SUPPLEMENT

Risk of first and recurrent serious infection in sarcoidosis: a Swedish register-based cohort study

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

Contents

SUPPLEMENTAL METHODS

Statistical analysis

Estimation of inverse probability of sarcoidosis weights

We estimated the probability of sarcoidosis (exposure) using a logistic regression model with the following covariates: age (continuous), sex, residential location, birth country, attained education, gross annual salary, civil status, calendar period, number of visits in the National Patient Register within two years before first sarcoidosis visit or corresponding date for comparators, and history of comorbidities including congestive heart disease, atrial fibrillation, hypertension, diabetes mellitus, dyslipidaemia, chronic obstructive pulmonary disease, asthma, acute myocardial infarction, stroke, autoimmune disease (in all analysis except for the stratification by history of autoimmune disease in the Discussion), primary immunodeficiency, and history of autoimmune disease or sarcoidosis in at least one first degree relative. We did not include history of serious infection in the year before inclusion in the model to avoid reverse causation bias as individuals with sarcoidosis were more likely to be in contact with healthcare at the time compared to their general population comparators. For more information, see [1, 2]. Variable definitions and encoding are available in **Table S1**. There were very few missing values (<1% in sarcoidosis and the general population) in a small number of covariates. Missing data were coded as a separate category.

Assuming consistency, positivity, and conditional exchangeability, stabilized inverse probability of exposure (sarcoidosis) weights for everyone in the dataset were then calculated using the formula $W_{st} = f(A)/f(A|L)$, where $f(A)$ denotes the proportion of exposed (with sarcoidosis) in our study population, and $f(A|L)$ the propensity score, i.e. the probability of exposure (sarcoidosis, yes/no) given confounders. To improve precision, we truncated (trimmed) the stabilized weights at the 1% of the extremes of their distribution [3]. That is, weights lower than the 1st percentile were set to the 1st-percentile weight and weights larger than the 99th percentile were set to the 99th percentile weight. We tested weight truncation in the range of 1% to 5% using our main analysis Cox model (with time to any serious infection as the outcome). We chose truncation at 1% after considering the variance-bias trade-off comparing deviation of the point estimate of the model with full and progressively truncated weights to that of the unadjusted (biased/unweighted model) [3].

We calculated standardized differences before and after applying the estimated stabilized and truncated inverse probability of sarcoidosis weights to check the balance of covariate distributions between exposed and unexposed (sarcoidosis vs. general population comparators) [4]. These are presented in **Table S2**. Standardized differences greater than 0.1 to 0.2 indicated covariate imbalance.

Cox proportional hazard models for first serious infection

We ran Cox proportional hazard models with time since diagnosis or matching (inclusion) as the time scale to estimate crude and adjusted hazard ratios and 95% confidence intervals comparing sarcoidosis (overall or either stratified by sarcoidosis treatment status around the time of diagnosis) to the general population comparators. To report marginal adjusted estimates, we incorporated

stabilized inverse-probability of sarcoidosis weights estimated as mentioned above. To account for the fact that some individuals may have been upweighted in adjusted (weighted) models even though weights were stabilized, we used robust (sandwich) standard errors were used to estimate 95% confidence intervals with better error coverage. The proportionality of hazards assumption was tested by inspecting Schoenfeld residuals plots for sarcoidosis (the only covariate in the models). As we *a priori* hypothesized, the proportionality assumption was violated, which lead to analyses performed with flexible parametric survival models described below.

Flexible parametric survival models for first serious infection

We used package *rstpm2* (version 1.5.1) [5] implemented in R (R Core Team, R Foundation for Statistical Computing, Vienna, Austria) to run flexible parametric survival models. Years since diagnosis or matching (for comparators) was used as the time scale in all models and the stabilized inverse-probability of sarcoidosis weights estimated as outlined above were integrated to the model. To identify the best model fit, we varied the degrees of freedom used in the natural cubic spline function to model the underlying (baseline) hazard function (on the log-cumulative hazard scale) and the time-varying effect of exposure (sarcoidosis or sarcoidosis, treated/not treated) from two to seven. A combination of model-derived hazard plots and the Akaike information criterion (AIC) was used in that order to choose the model with the best fit for the final analysis. Based on those, we used 4 degrees of freedom for modelling the baseline log-cumulative hazard and two degrees of freedom for modelling time-varying coefficients in all models.

Shared frailty models for recurrent serious infection

To model recurrent (up to six) serious infections and estimate adjusted within-individual hazard ratios and corresponding 95% confidence intervals comparing sarcoidosis to the general population, we ran a (random effects) shared frailty model using the package *survival* (version 3.1-8) in R (R Core Team, R Foundation for Statistical Computing, Vienna, Austria). In this model and in addition to stabilized inverse-probability of sarcoidosis weights estimated as explained above, a gamma frailty term was introduced to account for the dependence between recurrent events in individuals (correlation amongst survival times in individuals). We also ran a flexible parametric survival model using *rstpm2* as mentioned above but this time adding a gamma frailty term.

References

1. Rossides M, Kullberg S, Askling J, *et al.* Are infectious diseases risk factors for sarcoidosis or a result of reverse causation? Findings from a population-based nested case–control study. *Eur J Epidemiol* 2020.

2. Rossides M, Kullberg S, Eklund A, *et al.* Sarcoidosis diagnosis and treatment in Sweden: A register-based assessment of variations by region and calendar period. *Respir Med* 2020; 161: 105846.

3. Cole SR, Hernan MA. Constructing inverse probability weights for marginal structural models. *Am J Epidemiol* 2008; 168: 656–664.

4. Austin PC, Stuart EA. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects

in observational studies. *Stat Med* 2015; 34: 3661–3679.

5. Clements M, Liu X-R. rstpm2: Generalized Survival Models. R package version 1.5.1. (https://CRAN.R-project.org/package=rstpm2). Stockholm; 2017.

SUPPLEMENTAL TABLES

Table S1. Description of variables used in the analyses including International Classification of Disease (ICD) and Anatomical Therapeutic Chemical (ATC) codes used to define diseases in the National Patient, Cancer, and Prescribed Drug registers, respectively. Dispensation data in the Prescribed Drug Register were available starting July 1, 2005.

Disease/Variable	Definition/Data source	Encoding	ICD and/or ATC codes*
Sarcoidosis	≥2 inpatient or outpatient visits in the NPR listing an ICD code for sarcoidosis.	Yes/no	*ICD-10:* D86 *ICD-8/9:* 135
Sarcoidosis treated around the time of diagnosis	≥1 dispensation of systemic corticosteroids, methotrexate, or azathioprine ±3 months from the first sarcoidosis visit.	Treated/untreated	*ATC:* H02AB01; H02AB02; H02AB04; H02AB06; H02AB07; L01BA01; L04AX03; L04AX01
Haematopoietic or lung malignancy around the time of sarcoidosis diagnosis	≥1 registration in the Swedish Cancer Register ±6 months from the first sarcoidosis visit or corresponding date for comparators.	Yes/no	*ICD-7:* 162–163; 200–205
Serious infection (primary outcome)	≥1 hospitalisation in the NPR listing an ICD code as the primary discharge diagnosis.	Yes/no	*ICD-10:* A00–B99; D73.3; E06.0; E32.1; G00–G07; H00.0; H44.0; H60.0; H60.1; H60.2; H60.3; H66; H67; H70; I30.1; I40.0; J00–J22; J32; J34.0; J36; J38.3; J39.0; J39.1; J44.0; J85; J86; K04.4; K04.6; K04.7; K10.2; K11.3; K12.2; K14.0; K57.0; K57.2; K57.4; K57.8; K61; K63.0; K65.0; K65.9; L00–L08; L30.3; M00; M01; M46.2; M46.3; M46.4; M60.0; M64.5; M65.0; M71.0; M71.1; M72.6; M86; N10–N12; N13.6; N15.1; N15.9; N30.0; N30.8; N34.0; N39.0; N41.2; N43.1; N45; N48.2; N61; N70–N74; N75.1; O23; O26.4; O41.1; O75.3; O85; O86; O88.3; O91; O98
Serious infection excluding urinary tract infections (secondary outcome)	≥1 hospitalisation in the NPR listing an ICD code as the primary discharge diagnosis, excluding ICD codes for urinary tract infection.	Yes/no	*ICD-10:* N30.0; N39.0 (used for exclusion)
Serious infection excluding pneumonia (secondary outcome)	≥1 hospitalisation in the NPR listing an ICD code as the primary discharge diagnosis, excluding ICD codes for pneumonia.	Yes/no	*ICD-10:* J12–J18 (used for exclusion)

Table S1. (Continued).

Disease/Variable	Definition/Data source	Encoding	ICD and/or ATC codes*
Congestive heart disease	≥1 visit listing an ICD code in the NPR before first sarcoidosis visit or corresponding date for comparators.	Yes/no	*ICD-10:* I42; I50 *ICD-9:* 425; 428 *ICD-8:* 425; 427,0; 427,1; 428,9
Atrial fibrillation	≥1 visit listing an ICD code in the NPR before first sarcoidosis visit or corresponding date for comparators.	Yes/no	*ICD-10:* I48 *ICD-9:* 427D *ICD-8:* 427,92
Acute myocardial infarction	≥1 visit listing an ICD code in the NPR's inpatient component before first sarcoidosis visit or corresponding date for comparators.	Yes/no	*ICD-10:* I21 *ICD-8/9:* 410
Stroke	≥1 visit listing an ICD code in the NPR's inpatient component before first sarcoidosis visit or corresponding date for comparators.	Yes/no	*ICD-10:* I60; I61; I63; I64 *ICD-9:* 430; 431; 433; 434; 436 *ICD-8:* 430–434; 436
Hypertension	≥1 visit listing an ICD code in the NPR or ≥2 dispensations in the PDR before first sarcoidosis visit or corresponding date for comparators.	Yes/no	*ICD-10:* I10–I15 *ICD-9:* 400–405 *ICD-8:* 400–404 *ATC:* C02CA; C07; C08; C09
Diabetes mellitus	≥1 visit listing an ICD code in the NPR or ≥2 dispensations in the PDR before first sarcoidosis visit or corresponding date for comparators.	Yes/no	*ICD-10:* E10–E11 *ICD-8/9:* 250 *ATC:* A10
Dyslipidaemia	≥1 visit listing an ICD code in the NPR or ≥2 dispensations in the PDR before first sarcoidosis visit or corresponding date for comparators.	Yes/no	*ICD-10:* E78 *ICD-8/9:* 272 *ATC:* C10
Chronic obstructive pulmonary disease	≥1 visit listing an ICD code in the NPR before first sarcoidosis visit or corresponding date for comparators.	Yes/no	*ICD-10:* J41; J42; J43; J44.1; J44.8; J44.9 *ICD-9:* 491A; 491B; 491X; 492; 496 *ICD-8:* 491; 492
Asthma	≥1 visit listing an ICD code in the NPR before first sarcoidosis visit or corresponding date for comparators.	Yes/no	*ICD-10:* J45–J46 *ICD 8/9:* 493

Table S1. (Continued).

Disease/Variable	Definition/Data source	Encoding	ICD and/or ATC codes*
Autoimmune disease	≥2 visits listing an ICD code in the NPR before first sarcoidosis visit or corresponding date for comparators.	Yes/no	*ICD-10:* D51.0; D59.1; D68.6; D69.3; E05.0; E06.3; E10; E27.1; G35; G61.0; G70.0; K50; K51; K74.3; K90.0; L10; L12; L40; L63; M05–M09; M31.3; M31.5; M31.6; M32.1; M32.8; M32.9; M33; M34; M35.0; M35.1; M35.2; M45; [D86 in relatives] *ICD-9:* 136B; 242A; 245C; 255E; 281A; 283A; 287D; 340; 357A; 358A; 446E; 446F; 555; 556; 571G; 579A; 694E; 694F; 696; 704A; 710; 714; 720; [135 in relatives] *ICD-8:* 136,07; 242,00; 245,30; 255,10; 269,10; 281,0; 283,90; 287,10; 340; 357; 446,20; 446,30; 446,38; 563; 694; 696; 704,00; 712; 716; 733,00;734,0; 734,1; 734,9; [135 in relatives]
Autoimmune disease or sarcoidosis in ≥1 first degree relative	≥1 visit listing an ICD code in the NPR before first sarcoidosis visit or corresponding date for comparators in ≥1 first degree relative (biological mother, father, full sibling, or child) identified through the Multi-Generation Register.		
Primary immunodeficiency	≥1 visit listing an ICD code in the NPR before first sarcoidosis visit or corresponding date for comparators.	Yes/no	*ICD-10:* D80–D84 *ICD-9:* 279J; 279L; 279M; 279X
Systemic corticosteroids	≥1 dispensation in the PDR within six months first sarcoidosis visit or corresponding date for comparators.	Yes/no	*ATC:* H02AB
Other immunosuppressants	≥1 dispensation in the PDR within six months before first sarcoidosis visit or corresponding date for comparators.	Yes/no	*ATC:* L01BA01; L04AX03; L04AX01; L04AA13
Hydroxychloroquine	≥1 dispensation in the PDR within six months before first sarcoidosis visit or corresponding date for comparators.	Yes/no	*ATC:* P01BA02
Non-steroidal anti-inflammatory drugs	≥1 dispensation in the PDR within six months before first sarcoidosis visit or corresponding date for comparators.	Yes/no	*ATC:* M01A

Table S1. (Continued).

Disease/Variable	Definition/Data source	Encoding	ICD and/or ATC codes*
Inhaled corticosteroids	≥1 dispensation in the PDR within six months before first sarcoidosis visit or corresponding date for comparators.	Yes/no	*ATC:* R03BA
Antimicrobials	≥1 dispensation in the PDR within six months before first sarcoidosis visit or corresponding date for comparators.	Yes/no	*ATC:* J01; J02; J04; J05
Age	Data from the TPR.	Continuous	—
Region of residence	Data from the TRP. As registered the year before inclusion. Stockholm [Stockholm and Gotland counties]; Uppsala-Örebro [Uppsala, Södermanland, Värmland, Örebro, Västmanland, Dalarna, and Gävleborg]; West [Västra Götaland and Halland]; South [Skåne, Kronoberg, and Blekinge]; Southeast [Östergötland, Jönköping, and Kalmar]; North [Västernorrland, Jämtland, Västerbotten, and Norrbotten].	Stockholm (including Gotland), Uppsala-Örebro, West, South, Southeast, North	—
Birth country	Data from the TRP. Nordic [Sweden, Denmark, Norway, Finland, and Iceland].	Nordic, non-Nordic, missing	—
Attained education	Data from LISA. Completed education during the year before inclusion.	≤9, 10–12, ≥13 years, missing	—
Gross annual salary	Data from LISA. Earned during the year before first sarcoidosis visit or corresponding date for comparators and adjusted to 2014 inflation rate.	0–<100, 100–<300, 300–<600, ≥600 thousand SEK, missing	—
Married or in registered partnership	Data from LISA. Registered civil status during the year before inclusion.	Yes/no	—
Healthcare visits in the past two years	Sum of all visits in the NPR during the past two years before first sarcoidosis visit or corresponding date for comparators.	0, 1–2, ≥3 visits	—

ICD = International Classification of Disease; ATC = Anatomical Therapeutic Chemical; NPR = National Patient Register; PDR = Prescribed Drug Register; TPR = Total Population Register; LISA = Longitudinal Integrated Database for Health Insurance and Labour Market Studies.

*All sub-codes in the classification system are included if the full ICD or ATC code is not explicitly mentioned. The Swedish ICD classification system's 10^{th} revision was in use starting 1997, the 9^{th} revision between 1987 and 1996 (and 1997 in some healthcare practices) and the 8^{th} revision between 1969 and 1986.

Table S2. Demographic and clinical characteristics of individuals with sarcoidosis (n=8737) and their matched general population comparators (n=86 376) at baseline and standardized differences calculated before and after inverse probability of sarcoidosis weighting (IPSW).

	Before inverse probability weighting			After inverse probability weighting		
	Sarcoidosis	General population	Standardized difference*	Sarcoidosis	General population	Standardized difference*
Age in years, mean (SD)	49.8 (14.8)	49.8 (14.7)	0.003	48.9 (13.8)	49.8 (14.7)	-0.063
Age groups in years, %			—			—
18–44	42.5	42.5		45.5	42.5	
45–64	39.5	39.7		38.5	39.6	
65–85	18.1	17.8		16.1	17.9	
Female, %	44.5	44.6	0.001	42.8	44.6	0.036
Region of residence, %			<0.001			0.049
Stockholm	20.4	20.5		20.6	20.5	
Uppsala-Örebro	22.1	22.1		23.0	22.1	
West	18.0	18.0		18.2	18.0	
South	16.8	16.8		15.7	16.8	
Southeast	11.6	11.6		11.5	11.6	
North	11.1	11.1		11.0	11.0	
Birth country[†], %			0.064			0.064
Nordic	90.2	87.8		89.9	88.0	
Non-Nordic	9.5	11.8		9.7	11.6	
Missing	0.3	0.4		0.3	0.4	
Attained education in years, %			0.070			0.050
≤9	20.6	20.8		19.4	20.7	
10–12	49.1	46.4		47.9	46.7	
≥13	29.2	31.6		31.7	31.4	
Missing	1.2	1.2		1.0	1.2	

Table S2. (Continued).

	Before inverse probability weighting			After inverse probability weighting		
	Sarcoidosis	General population	Standardized difference	Sarcoidosis	General population	Standardized difference
Gross annual salary in 1000 SEK[‡], %			0.025			0.145
0–<100	40.2	39.3		37.8	39.4	
100–<300	30.3	30.0		31.1	30.0	
300–<600	26.0	27.3		27.6	27.2	
≥600	3.0	3.3		3.3	3.3	
Missing	0.5	<0.1		0.2	<0.1	
Married or in registered partnership, %	48.6	48.1	0.010	47.1	48.1	-0.020
Calendar period, %			<0.001			0.054
2003–2007	41.2	41.1		43.9	41.2	
2008–2013	58.8	58.9		56.1	58.8	
Health care visits in the past two years, %			0.875			0.101
0	16.9	50.9		42.6	47.8	
1–2	25.2	25.9		27.8	25.8	
≥3	57.9	23.2		29.6	26.4	
History of comorbidity, %						
Congestive heart disease	2.4	1.3	0.088	1.5	1.4	0.014
Atrial fibrillation	3.2	2.1	0.071	2.3	2.2	0.007
Acute myocardial infarction	2.1	1.8	0.025	1.8	1.8	0.000
Stroke	1.7	1.6	0.005	1.6	1.6	0.000
COPD	2.3	5.9	0.105	1.2	1.1	0.008
Asthma	4.6	2.4	0.119	2.8	2.6	0.013
Hypertension	21.4	15.9	0.142	16.4	16.4	-0.001
Diabetes mellitus	7.5	4.2	0.141	4.8	4.5	0.018
Dyslipidaemia	10.8	8.2	0.089	8.5	8.5	0.000
Autoimmune disease	7.9	4.3	0.152	5.2	4.6	0.029
Primary immunodeficiency	0.4	0.1	0.069	0.2	0.1	0.016

Table S2. (Continued).

	Before inverse probability weighting			After inverse probability weighting		
	Sarcoidosis	General population	Standardized difference	Sarcoidosis	General population	Standardized difference
Serious infection in the past year, %	3.9	0.8	—	2.3	0.9	—
≥1 first degree relative with autoimmune disease or sarcoidosis	1.1	0.5	0.069	0.7	0.5	0.026
≥1 medication dispensing in the past six months[§], %						
Systemic corticosteroids	18.7	2.9	—	15.0	3.0	—
Other immunosuppressants[‖]	1.2	0.7	—	0.8	0.8	—
Hydroxychloroquine	0.1	0.1	—	0.1	0.1	—
Inhaled corticosteroids	7.3	1.9	—	7.5	2.0	—
NSAIDs	26.3	9.7	—	27.5	10.0	—
Antimicrobials**	32.6	13.2	—	30.2	13.7	—

SD = standard deviation; SEK = Swedish krona; COPD = chronic obstructive pulmonary disease; NSAIDs = non-steroidal anti-inflammatory drugs.
Percentages may not sum to 100 owing to rounding.
*Standardized difference >0.2 indicates covariate imbalance.
†Nordic countries include Sweden, Denmark, Norway, Finland, and Iceland.
‡Adjusted to 2014 inflation level. 1.00 SEK ≈ 0.10 USD, 0.09 EUR or 0.08 GBP.
§Ascertained in individuals who entered the cohort starting Jan 1, 2006 for whom medication dispensations could be obtained from the Prescribed Drug Register (sarcoidosis, n=6723; general population, n=66 441).
‖Other immunosuppressants include methotrexate, azathioprine, and leflunomide.
**Antimicrobials include antibacterial, antimycobacterial, antifungal, and antiviral medications.

Table S3. Ten most common first serious infections after start of follow-up in individuals with sarcoidosis and their general population comparators.

Order	Disease (Swedish ICD-10 code)	n	% of individuals	% of all serious infections
Sarcoidosis			(n=8737)	(n=895)
1	Pneumonia, organism unspecified (J18)	138	1.6	15.4
2	Bacterial pneumonia (J15)	93	1.1	10.4
3	Urinary tract infection (N39)	56	0.6	6.3
4	Sepsis (A41)	52	0.6	5.8
5	Pyelonephritis (N10)	47	0.5	5.3
6	Erysipelas (A46)	44	0.5	4.9
7	Gastroenteritis and colitis (A09)	35	0.4	3.9
8	Acute upper respiratory infection (J06)	27	0.3	3.0
9	Viral intestinal infection (A08)	24	0.3	2.7
10	Bacterial intestinal infection (A04)	23	0.3	2.6
	Other			39.7
General population			(n=86 376)	(n=3881)
1	Pneumonia, organism unspecified (J18)	546	0.6	14.1
2	Urinary tract infection (N39)	346	0.4	8.9
3	Bacterial pneumonia (J15)	344	0.4	8.9
4	Sepsis (A41)	235	0.3	6.1
5	Erysipelas (A46)	205	0.2	5.3
6	Pyelonephritis (N10)	188	0.2	4.8
7	Gastroenteritis and colitis (A09)	159	0.2	4.1
8	Acute bronchitis (J20)	101	0.1	2.6
9	Unspecified infectious disease (B99)	85	0.1	2.2
10	Bacterial intestinal infection (A04)	84	0.1	2.2
	Other			40.8

ICD = International Classification of Diseases.

Table S4. Opportunistic first serious infections after inclusion in individuals with sarcoidosis and their general population comparators. Exact numbers for less than five events are not reported to eliminate the risk of identifying study participants.

	Sarcoidosis			General population		
Disease (Swedish ICD-10 code)	**n**	**% of individuals**	**% of all serious infections**	**n**	**% of individuals**	**% of all serious infections**
		(n=8737)	(n=895)		(n=86 376)	(n=3881)
Aspergillosis (B44)	≤5	—	—	≤5	—	—
Candidiasis (B37)	≤5	—	—	15	0.0	0.4
Tuberculosis (A15–A19)	11	0.1	1.2	13	0.0	0.4
Other mycobacterial infection (A31)	≤5	—	—	0	—	—
Pneumocystosis (B59)	≤5	—	—	7	0.0	0.2

ICD = International Classification of Diseases.

SUPPLEMENTAL FIGURES

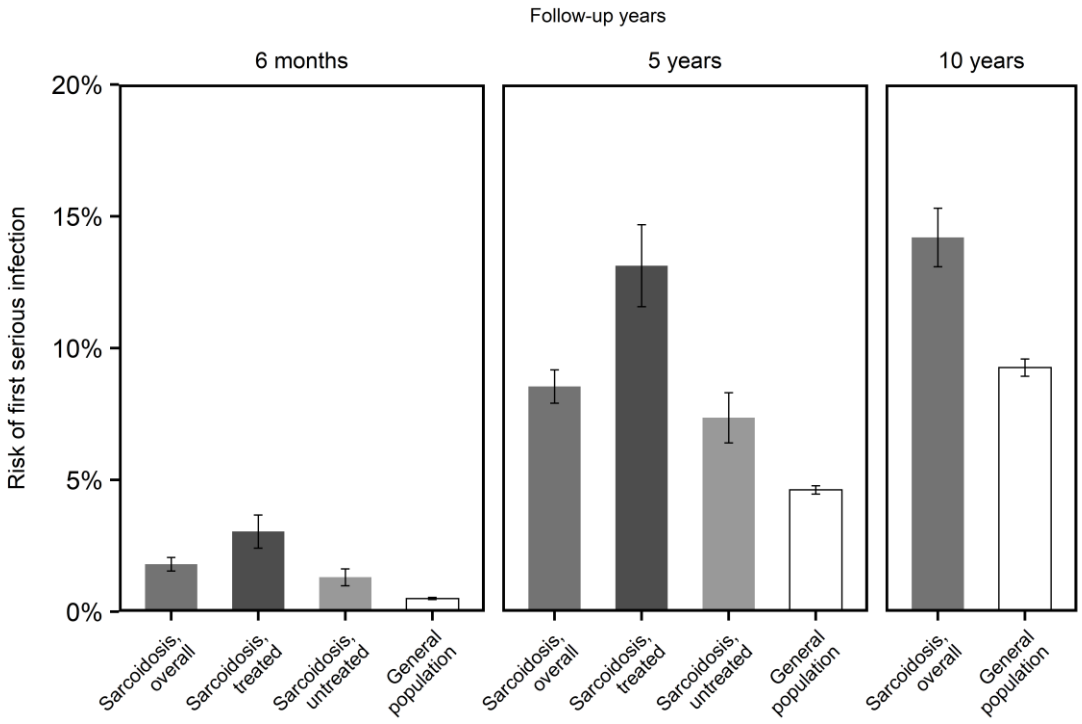

Figure S1. Risk of serious infection in sarcoidosis and the general population at six months, five, and ten years of follow-up. No risk estimates by treatment status could be obtained at 10 years because maximum follow-up in the treated and untreated groups was seven years.

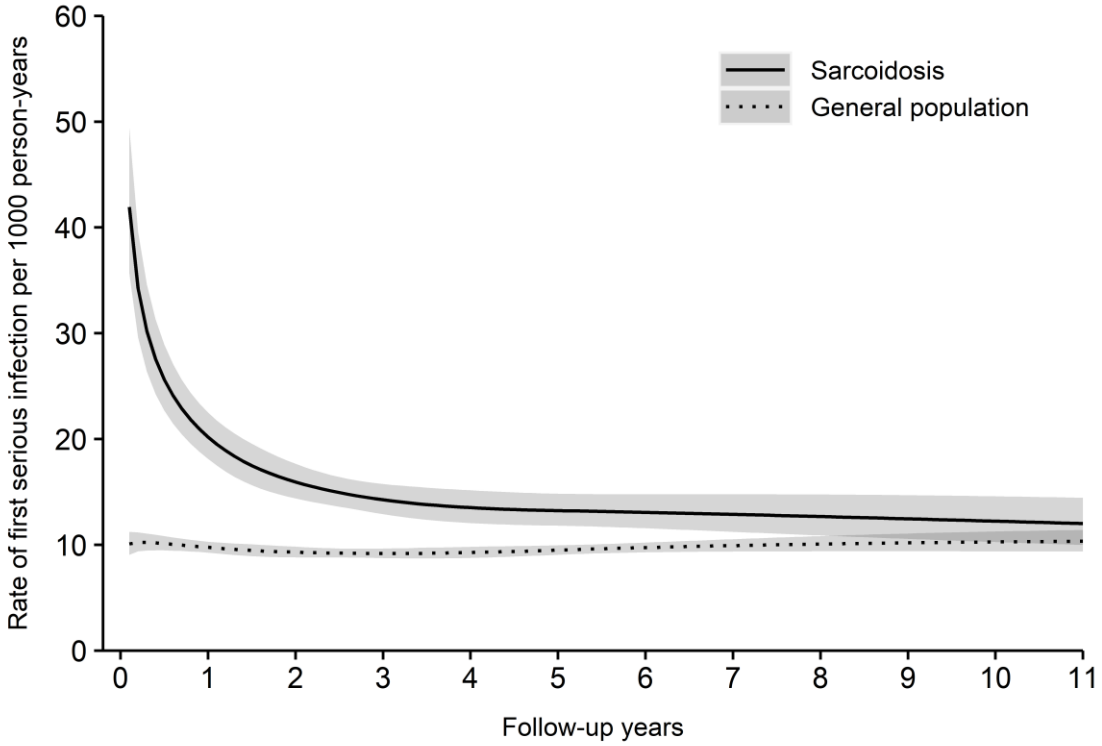

Figure S2. Adjusted rates of first serious infection in sarcoidosis and general population comparators by years of follow-up.

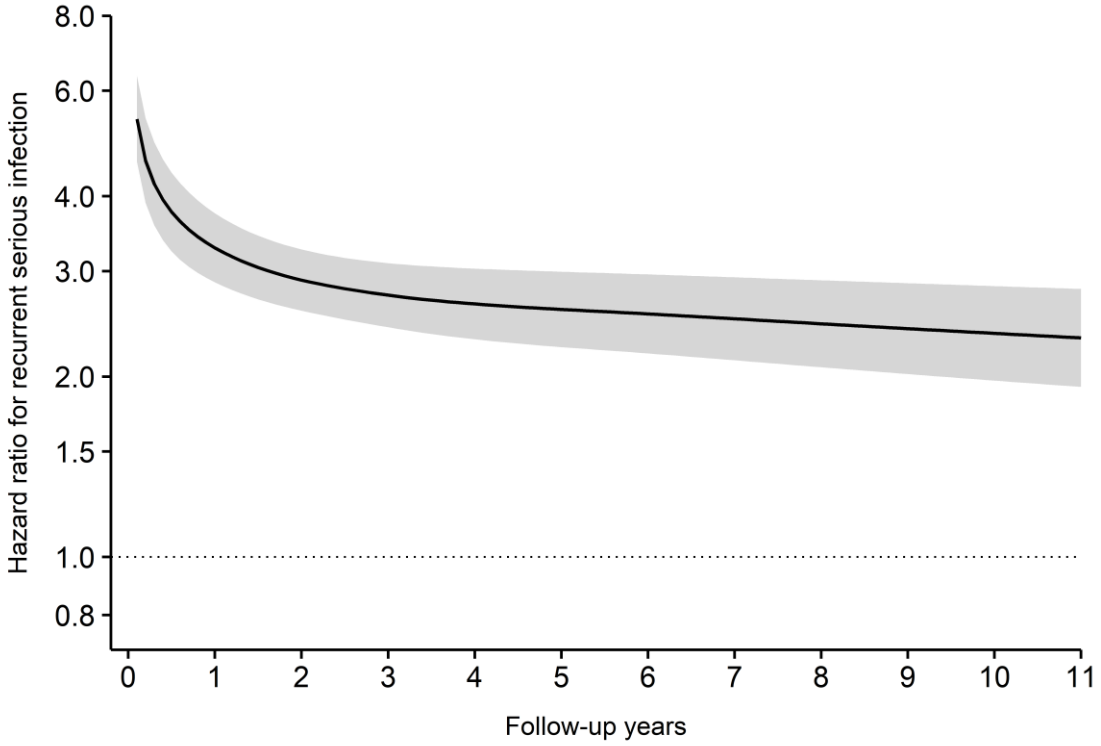

Figure S3. Adjusted within individual hazard ratios for recurrent serious infection by years of follow-up comparing sarcoidosis to the general population.