

# Use of Insurance Claims Data to Determine Prevalence and Confirm a Cluster of Sarcoidosis Cases in Vermont

---

LORI A. CRAGIN, MS<sup>a</sup>  
A. SCOTT LANEY, PhD<sup>a,b</sup>  
CORTLAND J. LOHFF, MD<sup>a</sup>  
BRENNAN MARTIN, MPH<sup>a</sup>  
JOHN A. PANDIANI, PhD<sup>a</sup>  
LYNN Z. BLEVINS, MD<sup>a,c</sup>

## SYNOPSIS

**Objectives.** In 2006, the Vermont Department of Health was asked to respond to a potential cluster of sarcoidosis cases related to a Vermont office building. Sarcoidosis prevalence has not been formally described for the United States. A range of <1–40/100,000 is commonly reported; however, we have not identified primary sources supporting this conclusion. Because of the wide prevalence range and lack of a local estimate, confirming existence of a cluster was difficult.

**Methods.** We ascertained the prevalence of sarcoidosis cases in Vermont by using insurance claims data to determine whether or not a cluster of sarcoidosis cases was related to the office building. We calculated county and state annual prevalence proportions for sarcoidosis for 2004 and 2005 and annual building prevalences for 1992–2006.

**Results.** The pooled sarcoidosis case prevalence for Vermont was 66.1/100,000. The pooled building annual prevalence (1,128/100,000) was statistically different from the county in which the building is located (odds ratio = 15.5, 95% confidence interval 3.0, 50.3).

**Conclusions.** We reported the first statewide sarcoidosis prevalence in the United States. This prevalence exceeded previous limited and unsubstantiated U.S. reports. Even with Vermont's elevated sarcoidosis prevalence, the presence of a cluster in this building was apparent.

---

<sup>a</sup>Vermont Department of Health, Burlington, VT

<sup>b</sup>Epidemic Intelligence Service, Centers for Disease Control and Prevention, Atlanta, GA

<sup>c</sup>Coordinating Office for Terrorism Preparedness and Emergency Response, Centers for Disease Control and Prevention, Atlanta, GA

Address correspondence to: Lori A. Cragin, MS, Vermont Department of Health, 108 Cherry St., PO Box 70, Burlington, VT 05402; tel. 802-651-1561; fax 802-863-7483; e-mail <lcragin@vdh.state.vt.us>.

©2009 Association of Schools of Public Health

A number of challenges are presented when investigating a potential disease cluster. These challenges increase substantially when the disease has an unknown etiology, a rare outcome, a variable case definition, or limited or no information regarding its expected occurrence.<sup>1</sup> Without a disease burden baseline, elevations or unusual aggregations are difficult or impossible to confirm.

The Vermont Department of Health was asked to respond to a potential cluster of sarcoidosis among occupants of a Vermont office building. Sarcoidosis is a non-caseating granulomatous disease of unknown etiology. This investigation, described by Laney et al., identified a high prevalence of sarcoidosis, asthma, and asthma-like symptoms among workers of a building with a history of water incursion and indoor environmental quality complaints.<sup>2</sup>

Although U.S. annual incidence rates of newly diagnosed sarcoidosis cases are available (2.8–10.9/100,000 for Caucasians),<sup>3–5</sup> prevalence is a more appropriate measure of disease burden for the analysis of disease clusters with cases occurring throughout multiple years, as our data indicated. While reviews of sarcoidosis cluster investigation exist,<sup>6,7</sup> to our knowledge, a nonscreening U.S. prevalence estimate has not been formally described. A U.S. prevalence range of <1–40/100,000 has been commonly reported,<sup>6,8–11</sup> however, we were unable to locate primary sources supporting this conclusion. Because of the wide prevalence range and lack of a local sarcoidosis disease burden estimate, determining whether these six sarcoidosis cases constituted a clear excess of normal expectancy was difficult.

## METHODS

### Case ascertainment: county and state

Sarcoidosis claims data were provided by five major health insurers, covering approximately 71% of Vermont's population. We defined a claim case as a person  $\geq 18$  years of age who was insured by one of the five insurers, and for whom an International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) code for sarcoidosis (ICD-9-CM code 135) had been assigned during 2004–2005. Claims case data contained date of birth, county of residence, and sex. Covered lives data contained sex and county of residence descriptors. Adults' covered lives were estimated by using census data to approximate the number of people  $\geq 18$  years of age. Claims case data did not include unique personal identifiers; therefore, we eliminated duplication among insurance databases by using date of birth, county of residence, and sex. This

was done separately for 2004–2005. In all, 50 suspected duplicates were identified and eliminated.

### Case ascertainment: building

We defined a sarcoidosis case associated with the Vermont office building as a person  $\geq 18$  years of age who worked  $\geq 20$  hours/week in the building and who self-reported a physician diagnosis of sarcoidosis during 1992–2006. X-ray findings were not available for independent review; however, biopsy confirmation was made in five of the six cases. The diagnosis had to have occurred after the person started working in the building. Institutional review board approval was obtained in accordance with the guidelines established at the Vermont Department of Health.

### Statistical analysis

We calculated county and state prevalence proportions for sarcoidosis by dividing the number of claims cases by the number of covered lives. We calculated prevalences for 2004 and 2005 by county of residence and by sex. We compared 2004 with 2005 county and state prevalences by using crude odds ratios (ORs), 95% confidence intervals (CIs), and *p*-values. To calculate pooled prevalences, we added 2004 and 2005 sarcoidosis cases and divided by the sum of the 2004 and 2005 covered lives. We calculated crude ORs, 95% exact CIs, and two-sided Fisher's exact *p*-values to compare building prevalence with the calculated Vermont state prevalence and with the prevalence for Bennington County, the county in which the building is located.

According to building management, the population was steady, with approximately 105 building occupants during 1992–1995, 130 during 1996–2003, and 136 during 2004–2006. To calculate building prevalence, we divided existing cases by the building population for each year during 1992–2006.

We calculated the incidence of sarcoidosis cases among building occupants for each year during 1992–2006 by dividing the number of new sarcoidosis cases by the number of employees in the building for that year. We calculated crude ORs, 95% CIs, and *p*-values to compare building incidence with the commonly reported U.S. incidence of 10.9/100,000.

## RESULTS

### County and state

The median age of sarcoidosis claims cases was 55 years (range: 18 to 87 years). Consistent with previous studies, patients were more likely to be female (58% in 2004 and 60% in 2005).<sup>8</sup> In 2004, 211 cases of sarcoidosis

occurred in Vermont among the 347,621 covered lives, and in 2005, 248 cases occurred among the 346,729 covered lives. This resulted in overall state prevalence proportions of 60.7/100,000 and 71.5/100,000 for 2004 and 2005, respectively. County prevalence proportions ranged from zero to 111.2 in 2004 and from 26.7 to 111.7 in 2005. Neither state, county, nor building prevalence proportions were statistically different during 2004–2005. As a result, pooled sarcoidosis prevalences were determined (Table). The pooled Bennington County prevalence was 72.8/100,000. The pooled sarcoidosis prevalence for the entire state was 66.1/100,000.

### Building

During 1992–2006, six cases of sarcoidosis were identified among 500 former and current occupants of the building, resulting in a 15-year prevalence of 1,200/100,000. At the time of the initial investigation, June 2006, sarcoidosis had been diagnosed in three of 136 building occupants, yielding a point prevalence of 2,206 cases/100,000. During 1992–2006, prevalent case counts ranged from one to four, with corresponding building prevalence proportions ranging from 952 to 2,941/100,000. The 2004 and 2005 pooled building prevalence (1,128/100,000) was statistically different from the Bennington County prevalence (OR=15.5, 95% CI 3.0, 50.3). In addition, we became aware of two new cases in 2006. Comparing the difference between the 2006 building prevalence and the 2006 Bennington County prevalence (OR=40.4, 95% CI 10.2, 117.0) reinforced the statistical difference and the appearance of an elevation.

Building incident cases ranged from zero to four. Resultant incidence rates ranged from zero to 1,471/100,000 (Figures 1 and 2).

### DISCUSSION

Characterization of a disease cluster requires knowledge of its normal expectancy. U.S. prevalence esti-

mates of sarcoidosis, a disease not routinely reported to health departments, have not been well-documented. Evidently, we reported the first statewide prevalence of sarcoidosis in the United States (66.1/100,000). Characterization of a disease cluster also requires knowledge of its normal expectancy among the population in which the cluster arises. Because sarcoidosis incidence rates vary by geography, race, and ethnicity, and because the etiology is unknown, generation of a local prevalence is a critical component of any cluster investigation.

In addition to reporting a previously unavailable prevalence, this example demonstrates that choosing the appropriate analytic metric of comparison on the basis of the epidemiology and natural history of the disease is key. In general, incidence is the most useful measure when disease is acute, curable, and of short duration. However, sarcoidosis is often nonacute and usually presents with periods of remission and reemergence, thereby extending its duration. These conditions are better measured by prevalence. When investigating clusters of rare diseases with limited case counts, annual incidence can appear unpredictable and statistically unstable. For this investigation, prevalence was a more dependable determination, consistently demonstrating an elevation and presence of a cluster.

### Limitations and strengths

Limitations of using insurance data exist because the primary purpose of insurance data is administrative (e.g., billing) and not research. This can potentially lead to misclassification. Although this study population represented approximately 71% of the Vermont population, specific groups (e.g., uninsured people, military personnel, and Medicare enrollees) were not represented. Thus, this prevalence might over- or underestimate the true sarcoidosis prevalence in Vermont. For example, people without sarcoidosis possibly received an ICD-9-CM code for sarcoidosis. However, because we were unable to eliminate duplication of covered lives among insurers, people who switched insurance providers midyear might be double-counted,

**Table. Comparison of 2004 and 2005 annual prevalence of sarcoidosis for Vermont, Bennington County, and a Bennington County office building**

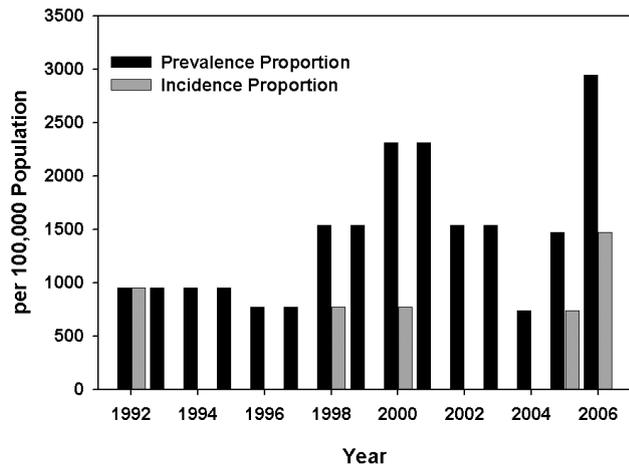
	Prevalence <sup>a</sup>			P-value <sup>b</sup>
	2004	2005	Pooled	
Office building	769.2	1,470.6	1,127.8	NA
Bennington County	67.4	78.3	72.8	0.68
State	60.7	71.5	66.1	0.08

<sup>a</sup>Per 100,000 population

<sup>b</sup>Comparison of 2004 and 2005 prevalence. P-values are two-sided Fisher's exact tests.

NA = not applicable

**Figure 1. Annual building incidence and prevalence rates of sarcoidosis from 1992 to 2006 per 100,000 people**



resulting in an inflated denominator and, therefore, an underestimated prevalence.

Additionally, our findings might not be generalizable to other populations because the study population was representative of the Vermont population, and therefore predominantly white. Conversely, because

racial disparities exist with sarcoidosis diagnoses,<sup>3,4</sup> this inherent restriction potentially controls confounding by race for our building comparison.

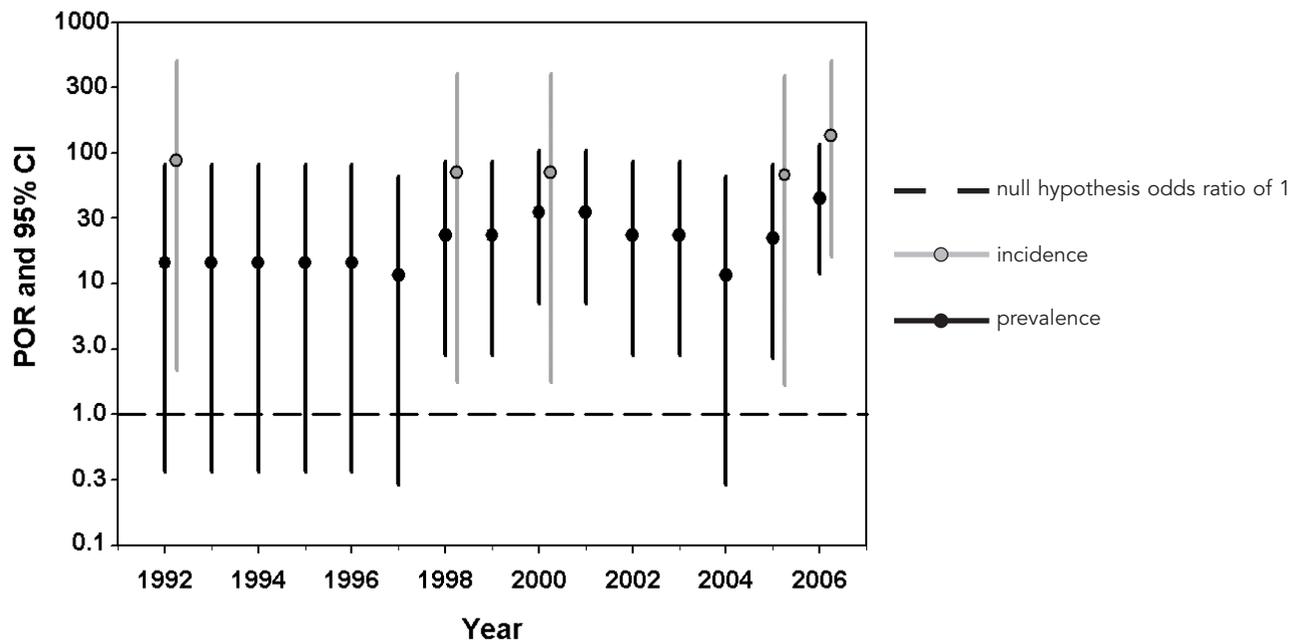
Furthermore, our estimates reflected diagnosed cases or cases with suspected clinical manifestation and might not be comparable to studies that included people who were asymptomatic (e.g., autopsies or screened population studies).<sup>9,12-18</sup>

Our primary purpose was to determine the background prevalence of disease for comparison in our cluster investigation. Although case ascertainment methodologies differed for building- and insurance-associated sarcoidosis cases, they both represented clinically symptomatic diagnosed cases. A major strength of our study was the use of a database representing a substantial portion of the Vermont population. In addition, by using data from multiple years, we were able to demonstrate stability in our estimates.

**CONCLUSIONS**

We reported the first statewide sarcoidosis prevalence in the United States. This prevalence exceeded previous limited and unsubstantiated U.S. reports. Even with Vermont's elevated sarcoidosis prevalence, the presence of a cluster in this building was apparent and, therefore, it was decided to relocate building occupants.

**Figure 2. Comparison of annual building incidence with U.S. incidence (10.9/100,000) of sarcoidosis and comparison of annual building prevalence with Vermont state prevalence (66.1/100,000) of sarcoidosis**



POR = prevalence odds ratio  
 CI = confidence interval

Given the relevance of reliable and population-specific prevalence estimates, additional U.S. population-based sarcoidosis prevalence studies are needed. A comprehensive understanding of the normal expectancy of sarcoidosis will not only aid in future cluster investigations, but also add to the understanding of the epidemiology, and potentially the etiology, of this disease.

The findings and conclusions in this article are those of the authors and do not necessarily represent the views of the Centers for Disease Control and Prevention.

## REFERENCES

1. Guidelines for investigating clusters of health events. *MMWR Recomm Rep* 1990;39(RR-11):1-23.
2. Laney AS, Cragin LA, Blevins LZ, Summer AD, Cox-Gasner JM, Kreiss K, et al. Sarcoidosis, asthma, and asthma-like symptoms among occupants of a historically water-damaged office building. *Indoor Air* 2009;19:83-90.
3. Henke CE, Henke G, Elveback LR, Beard CM, Ballard DJ, Kurland LT. The epidemiology of sarcoidosis in Rochester, Minnesota: a population-based study of incidence and survival. *Am J Epidemiol* 1986;123:840-5.
4. Rybicki BA, Major M, Popovich J Jr, Maliarik MJ, Iannuzzi MC. Racial differences in sarcoidosis incidence: a 5-year study in a health maintenance organization. *Am J Epidemiol* 1997;145:234-41.
5. Reich JM, Johnson RE. Incidence of clinically identified sarcoidosis in a northwest United States population. *Sarcoidosis Vasc Diffuse Lung Dis* 1996;13:173-7.
6. Newman LS, Rose CS, Maier LA. Sarcoidosis. *N Engl J Med* 1997;336:1224-34.
7. Hosoda Y, Yamaguchi M, Hiraga Y. Global epidemiology of sarcoidosis. What story do prevalence and incidence tell us? *Clin Chest Med* 1997;18:681-94.
8. American Thoracic Society. Statement on sarcoidosis. Joint Statement of the American Thoracic Society (ATS), the European Respiratory Society (ERS) and the World Association of Sarcoidosis and Other Granulomatous Disorders (WASOG) adopted by the ATS Board of Directors and by the ERS Executive Committee, February 1999. *Am J Respir Crit Care Med* 1999;160:736-55.
9. Hillerdal G, Nou E, Osterman K, Schmekel B. Sarcoidosis: epidemiology and prognosis. A 15-year European study. *Am Rev Respir Dis* 1984;130:29-32.
10. Fernandez Fabrellas E. [Epidemiology of sarcoidosis]. *Arch Bronconeumol* 2007;43:92-100.
11. Roberts SD, Mirowski GW, Wilkes D, Teague SD, Knox KS. Sarcoidosis. Part I: pulmonary manifestations. *J Am Acad Dermatol* 2004;51:448-51.
12. Reid JD. Sarcoidosis in coroner's autopsies: a critical evaluation of diagnosis and prevalence from Cuyahoga County, Ohio. *Sarcoidosis Vasc Diffuse Lung Dis* 1998;15:44-51.
13. Hagerstrand I, Linell F. The prevalence of sarcoidosis in the autopsy material from a Swedish town. *Acta Med Scand Suppl* 1964;425:171-4.
14. Iwai K, Sekiguti M, Hosoda Y, DeRemee RA, Tazelaar HD, Sharma OP, et al. Racial difference in cardiac sarcoidosis incidence observed at autopsy. *Sarcoidosis* 1994;11:26-31.
15. Kitaichi M. Prevalence of sarcoidosis around the world. *Sarcoidosis Vasc Diffuse Lung Dis* 1998;15:16-8.
16. Bauer HJ, Lofgren S. International study of pulmonary sarcoidosis in mass chest radiography. *Acta Med Scand Suppl* 1964;425:103-5.
17. Pollak B. Epidemiology of sarcoidosis in Canada. *Acta Med Scand Suppl* 1964;425:145.
18. Robins AB, Abeles H, Chaves AD. Prevalence and demographic characteristics of sarcoidosis in New York City. *Acta Med Scand Suppl* 1964;425:149-51.