ORIGINAL RESEARCH

Epidemiology of Sarcoidosis in U.S. Veterans from 2003 to 2019

Mohamed I. Seedahmed[1,2], Aaron D. Baugh[2], Mohamed T. Albirair[3], Yanting Luo[1,4], Jianhong Chen[1,2], Charles E. McCulloch[5], Mary A. Whooley[1,6,7], Laura L. Koth[2], and Mehrdad Arjomandi[1,2]

[1]San Francisco Veterans Affairs Medical Center, San Francisco, California; [2]Division of Pulmonary, Allergy, Critical Care, and Sleep Medicine, Department of Medicine, [4]Philip R. Lee Institute for Health Policy Studies, [5]Department of Epidemiology and Biostatistics, and [6]Department of Medicine, University of California San Francisco, San Francisco, California; [3]Department of Global Health, University of Washington, Seattle, Washington; and [7]Measurement Science Quality Enhancement Research Initiative, San Francisco Veterans Affairs Healthcare System, San Francisco, California

ORCID IDs: 0000-0002-7446-7346 (M.I.S.); 0000-0002-9527-691X (A.D.B.); 0000-0003-1132-6244 (M.T.A.); 0000-0001-8598-4800 (Y.L.); 0000-0001-7989-0074 (J.C.); 0000-0002-1279-6179 (C.E.M.); 0000-0002-5943-0078 (M.A.W.); 0000-0001-9541-3622 (L.L.K.); 0000-0002-0116-9217 (M.A.).

Abstract

Rationale: United States veterans represent an important population to study sarcoidosis. Their unique history of environmental exposures, wide geographic distribution, and long-term enrollment in a single integrated healthcare system provides an unparalleled opportunity to understand the incidence, prevalence, and risk factors for sarcoidosis.

Objectives: To determine the epidemiology, patient characteristics, geographic distribution, and associated risk factors of sarcoidosis among U.S. veterans.

Methods: We used data from the Veterans Health Administration (VHA) electronic health record system between 2003 and 2019 to evaluate the annual incidence, prevalence, and geographic distribution of sarcoidosis (defined using the International Classification of Diseases codes). We used multivariate logistic regression to examine patient characteristics associated with sarcoidosis incidence.

Results: Among more than 13 million veterans who received care through or paid for by the VHA, 23,747 (0.20%) incident diagnoses of sarcoidosis were identified. Compared with selected VHA control subjects using propensity score matching, veterans with sarcoidosis were more likely to be female (13.5% vs. 9.0%), of Black race (52.2% vs. 17.0%), and ever–tobacco users (74.2% vs. 64.5%). There was an increase in the annual incidence of sarcoidosis between 2004 and 2019 (from 38 to 52 cases/100,000 person-years) and the annual prevalence between 2003 and 2019 (from 79 to 141 cases/100,000 persons). In a multivariate logistic regression model, Black race (odds ratio [OR], 4.49; 95% confidence interval [CI], 4.33–4.65), female sex (OR, 1.64; 95% CI, 1.56–1.73), living in the Northeast compared with the western region (OR, 1.57; 95% CI, 1.48–1.67), history of tobacco use (OR, 1.36; 95% CI, 1.31–1.41), and serving in the Army, Air Force, or multiple branches compared with the Navy (OR, 1.08; 95% CI, 1.03–1.13; OR, 1.10; 95% CI, 1.04–1.17; OR, 1.27; 95% CI, 1.16–1.39, respectively) were significantly associated with incident sarcoidosis ($P < 0.0001$).

Conclusions: The incidence and prevalence of sarcoidosis are higher among veterans than in the general population. Alongside traditionally recognized risk factors such as Black race and female sex, we found that a history of tobacco use within the Veterans Affairs population and serving in the Army, Air Force, or multiple service branches were associated with increased sarcoidosis risk.

Keywords: incidence; prevalence; tobacco; sarcoidosis; veterans

Sarcoidosis is a multisystem disease of unknown etiology involving nonnecrotizing granuloma formation in affected organs, especially the lungs (90% of cases) (1). A recent comprehensive U.S. epidemiological study of sarcoidosis reported a prevalence near 60 cases/100,000 persons, with an incidence of approximately 10 cases/100,000 person-years (2). The incidence among Black cases was more than double that of non-Hispanic White cases (2). An analysis of National Center for Health Statistics data revealed increased sarcoidosis-related mortality over the last 3 decades, particularly among older Black females (3). These characteristics do not reflect the demographics of the veteran population (4, 5). Compared with the general population, veterans are more often older, male, and of non-Hispanic White race. Veteran cohorts have smaller but still persistent racial disparities in socioeconomic status and health outcomes compared with the general population (6). We anticipated these could yield important differences in understanding the veteran experience, given that previous literature has identified many of these factors as important in determining the course of sarcoidosis. Although sarcoidosis was reported as similarly female-predominant in a study of the active military (7), longer follow-up time among veterans could yield different epidemiology.

Exposures have always been theorized to play an essential role in sarcoidosis. Mechanistically, these exposures could alter disease incidence or trajectory (8–10). However, two barriers exist to understanding the epidemiology, risk factors, and role of exposure history in sarcoidosis. These are *1*) the need for large, longitudinal, relatively comprehensive datasets; and *2*) a diverse exposure history. The exposure patterns of veterans are distinct and may vary among military service branches (11–14). For instance, burn pit exposure is more common in the Army. This may be a sarcoidosis risk

factor given the prior association of organic fuel burning with sarcoidosis in small case series of veterans (15). In addition, the molecular phenotype of pulmonary sarcoidosis was recently found to differ between veterans and non-veterans (16); exposure to inorganic dust (e.g., silica, metals) could be a plausible explanation (9, 17). Despite this evidence suggesting sarcoidosis presents uniquely in veterans, to the best of our knowledge no previous study has attempted a comprehensive characterization of the disease in this population.

The Veterans Health Administration (VHA) is the largest integrated healthcare system in the United States. Therefore, in a national cohort of U.S. veterans, we aimed to examine the temporal trends in the annual incidence and prevalence of sarcoidosis, aiming to understand the effect of the International Classification of Diseases, tenth revision (ICD-10) system (10), which was implemented in 2015 (18). We also aimed to study patient characteristics and geographic distribution and identify risk factors for sarcoidosis. Portions of these results have been reported previously in abstract form (19).

Methods

Data Source and Study Design
This was a descriptive, retrospective, population-based, observational case–control study. We analyzed data from national VHA electronic health records (EHRs) that had been extracted to a central Corporate Data Warehouse and transformed into the Common Data Model of the Observational Medical Outcomes Partnership. Both provide access to records for approximately 14 million veterans enrolled in VHA who had an encounter at a Veterans Affairs (VA) facility since October 1999. Data analyses were conducted using the VA Informatics and Computing Infrastructure (20). This study was approved by the University of California

San Francisco Institutional Review Board and the Veterans Health Administration Research and Development Committee (institutional review board protocol #15-16660). There was no involvement of any patients or other members of the general public in our research design, conduct, reporting, or dissemination plans. A Health Insurance Portability and Accountability Act waiver was obtained to use identifiable data (including diagnosis dates).

We calculated sarcoidosis new diagnosis rate (hereafter: incidence) manifesting in any organ system between January 1, 2004, and December 31, 2019, using EHR data from veterans who received care at a VA facility or non-VA facility paid for by VA, between January 1, 2000, and December 31, 2020. The start date was selected to include patients receiving a diagnosis of sarcoidosis after the first international consensus statement on sarcoidosis (21).

Defining Case and Control Populations
Like prior EHR-based studies of sarcoidosis epidemiology, we considered patients to have sarcoidosis if they had one inpatient ICD-9 135, ICD-10 D86.*, or two outpatient ICD codes for sarcoidosis at least 1 month apart (2, 22) (*see* Table E1 in the data supplement). Prior work shows a single sarcoid-specific ICD code has a positive predictive value of 79% for sarcoidosis (23). If the encounter was inpatient, we reasoned that a single encounter would be sufficient because inpatient coding is known to be more accurate than outpatient coding. Per the 2020 American Thoracic Society's sarcoidosis clinical practice guideline (1), our study case definition required the absence of diagnostic codes for alternative diagnoses (Table E2). The window for alternative diagnoses was 2 years before the index date to 1 year after the index date. Any manifestations of sarcoidosis were sufficient for our definition. The first sarcoidosis-specific ICD code date

Correspondence and requests for reprints should be addressed to Mohamed I. Seedahmed, M.D., M.P.H., Division of Pulmonary, Critical Care, Allergy and Immunology, and Sleep, Department of Medicine, University of California San Francisco, Helen Diller Medical Center, 513 Parnassus Avenue, HSE 1314, Box 0111, San Francisco, CA 94143. E-mail: mohamed.seedahmed@ucsf.edu.

This article has a related editorial.

This article has a data supplement, which is accessible from this issue's table of contents at www.atsjournals.org.
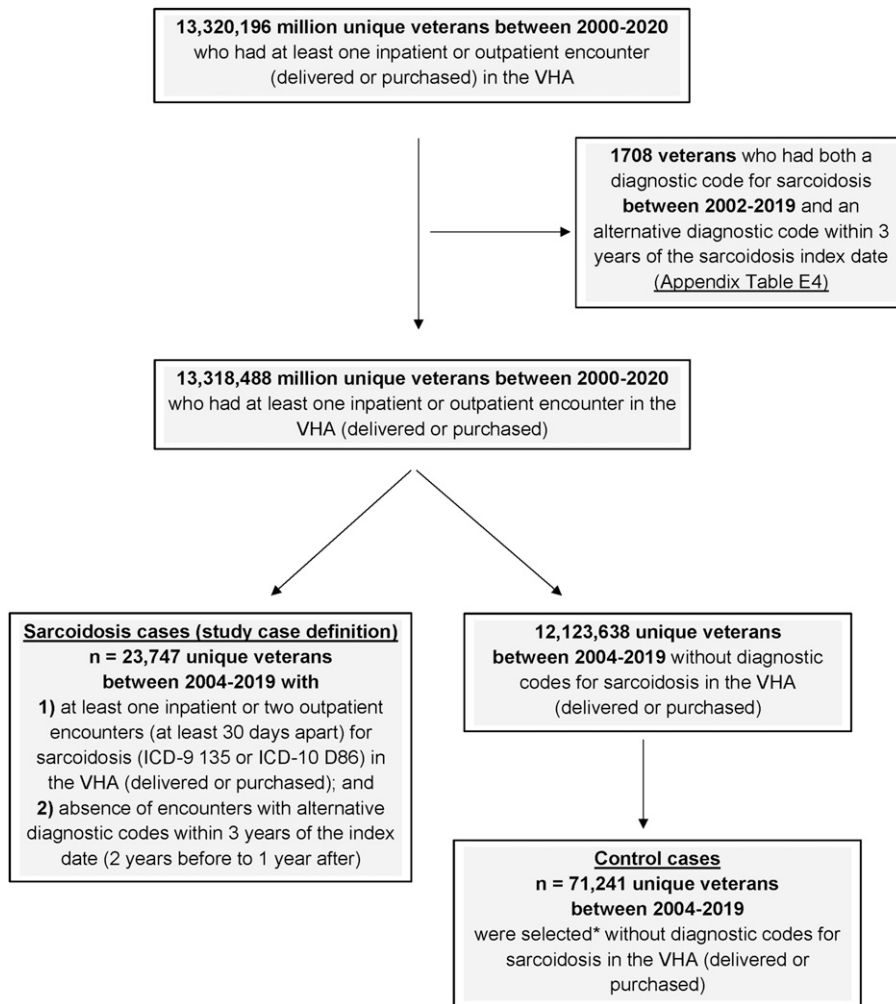
Figure 1. STROBE flowchart: sarcoidosis sample selection from the VHA database between January 1, 2000, and December 31, 2020. *Control cases were selected using incident density sampling and propensity score matching procedure (1:3). ICD = International Classification of Diseases; STROBE = Strengthening the Reporting of Observational Studies in Epidemiology; VHA = Veterans Health Administration.

in the record was defined as the index date for sarcoidosis diagnosis.

For the control group, we used an incidence density sampling procedure from the total population available through VA Corporate Data Warehouse without a diagnostic code for sarcoidosis by the end of each calendar year. First, we identified all veterans without sarcoidosis diagnostic codes who had at least one encounter in the VHA within the same calendar year of the index date for sarcoidosis. Second, we used propensity score matching to select three control cases for each sarcoidosis case. Propensity scoring used index year, birth year, and location of service (VA vs. non-VA). Control subjects and sarcoidosis cases may have had encounters in both VA and non-VA care facilities, provided at least one encounter is in the same system as the sarcoidosis index diagnosis.

Disease-related Variables

For all included participants, we collected demographic data on age, sex, race, ethnicity, residence regions (based on the 2010 Census Regions and Divisions of the United States [24]), service branch, and rurality (based on the VA definition of rurality and using the U.S. Department of Agriculture's Rural–Urban Community Areas system [25]). Service in military branches designated as "other" included any branches other than Army, Air Force, Navy, or Marines (e.g., Coast Guard, National Guard, etc.). Using data from the VA Observational Medical

Outcomes Partnership, tobacco usage history was defined dichotomously by any positive tobacco use as assessed from the health factor, dental, or Systematized Nomenclature of Medicine Clinical Terms domains, or any claims with procedural codes of tobacco use disorder, such as Procedural Terminology fourth Edition or the Healthcare Common Procedure Coding System (Table E3) (26, 27).

Statistical Analyses

All statistical analyses were performed using SAS software (SAS, version 9.4, SAS Institute, Inc). For sarcoidosis diagnosis, we calculated the annual diagnosis rate (hereafter: prevalence) from 2003 to 2019 and the annual new diagnosis rate (hereafter: incidence) from 2004 to 2019. To determine the denominator, we identified the number of unique patients with at least one inpatient or outpatient encounter delivered or purchased by the VHA during a calendar year. To determine the numerator, we identified the number of sarcoidosis cases as defined by our study definition during the calendar year. In addition, the numerator for the incidence required the absence of any prior sarcoidosis-specific diagnostic codes within the previous 2 years (to include only new cases in the same calendar year). The annual prevalence was calculated by dividing the number of patients with sarcoidosis diagnoses in a calendar year (numerator) by the number of patients with encounters (denominator) during that calendar year, then multiplied by 100,000 to be reported as per 100,000 persons. The annual incidence was calculated by dividing the number of new cases of sarcoidosis (numerator for the incidence) by the denominator for that calendar year, then multiplied by 100,000 to be reported as per 100,000 person-years.

In addition, we conducted univariate linear segmented regression using an interrupted time series to detect changes in incidence before versus after the introduction of ICD-10 in October 2015 (18). We fit a least-squares regression line to each segment of the independent variable, year ("index year"). We visually inspected the data to ensure linearity between the year and the incidence for each segment. Consistent with prior literature, we also compared the overall prevalence of sarcoidosis to race-specific prevalence by state. These comparisons were performed for 2019 because it was the closest to the present day available in our analysis. The overall state prevalence was calculated

Table 1. Characteristics of U.S. veterans with sarcoidosis and control subjects

	Control (n* = 71,241)	Sarcoidosis (n = 23,747)
Proportion of patients receiving non–VA-purchased care	5,937 (8.3)	1,979 (8.3)
Age, yr†	55 (47–63)	55 (47–63)
Sex		
Male	64,831 (91.0)	20,535 (86.5)
Female	6,396 (9.0)	3,202 (13.5)
Race		
White	45,932 (64.5)	9,678 (40.8)
Black	12,142 (17.0)	12,379 (52.1)
Asian and other	1,785 (2.5)	261 (1.1)
Unknown	11,382 (16)	1,429 (6.0)
Ethnicity		
Hispanic	3,710 (5.2)	680 (2.8)
Non-Hispanic	52,357 (81.9)	22,288 (93.9)
Unknown	9,174 (12.9)	779 (3.3)
Residence by region		
South	32,098 (45.1)	12,871 (54.2)
Northeast	8,699 (12.2)	3,271 (13.8)
Midwest	14,735 (20.7)	4,252 (18.0)
West	14,786 (20.8)	3,145 (13.2)
U.S. territories and non-U.S.	550 (0.7)	127 (0.5)
Unknown	373 (0.5)	81 (0.3)
Rural residence		
Rural	23,623 (33.2)	6,493 (27.3)
Urban	47,118 (66.1)	17,133 (72.2)
Unknown	500 (0.7)	121 (0.5)
Branch of service		
Air Force	7,656 (10.8)	3,129 (13.2)
Army	24,775 (34.8)	11,803 (49.7)
Marine	5,124 (7.2)	1,926 (8.1)
Navy	9,903 (13.9)	3,494 (14.7)
Other	21,180 (29.7)	2,213 (9.3)
Multiple	2,155 (3.0)	1,130 (4.8)
Unknown	448 (0.6)	52 (0.2)
Tobacco use		
No	25,315 (35.5)	6,115 (25.8)
Yes	45,926 (64.5)	17,632 (74.2)

Definition of abbreviation: VA = Veterans Affairs.
Data are presented as *n* (column %) unless otherwise noted. Control subjects were identified using incident density sampling from the total population available through VA Corporate Data Warehouse without a diagnostic code for sarcoidosis by the end of each calendar year. Then propensity score matching was applied to select three control cases for each sarcoidosis case.
*Number of cases.
†Median (25th–75th percentiles).

by identifying the number of unique veterans with sarcoidosis in each state divided by the total number of veterans living in that state. A similar approach using race-specific numerators and denominators was used in calculating the 2019 prevalence among Black versus White veterans for each state.

Descriptive statistics summarize the data for control and sarcoidosis cases. Categorical variables are presented as the number of cases and percentage frequency, and continuous variables are presented as the median and 25th–75th percentiles. We performed logistic regression to examine factors measured at the index date that were associated with the incidence. Through this multivariable analysis, we aimed to explore whether the incidence of sarcoidosis differed by age, sex, race, branch of service, geographic region, or history of tobacco use. These variables were selected based on prior literature using a directed acyclic graph (2, 3, 28–30). Missing data for categorical variables were coded as "unknown" categories. Statistical significance was defined as $P < 0.05$, and all hypothesis tests were two-sided.

Results

Patient Characteristics
We screened 13,320,196 veterans who received care through or paid for by the VHA between January 2000 and the end of December 2020 (Figure 1). We then excluded 1,708 veterans who had both a diagnostic code for sarcoidosis between 2002 and 2019 and an alternative diagnostic code within 3 years of the sarcoidosis diagnosis date. Among the remaining 13,318,488 veterans, we identified 23,747 patients between 2004 and 2019 with *1*) one inpatient or two outpatient diagnoses at least 1 month apart; and *2*) no alternative diagnoses. Compared with those without the disease (Table 1), including veterans with sarcoidosis were more likely to be female (13.5% vs. 9.0%), of Black race (52.2% vs. 17.0%), living in the South (54.2% vs. 45.1%), ever–tobacco users (74.2% vs. 64.5%), and Army veterans (49.7% vs. 34.8%).

Incidence and Prevalence of Sarcoidosis in the VHA
The annual prevalence of sarcoidosis diagnosis increased from 79 cases/100,000 persons in 2003 to 141 cases/100,000 persons in 2019 (Figure 2). The annual incidence increased from 38 cases/100,000 person-years in 2004 to 52 cases/100,000 person-years in 2019. A trend analysis of the incidence confirmed these findings. Between 2004 and 2015, we observed an absolute annual increase in the incidence of 0.72 cases/100,000 person-years (95% confidence interval [CI], 0.36–1.08; $P < .001$) compared with the preceding year. This annual trend in sarcoidosis incidence increased by 0.62, resulting in a slope of 1.34 throughout 2016–2019 (Figure 3). However, the difference between the two slopes is not statistically significant (95% CI, −0.75 to 1.99; $P = 0.35$; Table 2).

Risk Factors Associated with a New Diagnosis of Sarcoidosis
Black race; female sex; living in the Northeast, South, or Midwest regions; history of tobacco use; and serving in the Army, Air Force, or multiple service branches were significantly associated at $P < 0.0001$ with higher odds of incident sarcoidosis (Table 3). Black race had an odds ratio (OR) of 4.49 (95% CI, 4.33–4.65), female sex had an OR of 1.64 (95% CI, 1.56–1.73), living in the Northeast compared with the western region had the highest OR of 1.57 (95% CI, 1.48–1.67) compared with other regions, history of tobacco use had an OR of 1.36 (95% CI, 1.31–1.41), and serving in the Army, Air Force, or multiple service branches compared with the Navy had an OR of 1.08 (95% CI, 1.03–1.13), 1.10 (95%
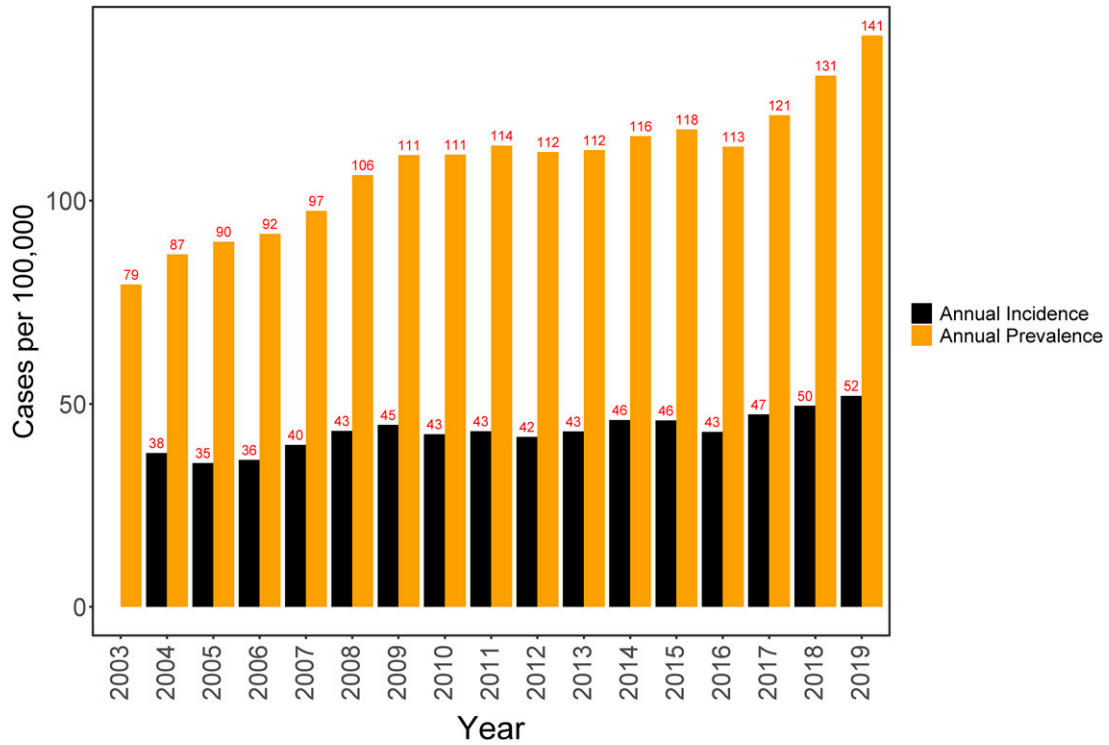
Figure 2. Annual incidence and prevalence of sarcoidosis diagnosis between 2003 and 2019.
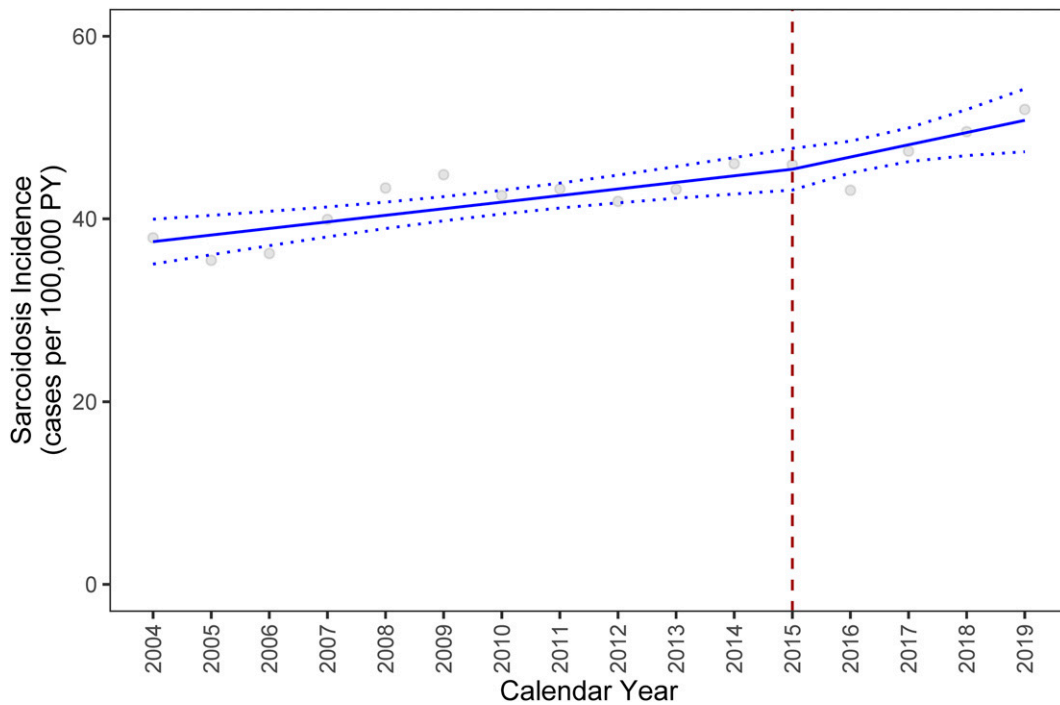
Figure 3. Trend analysis of the annual sarcoidosis incidence between 2004 and 2019 with fitted values (solid line) and 95% confidence intervals (dotted lines). Red dashed line represents implementation of International Classification of Diseases Tenth Revision coding system. The following model was specified to estimate the trend in the incidence of sarcoidosis across the two periods of the resultant: $Y(incidence) = \beta_0 (2004) + \beta_1 (Year\ 2005–2015) + \beta_2 (Year\ 2016–2019)$. PY = person-years.

Table 2. Parameter estimates from the linear segmented model using a time-interrupted series of the annual sarcoidosis incidence over the study period

Segmented Linear Regression		
Interval	Coefficient (95% Confidence Interval)	P Value
Intercept (2004)	37.50 (35.03 to 39.95)	<0.001
Index year 2005–2015	0.72 (0.36 to 1.08)	<0.001
Index year 2016–2019	0.62 (−0.75 to 1.99)	0.35

CI, 1.04–1.17), and 1.27 (95% CI, 1.16–1.39), respectively.

Geographic Distribution
The geographic burden of sarcoidosis was heterogeneous (Figure 4). The states with the highest 2019 prevalence include Mississippi, Alabama, and the South Atlantic region, except for Florida. The overall prevalence in 2019 ranged from 46 cases/100,000 persons in South Dakota and Hawaii (lowest) to 408 cases/100,000 persons in the District of Columbia (highest). Black veterans had the highest prevalence in racially stratified analysis. The highest prevalence among Black veterans was in New Hampshire (2,217 cases/100,000 persons), whereas, among White veterans, the highest prevalence was in Vermont (136 cases/100,000 persons) (Table E5). We found no marked regional concentrations of race-specific prevalence compared with the overall prevalence by states (Figure 4).

Discussion

Principal Findings
This study represents a comprehensive epidemiological evaluation of sarcoidosis among U.S. veterans. In 13,320,196 veterans, the annual prevalence and incidence of sarcoidosis between 2004 and 2019 were increasing and dramatically higher among veterans than in non-veteran cohorts (2, 31). The annual prevalence was about two times higher (141 vs. 60 cases/100,000 persons), and the annual incidence was about four to five times higher (52 vs. 10 cases/100,000 person-years). Independent risk factors for incident sarcoidosis included Black race, female sex, living in any region of the continental United States outside the West, history of tobacco use, and serving in the Army, Air Force, or multiple service branches. In 2019, the greatest regional prevalence of sarcoidosis was in the South Atlantic. The high burden of sarcoidosis among veterans, which may relate to their unique exposure profile, provides a rationale for VHA's investment in services like pulmonary rehabilitation. For clinicians, understanding sarcoidosis incidence and prevalence may aid in the diagnosis and management of potential cases and determine their associated disease trajectories, morbidities, and mortalities.

The notably higher annual incidence and prevalence of the disease than most previous non-veteran studies (2, 7, 31–34) suggests the following two hypotheses:

1. In offering comprehensive longitudinal care to its enrollees, the VHA may better capture actual disease incidence. Patient cost is often cited as a factor in the loss of follow-up, especially where interventions require multiple visits (35). The complex process of securing the diagnosis may be particularly vulnerable to such dynamics but is addressed by universal access within the VA system (36). Weighing against this proposition is that this incidence is also substantially higher than that reported in countries with universal healthcare systems (33, 37).

Table 3. Multivariable model evaluating risk factors associated with new-onset sarcoidosis in U.S. veterans

	Logistic Regression	
	Sarcoidosis (n* = 23,747) Control (n = 71,241)	
	OR (95% CI)	P Value
Sex		<0.0001
Male	Reference	
Female	1.64 (1.56–1.73)	
Race		<0.0001
White	Reference	
Black	4.49 (4.33–4.65)	
Asian and other	0.78 (0.68–0.89)	
Unknown	0.71 (0.66–0.75)	
Residence by region		<0.0001
West	Reference	
South	1.24 (1.18–1.30)	
Northeast	1.57 (1.48–1.67)	
Midwest	1.20 (1.13–1.27)	
US territories	1.08 (0.87–1.33)	
Unknown	1.03 (0.78–1.38)	
Rural residence		0.41
Urban	Reference	
Rural	1.01 (0.98–1.05)	
Unknown	0.88 (0.69–1.11)	
Branch of service		<0.0001
Navy	Reference	
Army	1.08 (1.03–1.13)	
Marine	0.96 (0.90–1.03)	
Air Force	1.10 (1.04–1.17)	
Other	0.29 (0.28–0.31)	
Multiple	1.27 (1.16–1.39)	
Unknown	0.39 (0.29–0.53)	
Tobacco use		<0.0001
No	Reference	
Yes	1.36 (1.31–1.41)	

Definition of abbreviations: CI = confidence interval; OR = odds ratio.
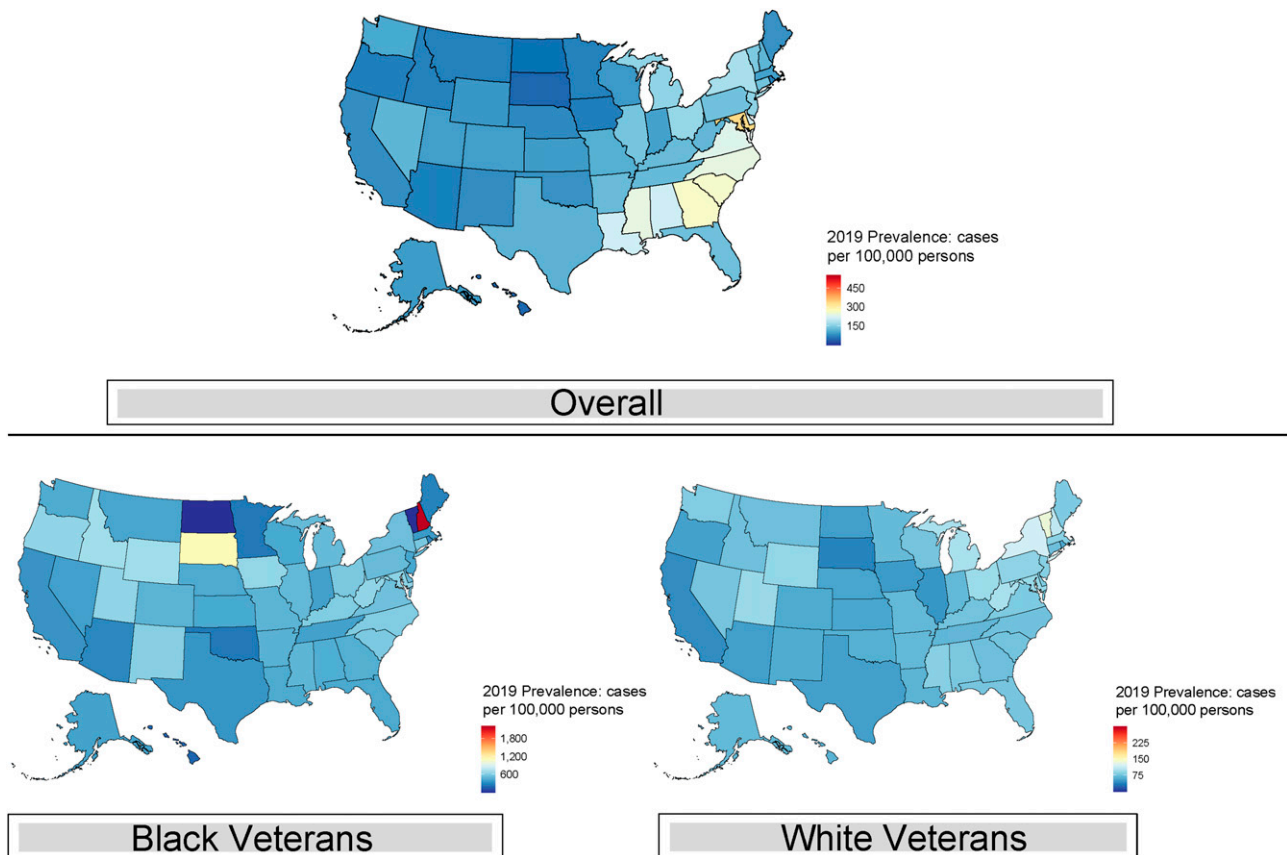*Number of cases.

Figure 4. Geographic distribution of prevalent sarcoidosis cases (overall and race specific) among U.S. veterans in 2019.

2. The incidence may truly be higher among veterans than among the general population. This epidemiological pattern is also seen in veterans with other lung autoimmune diseases, such as idiopathic pulmonary fibrosis, systemic lupus erythematosus, and rheumatoid arthritis (13, 38). Environmental exposures have long been theorized as necessary for sarcoidosis. Military service, which is associated with a range of potentially important and unusual exposures, might be one plausible causative pathway (14). Regardless, as with combat-related injuries, our findings suggest that veterans are a population with unique healthcare needs that may benefit from specialized, dedicated care.

Our study also allowed the assessment of temporal trends in annual incidence and prevalence. The first joint statement on sarcoidosis was released in 1999 to improve patient case detection and clinical care (21). We found that annual incidence increased continuously from 2004 to 2009, plateaued (around 45 cases/100,000 person-years) between 2009 and 2015, and then increased again (to 52 cases/100,000 person-years in 2019) after the October 2015 implementation of ICD-10. It is plausible that the Joint Statement generated new interest in this previously underdiagnosed disease, similar to other highly underdiagnosed respiratory disorders (39, 40). The unique structure of the VHA could have ideally positioned it to capture this effect. However, we did not find a significant change in the annual incidence trend after the implementation of ICD-10. The lack of change in our study also suggests the Veterans Access Choice and Accountability Act of 2014 (41) may not have greatly impacted the incidence of sarcoidosis.

Despite an overall study population that reflects U.S. veterans, our results agreed with prior studies in identifying a greater proportion of cases among Black and female veterans than control subjects, with both demographic traits as risk factors for sarcoidosis (2, 5, 42–44). Moreover, in 2019, Black veterans had a higher prevalence than White veterans. Our study also agrees with the most recent sarcoidosis population-based studies in identifying an older age of diagnosis (2, 31, 45). These changes in the age-specific distribution of sarcoidosis could be related to improved sarcoidosis detection. This focus on an older population also aligns with recent research suggesting the timing of menopause may be an influential risk factor in females from minority racial and/or ethnic backgrounds (46).

The increased odds of sarcoidosis in patients with a history of tobacco use highlights essential issues about the complexity of the disease process. Most previous cohorts have reported higher risk among nonsmokers (28, 29, 47). However, smoking is associated with increased gallium uptake, a common index of sarcoidosis activity (48). A prior Japanese cohort found an increased risk among smokers and theorized mechanistically gene–environment interaction (49). Supporting this logic, a large study of genetic risk in sarcoidosis found that the strongest association among those with

non-Lofgren's disease was a single-nucleotide polymorphism that conferred increased risk on smokers (50). Although we did not have the genetic data necessary to explore these hypotheses, our analyses add to a growing body of evidence suggesting an adverse role of tobacco use in sarcoidosis.

As in previous population-based studies, we found notable variation in the geographic distribution of sarcoidosis, with the highest prevalence in the South Atlantic (2, 43). Although previous studies suggest this may be due to local differences in environmental exposures (2), this is difficult to disentangle from differences in regional expertise and resources available for sarcoidosis, as these factors tend to become interrelated over time (i.e., areas with heavier disease burden gain more experience in management).

Like many previous studies, we found an increased disease risk among our Black participants (51). We treated this finding as more sociological than biological because under the historic "one-drop rule," individuals with even a single Black ancestor often self-identify as Black race (52). This is especially the case in handling EHR data, which are marked by inconsistency and irregular practices in recording race and ethnicity data even when standardized policies are in place (52). The increased odds among Black veterans may suggest unique gene–environment interactions that increase risk or something about the sociological environment for marginalized populations. To exemplify the latter point, recent work suggests race may capture and obscure the health-related consequences of racism when listed as a disease or physiological determinant in clinical medicine (53). Given that neighborhood deprivation is also associated with incident sarcoidosis (54) and that Black individuals are disproportionately housed in high-deprivation neighborhoods

because of historically discriminatory policies (55), the influence of structural racism on the incidence and course of sarcoidosis is an urgent research priority.

Our work identified that sarcoidosis risk varied by the branch of service, which has been considered a proxy for exposure history as it is less influenced by recall bias (56). In other idiopathic diseases, the differing exposure histories between service branches have been independently associated with the odds of disease incidence and survival (56, 57). In some cases, plausible mechanisms are readily apparent. Beryllium exposure is common during airplane construction and repair; because it also causes sarcoidosis-like disease, this could plausibly contribute to a higher incidence in Air Force veterans (58). The meaning of other associations is less clear. An interesting avenue of analysis might be whether combat theater or service generation may be a similarly informative proxy for different exposure histories. It may elucidate the increasing incidence and prevalence of disease, together with a better understanding of mortality trends. Although suggestive data about improving mortality have appeared regarding cardiac sarcoidosis (13), a comprehensive examination across all manifestations awaits.

Limitations

First, although diagnostic claims to identify sarcoidosis cases in the EHR have a misclassification risk, our case definition was more restrictive than commonly used approaches (2, 22) or previously validated algorithms (23). Any misclassification of patients with sarcoidosis to the control group would have resulted in a falsely lower (not higher) incidence and prevalence of sarcoidosis. We did not include biopsy-related codes because a sizable proportion of histopathological information is in the unstructured domain (23). Second, our data

collection period after the end of major overseas wars was limited. Preclinical undiagnosed disease of as long as 5 years' duration is a known phenomenon in sarcoidosis (59). Latency periods in sarcoidosis are poorly studied and require better characterization (17). We may therefore underestimate the true peak of sarcoidosis incidence and prevalence if there are additional exposure-related cases to be captured from these events. Third, previous early work in veteran cohorts suggests that the regional pattern of sarcoidosis prevalence is robust to the effects of race (60). Further research is needed to elucidate the primary unmeasured drivers of this geographic variability.

Conclusions

This epidemiological characterization of sarcoidosis among U.S. veterans with VHA-delivered or -purchased care identified a higher annual incidence and prevalence of sarcoidosis than general population studies. Alongside traditionally recognized risk factors, we identified relatively novel increased odds of sarcoidosis associated with tobacco use within the VA population and varying odds of sarcoidosis by a veteran's branch of service. Because of those findings and the observed regional variance, the study serves as an important stimulus for future research into the environmental exposures that may influence the cause and course of sarcoidosis as well as define the important clinical phenotypes. ∎

References

1 Crouser ED, Maier LA, Wilson KC, Bonham CA, Morgenthau AS, Patterson KC, *et al.* Diagnosis and detection of sarcoidosis. An official American Thoracic Society clinical practice guideline. *Am J Respir Crit Care Med* 2020;201:e26–e51.

2 Baughman RP, Field S, Costabel U, Crystal RG, Culver DA, Drent M, *et al.* Sarcoidosis in America: analysis based on health care use. *Ann Am Thorac Soc* 2016;13:1244–1252.

3 Swigris JJ, Olson AL, Huie TJ, Fernandez-Perez ER, Solomon J, Sprunger D, *et al.* Sarcoidosis-related mortality in the United States from 1988 to 2007. *Am J Respir Crit Care Med* 2011;183:1524–1530.

4 National Center for Veterans Analysis and Statistics. Profile of veterans: 2017 report. 2019 [accessed 2022 Feb 5]. Available from: https://www.va.gov/vetdata/docs/SpecialReports/Profile_of_Veterans_2017.pdf.

5 National Center for Veteran Analysis and Statistics. Department of Veterans Affairs statistics at a glance. 2020 [updated 2019 Dec 31; accessed 2021 Jan 1]. Available from: https://www.va.gov/vetdata/docs/Quickfacts/Stats_at_a_glance_4_6_20.PDF.

6 Sheehan CM, Hayward MD. Black/white differences in mortality among veteran and non-veteran males. *Soc Sci Res* 2019;79:101–114.

7 Parrish SC, Lin TK, Sicignano NM, Lazarus AA. Sarcoidosis in the United States military health system. *Sarcoidosis Vasc Diffuse Lung Dis* 2018;35:261–267.

8   Judson MA. Environmental risk factors for sarcoidosis. *Front Immunol* 2020;11:1340.

9   Newman KL, Newman LS. Occupational causes of sarcoidosis. *Curr Opin Allergy Clin Immunol* 2012;12:145–150.

10 Izbicki G, Chavko R, Banauch GI, Weiden MD, Berger KI, Aldrich TK, *et al.* World Trade Center "sarcoid-like" granulomatous pulmonary disease in New York City Fire Department rescue workers. *Chest* 2007; 131:1414–1423.

11 Olenick M, Flowers M, Diaz VJ. US veterans and their unique issues: enhancing health care professional awareness. *Adv Med Educ Pract* 2015;6:635–639.

12 Sheikh G, Daouk S. Sarcoidosis in veterans: environmental exposures in deployment [abstract]. *Chest* 2019;156:A1674.

13 Ying D, Schmajuk G, Trupin L, Blanc PD. Inorganic dust exposure during military service as a predictor of rheumatoid arthritis and other autoimmune conditions. *ACR Open Rheumatol* 2021;3:466–474.

14 Geretto M, Ferrari M, De Angelis R, Crociata F, Sebastiani N, Pulliero A, *et al.* Occupational exposures and environmental health hazards of military personnel. *Int J Environ Res Public Health* 2021;18:5395.

15 Kajdasz DK, Lackland DT, Mohr LC, Judson MA. A current assessment of rurally linked exposures as potential risk factors for sarcoidosis. *Ann Epidemiol* 2001;11:111–117.

16 Banoei MM, Iupe I, Bazaz RD, Campos M, Vogel HJ, Winston BW, *et al.* Metabolomic and metallomic profile differences between veterans and civilians with pulmonary sarcoidosis. *Sci Rep* 2019;9:19584.

17 Ronsmans S, De Ridder J, Vandebroek E, Keirsbilck S, Nemery B, Hoet PHM, *et al.* Associations between occupational and environmental exposures and organ involvement in sarcoidosis: a retrospective case-case analysis. *Respir Res* 2021;22:224.

18 Centers for Medicare & Medicaid Services. ICD-10. 2021 [accessed 2022 Mar 29]. Available from: https://www.cms.gov/Medicare/Coding/ICD10.

19 Seedahmed MI, Chen J, Potter J, Mcculloch C, Whooley M, Koth L, *et al.* The epidemiology of sarcoidosis among US veterans, 2000–2020: improving care for patients with lung disease or critical illness through training and evaluation [abstract]. *Am J Respir Crit Care Med* 2022; 205:A1003.

20 U.S. Department of Veterans Affairs. VA Informatics and Computing Infrastructure (VINCI). [updated 2022 Mar 16; accessed 2021 Jun 9]. Available from: https://www.research.va.gov/programs/vinci/default.cfm.

21 Costabel U, Hunninghake GW; Sarcoidosis Statement Committee. American Thoracic Society. European Respiratory Society. World Association for Sarcoidosis and Other Granulomatous Disorders. ATS/ ERS/WASOG statement on sarcoidosis. *Eur Respir J* 1999;14:735–737.

22 Ceder S, Rossides M, Kullberg S, Eklund A, Grunewald J, Arkema EV. Positive predictive value of sarcoidosis identified in an administrative healthcare registry: a validation study. *Epidemiology* 2021;32:444–447.

23 Seedahmed MI, Mogilnicka I, Zeng S, Luo G, Whooley MA, McCulloch CE, *et al.* Performance of a computational phenotyping algorithm for sarcoidosis using diagnostic codes in electronic medical records: case validation study from 2 Veterans Affairs medical centers. *JMIR Form Res* 2022;6:e31615.

24 United States Census Bureau. 2010 Census regions and divisions of the United States. 2021 [accessed 2021 Dec 10]. Available from: https:// www.census.gov/geographies/reference-maps/2010/geo/2010-census-regions-and-divisions-of-the-united-states.html.

25 U.S. Department of Agriculture Economic Research Service. Rural-urban commuting area codes. 2020 [accessed 2021 Dec 10]. Available from: https://www.ers.usda.gov/data-products/rural-urban-commuting-area-codes/.

26 Calhoun PS, Wilson SM, Hertzberg JS, Kirby AC, McDonald SD, Dennis PA, *et al.*; VA Mid-Atlantic MIRECC Workgroup. Validation of Veterans Affairs electronic medical record smoking data among Iraq- and Afghanistan-era veterans. *J Gen Intern Med* 2017;32:1228–1234.

27 Melzer AC, Pinsker EA, Clothier B, Noorbaloochi S, Burgess DJ, Danan ER, *et al.* Validating the use of veterans affairs tobacco health factors for assessing change in smoking status: accuracy, availability, and approach. *BMC Med Res Methodol* 2018;18:39.

28 Newman LS, Rose CS, Bresnitz EA, Rossman MD, Barnard J, Frederick M, *et al.*; ACCESS Research Group. A case control etiologic study of sarcoidosis: environmental and occupational risk factors. *Am J Respir Crit Care Med* 2004;170:1324–1330.

29 Ungprasert P, Crowson CS, Matteson EL. Smoking, obesity and risk of sarcoidosis: a population-based nested case-control study. *Respir Med* 2016;120:87–90.

30 Jajosky P. Sarcoidosis diagnoses among U.S. military personnel: trends and ship assignment associations. *Am J Prev Med* 1998;14:176–183.

31 Ungprasert P, Carmona EM, Utz JP, Ryu JH, Crowson CS, Matteson EL. Epidemiology of sarcoidosis 1946-2013: a population-based study. *Mayo Clin Proc* 2016;91:183–188.

32 Fidler LM, Balter M, Fisher JH, To T, Stanbrook MB, Gershon A. Epidemiology and health outcomes of sarcoidosis in a universal healthcare population: a cohort study. *Eur Respir J* 2019;54:1900444.

33 Arkema EV, Cozier YC. Epidemiology of sarcoidosis: current findings and future directions. *Ther Adv Chronic Dis* 2018;9:227–240.

34 Dumas O, Abramovitz L, Wiley AS, Cozier YC, Camargo CA Jr. Epidemiology of sarcoidosis in a prospective cohort study of U.S. women. *Ann Am Thorac Soc* 2016;13:67–71.

35 Gao X, Obeid A, Aderman CM, Talcott KE, Ali FS, Adam MK, *et al.* Loss to follow-up after intravitreal anti-vascular endothelial growth factor injections in patients with diabetic macular edema. *Ophthalmol Retina* 2019;3:230–236.

36 Judson MA, Thompson BW, Rabin DL, Steimel J, Knatterud GL, Lackland DT, *et al.*; ACCESS Research Group. The diagnostic pathway to sarcoidosis. *Chest* 2003;123:406–412.

37 Deubelbeiss U, Gemperli A, Schindler C, Baty F, Brutsche MH. Prevalence of sarcoidosis in Switzerland is associated with environmental factors. *Eur Respir J* 2010;35:1088–1097.

38 Kaul B, Lee JS, Zhang N, Vittinghoff E, Sarmiento K, Collard HR, *et al.* Epidemiology of idiopathic pulmonary fibrosis among US veterans, 2010–2019. *Ann Am Thorac Soc* 2022;19:196–203.

39 Çolak Y, Afzal S, Nordestgaard BG, Vestbo J, Lange P. Prognosis of asymptomatic and symptomatic, undiagnosed COPD in the general population in Denmark: a prospective cohort study. *Lancet Respir Med* 2017;5:426–434.

40 Aaron SD, Boulet LP, Reddel HK, Gershon AS. Underdiagnosis and overdiagnosis of asthma. *Am J Respir Crit Care Med* 2018;198: 1012–1020.

41 U.S. Department of Veterans Affairs. Veterans Access, Choice and Accountability Act of 2014 ("Choice Act"). [created 2014 Sep 26; accessed 2022 Nov 16]. Available from: https://www.va.gov/opa/ choiceact/documents/choice-act-summary.pdf.

42 Mirsaeidi M, Machado RF, Schraufnagel D, Sweiss NJ, Baughman RP. Racial difference in sarcoidosis mortality in the United States. *Chest* 2015;147:438–449.

43 Cozier YC, Berman JS, Palmer JR, Boggs DA, Serlin DM, Rosenberg L. Sarcoidosis in Black women in the United States: data from the Black Women's Health Study. *Chest* 2011;139:144–150.

44 Baughman RP, Teirstein AS, Judson MA, Rossman MD, Yeager H Jr, Bresnitz EA, *et al.*; Case Control Etiologic Study of Sarcoidosis (ACCESS) research group. Clinical characteristics of patients in a case control study of sarcoidosis. *Am J Respir Crit Care Med* 2001;164: 1885–1889.

45 Sawahata M, Sugiyama Y, Nakamura Y, Nakayama M, Mato N, Yamasawa H, *et al.* Age-related and historical changes in the clinical characteristics of sarcoidosis in Japan. *Respir Med* 2015;109:272–278.

46 Cozier YC, Berman JS, Palmer JR, Boggs DA, Wise LA, Rosenberg L. Reproductive and hormonal factors in relation to incidence of sarcoidosis in US Black women: the Black Women's Health Study. *Am J Epidemiol* 2012;176:635–641.

47 Murin S, Bilello KS, Matthay R. Other smoking-affected pulmonary diseases. *Clin Chest Med* 2000;21:121–137, ix.

48 Valeyre D, Soler P, Clerici C, Pré J, Battesti JP, Georges R, *et al.* Smoking and pulmonary sarcoidosis: effect of cigarette smoking on prevalence, clinical manifestations, alveolitis, and evolution of the disease. *Thorax* 1988;43:516–524.

49 Hattori T, Konno S, Shijubo N, Ohmichi M, Nishimura M. Increased prevalence of cigarette smoking in Japanese patients with sarcoidosis. *Respirology* 2013;18:1152–1157.

50 Rivera NV, Patasova K, Kullberg S, Diaz-Gallo LM, Iseda T, Bengtsson C, *et al.* A gene-environment interaction between smoking and gene polymorphisms provides a high risk of two subgroups of sarcoidosis. *Sci Rep* 2019;9:18633.

51 Mersha TB, Abebe T. Self-reported race/ethnicity in the age of genomic research: its potential impact on understanding health disparities. *Hum Genomics* 2015;9:1.

52 Lee SJC, Grobe JE, Tiro JA. Assessing race and ethnicity data quality across cancer registries and EMRs in two hospitals. *J Am Med Inform Assoc* 2016;23:627–634.

53 Baugh AD, Shiboski S, Hansel NN, Ortega V, Barjaktarevic I, Barr RG, *et al.* Reconsidering the utility of race-specific lung function prediction equations. *Am J Respir Crit Care Med* 2022;205: 819–829.

54 Li X, Sundquist J, Hamano T, Sundquist K. Neighborhood deprivation and risks of autoimmune disorders: a national cohort study in Sweden. *Int J Environ Res Public Health* 2019;16:3798.

55 Mujahid MS, Gao X, Tabb LP, Morris C, Lewis TT. Historical redlining and cardiovascular health: the Multi-Ethnic Study of Atherosclerosis. *Proc Natl Acad Sci USA* 2021;118:e2110986118.

56 Beard JD, Engel LS, Richardson DB, Gammon MD, Baird C, Umbach DM, *et al.* Military service, deployments, and exposures in relation to amyotrophic lateral sclerosis etiology. *Environ Int* 2016;91:104–115.

57 Beard JD, Engel LS, Richardson DB, Gammon MD, Baird C, Umbach DM, *et al.* Military service, deployments, and exposures in relation to amyotrophic lateral sclerosis survival. *PLoS One* 2017;12:e0185751.

58 Blanc PD, Annesi-Maesano I, Balmes JR, Cummings KJ, Fishwick D, Miedinger D, *et al.* The occupational burden of nonmalignant respiratory diseases: an official American Thoracic Society and European Respiratory Society statement. *Am J Respir Crit Care Med* 2019;199:1312–1334.

59 Arkema EV, Eklund A, Grunewald J, Bruze G. Work ability before and after sarcoidosis diagnosis in Sweden. *Respir Med* 2018;144S:S7–S12.

60 Gentry JT, Nitowsky HM, Michael M Jr. Studies on the epidemiology of sarcoidosis in the United States: the relationship to soil areas and to urban-rural residence. *J Clin Invest* 1955;34:1839–1856.