The prevalence and geographic distribution of sarcoidosis in the United States

To the Editor: Sarcoidosis is a systemic granulomatous disease with unknown etiology with an estimated prevalence of 60 per 100,000. Previous studies report higher incidence of sarcoidosis in the Midwest and Northeast compared to the Southwest United States.[1] It has been hypothesized that components of atmospheric particulate matter or infectious causes could serve as immunogenic triggers, causing a systemic granulomatous response.[2] Further, prevalence rates for sarcoidosis are elevated among those with Scandinavian or Black/African American ancestry.[1,3,4] African Americans were 3 times as likely to be diagnosed with sarcoidosis when compared to Caucasians, regardless of geographical location.[3] The purpose of this study is to explore the spatial variation of sarcoidosis prevalence to identify clusters of disease that may be explained by geographical factors.

Data were obtained from IBM MarketScan Research Databases containing de-identified health care claims for continuously enrolled and privately insured individuals in the United States from 2012 to 2018. The 7-year prevalence of sarcoidosis was calculated by identifying the total number of adult patients with 2 or more health care claims for sarcoidosis separated by at least 1 year over the total population and stratified by core-based statistical area, a geographical area typically comprised of 1 or more adjacent counties within the United States and located in metropolitan areas. SaTScan v9.6 was used to identify geographical clusters of sarcoidosis. Relative risks and *P* values were calculated and exported into Esri ArcGIS Desktop 10.8 for mapping. This process was repeated while controlling for race by including the percentage of Black/African Americans in the core-based statistical area population based on the 2010 US Census.

In total, 37,688 individuals with sarcoidosis were identified with a mean (SD) age of 50.5 (8.50) years old, 59.3% female. The estimated 7-year prevalence of sarcoidosis in the United States was 35.22 per 100,000. 13 of 17 clusters were significant (Table I) and were located on the East Coast of the United States (Fig 1). After controlling for race, only 1 cluster remained (not statistically significant), suggesting that geographical racial variation in the United States may explain the spatial clustering of sarcoidosis prevalence across the United States.

Limitations include that prevalence was based on health insurance claims and therefore some misclassification errors could be present. Furthermore, only those with private insurance and residing in a core-based statistical area were included in the study, which could contribute to selection bias. Despite these limitations, the Optum health care database was queried from 2010 to 2013, 14,482 (73.93%) patients with sarcoidosis had commercial insurance, with 5107 (26.07%) having Medicare supporting the generalizability of these findings in terms of insurance coverage.[4] Given that significant clusters were identified (regardless of race) near metropolitan areas, the presence of these clusters could be impacted by other social, geographical, or environmental characteristics. Future studies are needed to determine if there are other factors influencing the incidence of sarcoidosis.

Cluster analysis allows us to present a novel visual approach to identify patterns in sarcoidosis prevalence in the United States (Fig 1). Significant clusters of sarcoidosis were detected on the East Coast of the United States and near urban areas, implying that racial, socioeconomic, and environmental factors may potentially drive this geographical distribution.

Hannah H. Nam, BA,[a] Abigail Washington, BA,[a,b] Melissa Butt, DrPH,[b,c,d] Steven Maczuga, MS,[c] Daniel Guck, DO,[e] Jeff D. Yanosky, ScD, MS,[b] and Matthew F. Helm, MD[c]

From the Department of Dermatology, Penn State College of Medicine, Hershey, Pennsylvania[a]; Department of Public Health Sciences, Penn State College of Medicine, Hershey, Pennsylvania[b]; Department of Dermatology, Penn State Health, Hershey, Pennsylvania[c]; Department of Family and Community Medicine, Penn State Health, Hershey, Pennsylvania[d]; and Division of Pulmonary and Critical Care Medicine, Penn State Health, Hershey, Pennsylvania.[e]

Funding sources: Dr Butt was supported by the Health Resources and Services Administration (HRSA) of the US Department of Health and Human Services (HHS) as part of an award totaling $381,932 with 14% financed with nongovernmental sources. The contents are those of the author(s) and do not necessarily represent the official views of, nor an endorsement by, HRSA, HHS, or the US Government. For more information, please visit HRSA.gov.

Table I. SaTScan cluster analysis results

Cluster number	Location	Radius (km)	Cases per 100,000	Relative risk	*P* value
Not controlling for race					
1	Cape May County, NJ	410.40	50.6	1.68	<.0001
2	West Princess Anne, MD	181.38	69.4	2.22	<.0001
3	Waterford Township, PA	173.50	54.4	1.76	<.0001
4	Huron Township. OH	255.84	47.3	1.54	<.0001
5	Dothan, AL	450.35	46.5	1.52	<.0001
6	Tuscaloosa County, AL	306.78	57.1	1.82	<.0001
7	Bloomingdale, GA	277.49	51.9	1.66	<.0001
8	Queensbury, NY	299.32	47.7	1.53	<.0001
9	Jackson, MI	124.46	57.7	1.82	<.0001
10	Forsyth County, NC	167.86	42.6	1.34	<.0001
11	Fort Thomas, KY	168.87	37.7	1.19	<.0001
12	Little Rock, AR	350.61	39.8	1.24	.0005
13	Lexington, KY	110.94	38.7	1.21	.0006
14	Sumter County, FL	0	13	3.51	.142
15	Highlands County, FL	75.89	109	1.40	.439
16	Shreveport, LA	246.98	237	1.22	.772
17	Scipio Township, IN	56.25	98	1.32	.976
Controlling for race					
1	Natrona County, WY	0	3	1.00	1.00

Clusters were identified using the number of cases per core-based statistical area (*CBSA*) with the total CBSA population and using a latitude and longitude coordinate system. The cluster analysis was run using a discrete Poisson probability model, with 10% of the population at risk set as the maximum spatial cluster size.

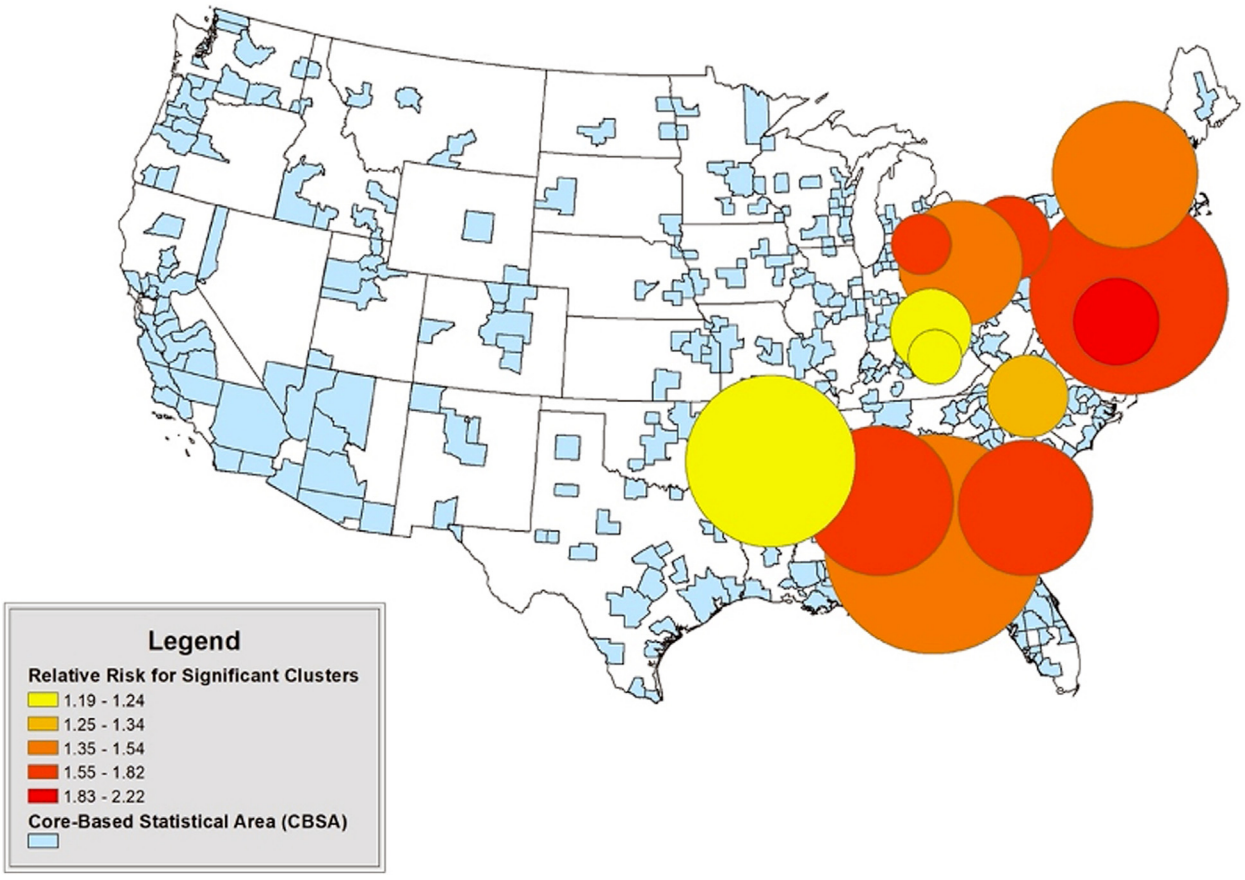

Fig 1. Relative risk for significant clusters of sarcoidosis in the United States.

IRB approval status: Health services research studies using the MarketScan Research Database were reviewed and approved by the Pennsylvania State Human Research Protection Program.

Key words: cluster analysis; epidemiology; geography; granulomatous disease; MarketScan; prevalence; sarcoid; sarcoidosis.

Correspondence to: Matthew F. Helm, MD, Department of Dermatology, Penn State Milton S. Hershey Medical Center, HU14, 500 University Dr, Suite 100, Hershey, PA 17033

E-mail: mhelm2@pennstatehealth.psu.edu

Conflicts of interest

None declared.

REFERENCES

1. Brito-Zerón P, Kostov B, Superville D, Baughman RP, Ramos-Casals M. Autoimmune Big Data Study Group. Geo-epidemiological big data approach to sarcoidosis: geographical and ethnic determinants. *Clin Exp Rheumatol*. 2019;37(6):1052-1064.
2. Starshinova AA, Malkova AM, Basantsova NY, et al. Sarcoidosis as an autoimmune disease. *Front Immunol*. 2020;10:2933. https://doi.org/10.3389/fimmu.2019.02933
3. Rybicki BA, Major M, Popovich J, Maliarik MJ, Iannuzzi MC. Racial differences in sarcoidosis incidence: a 5-year study in a health maintenance organization. *Am J Epidemiol*. 1997;145(3):234-241. https://doi.org/10.1093/oxfordjournals.aje.a009096
4. Baughman RP, Field S, Costabel U, et al. Sarcoidosis in America. Analysis based on health care use. *Ann Am Thorac Soc*. 2016;13(8):1244-1252. https://doi.org/10.1513/AnnalsATS.201511-760OC

https://doi.org/10.1016/j.jdin.2022.07.006