better

understanding the population burden and epidemiology of IPF has important implications for clinical management and outcomes.

Establishing the epidemiology of IPF has been challenging owing to disease complexity, relatively low prevalence, and heterogeneity of populations sampled. Early registry-based studies, in which IPF was likely underdiagnosed, reported incidence rates of 0.22–8.8 per 100,000 person-years and prevalence rates of 0.5–27.9 per 100,000 (4). However, more recent studies that have used large administrative databases to identify population-based cohorts have noted substantially higher incidence and prevalence rates. One study of Medicare beneficiaries over the age of 65 reported a cumulative prevalence of 494 cases per 100,000 in 2011 and an annual incidence of 94 cases per 100,000 person-years (5).

Very little is known about the population burden of IPF among U.S. Veterans. The Veterans Health Administration (VHA) is the largest integrated healthcare system in the United States and provides longitudinal comprehensive care to Veterans. It consists of 130 hospitals, more than 1,000 community-based outpatient clinics, and a network of purchased care from community providers, which together serve over 9 million patients annually (6, 7). In 2017, approximately 91% of Veterans were male and the mean age was 65 years. The mean age among the non-Veteran population was 42 years old, and less than half were male (7). Because IPF is a disease most frequently seen in older males (8), and because observational data has implicated exposures in the pathogenesis of disease (9–11), we hypothesized that the prevalence of IPF among Veterans may be higher than what has been reported in the general population.

In this study, we sought to describe the epidemiology (incidence, prevalence, and geographic distribution) of IPF among U.S. Veterans who receive care through the VHA and identify risk factors associated with IPF in the Veteran population.

Methods

The University of California San Francisco and the San Francisco Veterans Affairs (VA) Institutional Review Boards approved this study (IRB 20–30063).

Data Source and Patient Identification

We analyzed electronic health record data from Veterans who were enrolled in the VHA and had at least one inpatient or outpatient encounter at a VA facility or a non-VA facility paid for by the VA, between 2010 and 2019. We identified all patients who had an *International Classification of Disease* (ICD) diagnosis code for IPF (ICD-9-CM code 515, 516.3, 516.31 or ICD-10-CM code J84.111, J84.112, J84.89, J84.9, J84.10, J84.17) recorded between January 1, 2010 and December 31, 2019 (*see* Table E1 in the online supplement). We included the broader ICD-9 diagnosis code 515 and its ICD-10 equivalents (J84.89, J84.9, J84.10, J84.17) to improve sensitivity because these codes are often used during the initial workup of IPF (12). In our clinical experience with the VA in particular, Veterans with IPF may receive only these broader fibrosis codes.

To refine the initial cohort, we used an algorithm similar to those used in other IPF population-based studies (5, 12, 13). Patients were considered to have IPF if they did not have any other diagnosis code for an alternative interstitial lung disease (Table E1) after the first IPF diagnosis code. This subgroup was labeled "broad case definition." Among these Veterans, those who had procedure code for a lung biopsy or a computed tomographic (CT) scan of the thorax (Table E2) before the last IPF diagnosis were labeled "narrow case definition" (Figure 1). Of note, Veterans have several options for healthcare coverage, including dual enrollment, through both VA and non-VA (Medicare, Medicaid, or employer sponsored) health insurance plans. Estimates of dual use have ranged from 20% to 56% depending on patient demographics and services provided (14–16). Thus, for dual users, if CT scans or lung biopsies were conducted outside the VHA, these patients may not be captured by the narrow case definition.

Covariates including age, sex, race, ethnicity, rural versus urban residence, location of medical care (VA medical facility, state, and region), and smoking history were obtained from electronic heath records. The VA defines rurality using the U.S. Department of Agriculture's Rural-Urban Community Areas (RUCA) system (17). RUCA codes classify U.S. census tracts using measures of population density, urbanization, and daily commuting. All Veterans were categorized into rural versus urban residence based on home address at time of IPF diagnosis. A positive smoking history was defined as an ICD-9 or ICD-10 diagnosis code for tobacco use disorder any time prior to IPF diagnosis. We also included a positive tobacco use assessment from the health factor database as indicative of prior smoking history.

Statistical Analysis

We calculated the annual incidence and prevalence of IPF among Veterans between 2010 and 2019 using both the broad and narrow case definitions. Annual incidence was calculated by dividing the number of new cases of IPF by the number of patients without a prior diagnosis who had at least one encounter with the healthcare system during that calendar year. For annual prevalence, we first identified the number of patients who had at least one encounter with the healthcare system during the calendar year. From that denominator, we identified the number of patients who had an ICD diagnosis for IPF within the prior 5 years (numerator). IPF prevalence among each state's Veteran population was calculated in 2014 (midpoint) and 2019 (most recent). We chose 2014 for midpoint prevalence to capture data prior to the ICD-9 to ICD-10 transition that occurred in October 2015.

To examine factors associated with IPF among the Veteran population, we performed multivariate logistic regression including age, race, ethnicity, smoking history, and rural residence as explanatory variables. Because of nonlinearity of the continuous variable age, we presented stratified odds ratios by decade. Because of interaction between sex and multiple variables (age, ethnicity, and tobacco use), and owing to the smaller proportion of females in this cohort, we stratified the multivariable models by sex and estimated marginal differences in IPF prevalence between males and females using regression standardization. All analyses were conducted using StataCorp Analysis Software version 16.1.

Results

Approximately 10.7 million Veterans who received care through the VHA between 2010 and 2019 were eligible for inclusion in this study (Figure 1). Among these Veterans, we identified 139,116 incident diagnoses of IPF using the broad case definition (1.26%) and 82,557 incident cases of IPF using the narrow case definition (0.77%). Demographics were similar between the

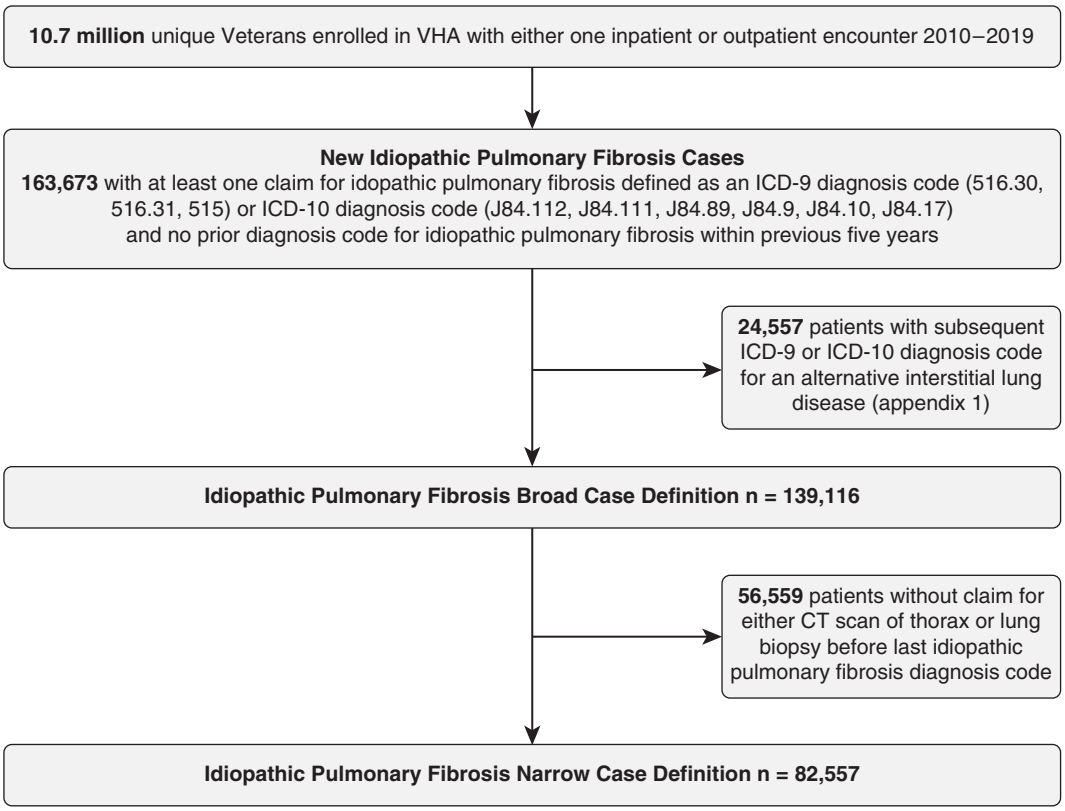

Figure 1. Veterans Health Administration idiopathic pulmonary fibrosis sample selection. VHA = Veterans Health Administration; ICD = *International Classification of Disease*; CT = computed tomography.

Table 1. Characteristics of U.S. Veterans with incident idiopathic pulmonary fibrosis (2010–2019)

	IPF Cohort		
	Broad Case Definition	**Narrow Case Definition**	**Control Population***
Total cases (*n*)	139,116	82,557	10,650,206
Demographics			
Mean age at diagnosis in yr, (SD)	70.7 (11.8)	70.2 (11.0)	61.5 (18.9)
Sex, *n* (%)			
Male	129,598 (93%)	76,779 (93%)	9,767,864 (91%)
Female	9,513 (7%)	5,776 (7%)	918,867 (9%)
Race, *n* (%)			
White	108,071 (78%)	64,435 (78%)	7,068,839 (66%)
Black or African American	14,338 (10%)	8,989 (11%)	1,479,129 (14%)
Asian/Native Hawaiian or Pacific Islander/American Indian	2,472 (2%)	1,468 (2%)	264,775 (3%)
Unknown	14,235 (10%)	7,665 (9%)	1,874,966 (18%)
Ethnicity, *n* (%)			
Hispanic or Latino	5,651 (4%)	3,544 (4%)	556,823 (5%)
Not Hispanic or Latino	120,343 (87%)	72,036 (87%)	8,347,958 (78%)
Unknown	13,122 (9%)	6,977 (8%)	1,782,928 (17%)
Rural, *n* (%)	55,598 (43%)	32,252 (42%)	3,463,456 (34%)
Region, *n* (%)			
West	27,900 (20%)	15,074 (18%)	2,369,694 (22%)
Midwest	36,531 (26%)	19,131 (23%)	2,439,238 (23%)
South	54,512 (39%)	30,576 (37%)	4,032,728 (38%)
Northeast	20,096 (14%)	17,776 (22%)	1,840,928 (17%)
Surgical or Transbronchial Lung Biopsy	3,029 (2%)	2,895 (4%)	14,502 (0.1%)
Tobacco Use	109,895 (79%)	69,144 (84%)	5,543,334 (52%)

Definition of abbreviations: IPF = idiopathic pulmonary fibrosis; SD = standard deviation.
*Control population defined as Veterans with no IPF diagnosis codes. All *P* values less than 0.001 comparing broad case definition to control population. *P* values were calculated with *t* test for continuous variables and chi-square test for categorical variables.

broad and narrow case definition cohorts (Table 1). Compared to Veterans without IPF, Veterans with IPF were older and more likely to be non-Hispanic, White, and male. Approximately 80% of Veterans with IPF were current or former smokers.

VHA Incidence and Prevalence of Idiopathic Pulmonary Fibrosis

The annual incidence and prevalence of IPF increased between 2010 and 2019 (Figure 2). Among the broad case definition subgroup, incidence increased from 141 cases per 100,000 person-years in 2010 to 331 cases per 100,000 person-years in 2019. Annual prevalence also rose from 582 cases per 100,000 in 2010 to 1,160 cases per 100,000 in 2019. The incidence and prevalence of IPF was lower in the narrow case definition cohort; however, the increase in incidence and prevalence rates over time mirrored that seen in the broad case definition. Among the narrow case definition subgroup, incidence increased from 73 cases per 100,000 person-years in 2010 to 210 cases per 100,000 person-years in 2019. Annual prevalence rose from 276 cases per 100,000 in 2010 to 725 cases per 100,000 in 2019.

Risk Factors Associated with IPF Diagnosis

Older age, White race, a history of tobacco use, and rural residence were associated with higher odds of incident IPF for both male and female Veterans (Tables 2 and 3). The risk of incident IPF increased by decade until age 80, after which a plateau effect was observed. Among female Veterans, non-Hispanic ethnicity was associated with slightly higher odds of IPF. These trends held true for both the broad and narrow case definition. The average marginal difference in prevalence of IPF between the male and female sex, accounting for interactions, was small (0.0048), and in favor of the female sex.

Geographic Distribution

The geographic burden of IPF was heterogeneous across the United States (Figure 3), and Veterans with IPF were more likely than Veterans without IPF to live in rural areas. The 2019 standardized prevalence of IPF among the state's Veteran population using the narrow case definition ranged from 430 cases per 100,000 in Utah (lowest) to 1,469 cases per 100,000 in Montana (highest). Compared with the midpoint prevalence in 2014, the prevalence of IPF among the Veteran population increased across all states in 2019; however, the states with the highest standardized prevalence rates (Iowa, Kentucky, Montana, and West Virginia) remained consistent between 2014 and 2019. A similar geographic distribution was seen when the broad case definition was used.

Discussion

This study represents the first comprehensive epidemiological analysis of IPF in the U.S. Veteran population. Among 10.7 million U.S. Veterans who received care through the VHA during the last decade, the annual incidence and prevalence of IPF substantially increased. Between 2010 and 2019, the prevalence doubled (from 582 to 1,160 cases per 100,000 using the broad case definition and from 276 to 725 cases per 100,000 using the narrow case definition), and incidence more than doubled (from 141 to 331 cases per 100,000 person-years using the broad case definition and from 73 to 210 cases per 100,000 person-years using the narrow case

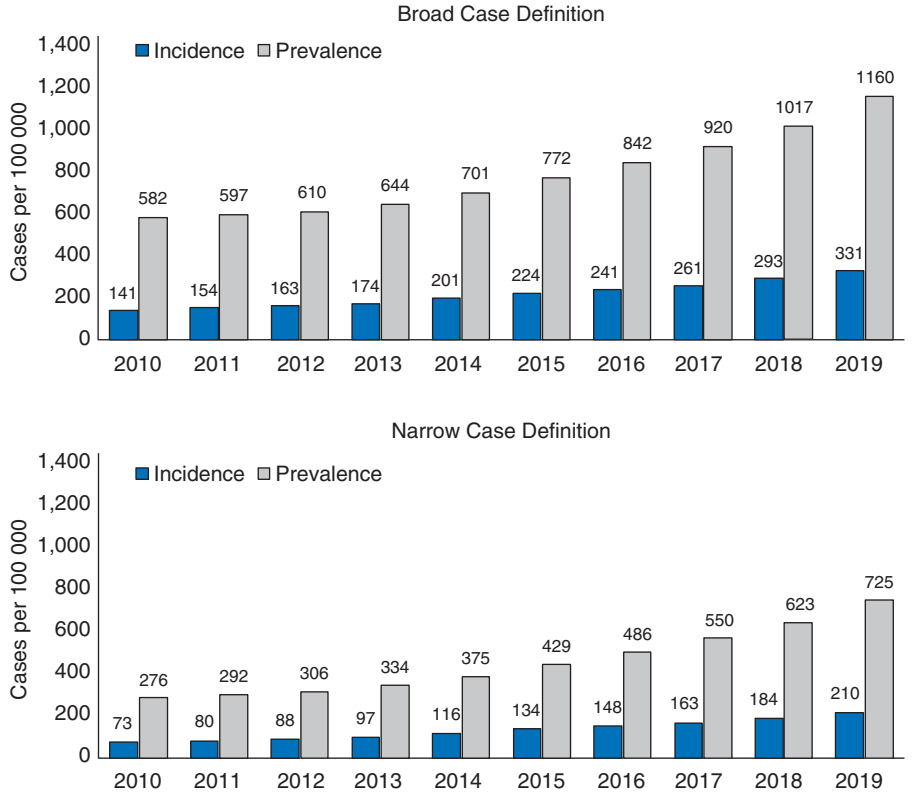

Figure 2. Annual incidence and prevalence of idiopathic pulmonary fibrosis among U.S. Veterans, 2010–2019.

Table 2. Risk factors associated with incident idiopathic pulmonary fibrosis among male Veterans

	Broad Case Definition IPF: *n* = 129,598 Non-IPF: *n* = 9,767,864		Narrow Case Definition IPF: *n* = 76,779 Non-IPF: *n* = 9,767,864	
	Odds Ratio (95% CI)	*P* Value	Odds Ratio (95% CI)	*P* Value
Age				
⩽60 yr	(Reference)		(Reference)	
>60–70 yr	2.83 (2.78–2.89)	<0.001	3.11 (3.04–3.19)	<0.001
>70–80 yr	4.39 (4.30–4.47)	<0.001	4.77 (4.65–4.89)	<0.001
>80 yr	3.21 (3.15–3.28)	<0.001	2.86 (2.78–2.94)	<0.001
Race				
White	(Reference)		(Reference)	
Black or African American	0.90 (0.88–0.91)	<0.001	0.93 (0.90–0.95)	<0.001
Asian/Native Hawaiian or Pacific Islander/American Indian	0.86 (0.82–0.90)	<0.001	0.86 (0.81–0.91)	<0.001
Ethnicity				
Hispanic or Latino	(Reference)		(Reference)	
Not Hispanic or Latino	0.99 (0.96–1.02)	0.44	0.94 (0.90–0.97)	0.001
Tobacco Use	2.94 (2.89–2.98)	<0.001	4.02 (3.93–4.11)	<0.001
Rural Residence	1.26 (1.25–1.28)	<0.001	1.21 (1.19–1.23)	<0.001

Definition of abbreviations: CI = confidence interval; IPF = idiopathic pulmonary fibrosis.

definition). Risk factors associated with IPF included older age, White race, rural residence, and a history of tobacco use. The effect of sex on odds of IPF was attenuated once accounting for other risk factors. We also found significant geographic heterogeneity of disease across the United States with a greater prevalence of IPF among Veterans living in rural compared with urban areas.

Prior studies that have examined the incidence and prevalence of IPF in the Medicare population (5) and among U.S. adults aged 18 to 64 (13) also demonstrated increasing prevalence over time, but little is known about the epidemiology of IPF among U.S. Veterans. One study of 760,621 Veterans who were deployed to combat operations in Iraq and Afghanistan reported that 0.3% were diagnosed with interstitial lung disease between 2002 and 2011 (18). However, the study focused exclusively on younger patients (mean age of 40 years old) with non-IPF interstitial lung disease. IPF is one of the most common interstitial lung diseases, and because prior literature has implicated age, male sex, and tobacco use in the pathogenesis, we hypothesized that incidence and prevalence may be higher when evaluating the Veteran population at large. Our estimates of incidence and prevalence are similar to what has been reported in Medicare data, suggesting that among cohorts demographically enriched for older patients, the population burden of IPF is likely higher than what has been described

Table 3. Risk factors associated with incident idiopathic pulmonary fibrosis among female Veterans

	Broad Case Definition IPF: *n* = 9,513 Non-IPF: *n* = 918,867		Narrow Case Definition IPF: *n* = 5,776 Non-IPF: *n* = 918,867	
	Odds Ratio (95% CI)	*P* Value	Odds Ratio (95% CI)	*P* Value
Age				
⩽60 yr	(Reference)		(Reference)	
>60–70 yr	4.19 (3.95–4.44)	<0.001	4.72 (4.38–5.09)	<0.001
>70–80 yr	5.61 (5.15–6.10)	<0.001	6.33 (5.69–7.03)	<0.001
>80 yr	3.70 (3.35–4.09)	<0.001	2.99 (2.58–3.45)	<0.001
Race				
White	(Reference)		(Reference)	
Black or African American	0.86 (0.80–0.92)	<0.001	0.87 (0.80–0.95)	0.002
Asian/Native Hawaiian or Pacific Islander/American Indian	0.75 (0.64–0.89)	0.001	0.81 (0.66–0.99)	0.047
Ethnicity				
Hispanic or Latino	(Reference)		(Reference)	
Not Hispanic or Latino	1.20 (1.05–1.37)	0.007	1.18 (1.0–1.4)	0.051
Tobacco Use	2.79 (2.64–2.96)	<0.001	3.47 (3.22–3.74)	<0.001
Rural Residence	1.24 (1.17–1.31)	<0.001	1.20 (1.12–1.29)	<0.001

Definition of abbreviations: CI = confidence interval; IPF = idiopathic pulmonary fibrosis.

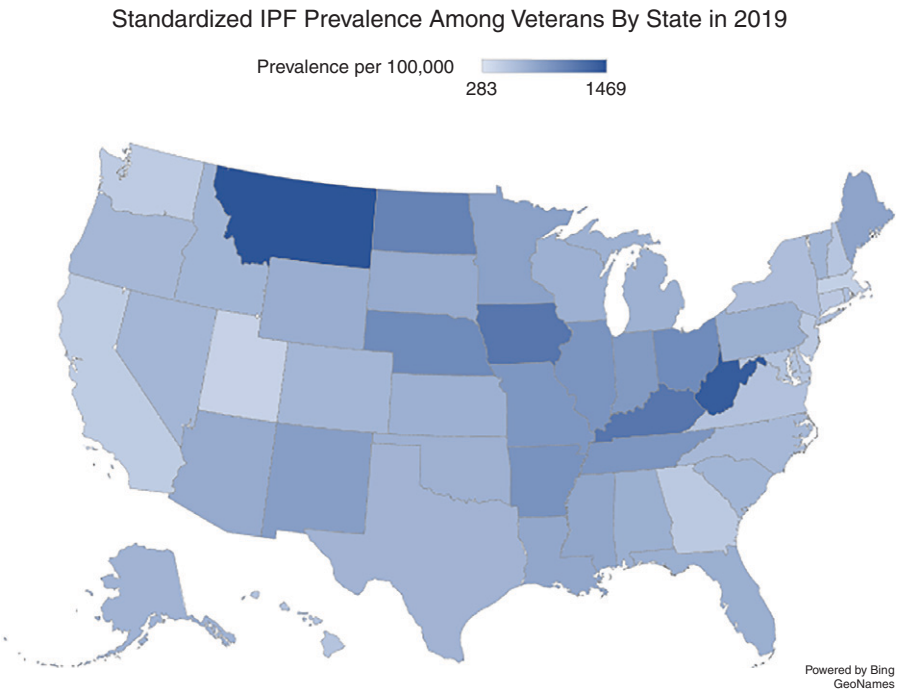

Figure 3. Geographic distribution of prevalent idiopathic pulmonary fibrosis cases among U.S. Veterans (narrow case definition). State prevalence was calculated by identifying the number of unique Veterans with IPF in each state divided by total number of Veterans living in the state. IPF = idiopathic pulmonary fibrosis.

in registry-based studies, which have historically relied on individual patient recruitment, which has likely led to underestimation of true disease burden owing to selection bias of the referral base.

The rising incidence and prevalence of IPF among Veterans is likely due to a combination of factors, including increasing disease awareness and treatment motivation with the approval of antifibrotic therapies in 2014, an aging Veteran population, and the implementation of low-dose CT scans for lung cancer screening among smokers, which has been shown to detect incidental or subclinical pulmonary disease in up to 40% of Veterans (19). It is possible that the transitions from ICD-9 to ICD-10 codes in October 2015 also contributed to changes in coding practices, resulting in increasing incidence and prevalence rates. However, this is unlikely to be the sole driver as IPF rates were increasing before and after the transition year. With the recent expansion of lung cancer screening guidelines to include patients 50–80 years old, we hypothesize that detection of subclinical or early pulmonary fibrosis will further increase. This has important implications on healthcare systems that will need to develop the infrastructure to meet the specialty care needs of patients, both Veterans and non-Veterans, with pulmonary fibrosis.

Among Veterans, we found substantial variation in the geographic distribution of IPF across the United States. This has been noted in other large database studies, although the factors behind the geographic variability have not been elucidated (5). We also observed a higher prevalence of IPF in rural compared with urban areas. Rurality is not yet an established risk factor for IPF, but meta-analysis data has suggested an association between IPF and environmental factors such as agriculture, farming, livestock, wood dust, and metal dust (20). Exposures, whether combat, occupational, or residential, are of particular relevance to the Veteran population. Further work is needed to determine the extent to which exposures, underlying patient demographics, recognition, and/or coding differences contribute to the geographic heterogeneity of IPF and the higher prevalence of IPF in rural areas.

In our cohort, we noted that after accounting for interactions with age, race/ethnicity, and tobacco use history, the average marginal difference in prevalence between males and females was small (0.0048) and slightly favored females. Early epidemiological studies have previously demonstrated higher incidence and prevalence of IPF among men. This difference has been hypothesized to be driven by genetic and biologic predispositions or perhaps in part owing to differential exposures between men and women. However, few studies have used multivariable models to examine the association between male sex and incident IPF while controlling for other risk factors such as age and smoking. It is possible that once controlling for these risk factors, the association between male sex and IPF is attenuated. Alternatively, female Veterans may represent a unique subgroup that may not be generalizable to the non-Veteran population. The association between sex and IPF deserves further study.

This study used the strength of the VHA's integrated healthcare system and data repository to identify IPF cases. Accurate identification of IPF cases using billing-code-based algorithms depends heavily on the characteristics of the underlying source population (21). For example, algorithms that successfully identify IPF cases in older patients seen at tertiary care centers may not perform as well in younger community-based cohorts because of an overall lower

likelihood of disease. The algorithms used in our study are adapted from prior literature that has defined their performance across varying cohorts (5, 12, 13). The use of ICD-9 code 515 and its ICD-10 equivalents (J84.89, J84.9, J84.10, and J84.17) has been a particular point of discussion. A study that cross-referenced code-based diagnosis with review of electronic health records found that among a network of Mid-Atlantic Veterans, ICD-9 codes 516.3, 516.31, and 515 were commonly used for patients with pulmonary fibrosis (22). We thus included the 515 code to optimize sensitivity in our study and believe that it more closely approximates the true population burden of disease among the Veteran population. Because the underlying VHA population is predominantly older and White, with a high prevalence of tobacco use, there is a higher pre-test probability of IPF and likely a greater diagnostic specificity of the 515 code.

Limitations

Our study has a number of limitations. First, although previously validated in other cohorts, this ICD-based algorithm has not specifically been case validated in the VHA. Reassuringly, the broad and narrow case cohorts in the VHA were demographically consistent with IPF and the conclusions internally consistent with both approaches. Additionally, the estimates of incidence and prevalence are similar to what has been observed in other large population-based studies like Medicare. Second, we refined the algorithm by requiring either a CT scan or surgical lung biopsy prior to the last IPF diagnosis code (narrow case definition). Patients who have both VA and non-VA (Medicare, Medicaid, or employer-sponsored) health insurance may have completed a CT scan or lung biopsy outside the VA Healthcare System. If these scans or biopsies were paid for by the VA, they would be captured by our algorithm; however, if they were paid for by non-VA insurance, they would not be captured by our restrictive algorithm. Prior studies have suggested that the percentage of dual users of VA and non-VA health insurance ranges from 20% to 56% depending on patient demographics and services provided (14–16). For example, most Veterans over the age of 65 qualify for Medicare and although cost-sharing is usually lower with VA insurance, patients with complex medical conditions are more likely to be dual users. We thus suspect that the narrow case definition likely underestimates the prevalence of IPF among Veterans and that the "true" prevalence lies in between the broad and narrow case definitions.

Conclusions

The past decade has seen substantial progress in our understanding and treatment of IPF. However, the efficiency and speed with which we have been able to answer research questions has been limited by access to sufficiently large, demographically and clinically diverse patient cohorts. Like many relatively rare diseases, IPF has historically been studied primarily through the use of discrete tertiary-care-based patient cohorts that take years to accumulate, number at best in the hundreds, are methodologically incompatible, and lack generalizability. Clinical data sets from electronic health records and other registry sources represent a new and exciting opportunity to change the research paradigm by providing direct access to large, real-world, population-based patient cohorts and comprehensive data sets. They are also a critical component of a transformative, practice-based model of discovery and implementation known as the learning healthcare system that has been promoted by the National Academy of Medicine, the Agency of Healthcare Research and Quality, and the Food and Drug Administration (23, 24).

In this study, we used the strength of the VA learning healthcare system to establish a population-based cohort that can serve as the foundation for future work that uses data generated from patient–healthcare system interactions to more efficiently answer questions related to epidemiology, comparative effectiveness, and predictive analytics in IPF with the ultimate goal of improving care for Veterans with interstitial lung disease. ■

Author disclosures are available with the text of this article at www.atsjournals.org.

References

1 Raghu G, Remy-Jardin M, Myers JL, Richeldi L, Ryerson CJ, Lederer DJ, *et al.*; American Thoracic Society, European Respiratory Society, Japanese Respiratory Society, and Latin American Thoracic Society. Diagnosis of idiopathic pulmonary fibrosis. An official ATS/ERS/JRS/ALAT clinical practice guideline. *Am J Respir Crit Care Med* 2018;198:e44–e68.

2 Ley B, Collard HR, King TE Jr. Clinical course and prediction of survival in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2011;183:431–440.

3 Fernández Pérez 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;137:129–137.

4 Kaunisto J, Salomaa ER, Hodgson U, Kaarteenaho R, Myllärniemi M. Idiopathic pulmonary fibrosis–a systematic review on methodology for the collection of epidemiological data. *BMC Pulm Med* 2013;13:53.

5 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;2:566–572.

6 Veterans Health Administration. Providing Health Care for Veterans. U.S. Department of Veterans Affairs [updated 2019 Oct 25; accessed 2021 Jan 11]. Available from: https://www.va.gov/health/.

7 National Center for Veteran Analysis and Statistics. U.S. Department of Veterans Affairs [updated 2019 Nov 5; accessed 2021 Jan 12]. Available from: https://www.va.gov/vetdata/.

8 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;192:1200–1207.

9 Baumgartner KB, Samet JM, Coultas DB, Stidley CA, Hunt WC, Colby TV, *et al.*; Collaborating Centers. Occupational and environmental risk factors for idiopathic pulmonary fibrosis: a multicenter case-control study. *Am J Epidemiol* 2000;152:307–315.

10 Baumgartner KB, Samet JM, Stidley CA, Colby TV, Waldron JA. Cigarette smoking: a risk factor for idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 1997;155:242–248.

11 Kim JS, Podolanczuk AJ, Borker P, Kawut SM, Raghu G, Kaufman JD, *et al.* Obstructive sleep apnea and subclinical interstitial lung disease in the Multi-Ethnic Study of Atherosclerosis (MESA). *Ann Am Thorac Soc* 2017;14:1786–1795.

12 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;14:880–887.

13 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;48:179–186.

14 Hynes DM, Koelling K, Stroupe K, Arnold N, Mallin K, Sohn MW, *et al.* Veterans' access to and use of Medicare and Veterans Affairs health care. *Med Care* 2007;45:214–223.
15 Thorpe CT, Gellad WF, Mor MK, Cashy JP, Pleis JR, Van Houtven CH, *et al.* Effect of dual use of Veterans Affairs and Medicare Part D drug benefits on antihypertensive medication supply in a national cohort of Veterans with dementia. *Health Serv Res* 2018;53:5375–5401.
16 Thorpe JM, Thorpe CT, Schleiden L, Cashy J, Carico R, Gellad WF, *et al.* Association between dual use of Department of Veterans Affairs and Medicare Part D drug benefits and potentially unsafe prescribing. *JAMA Intern Med* 2019;179:1584–1586.
17 Economic Research Service. Rural-urban commuting codes. U.S. Department of Agriculture [updated 2019 Oct 24; accessed 2021 Jan 11]. Available from: https://www.ers.usda.gov/data-products/rural-urban-commuting-area-codes/.
18 Pugh MJ, Jaramillo CA, Leung KW, Faverio P, Fleming N, Mortensen E, *et al.* Increasing prevalence of chronic lung disease in Veterans of the wars in Iraq and Afghanistan. *Mil Med* 2016;181:476–481.
19 Kinsinger LS, Anderson C, Kim J, Larson M, Chan SH, King HA, *et al.* Implementation of lung cancer screening in the Veterans Health Administration. *JAMA Intern Med* 2017;177:399–406.
20 Taskar VS, Coultas DB. Is idiopathic pulmonary fibrosis an environmental disease? *Proc Am Thorac Soc* 2006;3:293–298.
21 Ley B, Collard HR. House of cards? Testing fundamental assumptions in idiopathic pulmonary fibrosis epidemiology. *Am J Respir Crit Care Med* 2015;192:1147–1148.
22 Bedoya A, Pleasants RA, Boggan JC, Seaman D, Reihman A, Howard L, *et al.* Interstitial lung disease in a Veterans Affairs regional network; a retrospective cohort study. *PLoS One* 2021;16:e0247316.
23 Sherman RE, Anderson SA, Dal Pan GJ, Gray GW, Gross T, Hunter NL, *et al.* Real-world evidence - what is it and what can it tell us? *N Engl J Med* 2016;375:2293–2297.
24 Bindman AB. The Agency for Healthcare Research and Quality and the development of a learning health care system. *JAMA Intern Med* 2017; 177:909–910.

Online Data Supplement

Epidemiology of Idiopathic Pulmonary Fibrosis Among U.S. Veterans, 2010 – 2019

Bhavika Kaul MD, Joyce S. Lee MD, Ning Zhang MS, Eric Vittinghoff PhD, Kathleen Sarmiento MD, Harold R. Collard MD, Mary A. Whooley MD

Table E1. Idiopathic Pulmonary Fibrosis Diagnostic Codes

INCLUSION CODES		
ICD-9-CM Code	**ICD-10-CM Code**	**Description**
516.3	J84.111	Idiopathic interstitial pneumonia, not otherwise specified
516.31	J84.112	Idiopathic pulmonary fibrosis, Familial idiopathic pulmonary fibrosis
515	J84.89, J84.9, J84.10, J84.17	Post-inflammatory pulmonary fibrosis, other specified interstitial pulmonary disease, interstitial pulmonary disease unspecified, pulmonary fibrosis unspecified

EXCLUSION CODES		
ICD-9-CM Code	**ICD-10 Code**	**Description**
516.32-516.37	J84.113-J84.117, J84.2	Idiopathic nonspecific interstitial pneumonia, acute interstitial pneumonia, respiratory bronchiolitis interstitial pneumonia, idiopathic lymphoid interstitial pneumonia, cryptogenic organizing pneumonia, and desquamative interstitial pneumonia
495	J67.0-J67.9	Extrinsic allergic alveolitis, Hypersensitivity Pneumonitis
516.8	J84.09	Other specified alveolar and parietoalveolar pneumonopathies
516.9	J84.09	Unspecified alveolar and parietoalveolar pneumonopathies
714.81	M05.10	Rheumatoid lung
277.8	C96.6	Other specified disorders of metabolism
272.7	E75.21-75.22	Lipidoses
277.3	E85.9	Amyloidosis
518.3	J82	Pulmonary eosinophilia
516.0	J84.01	Pulmonary alveolar proteinosis
516.2	J84.02	Pulmonary alveolar microlithiasis
516.1	J84.03	Idiopathic pulmonary hemosiderosis
517.8	J99	Lung involvement in other diseases classified elsewhere
446.21	M31.0	Goodpasture's syndrome
446.4	M31.30	Granulomatosis with polyangiitis (Wegener's)
446.4	M30.1	Churg-Strauss
446.4	M31.7	Microscopic polyangiitis
710.0	M32.10	Systemic lupus erythematosus
710.4	M33.20	Polymyositis
710.3	M33.90	Dermatomyositis
710.1	M34.0	Systemic sclerosis

710.1, 517.2	M34.81	Lung involvement in systemic sclerosis
710.2	M35.00	Sjögren's disease
720	M45.9	Ankylosing spondylitis
237.7	Q85.00	Neurofibromatosis
759.5	Q85.1	Tuberous sclerosis
135	D86.0, D86.2, D86.9	Sarcoidosis, Sarcoidosis of lung with sarcoidosis of lymph nodes
555	K5000, K5010, K5080, K5090	Regional Enteritis (Chron's Disease)
500	J60	Coal workers' pneumoconiosis
501	J61	Asbestosis
502	J62.8	Pneumoconiosis due to other silica or silicates
503	J63.0-63.6	Pneumoconiosis due to other inorganic dust
504	J66.8	Pneumonopathy due to inhalation of other dust
505	J64	Pneumoconiosis, unspecified
506.4	J68.4	Chronic respiratory conditions due to fumes or vapors
508.8	J70.8	Respiratory conditions due to other specified external agents
508.1	J70.1	Chronic and other pulmonary manifestations due to radiation

Table E2. Procedure Codes

Procedure	ICD 9-CM	ICD 10-CM	CPT-4
Chest Computed Tomography	87.41	BW24, BW25, BB24, BP2W, B32T, B32S	71250, 71260, 71270, 71275
Surgical Lung Biopsy	33.20, 33.28, 34.21	0BBM, 0BBL, 0BBJ, 0BBF, 0BBD, 0BBK, 0BBG, 0BBC, 0BBH 0BB9, 0BBB, 0BB6, 0BB7, 0BB3, 0BB5, 0BB8, 0BB4	32095-97, 32100-32160, 32602, 32607-8
Transbronchial Lung Biopsy	33.27	0BBK, 0BBL	31628, 31629, 31632

Table E3. Idiopathic Pulmonary Fibrosis Among U.S. Veterans 2019 State Prevalence.

State	IPF Cases (Broad)	IPF Cases (Narrow)	Total Veterans	Prevalence* (Broad)	Prevalence* (Narrow)
ALABAMA	2,390	1,154	160,139	1,492	721
ALASKA	383	246	36,088	1,061	682
ARIZONA	3,198	1,871	247,755	1,291	755
ARKANSAS	2,237	1,131	119,128	1,878	949
CALIFORNIA	5,128	3,068	647,017	793	474
COLORADO	1,770	1,065	162,338	1,090	656
CONNECTICUT	570	380	75,051	759	506
DELAWARE	237	149	27,472	863	542
DISTRICT OF COLUMBIA	106	83	29,330	361	283
FLORIDA	9,091	5,523	767,299	1,185	720
GEORGIA	3,101	1,582	316,405	980	500
HAWAII	476	235	43,584	1,092	539
IDAHO	704	436	65,272	1,079	668
ILLINOIS	4,182	2,454	263,152	1,589	933
INDIANA	3,027	1,596	178,051	1,700	896
IOWA	2,405	1,158	97,292	2,472	1,190
KANSAS	1,035	660	92,977	1,113	710
KENTUCKY	2,979	1,728	145,225	2,051	1,190
LOUISIANA	2,030	984	126,182	1,609	780
MAINE	693	454	56,329	1,230	806
MARYLAND	1,217	746	137,353	886	543
MASSACHUSETTS	1,015	565	128,312	791	440
MICHIGAN	2,625	1,512	214,766	1,222	704
MINNESOTA	2,418	1,332	160,737	1,504	829
MISSISSIPPI	1,318	754	94,791	1,390	795
MISSOURI	3,356	1,838	200,881	1,671	915
MONTANA	1,439	742	50,527	2,848	1,469
NEBRASKA	1,342	716	70,233	1,911	1,019
NEVADA	1,180	751	113,936	1,036	659
NEW HAMPSHIRE	425	252	47,420	896	531
NEW JERSEY	1,066	646	126,739	841	510
NEW MEXICO	1,141	644	75,613	1,509	852
NEW YORK	3,041	1,977	331,439	918	596
NORTH CAROLINA	3,589	2,149	335,087	1,071	641
NORTH DAKOTA	560	333	30,577	1,831	1,089
OHIO	5,353	3,242	321,002	1,668	1,010

OKLAHOMA	2,022	960	135,836	1,489	707
OREGON	1,486	926	141,853	1,048	653
PENNSYLVANIA	3,611	2,236	310,755	1,162	720
RHODE ISLAND	245	169	29,125	841	580
SOUTH CAROLINA	2,226	1,309	195,142	1,141	671
SOUTH DAKOTA	598	337	44,505	1,344	757
TENNESSEE	3,717	1,994	216,019	1,721	923
TEXAS	8,649	4,654	684,085	1,264	680
UTAH	382	236	54,889	696	430
VERMONT	224	143	21,356	1,049	670
VIRGINIA	2,332	1,439	258,239	903	557
WASHINGTON	1,596	966	200,832	795	481
WEST VIRGINIA	1,803	1,114	78,999	2,282	1,410
WISCONSIN	1,930	1,158	162,680	1,186	712
WYOMING	401	218	29,641	1,353	735

Definition of Abbreviations: IPF = Idiopathic Pulmonary Fibrosis.
Prevalence defined as cases per 100,000. Total Veterans represents number of Veterans enrolled in Veterans Health Administration by state.