ERJ Express. Published on June 25, 2018 as doi: 10.1183/13993003.00385-2018

EUROPEAN RESPIRATORY *journal*

FLAGSHIP SCIENTIFIC JOURNAL OF ERS

Early View

Original article

Familial Aggregation and Heritability of Sarcoidosis: A Swedish Nested Case-Control Study

Marios Rossides, Johan Grunewald, Anders Eklund, Susanna Kullberg, Daniela Di Giuseppe, Johan Askling, Elizabeth V. Arkema

Please cite this article as: Rossides M, Grunewald J, Eklund A, *et al*. Familial Aggregation and Heritability of Sarcoidosis: A Swedish Nested Case-Control Study. *Eur Respir J* 2018; in press (https://doi.org/10.1183/13993003.00385-2018).

Familial Aggregation and Heritability of Sarcoidosis: A Swedish Nested Case-Control Study

Marios Rossides[1], Johan Grunewald[2,3,4], Anders Eklund[2,3,4], Susanna Kullberg[2,3,4], Daniela Di Giuseppe[1], Johan Askling[1,5], Elizabeth V. Arkema[1]

[1] Clinical Epidemiology Division, Department of Medicine Solna, Karolinska Institutet, Stockholm, Sweden
[2] Respiratory Medicine Division, Department of Medicine Solna, Karolinska Institutet, Stockholm, Sweden
[3] Center for Molecular Medicine, Karolinska Institutet, Stockholm, Sweden
[4] Department of Respiratory Medicine, Karolinska University Hospital, Stockholm, Sweden
[5] Department of Rheumatology, Karolinska University Hospital, Stockholm, Sweden

Correspondence
Marios Rossides
Karolinska Institutet, Department of Medicine Solna, Clinical Epidemiology Division, T2, 171 76 Stockholm, Sweden
Email: marios.rossides@ki.se
Office: +46 8 517 794 06

Social media summary: Having relatives with sarcoidosis is a strong risk factor for the disease; 39% of the susceptibility is due to genetics

ABSTRACT

Sarcoidosis is believed to be caused by both genetic and environmental risk factors, but the proportion of the susceptibility to sarcoidosis which is mediated by genetics remains unknown. We aimed to estimate the familial aggregation and heritability of sarcoidosis using a case-control-family study design and population-based Swedish registers.

We identified 23880 individuals with visits for sarcoidosis in the National Patient Register using ICD codes (1964–2013). Information Löfgren's syndrome was available for a subset diagnosed at Karolinska University Hospital. General population controls were matched to cases (10:1). Relatives of cases and controls were identified from the Multi-Generation Register and ascertained for sarcoidosis in the Patient Register. We estimated familial relative risks (RR) for sarcoidosis using conditional logistic regression and the heritability using biometric models.

Having ≥1 first degree relative with sarcoidosis was associated with a 3.7-fold increase in the risk for sarcoidosis (95% CI 3.4–4.1). The RR increased in those with ≥2 relatives (RR 4.7) and in Löfgren's syndrome (RR 4.1). The heritability was 39% (95% CI 12–65).

This large investigation showed that having a relative with sarcoidosis is a very strong risk factor for the disease. Genetic variation is an important, albeit partial, contributing factor to the risk for sarcoidosis.

INTRODUCTION

Sarcoidosis is a systemic granulomatous disease affecting mostly the pulmonary and lymphatic systems. Granulomatous formations suggest that an exogenous agent causes an immunological response in a genetically predisposed individual [1]. In fact, over the years, several loci in the human genome have been implicated in the pathophysiology of the disease [2-8]. However, what proportion of the susceptibility to sarcoidosis is mediated by genetics remains unknown.

Despite several anecdotal reports of familial clustering [9-14] only a few studies reported familial relative risks for sarcoidosis (**Table 1**). The largest of those studies, the Case-Control Etiologic Study of Sarcoidosis (ACCESS), showed that having first degree relatives with the disease increased the risk of sarcoidosis by almost fourfold overall, and 17-fold for White Americans in particular [15]. In another investigation, the familial RR was estimated to be 2.5 in an African American population [16] while it ranged from 36 to 73 in a study from the UK [17]. All previous studies used questionnaires to identify relatives and ascertain their sarcoidosis status, which may have introduced bias. Some earlier investigations did not use active control groups or used unconventional statistical methods.

Familial aggregation estimates for different sarcoidosis phenotypes, such as Löfgren's vs. non-Löfgren's disease, remain unknown. Löfgren's syndrome is a distinct sarcoidosis phenotype that presents with fever, bilateral hilar lymphadenopathy, arthritis, and erythema nodosum [1]. It is characterized by acute onset, distinct pathophysiology [7, 18-20], and more favourable prognosis compared to other disease presentations [1, 18].

Familial clustering of disease is suggestive of shared genetic or environmental factors amongst family members. Heritability is a population statistic used to summarize the proportion of variation in a phenotypic trait predicted by genotypic variation in a population. Heritability of sarcoidosis was estimated to be 60–70% in two very small studies (**Table 1**) [11, 21], indicating that genetics play a

major role in the development of the disease. However, much lower heritability estimates were reported from genome-wide association studies, in which identified genetic loci accounted for less than 5% of the susceptibility to sarcoidosis [6, 7].

Investigating the familial aggregation and heritability of sarcoidosis is valuable for better understanding the aetiology of sarcoidosis and guiding future epidemiological and molecular research. It is also clinically relevant, as familial relative risks could be used for patient counselling and aid differential diagnosis. The specific objectives of our study were to estimate (a) familial aggregation estimates, overall and stratified by age at diagnosis, sex, kinship, number of affected relatives, and sarcoidosis phenotype, and (b) the heritability of sarcoidosis, using population-based Swedish register data.

METHODS

Study design and data sources

We conducted a case-control-family study nested in Swedish population-based registers. In Sweden, residents are assigned a personal identification number that can be used to link their records in registers. Access to health care is universal and largely provided by tax-funded hospitals. Interactions with the health care system are captured by the National Patient Register (NPR). ICD-coded discharge diagnoses for hospitalizations are recorded since 1964 (nationwide coverage since 1987) and for outpatient non-primary care visits since 2001. Missing visits in the inpatient component are negligible, but the coverage of the outpatient component was less complete with about 10% of the visits not registered [22]. The NPR is expected to have good coverage for sarcoidosis as most patients with the disease receive care in publicly funded outpatient clinics. The Multi-Generation Register (MGR) links registered residents from 1961 and onwards to their biological or adoptive parents [23]. It is an excellent resource to identify relatives of individuals born in Sweden.

Proband cases and controls

Proband cases—individuals with sarcoidosis whose relatives were ascertained—were identified from the NPR using ICD codes (ICD-8/9 135, ICD-10 D86 including all subcategories). They were included if they were hospitalized at least once or had at least two outpatient visits for sarcoidosis from (1964–2013). The date of hospitalization or second outpatient visit was the index date. For a subset of cases (n=983) who were diagnosed by pulmonologists at Karolinska University Hospital, we could obtain information on sarcoidosis phenotype (Löfgren's vs. non-Löfgren's disease).

Controls without sarcoidosis were sampled from the general population (Total Population Register) and were individually matched 10:1 to cases on birth year, sex, county of residence, and date of inclusion. To reduce the likelihood of sarcoidosis misclassification, we excluded individuals younger than 18 years at inclusion and those diagnosed with a haematological or lung malignancy within 6

months of the index date in the Swedish Cancer Register (ICD-7 162, 163, 200–205). As coverage of the MGR is incomplete for foreign-born individuals, we excluded cases and controls born outside the country to ensure consistent identifiability of their relatives and increase the ethnic homogeneity of our study population.

Identification of relatives and ascertainment of sarcoidosis

First and second degree relatives of proband cases and controls were identified from the MGR. In our study, parents, full siblings, and offspring of probands composed first degree kinships (50% genetic similarity), and half siblings composed second degree kinships (25% genetic similarity). Full siblings were defined as those sharing the same biological parents, whereas half siblings shared only one biological parent. As disease aggregation is a function of family size, the similarity of familial structures of proband cases and controls was confirmed to ensure fair comparisons. The same definition of sarcoidosis used to identify proband cases was used to ascertain sarcoidosis in relatives of cases and controls. Ethical approval for the study was obtained from the Regional Ethics Review Board in Stockholm (DNR 2014/230-31).

Statistical analysis

To measure the degree of familial aggregation of sarcoidosis, we used conditional logistic regression models estimating odds ratios, which we interpreted as familial relative risks (RRs). Unless otherwise stated, the familial RR in our study indicates the relative risk of sarcoidosis associated with having relatives with the disease irrespective of the timing of diagnosis of proband and relative. Two different modelling strategies were employed. In the first, each proband-relative relationship contributed a separate observation in our dataset and familial RRs were reported for different kinships. To account for the lack of statistical independence due to family clustering, robust sandwich variance estimators were used to construct 95% confidence intervals (CIs). In the second modelling approach, we compared proband cases and controls who had at least one (or at least two)

first degree relatives diagnosed with sarcoidosis at any time point during the ascertainment period (1964–2013).

We have previously shown that the average age at disease diagnosis in Sweden is 50 years, but there is great variation by sex, with incidence peaking 10 years later in females compared to males [24]. To examine if such variations could be attributed to familial aggregation (and potentially genetics), we stratified our models by age at diagnosis and sex of the proband and relative. The model was also stratified by disease phenotype (Löfgren's vs. non-Löfgren's disease) for cases captured in the Karolinska clinical cohort.

We estimated the heritability of sarcoidosis using a threshold-liability model [25] incorporating a time aspect [26]. To estimate the proportion of variance attributable to additive genetic and shared environmental factors, we used probit variance component analysis. Methods of weight and prevalence estimation are extensively described elsewhere [26-29]. We included full and half siblings born within 10 years of the proband case or control (i.e. assumed to be reared together), and we right-censored individuals at age 80 years to account for the lifetime risk for sarcoidosis.

We performed several sensitivity analyses described in detail in the **Online Supplement**. Data management and analyses were performed using SAS software version 9.4 (SAS Institute Inc., Cary, NC, USA) and R version 3.4.2 (R Foundation for Statistical Computing, Vienna, Austria).

RESULTS

Our study included 23880 proband cases with prevalent sarcoidosis and 171891 non-sarcoidosis controls from the general population with at least one first or second degree relative with sarcoidosis. The mean age of probands at inclusion was 50 years (SD 16.0) and 47% were female. The family structures of cases and controls were very similar with an average of four first degree relatives identified per case or control (**Table 2**). The age and sex distributions of relatives of proband cases and controls were also similar (**Table E1**).

Familial relative risk for sarcoidosis

Table 3 provides a summary of the relative risks for sarcoidosis associated with having relatives of different degree diagnosed with sarcoidosis. Approximately 4% of proband cases had at least one first degree relative diagnosed with the disease, whereas only 1% of proband controls had at least one first degree relative diagnosed with sarcoidosis. The familial RR associated with having at least one first degree relative with the disease was 3.73 (95% CI 3.43–4.06). The familial RR for sarcoidosis increased to 4.69 (95% CI 2.93–7.51) when at least two first degree relatives were diagnosed with sarcoidosis in a family and was attenuated to 1.50 (95% CI 0.98–2.30) when a second degree relative (a half sibling) was diagnosed with the disease. No great differences in the familial RR amongst first degree kinships were observed (**Table 3**).

Stratifications of the familial RR associated with having at least one first degree relative with the disease by age of the proband, sex of proband and relative and Löfgren's syndrome are shown in **Figure 1**. The RR for sarcoidosis was higher for probands aged 18–49 years at the time of inclusion to the study compared to those 50 years or older (RR 3.99 [95% CI 3.55–4.48] vs. 3.48 [95% CI 3.08–3.92], respectively). The RR was higher for male proband to female relative relationships (RR 4.10 [95 CI 3.49–4.82]) compared to other sex combinations for which the RR was centred on 3.5. In addition, there was an indication of increased familial RR for individuals in the clinical cohort whose disease

manifested as Löfgren's compared to non-Löfgren's disease manifestations (RR 4.14 [95% CI 2.21–7.75] vs. 3.32 [1.98–5.56], respectively). Stratified analyses by age and sex amongst different kinships are provided in **Table E2** and **Table E3**, respectively.

Restricting our analyses to probands with a history of disease at least a year before inclusion yielded similar results (**Table E4**). In addition, the familial RR from the main analyses proved robust in a probabilistic bias analysis where our register-based definition of sarcoidosis was subjected to extreme misclassification and when we used a stricter definition for sarcoidosis, requiring at least two visits in the NPR listing sarcoidosis (**Table E5**).

Heritability of sarcoidosis

The heritability of the disease was estimated to be 39% (95% CI 12–65) from an additive genetic and non-shared environmental variance model (**Figure 2**). No modification of the heritability was observed by sex of the proband. No variance attributable to shared environment between siblings could be identified in the data. Estimations of the ceiling heritability using Falconer's formulae provided similar results (**Table E4**).

DISCUSSION

To our knowledge, this case-control-family study represents the largest investigation to date aiming to estimate the familial aggregation of sarcoidosis. Using population-based Swedish registers with coverage spanning three generations, we show that familial exposure is likely the strongest risk factor for sarcoidosis described to date; associated with an almost fourfold increased risk for developing the disease. The variation in familial RRs by number of affected relatives, type of kinship, and age at sarcoidosis diagnosis, further provides evidence that genetics influence, at least partly, the development of sarcoidosis. Using biometric model analysis, we show that the heritability is 39%.

In previous studies, reported familial relative risks for first degree relatives range from 2.5 to 73.0 [15-17]. The absence of a control group and the small sample size of most previous investigations [16, 17] make comparability to our study difficult. Our findings are however comparable to those reported by the ACCESS study comprising of 701 White and Black American cases and controls [15]. The overall familial relative risk for first degree relatives is very similar to ours (odds ratio 3.8), but they reported higher risks for siblings and second degree relatives (odds ratios 5.8 and 5.2, respectively). When restricting to White American probands in the ACCESS study, the odds ratio was 16.6 [15]. It is difficult to disentangle the reasons behind the observed dissimilarities. There are evident differences in the methodology used for disease ascertainment and, importantly, the ethnic composition of the ACCESS participants and our study population. Differences in the analytical approach and the limited power of some analyses in the ACCESS study could be alternative explanations for some of the discrepancies.

Contrary to the ACCESS study [15] but in line with another study [16], we observed a higher familial RR for those diagnosed before the age of 50 years compared to those older than 50 years and for male-female proband-relative relationships compared to other relationships. Although the observed

weak effect measure modification by age might be interpreted as an indication of genetics' involvement in disease aetiology, we cannot exclude the possibility that this resulted from a slight overrepresentation of male probands in the younger age group, for whom sarcoidosis prevalence is higher as we have previously shown [24]. Further stratification by sex of the proband, however, did not eliminate the effect of age (data not shown). On the other hand, the observed higher RR for male-female proband-relative relationships did not persist when we stratified by kinship, suggesting that the observed modification by sex was merely a feature of our data.

Despite small numbers, analyses using information on sarcoidosis phenotype from our clinical cohort at Karolinska University Hospital indicated stronger familial associations for disease manifesting as Löfgren's compared to non-Löfgren's disease. This is in line with previous investigations describing Löfgren-like disease being somewhat more prevalent in familial than non-familial disease [13, 30]. While this finding needs more exploration and replication by other studies, it may be suggestive of a stronger genetic association in Löfgren's compared to non-Löfgren's disease [7], providing also a reasonable, albeit partial, explanation for the differences between the two manifestations of sarcoidosis in terms of phenotype and prognosis [18-20].

Thirty-nine percent of the liability to sarcoidosis was attributed to additive genetic factors in our population, while the contribution of shared environmental factors and the effect of sex were negligible. The prominent role of environmental factors in the aetiology of sarcoidosis is not surprising. Factors extending from infectious agents [31] and mould [32, 33] to nanoparticles [34] have been implicated in sarcoidosis. The specific role of these factors and their potential to explain the environmental component of sarcoidosis susceptibility remains to be fully elucidated.
The heritability of 39% we found in our population contrasts with the notion that genetics might be more influential than environmental factors for the liability to sarcoidosis as suggested by earlier reports from the US [11] and Finland/Denmark [21]. Despite differences in definition of sarcoidosis

and the ethnic composition of our populations, we believe that low power is likely the primary reason for the inflated heritability estimates in previous studies. Also, under certain circumstances, Falconer's method and the twin-study design used in those studies are prone to overestimation of heritability [35]. Nevertheless, we should not disregard the fact that even amongst more similar populations and environments in Sweden, Denmark, and Finland, particularities of the microenvironment and local genetic variation (e.g. founder effects) may alter to some extent the balance between the genetic and environmental components of sarcoidosis susceptibility.

The heritability estimates reported in our study are also different from estimates originating from genome-wide association studies [6, 7]. A number of common sarcoidosis single nucleotide polymorphisms accounted only for 5% overall [6], and 4% and 2% of the phenotypic variability of Löfgren's and non-Löfgren's disease, respectively [7]. These low numbers highlight the fact that genome-wide association studies cannot yet capture all genetic variance in the population ("missing heritability"), likely due to small sample sizes or the inability to identify certain genetic effects [36, 37], such as rare gene variants that have been identified by other methods [38].

A limitation of our study originates from the change in the availability of data in the NPR to define sarcoidosis over time. Our capacity to ascertain sarcoidosis increased considerably since 2001 with the capture of outpatient visits by the NPR. Before then, only hospitalization data were available, hence cases and relatives identified in the earlier period likely represent a smaller group of cases with more severe disease. It is also possible that a few controls had asymptomatic disease when they were sampled or developed sarcoidosis later. This may have resulted in a slight underestimation of familial RRs and heritability. In addition, in the absence of histological or clinical confirmation of cases, using a register-based definition might have led to some misclassification. However, familial RRs remained robust when our sarcoidosis definition was subjected to extreme misclassification using probabilistic bias analysis methods and when a stricter definition was used.

Our heritability estimations were based on a set of assumptions that should be considered when interpreting the results. We assumed that the liability to sarcoidosis was normally distributed, the prevalence of sarcoidosis was correctly specified, that there was no assortative mating, genetic effects were additive (no epistasis or dominance amongst loci), and that there were no gene-environment interactions. While such assumptions may not be perceived extreme, they are indeed difficult to test. Last, we cannot assume that our familial RR and heritability estimates are readily generalizable to all populations as they reflect the genetic composition and environmental exposures in the Swedish context.

Our study has several strengths. By using population-based registers, we estimated valid and precise familial aggregation and heritability estimates. Our study was also well powered to obtain informative stratified estimates of familial RRs. In addition, having access to a well-described clinical cohort allowed us to complement our findings with an estimate of the familial aggregation of Löfgren's and non-Löfgren's sarcoidosis. Future research efforts should focus on the familial aggregation of other disease phenotypes, disease complications, and disease prognosis.

In summary, familial exposure to sarcoidosis is a very strong risk factor for the disease. Having at least one first degree relative with sarcoidosis is associated with an almost fourfold increased risk for the disease. There was a trend towards increased familial RRs with higher number of affected relatives, if disease is diagnosed before the age of 50, and for Löfgren's syndrome. The heritability of the disease was 39%, suggesting a stronger implication of environmental factors in the development of sarcoidosis than what was previously perceived. Future research efforts should focus on identifying environmental agents implicated in the aetiology of sarcoidosis and unidentified genetic factors that could account for the missing heritability of sarcoidosis.

Acknowledgements

The authors would like to thank patients and health care personnel contributing data to the clinical cohort at Karolinska University Hospital.

Contributors

Conception and design of the study: MR, EVA; data analysis: MR; interpretation of data: all authors; drafting of first manuscript: MR; critical revision for important intellectual content: all authors; final approval of the version to be submitted: all authors; agreement to be accountable for all aspects of this study: all authors.

Conflict of interest

EVA reports funding from the Swedish Heart and Lung Foundation and the Swedish Society of Medicine. MR, JG, AE, SK, DDG, and JA have no conflicts of interest to declare.

Support statement

This study received funding from the Swedish Heart-Lung Foundation. The register linkage used in this study was also funded by the Swedish Society of Medicine. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. Research on sarcoidosis at Karolinska Institutet is supported by the Swedish Research Council, the Swedish Foundation for Strategic Research, the Strategic Research Area in Epidemiology (SfoEpi), a regional agreement on medical training and clinical research (ALF) between the Stockholm County Council and Karolinska Institutet, King Oscar's II Jubilee Foundation, the King Gustaf V's and Queen Victoria´s Freemasons' Foundation, and Karolinska Institutet's Research Foundation.

REFERENCES

1. 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: 736-755.
2. Schürmann M, Reichel P, Müller-Myhsok B, et al. Results from a genome-wide search for predisposing genes in sarcoidosis. *Am J Respir Crit Care Med* 2001: 164: 840-846.
3. Iannuzzi MC, Iyengar SK, Gray-McGuire C, et al. Genome-wide search for sarcoidosis susceptibility genes in African Americans. *Genes Immun* 2005: 6: 509-518.
4. Franke A, Fischer A, Nothnagel M, et al. Genome-wide association analysis in sarcoidosis and Crohn's disease unravels a common susceptibility locus on 10p12.2. *Gastroenterology* 2008: 135: 1207-1215.
5. Adrianto I, Lin CP, Hale JJ, et al. Genome-wide association study of African and European Americans implicates multiple shared and ethnic specific loci in sarcoidosis susceptibility. *PLoS One* 2012: 7: e43907.
6. Fischer A, Ellinghaus D, Nutsua M, et al. Identification of Immune-Relevant Factors Conferring Sarcoidosis Genetic Risk. *Am J Respir Crit Care Med* 2015: 192: 727-736.
7. Rivera NV, Ronninger M, Shchetynsky K, et al. High-Density Genetic Mapping Identifies New Susceptibility Variants in Sarcoidosis Phenotypes and Shows Genomic-driven Phenotypic Differences. *Am J Respir Crit Care Med* 2016: 193: 1008-1022.
8. Wolin A, Lahtela EL, Anttila V, et al. SNP Variants in Major Histocompatibility Complex Are Associated with Sarcoidosis Susceptibility-A Joint Analysis in Four European Populations. *Front Immunol* 2017: 8: 422.
9. Wiman LG. Familial occurrence of sarcoidosis. *Scand J Respir Dis Suppl* 1972: 80: 115-119.
10. Familial associations in sarcoidosis. A report to the Research Committee of the British Thoracic and Tuberculosis Association. *Tubercle* 1973: 54: 87-98.
11. Headings VE, Weston D, Young RC, Jr., et al. Familial sarcoidosis with multiple occurrences in eleven families: a possible mechanism of inheritance. *Ann N Y Acad Sci* 1976: 278: 377-385.
12. Brennan NJ, Crean P, Long JP, et al. High prevalence of familial sarcoidosis in an Irish population. *Thorax* 1984: 39: 14-18.
13. Fité E, Alsina JM, Antó JM, et al. Sarcoidosis: family contact study. *Respiration* 1998: 65: 34-39.
14. Pietinalho A, Ohmichi M, Hirasawa M, et al. Familial sarcoidosis in Finland and Hokkaido, Japan – a comparative study. *Respir Med* 1999: 93: 408-412.
15. Rybicki BA, Iannuzzi MC, Frederick MM, et al. Familial aggregation of sarcoidosis. A case-control etiologic study of sarcoidosis (ACCESS). *Am J Respir Crit Care Med* 2001: 164: 2085-2091.
16. Rybicki BA, Kirkey KL, Major M, et al. Familial risk ratio of sarcoidosis in African-American sibs and parents. *Am J Epidemiol* 2001: 153: 188-193.
17. McGrath DS, Daniil Z, Foley P, et al. Epidemiology of familial sarcoidosis in the UK. *Thorax* 2000: 55: 751-754.
18. Grunewald J, Eklund A. Löfgren's syndrome: human leukocyte antigen strongly influences the disease course. *Am J Respir Crit Care Med* 2009: 179: 307-312.
19. Grunewald J. HLA associations and Löfgren's syndrome. *Expert Rev Clin Immunol* 2012: 8: 55-62.

20. Sato H, Woodhead FA, Ahmad T, et al. Sarcoidosis HLA class II genotyping distinguishes differences of clinical phenotype across ethnic groups. *Hum Mol Genet* 2010: 19: 4100-4111.

21. Sverrild A, Backer V, Kyvik KO, et al. Heredity in sarcoidosis: a registry-based twin study. *Thorax* 2008: 63: 894-896.

22. Socialstyrelsen. Kvalitet och innehåll i patientregistret. Utskrivningar från slutenvården 1964–2007 och besök i specialiserad öppenvård (exklusive primärvårdsbesök) 1997–2007 (Quality and contents of the patient register) [In Swedish]. Stockholm: Socialstyrelsen (The National Board of Health and Welfare); 2009.

23. Statistics Sweden. Multi-generation register 2010. A description of contents and quality. Örebro: Department of Population and Welfare Statistics, Statistics Sweden; 2011.

24. Arkema EV, Grunewald J, Kullberg S, et al. Sarcoidosis incidence and prevalence: a nationwide register-based assessment in Sweden. *Eur Respir J* 2016: 48: 1690-1699.

25. Falconer DS, Mackay TF. Introduction to Quantitative Genetics. 4th ed. Pearson, Harlow, UK, 1996.

26. Cederkvist L, Holst KK, Andersen KK, et al. Incorporation of the time aspect into the liability-threshold model for case-control-family data. *Stat Med* 2017: 36: 1599-1618.

27. Scheike TH, Holst KK, Hjelmborg JB. Estimating heritability for cause specific mortality based on twin studies. *Lifetime Data Anal* 2014: 20: 210-233.

28. Holst KK, Scheike TH, Hjelmborg JB. The liability threshold model for censored twin data. *Comput Stat Data An* 2016: 93: 324-335.

29. Javaras KN, Laird NM, Hudson JI, et al. Estimating disease prevalence using relatives of case and control probands. *Biometrics* 2010: 66: 214-221.

30. Pacheco Y, Calender A, Israel-Biet D, et al. Familial vs. sporadic sarcoidosis: BTNL2 polymorphisms, clinical presentations, and outcomes in a French cohort. *Orphanet J Rare Dis* 2016: 11: 165.

31. Suzuki Y, Uchida K, Takemura T, et al. Propionibacterium acnes-derived insoluble immune complexes in sinus macrophages of lymph nodes affected by sarcoidosis. *PLoS One* 2018: 13: e0192408.

32. Newman LS, Rose CS, Bresnitz EA, et al. A case control etiologic study of sarcoidosis: environmental and occupational risk factors. *Am J Respir Crit Care Med* 2004: 170: 1324-1330.

33. Tercelj M, Salobir B, Harlander M, et al. Fungal exposure in homes of patients with sarcoidosis - an environmental exposure study. *Environ Health* 2011: 10: 8.

34. Huizar I, Malur A, Midgette YA, et al. Novel murine model of chronic granulomatous lung inflammation elicited by carbon nanotubes. *Am J Respir Cell Mol Biol* 2011: 45: 858-866.

35. Tenesa A, Haley CS. The heritability of human disease: estimation, uses and abuses. *Nat Rev Genet* 2013: 14: 139-149.

36. Manolio TA, Collins FS, Cox NJ, et al. Finding the missing heritability of complex diseases. *Nature* 2009: 461: 747-753.

37. Timpson NJ, Greenwood CMT, Soranzo N, et al. Genetic architecture: the shape of the genetic contribution to human traits and disease. *Nat Rev Genet* 2018: 19: 110-124.

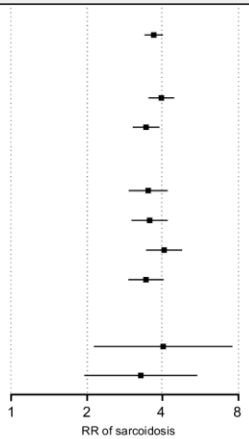
38. Calender A, Rollat Farnier PA, Buisson A, et al. Whole exome sequencing in three families segregating a pediatric case of sarcoidosis. *BMC Med Genomics* 2018: 11: 23.

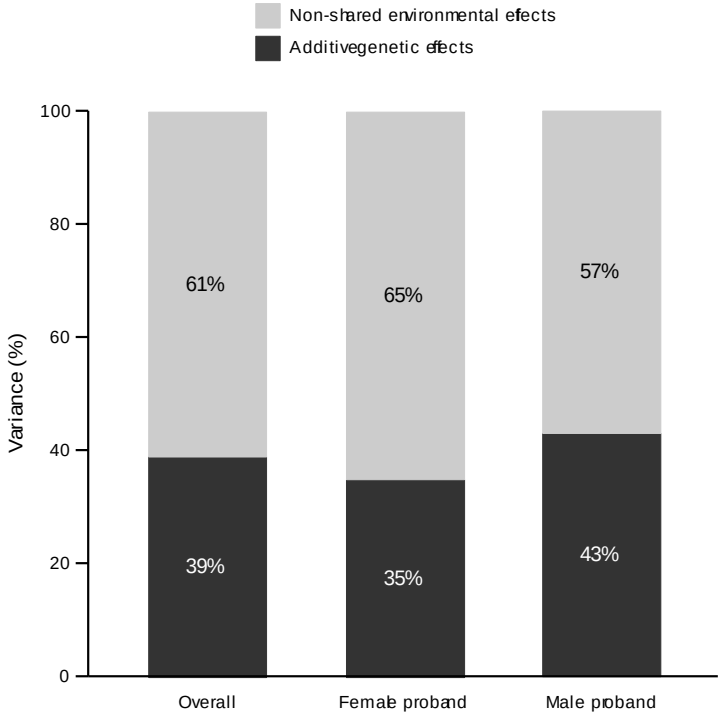
FIGURE LEGENDS

Figure 1. Relative risk of sarcoidosis associated with having at least one first degree relative with sarcoidosis, stratified by age of the proband at inclusion, sex of proband and relative, and Löfgren's syndrome.

Figure 2. Additive genetic (heritability) and non-shared environmental effects contributing to the susceptibility for sarcoidosis, overall and stratified by the sex of the proband.

≥1 First Degree Relative with Sarcoidosis	Cases Exposed/Total (%)	Controls Exposed/Total (%)	Familial RR (95% CI)
Overall	831/20 322 (4.1)	1907/164 628 (1.2)	3.73 (3.43–4.06)
Age of proband			
18–49 years	438/10 138 (4.3)	975/85 445 (1.1)	3.99 (3.55–4.48)
≥50 years	393/10 184 (3.9)	932/79 183 (1.2)	3.48 (3.08–3.92)
Sex of proband/relative			
Female/Female	186/8823 (2.1)	434/70 970 (0.6)	3.56 (2.98–4.24)
Female/Male	217/8744 (2.5)	505/70 601 (0.7)	3.58 (3.04–4.22)
Male/Female	228/10 081 (2.3)	470/83 218 (0.6)	4.10 (3.49–4.82)
Male/Male	220/9943 (2.2)	524/82 409 (0.6)	3.49 (2.97–4.09)
Löfgren's syndrome			
Yes	15/356 (4.2)	31/2812 (1.1)	4.14 (2.21–7.75)
No	21/627 (3.4)	49/4843 (1.0)	3.32 (1.98–5.56)




Non-shared environmental effects
Additive genetic effects
Variance (%)
100
80
60
40
20
0
61%
39%
65%
35%
57%
43%
Overall
Female proband
Male proband

TABLES

Table 1. Summary of previous studies on familial aggregation and heritability of sarcoidosis. Only familial aggregation studies with a control group were considered for the table.

First author, year of publication	Country	N probands		N relatives	Mean age at inclusion, years		Sarcoidosis ascertainment		Estimates of familial aggregation/heritability
		Cases	Controls		Cases	Relatives	Probands	Relatives	
Headings, 1976 [11]	USA	80	Prevalence estimate	523 parents and full siblings	34	Not reported	Biopsy proven	Reported by probands	7.2% of proband cases with ≥1 FDR with sarcoidosis Heritability: 60–70%
McGrath, 2000 [17]	UK	406	Prevalence estimate	Not reported	40	Not reported	Biopsy and non-biopsy proven	Reported by proband, confirmed by family doctor	5.9% of proband cases with ≥1 family member with sarcoidosis RR in siblings: 36–73
Rybicki, 2001 [15]	USA	706	706	10 862 FDRs 17 047 SDRs	42 (median)	Not reported	Biopsy proven	Reported by proband	OR in FDRs: 3.8 (1.9–7.6) OR in SDRs: 5.2 (1.5–18.2)
Rybicki, 2001 [16]	USA	179	Prevalence estimate and simulation	488 parents and full siblings	43	66 in parents, 45 in siblings	Biopsy and non-biopsy proven	Reported by relative	RR in FDRs: 2.5 (1.6–3.7)
Sverrild, 2008 [21]	Denmark & Finland	210	NA	210 twin pairs (50 MZ, 160 DZ)	Not reported	NA	Register-based	NA	Concordance rates: 0.148 in MZ twins, 0.012 in DZ twins Heritability: 66% (45–80%)

FDR = first degree relative; SDR = second degree relative; NA = not applicable; MZ = monozygotic; DZ = dizygotic.
* In 114 cases with biopsy-proven sarcoidosis diagnosis.

Table 2. Family structure of proband cases and controls with at least one first or second degree relative identified from the Swedish Multi-Generation Register.

	Cases*		Controls*	
	N relatives	**Mean N relatives per case (SD)**	**N relatives**	**Mean N relatives per control (SD)**
First degree relatives				
All relatives	86 136	4.2 (2.2)	713 708	4.3 (2.2)
Parents	28 835	1.4 (0.9)	245 054	1.5 (0.8)
Full siblings	24 104	1.2 (1.4)	203 989	1.2 (1.4)
Offspring	33 197	1.6 (1.4)	264 665	1.6 (1.4)
Second degree relatives				
Half siblings	7511	2.1 (1.6)	15 202	2.1 (1.5)

* No. probands with ≥1 first degree relative: 20 322 cases, 164 628 controls. No. probands with ≥1 second degree relative: 3558 cases, 7263 controls.

Table 3. Relative risk of sarcoidosis associated with having relatives with the disease.

	Cases		Controls		RR (95% CI)
	N exposed/ N total	% exposed	N exposed/ N total	% exposed	
First degree relatives					
≥1 relative*	831/20 332	4.1	1907/164 628	1.2	3.73 (3.43–4.06)
≥2 relatives*	28/20 332	0.1	49/164 628	<0.1	4.69 (2.93–7.51)
Parents†	271/28 835	0.9	638/245 054	0.3	3.68 (3.23–4.19)
Full siblings†	330/24 104	1.4	696/203 989	0.3	4.08 (3.45–4.83)
Offspring†	262/33 197	0.8	623/264 665	0.2	3.23 (2.78–3.76)
Second degree relatives					
Half siblings†	44/7511	0.6	49/15 202	0.3	1.50 (0.98–2.30)

* For these analyses, an indicator for having ≥1 or ≥2 first degree relatives with sarcoidosis was created for each case and control.

† For these analyses, each proband-relative relationship contributed a unique observation in the dataset. The confidence intervals were adjusted for autocorrelation arising from family clustering using robust estimates of the variance.

ONLINE SUPPLEMENT

Supplementary Methods

Sensitivity analysis

We conducted several sensitivity analyses. First, we were concerned that familial aggregation estimates might have been influenced by a more thorough ascertainment of the proband case's relatives due to the history of sarcoidosis in the proband. To address this, we repeated our familial RR analysis and regarded probands to be exposed only if relatives received their diagnosis of sarcoidosis at least a year before the proband. Second, we examined whether the familial RR was biased owing to misclassification of the sarcoidosis definition. As outlined below, we used probabilistic bias analysis methods [1, 2] to re-estimate the familial RRs under predetermined bias assumptions. We also tested a stricter definition for sarcoidosis requiring at least two visits listing sarcoidosis in the National Patient Register.

Last, to examine the robustness of heritability estimates, we calculated the ceiling heritability of sarcoidosis using Falconer's method [3]. Heritability (including potential influence of common environmental effects) equals twice the tetrachoric correlations between probands and first degree relatives [3]. For the calculations, we used a liberal and a strict sarcoidosis prevalence estimate for the Swedish population [4].

Probabilistic bias analysis for sarcoidosis definition misclassification

We followed the methods described by Lash et al. [2] and Bollaerts et al. [1] to test the robustness of the familial relative risk from the main analysis against potential misclassification of our register-based definition used for the ascertainment of probands and relatives. We defined exposure as having ≥1 first degree relative with sarcoidosis and used the numbers of exposed and unexposed cases and controls from the main analyses to calculate the crude odds ratio (interpreted as the familial relative risk):

	Cases	**Controls**
≥1 first degree relative with sarcoidosis	831	1907
No first degree relatives with sarcoidosis	19491	162721

We then assigned probability distributions for bias parameters as follows:

$$\text{Positive Predictive Value (PPV)} \sim \text{Beta}(15,4)$$

$$\text{Negative Predictive Value (NPV)} \sim \text{Uniform}(1)$$

The PPV values centred at 80% and NPV was uniformly defined to be 100% considering the rarity of sarcoidosis in the general population.

We performed Monte Carlo simulations with 10 million repetitions sampling from the above distributions to define the bias parameters. We sequentially misclassified the outcome definition (sarcoidosis in the proband) and the exposure (sarcoidosis in the relative). The final estimate for the familial relative risk represents the 50th percentile of the re-estimated familial relative risks. We used the residual error from each simulation to calculate 95% simulation confidence intervals (2.5th and 97.5th percentiles). Of note, the resulting simulation confidence intervals did not account for the residual random error from the original analyses.

The crude odds ratio was calculated using the information above to be 3.64**.** Adjusting for sarcoidosis definition misclassification in simulations resulted in a largely similar odds ratio (3.61 [95% simulation confidence interval 3.33–3.93]). The familial relative risk from main analysis estimated from a logistic regression model was 3.73 (95% confidence interval 3.43–4.06).

Supplementary References

1. Bollaerts K, Shinde V, Dos Santos G, et al. Application of Probabilistic Multiple-Bias Analyses to a Cohort- and a Case-Control Study on the Association between Pandemrix and Narcolepsy. *PLoS One* 2016: 11: e0149289.
2. Lash TL, Fox MP, Fink AK. Applying Quantitative Bias Analysis to Epidemiologic Data. Springer-Verlag, New York, 2009.
3. Falconer DS, Mackay TF. Introduction to Quantitative Genetics. 4th ed. Pearson, Harlow, UK, 1996.
4. Arkema EV, Grunewald J, Kullberg S, et al. Sarcoidosis incidence and prevalence: a nationwide register-based assessment in Sweden. *Eur Respir J* 2016: 48: 1690-1699.

Supplementary Tables

Table E1. Distribution of age at inclusion and sex amongst first and second degree relatives of proband cases and controls.

	Mean age at inclusion (SD)		Female, %	
	Relatives of cases	**Relatives of controls**	**Relatives of cases**	**Relatives of controls**
First degree relatives				
All	48.0 (25.2)	48.0 (25.5)	50	50
Parents	74.3 (15.7)	74.2 (15.5)	51	51
Full siblings	45.3 (14.2)	45.0 (14.3)	49	49
Offspring	27.1 (15.9)	26.1 (15.7)	49	49
Second degree relatives				
Half siblings	40.5 (15.8)	39.5 (15.6)	48	49

Table E2. Relative risk of sarcoidosis associated with having one or more first degree relatives with sarcoidosis, stratified by probands' age at inclusion and sex.

	Age 18–49 years at inclusion			Age ≥50 years at inclusion		
	N exposed/N total (%)			N exposed/N total (%)		
	Cases	Controls	RR (95% CI)	Cases	Controls	RR (95% CI)
≥1 first degree relative	438/10 138 (4.3)	975/85 445 (1.1)	3.99 (3.55–4.48)	393/10 184 (3.9)	932/79 183 (1.2)	3.48 (3.08–3.92)
Female proband	177/3765 (4.7)	392/31 667 (1.2)	4.02 (3.34–4.82)	221/5959 (3.7)	535/45 959 (1.2)	3.41 (2.90–4.00)
Male proband	261/6373 (4.1)	583/53 778 (1.1)	3.97 (3.42–4.61)	172/4225 (4.1)	397/33 224 (1.2)	3.57 (2.98–4.29)
≥2 first degree relatives	16/10 138 (0.2)	22/85 445 (<0.1)	6.01 (3.13–11.55)	12/10 184 (0.2)	27/79 183 (<0.1)	3.62 (1.81–7.22)

Table E3. Relative risk of sarcoidosis associated with having a first degree relative with the disease, stratified by kinship and sex of proband and relative.

Sex of proband/ sex of relative	N exposed/N total (%)		RR (95% CI)
	Cases	Controls	
Parents			
Female/Female	63/6186 (1.0)	163/52 536 (0.3)	3.30 (2.56–4.27)
Female/Male	51/5853 (0.9)	112/49 708 (0.2)	3.91 (2.91–5.25)
Male/Female	98/8577 (1.1)	227/72 996 (0.3)	3.74 (3.04–4.61)
Male/Male	59/8219 (0.7)	136/69 814 (0.2)	3.61 (2.74–4.75)
Full siblings			
Female/Female	62/4903 (1.3)	131/42 221 (0.3)	4.79 (3.15–7.28)
Female/Male	80/5180 (1.5)	172/43 585 (0.4)	3.97 (2.75–5.74)
Male/Female	82/6900 (1.2)	156/57 513 (0.3)	4.55 (2.13–6.63)
Male/Male	106/7121 (1.5)	237/60 670 (0.4)	3.35 (2.54–4.42)
Offspring			
Female/Female	61/8908 (0.7)	149/69 143 (0.2)	2.76 (2.00–3.81)
Female/Male	93/9085 (1.0)	225/72 538 (0.3)	3.16 (2.41–4.15)
Male/Female	50/7520 (0.7)	92/59 955 (0.2)	3.77 (2.52–5.64)
Male/Male	58/7684 (0.8)	157/63 029 (0.3)	3.02 (2.21–4.14)

Table E4. Relative risk of sarcoidosis associated with having relatives diagnosed with sarcoidosis at least a year before the proband case.

	Cases		Controls		
	N exposed/ N total	% exposed	N exposed/ N total	% exposed	RR (95% CI)
First degree relatives					
≥1 relative*	454/20 094	2.3	997/161 667	0.6	3.89 (3.47–4.36)
≥2 relatives*	12/20 094	0.1	12/161 667	<0.1	8.09 (3.61–18.13)
Parents†	212/28 835	0.7	454/243 939	0.2	3.96 (3.41–4.60)
Full siblings†	180/23 699	0.8	362/199 486	0.2	4.39 (3.46–5.58)
Offspring†	77/23 761	0.3	193/181 114	0.1	3.20 (2.42–4.23)
Second degree relatives					
Half siblings†	23/6935	0.3	21/13 297	0.2	1.89 (0.88–4.06)

* For these analyses, an indicator for having ≥1 or ≥2 first degree relatives with sarcoidosis was created for each case and control.

† For these analyses, each proband-relative relationship contributed a unique observation in the dataset. The confidence intervals were adjusted for autocorrelation arising from family clustering using robust estimates of the variance.

Table E5. Relative risk of sarcoidosis associated with having relatives with the disease using a stricter definition for sarcoidosis ascertainment (≥2 ICD-coded visits in the National Patient Register, 1964–2013).

	Cases		Controls		
	N exposed/ N total	% exposed	N exposed/ N total	% exposed	RR (95% CI)
First degree relatives					
≥1 relative*	482/14 936	3.2	1045/120 068	0.9	3.86 (3.46–4.32)
≥2 relatives*	10/14 936	0.1	19/120 068	<0.1	4.22 (1.94–9.22)
Parents†	147/22 647	0.7	312/188 961	0.2	3.99 (3.33–4.78)
Full siblings†	202/18 690	1.1	397/156 057	0.3	4.26 (3.42–5.30)
Offspring†	143/22 696	0.6	355/182 051	0.2	3.08 (2.52–3.76)
Second degree relatives					
Half siblings†	28/6017	0.5	33/12 021	0.3	1.30 (0.79–2.15)

* For these analyses, an indicator for having ≥1 or ≥2 first degree relatives with sarcoidosis was created for each case and control.

† For these analyses, each proband-relative relationship contributed a unique observation in the dataset. The confidence intervals were adjusted for autocorrelation arising from family clustering using robust estimates of the variance.

Table E6. Heritability estimates for sarcoidosis using Falconer's methods based on a liability-threshold model for the disease. Liberal and strict estimates of sarcoidosis prevalence in our population obtained from Arkema et al., 2016 were used to account for the case-control study design.

	Cases with relatives with sarcoidosis	Cases with relatives without sarcoidosis	Familial relative risk*	Liberal prevalence assumption		Strict prevalence assumption	
				Sarcoidosis prevalence per 100,000	Heritability (95% CI)	Sarcoidosis prevalence per 100,000	Heritability (95% CI)
≥1 first degree relative	831	19 491	3.73	160	35% (33–37)	64	32% (30–34)
Female/Female†	186	8637	3.56	141	31% (27–35)	55	28% (24–32)
Female/Male†	217	8527	3.58	141	31% (27–35)	55	29% (26–32)
Male/Female†	228	9853	4.10	179	35% (31–39)	73	32% (29–35)
Male/Male†	220	9723	3.49	179	31% (27–35)	73	29% (26–32)
Parents	271	28 564	3.68	160	29% (26–32)	64	27% (24–30)
Full siblings	330	23 774	4.08	160	33% (30–36)	64	30% (27–32)
Offspring	262	32 935	3.23	160	26% (23–29)	64	24% (21–27)

* Estimates obtained from the main analysis in this study.

† Sex of proband/sex of relative.