Clinical phenotypes of extrapulmonary sarcoidosis: an analysis of a French, multi-ethnic, multicentre cohort

Raphael Lhote[1,2], Isabella Annesi-Maesano[1], Hilario Nunes[3], David Launay [4], Raphael Borie [5], Karim Sacré[6], Nicolas Schleinitz[7], Mohamed Hamidou[8], Matthieu Mahevas[9], Hervé Devilliers [10], Philippe Bonniaud[11], François Lhote[12], Julien Haroche[2], Pierre Rufat[13], Zahir Amoura[2], Dominique Valeyre[3] and Fleur Cohen Aubart[2]

ABSTRACT Sarcoidosis is a rare disease of unknown cause with wide heterogeneity in clinical features and outcomes. We aimed to explore sarcoidosis phenotypes and their clinical relevance with particular attention to extrapulmonary subgroups.

The Epidemiology of Sarcoidosis (EpiSarc) study is a French retrospective multicentre study. Sarcoidosis patients were identified through national hospitalisation records using appropriate codes from 11 hospital centres between 2013 and 2016 according to a standardised protocol. Medical charts were reviewed. The phenotypes of sarcoidosis were defined using a hierarchical cluster analysis.

A total of 1237 patients were included (562 men and 675 women). The mean age at sarcoidosis diagnosis was 43.5±13 years. Hierarchical cluster analysis identified five distinct phenotypes according to organ involvement and disease type and symptoms: 1) erythema nodosum, joint involvement and hilar lymph nodes (n=180); 2) eye, neurological, digestive and kidney involvement (n=137); 3) pulmonary involvement with fibrosis and heart involvement (n=630); 4) lupus pernio and a high percentage of severe involvement (n=41); and 5) hepatosplenic, peripheral lymph node and bone involvement (n=249). Phenotype 1 was associated with being European/Caucasian and female and with non-manual work, phenotype 2 with being European/Caucasian, and phenotypes 3 and 5 with being non-European/ Caucasian. The labour worker proportion was significantly lower in phenotype 5 than in the other phenotypes.

This multicentre study confirms the existence of distinct phenotypes of sarcoidosis, with a non-random distribution of organ involvement. These phenotypes differ according to sex, geographical origin and socioprofessional category.

Introduction

Sarcoidosis is a rare heterogeneous multisystemic granulomatous disease of unknown cause. It preferentially affects the lungs and intrathoracic lymph nodes. Extrapulmonary involvement is observed in 30%–50% of patients and occurs more frequently in skin, eye, lymphatic organ, liver and spleen involvements [1]. Sarcoidosis has a variable clinical presentation and heterogeneous outcomes, hitherto unexplained.

The first clinical phenotype of sarcoidosis was described by the Swedish pulmonologist Sven Löfgren in 1946 [2]. So-called Löfgren syndrome is mostly seen in young European/Caucasian patients with acute-onset erythema nodosum, bilateral hilar lymphadenopathy, fever or migratory polyarthritis and typically has a favourable outcome. Several studies have since described the heterogeneity of sarcoidosis according to ethnicity, sex and age [3–7] and identified phenotypes, but other phenotypes are yet to be recognised. The determination of sarcoidosis phenotypes was recently supported through a clustering approach by the Genotype–Phenotype Relationship in Sarcoidosis (GenPhenReSa) study of Caucasian patients in an international panel of centres [8]. This approach identified five subgroups of patients characterised by pulmonary and extrapulmonary organ involvement. Subgroups of extrapulmonary involvement have also been reported in a significant American cohort, in which skin lesions were the most important feature [9]. More recently, ^{18}F-2-fluoro-2-deoxy-D-glucose (FDG) positron emission tomography (PET) coupled with cluster analysis succeeded in identifying in a small sample of patients an ordered stratification into four phenotypes [10]. However, other studies are necessary to confirm these findings and allow a better understanding of sarcoidosis, phenotype expression and treatment guidance [11].

The aim of the Epidemiology of Sarcoidosis (EpiSarc) study was to identify and explore sarcoidosis phenotypes using an unsupervised analytical approach in a well-characterised cohort of French sarcoidosis patients with extrathoracic involvement.

Patients and methods

The EpiSarc study was a multicentre (11 French centres) retrospective transversal study. Patients were identified using the French hospitalisations database (Programme de Médicalisation des Systèmes d'Information), which records the medical codes of hospitalisations (according to the International Classification of Diseases, 10th revision). Selected patients were hospitalised between January 2013 and December 2016 with a medical code corresponding to sarcoidosis (D86, G532 and/or M633), as newly diagnosed or chronic disease cases. Medical charts were manually reviewed for the verification of inclusion criteria. The inclusion criteria were as follows: 1) a sarcoidosis diagnosis according to the American Thoracic Society and World Association for Sarcoidosis and Other Granulomatous Diseases (WASOG) recommendations, including compatible clinico-radiological findings, histological evidence of noncaseating granuloma and the exclusion of other granulomatous diseases [12]; and 2) the involvement of at least one extrapulmonary organ. Patients with Löfgren syndrome, defined by the association of fever, erythema nodosum, arthralgia and bilateral hilar lymphadenopathy with or without histological evidence, were also included. Demographic, clinical, biological and imaging data were extracted from medical records. Organ involvements were recorded according to the WASOG organ assessment instrument and were classified as present if they were observed at any time in the sarcoidosis history [13]. We studied organ involvement

Affiliations: ^{1}Epidémiologie des Maladies Allergiques et Respiratoires (EPAR), Faculté de Médecine Saint-Antoine, Sorbonne Université, Institut Pierre Louis d'Epidémiologie et de Santé Publique (IPLESP UMRS 1136), Paris, France. ^{2}Service de Médecine Interne 2, Centre National de Référence Maladies Systémiques Rares, Lupus et Syndrome des Anticorps Antiphospholipides, Centre National de Référence Histiocytoses, Sorbonne Université, Assistance Publique Hôpitaux de Paris, Hôpital de la Pitié-Salpêtrière, Paris, France. ^{3}Service de Pneumologie, Assistance Publique Hôpitaux de Paris, Hôpital Avicenne, Bobigny, France. ^{4}Dépt de Médecine Interne et Immunologie Clinique, Centre de Référence des Maladies Auto-immunes Systémique Rares du Nord et du Nord-Ouest de France (CeRAINO), Université de Lille, INSERM U995 – LIRIC – Lille Inflammation Research International Center, CHRU de Lille, Lille, France. ^{5}Service de Pneumologie A, INSERM U1152, Université Paris Diderot, Assistance Publique Hôpitaux de Paris, Hôpital Bichat, Paris, France. ^{6}Service de Médecine Interne, Université Paris Diderot, Assistance Publique Hôpitaux de Paris, Hôpital Bichat, Paris, France. ^{7}Service de Médecine Interne, Université Aix Marseille, Assistance Publique Hôpitaux de Marseille, Hôpital de la Timone, Marseille, France. ^{8}Service de Médecine Interne, Université de Nantes, Centre Hospitalier Universitaire, Nantes, France. ^{9}Service de Médecine Interne, Assistance Publique Hôpitaux de Paris, Hôpital Mondor, Créteil, France. ^{10}Service de Médecine Interne 2, Centre Hospitalier Dijon-Bourgogne, Dijon, France. ^{11}Service de Pneumologie et Soins Intensifs Respiratoires, Centre de Référence Constitutif des Maladies Pulmonaires Rares de l'Adulte, Centre Hospitalier Dijon-Bourgogne, Dijon, France. ^{12}Service de Médecine Interne, Hôpital Delafontaine, St-Denis, France. ^{13}Dépt d'Information Médicale et Biostatistiques, Sorbonne Université, Assistance Publique Hôpitaux de Paris, Hôpital de la Pitié-Salpêtrière, Paris, France.

Correspondence: Fleur Cohen Aubart, Service de Médecine Interne 2, Groupe Hospitalier Pitié-Salpêtrière, 47-83 Boulevard de l'Hôpital, 75 651 Paris Cedex 13, France. E-mail: fleur.cohen@aphp.fr

listed in the extrapulmonary Physician Organ Severity Tool (ePOST) score and pulmonary involvement [14]. In addition, the analysis also considered the sociodemographic characteristics of the patients (age, self-reported geographical origin and occupation at the time of diagnosis, classified according to the Institut National de la Statistique et des Etudes Economiques (INSEE) tool (www.insee.fr/fr/information/ 2497958) and to the International Standard Classification of Occupation (ISCO) 2008), comorbidities (thromboembolism disease, neoplasm, autoimmune disease, chronic infectious disease), treatments and outcomes. Occupations were categorised into six groups according to INSEE classification and ISCO 2008: craft and trade-related workers (ISCO 2008 group 7); labour workers (ISCO 2008 groups 8 and 9); farmers and skilled agricultural workers (ISCO 2008 group 6); clerk support workers (ISCO 2008 group 4); intermediary occupations (ISCO 2008 groups 0, 3 and 5); and upper class, framework and higher intellectual professions (ISCO 2008 groups 1 and 2).

Statistical analysis

Variables are presented as percentages (categorical variables) or mean±SD (continuous variables) as appropriate. The phenotypes of sarcoidosis were defined using a hierarchical cluster analysis. For hierarchical ascendant clustering, preferential variable associations were obtained after including the variables of organ involvement listed in the ePOST score, in addition to the following, referring to intrathoracic involvement: mediastinal lymph nodes, parenchymal involvement and lung fibrosis. Clustering was represented in a dendrogram plot ("pvclust packages" in R version 3.3, R Project for Statistical Computing, Vienna, Austria) generated by Ward's method. Principal component analysis was applied to select the final clinical variables to include in the model to identify phenotypes according to statistical significance. The population was then separated into clusters ("factoMineR packages" in R version 3.3) [15]. Unsupervised hierarchical cluster analysis was performed using the Euclidian method to create the distance matrix and Ward's minimum variance method. Differences in sarcoidosis phenotypes were compared using Chi-squared test or Fisher's exact test when necessary (categorical variables) and a t-test, ANOVA, Wilcoxon test or Kruskal–Wallis test (continuous variables) when appropriate. In addition, multinomial logistic regression was used to confirm the relationships of the different clusters to the set of independent variables. The multinomial logistic regression models were adjusted for sex, geographical origin and socioprofessional activities. All tests were bilateral, with Type I errors of 5% and 95% confidence intervals. The analyses were performed in R software, version 3.3. This study was approved according to the French legislation by the Commission Nationale Informatique et Libertés organisation.

Results

Demographic and clinical characteristics

During the study period, 2090 patients were screened. Among them, 853 patients were excluded (figure 1) because of the absence of histological documentation (n=321, 15%), the absence of extrapulmonary organ involvement (n=275, 13%), insufficient data in medical records (n=95, 5%), the absence of a sarcoidosis diagnosis (n=155, 7%) and redundant follow-up in two participating centres (n=7, 0.05%). The final study included 1237 patients, 562 (44.5%) male and 675 (54.5%) female. The mean age at sarcoidosis diagnosis was 43.5±13 years. Most of the patients were European/Caucasian (n=541, 44%), followed by Afro-Caribbean (n=314, 25.5%). Detailed clinical characteristics of the population are described in table 1.
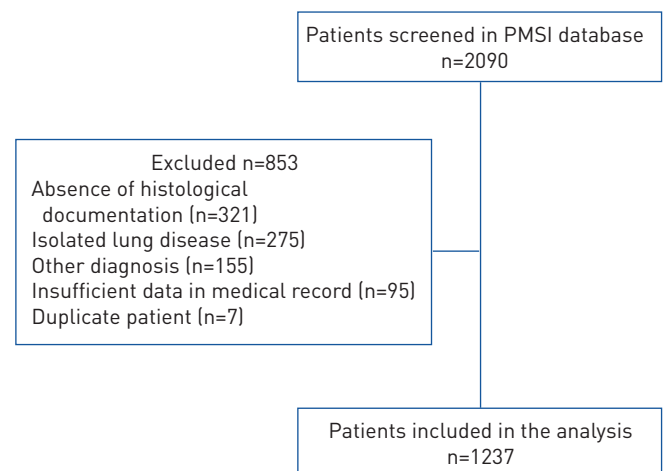

FIGURE 1 Flow chart. Description of the EpiSarc study sample enrolled from the French hospitalisation database (Programme de Médicalisation des Systèmes d'Information (PMSI)) in the 11 participating centres, from January 2013 to December 2016.

TABLE 1 Clinical presentation of patients in the EpiSarc study

	All	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	p-value
Subjects n	1237	180	137	630	41	249	
Male	562 (45)	61 (34)	60 (44)	310 (49)	13 (32)	118 (47)	<0.002
Age at sarcoidosis diagnosis years	43.5±13	42±10.2	43.2±16.7	44.5±13	42.2±10.5	42.2±13.1	0.8
Number of organs affected	3.4±1.5	2.8±1.1	2.7±1.5	3.2±1.2	4.4±1.7	4.8±1.4	<0.001
Intrathoracic involvement	1143 (92)	178 (99)	45 (33)	630 (100)	41 (100)	249 (100)	<0.001
Intrathoracic nodes	1115 (90)	178 (99)	23 (17)	627 (99.5)	41 (100)	246 (99)	<0.001
Parenchymal involvement	792 (64)	71 (39)	27 (20)	453 (72)	36 (88)	205 (82)	<0.001
Fibrosis	176 (14)	2 (1)	3 (2)	120 (19)	12 (29)	39 (16)	<0.001
Peripheral nodes	518 (42)	43 (24)	49 (36)	204 (32)	11 (27)	211 (85)	<0.001
Skin involvement	396 (32)	27 (15)	26 (19)	235 (37)	39 (95)	69 (28)	<0.001
Cutaneous sarcoids	357 (29)	26 (14)	26 (19)	234 (37)	2 (5)	69 (28)	<0.001
Lupus pernio	38 (3)	0 (0)	0 (0)	0 (0)	38 (93)	0 (0)	<0.001
Erythema nodosum	136 (11)	120 (67)	3 (2)	3 (0.5)	2 (5)	8 (3)	<0.001
Joint involvement	350 (28)	162 (90)	20 (15)	96 (15)	18 (44)	54 (22)	<0.001
Liver involvement	301 (24)	11 (6)	32 (23)	85 (13)	4 (10)	169 (68)	<0.001
Eye involvement	297 (24)	24 (13)	35 (26)	184 (29)	6 (15)	48 (19)	<0.001
Conjunctival nodules	24 (2)	1 (0.5)	2 (1.5)	14 (2)	0 (0)	7 (3)	0.5
Uveitis	265 (21)	22 (12)	33 (24)	166 (26)	4 (10)	40 (16)	<0.001
Anterior uveitis[#]	125 (11)	13 (7)	15 (12)	77 (13)	1 (2)	19 (8)	0.04
Posterior uveitis[#]	36 (3)	6 (3)	6 (5)	22 (4)	0 (0)	2 (1)	0.1
Panuveitis[#]	57 (5)	4 (2)	9 (7)	34 (6)	1 (2)	9 (4)	0.2
Ear, nose and throat involvement	280 (23)	23 (13)	25 (18)	132 (21)	28 (68)	72 (29)	<0.001
Nasosinusal involvement	131 (11)	6 (3)	8 (6)	72 (11)	20 (50)	25 (10)	<0.001
Laryngopharyngeal involvement	28 (2)	0 (0)	5 (4)	14 (2)	4 (10)	5 (2)	0.003
Parotid and lacrimal gland involvement	137 (11)	16 (9)	13 (10)	55 (9)	3 (7)	50 (20)	<0.001
Neurological involvement	271 (22)	13 (7)	43 (31)	181 (28)	10 (24)	24 (10)	<0.001
Cranial nerve involvement	130 (11)	12 (7)	19 (14)	87 (14)	3 (7)	9 (4)	<0.001
Optical nerve involvement	55 (4)	4 (2)	9 (7)	39 (6)	1 (2.5)	2 (1)	0.003
Medullary involvement	37 (3)	2 (1)	7 (5)	23 (4)	2 (5)	3 (1)	0.07
Parenchymal involvement	97 (8)	2 (1)	20 (15)	67 (11)	4 (10)	4 (2)	<0.001
Pituitary infiltration	33 (3)	0 (0)	6 (5)	21 (4)	2 (5)	4 (2)	0.051
Cerebrospinal fluid inflammation	103 (8)	4 (2)	16 (12)	73 (12)	3 (7)	7 (3)	<0.001
Neuromuscular involvement	42 (3)	2 (1)	9 (7)	21 (3)	2 (5)	8 (3)	0.1
Spleen involvement	195 (16)	3 (2)	10 (7)	11 (2)	3 (7)	168 (67)	<0.001
Cardiac involvement	180 (15)	11 (6)	8 (6)	120 (19)	7 (17)	34 (14)	<0.001
Bone involvement	118 (10)	7 (4)	8 (6)	20 (3)	10 (24)	73 (29)	<0.001
Kidney involvement	75 (6)	3 (2)	16 (12)	37 (6)	1 (2.5)	18 (7)	0.004
Digestive involvement	41 (3)	1 (0.5)	33 (24)	2 (0.5)	0 (0)	5 (2)	<0.001
Hypercalcaemia	99 (8)	7 (4)	6 (4)	51 (8)	3 (7)	31(12)	0.01

Data are presented as n (%) or mean±SD, unless otherwise stated. [#]: out of 1188 patients (among the 265 patients with uveitis, the characteristics of uveitis were not available in 49).

Clinical characteristics of the patients are described in table 2. Organ involvement mainly consisted of hilar lymphadenopathy (n=1115, n=90%), parenchymal lung involvement (n=792, 64%), pulmonary fibrosis (n=176, 14%), peripheral lymphadenopathy (n=518, 42%), skin involvement (n=396, 32%) and joint involvement (n=350, 28%). The mean number of organs involved was 3.4±1.5, and 520 patients (42%) had more than three organs involved.

Sarcoidosis phenotypes
Hierarchical cluster analysis identified five groups of patients, corresponding to five sarcoidosis phenotypes. The clinical characteristics of the population according to these five phenotypes are described in table 2. The five phenotypes were as follows: 1) a higher frequency of hilar lymph nodes (99%) and joint involvement (90%), a lower frequency of parenchymal lung involvement (39%) and the presence of erythema nodosum (67%) (n=180); 2) a higher frequency of neurological (31%), digestive (24%) and kidney (12%) involvement and a lower frequency of mediastino-pulmonary involvement (33%) (n=137); 3) a higher frequency of parenchymal lung involvement (72%), including pulmonary fibrosis (19%), cardiac involvement (19%) and cutaneous sarcoids (37%) (n=630); 4) an aggregation of all patients with lupus pernio (95%) and a high frequency of ear, nose and throat (ENT) (68%) and parenchymal lung involvement (88%), including

TABLE 2 Occupational activities, geographical origins and outcomes of patients in the EpiSarc study

	All	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	p-value
Subjects n	1237	180	137	630	41	249	
Occupational status	964 (78)	137 (76)	102 (74)	493 (78)	27 (66)	205 (82)	
Employed	887 (92)	125 (92)	88 (86)	461 (94)	25 (93)	188 (92)	<0.001
Upper class: managers and professionals	66 (8)	9 (7)	9 (10)	31 (7)	3 (12)	14 (8)	
Clerk support workers	259 (29)	41 (33)	22 (25)	116 (25)	11 (44)	69 (37)	
Intermediary occupations	168 (19)	37 (30)	12 (14)	85 (18)	4 (16)	30 (16)	
Craftsmen and trade-related workers	26 (3)	1 (0.5)	5 (6)	12 (3)	0 (0)	8 (4)	
Labour workers	360 (41)	37 (30)	39 (44)	211 (46)	7 (28)	66 (35)	
Skilled agricultural workers	8 (1)	0 (0)	1 (1)	6 (1)	0 (0)	1 (0.5)	
Student	14 (1.5)	1 (0.5)	7 (7)	2 (0.5)	0 (0)	4 (2)	
Unemployed	63 (6.5)	11 (8)	7 (7)	30 (6)	2 (7)	13 (6)	
Geographical origins	123 (99.5)	179 (99)	136 (99)	627 (99)	40 (98)	248 (99.5)	0.01
European/Caucasian	541 (44)	91 (51)	85 (62.5)	264 (42)	8 (20)	93 (37.5)	
Northern African	305 (25)	47 (26)	24 (18)	159 (25)	15 (37.5)	60 (24)	
Sub-Saharan Africa	208 (17)	26 (15)	17 (12.5)	106 (17)	6 (15)	53 (21.5)	
Overseas departments[#]	106 (8.5)	9 (5)	5 (3.5)	57 (9)	8 (20)	27 (11)	
Others	70 (5.5)	6 (3)	5 (3.5)	41 (7)	3 (7.5)	15 (6)	
Familial sarcoidosis[¶]	62 (8)	4 (4)	6 (7)	35 (10)	3 (11)	14 (8)	0.4
Follow-up duration years	8.1±8	6.3±7.1	7±7.1	8.8±8.6	14.6±9.7	7.5±8	0.03
Death	23 (2)	0 (0)	2 (1)	18 (3)	2 (5)	1 (0.5)	0.01
Treatments							
Corticosteroids	906 (73)	95 (53)	98 (72)	492 (78)	36 (88)	185 (74)	<0.001
Methotrexate	388 (31)	34 (19)	40 (22)	209 (33)	31 (76)	74 (30)	<0.001
Azathioprine	161 (13)	13 (7)	10 (7)	85 (14)	14 (34)	39 (16)	<0.001
Hydroxychloroquine	283 (23)	43 (24)	20 (15)	138 (22)	20 (49)	62 (25)	<0.001
TNF-α blockers	113 (9)	7 (4)	15 (11)	62 (10)	17 (41)	12 (5)	<0.001
Cyclophosphamide	61 (5)	3 (2)	13 (9)	36 (6)	3 (7)	6 (2)	0.003

Data are presented as n (%) or mean±SD, unless otherwise stated. TNF: tumor necrosis factor. [#]: the French overseas departments include the Caribbean and Indian Ocean isles; these populations are mainly from African American descendants. [¶]: data presented are from 775 patients (for the others, this information was not available).

pulmonary fibrosis (29%) (n=41); and 5) a high frequency of parenchymal lung involvement (82%), peripheral nodes (68%) and hepatic (68%), splenic (68%) and bone (29%) involvement (n=249). Parenchymal lung involvement, including pulmonary fibrosis, was present in clusters 3, 4 and 5. Cluster 3 associated neurological and abdominal involvement, cluster 4 aggregated cases of lupus pernio and ENT involvement, and cluster 5 comprised peripheral lymph nodes and hepatosplenic and bone involvement, in addition to parenchymal localisation. There was no difference in age at diagnosis regarding phenotype status. To confirm reproducibility and the stability of the results, the cluster analysis was performed twice, once after including patients from seven centres (corresponding to 1081 patients), then after including all 11 centres (1237 patients), leading to a concordance of 85% of the results, with the same clustering.

In addition to cluster separation, non-supervised preferential associations of organ involvement are shown in figure 2. Erythema nodosum was associated with joint involvement. Peripheral nodes were associated with bone and hepatosplenic involvement. Cardiac localisations were associated with parenchymal lung involvement and fibrosis. Lupus pernio was associated with ENT manifestations, uveitis and neurological involvement. These associations strengthened the results of the cluster distributions.

Treatments and outcomes

Treatments and outcomes according to the five phenotypes are described in table 2. The mean duration of follow-up was 8.1±8 years, and was longest in cluster 4. The most common treatment regimens were corticosteroids (n=907, 73%), methotrexate (n=388, 31%), hydroxychloroquine (n=283, 23%), azathioprine (n=161, 13%), tumor necrosis factor-α (TNF-α) antagonists (n=113, 9%) and cyclophosphamide (n=61, 5%). During follow-up, 23 patients died (2%).

Treatment regimens (*i.e.* corticosteroids, methotrexate, azathioprine, TNF-α antagonists and/or cyclophosphamide) were significantly different among the five phenotypes.

Phenotype 1 was treated less than other phenotypes with corticosteroids (n=95, 53% *versus* 77%, p<0.001), with methotrexate (n=34, 19% *versus* 34%, p<0.001), with azathioprine (n=13, 7% *versus* 14%, p=0.02),
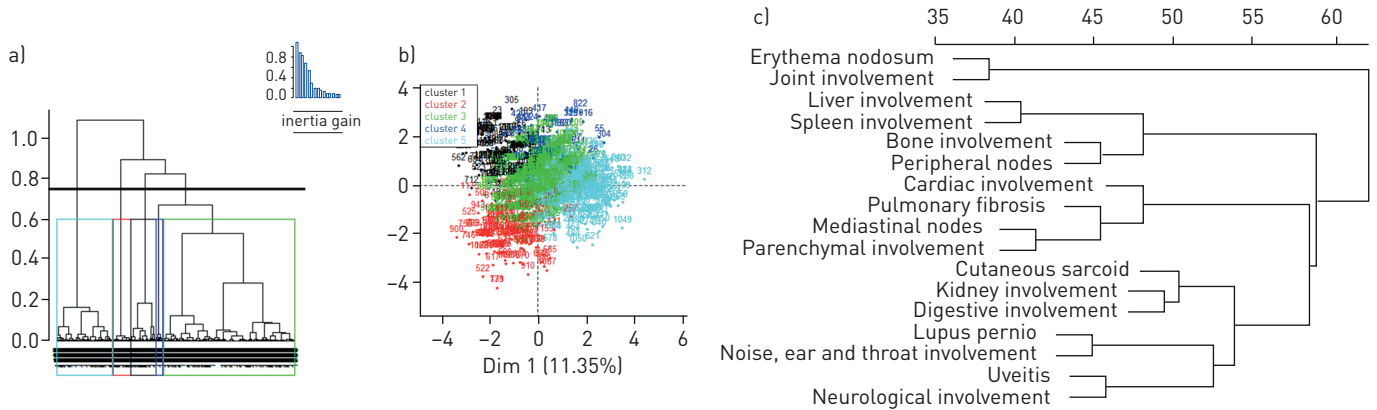
FIGURE 2 Preferential association of organ involvement in sarcoidosis in the EpiSarc study. a) Hierarchical cluster analysis. The bold horizontal line represents the section of the population allowing the smallest number of statistically valid clusters. b) Factor map. c) Dendrogram, which includes organ involvement listed in the extrapulmonary Physician Organ Severity Tool (ePOST) score and pulmonary involvement.

with TNF-α antagonists (n=7, 4% *versus* 10%, p=0.01) and with cyclophosphamide (n=3, 2% *versus* 5%, p=0.046). Phenotype 2 was treated more than other phenotypes with cyclophosphamide (n=15, 11% *versus* 9%, p=0.02) and treated less with azathioprine (n=10, 7% *versus* 14%, p=0.046). Phenotype 3 was treated more than other phenotypes with corticosteroids (n=493, 78% *versus* 68%, p<0.001). Phenotype 4 was treated more than other phenotypes with corticosteroids (n=36, 88% *versus* 73%, p=0.05), methotrexate (n=31, 76% *versus* 30%, p<0.001), azathioprine (n=14, 34% *versus* 12%, p<0.001) and TNF-α antagonists (n=17, 41% *versus* 8%, p<0.001). Phenotype 5 was treated less than the other clusters with TNF-α antagonists (n=12, 5% *versus* 10%, p=0.01). The number of treatments was also different between phenotypes (p<0.001) (figure 3). Patients received a lower number of treatments with phenotype 1 (1.1±1.3) than with phenotype 2 (1.6±1.3), phenotype 3 (1.7±1.2) and phenotype 5 (1.6±1.2). The number of treatments was highest among patients with phenotype 4 (3.6±1.7). The number of deaths was also significantly different in the clusters, with no deaths from phenotype 1; 0.5%–3% with phenotypes 2, 3 and 5; and 5% with phenotype 4 (p=0.01).
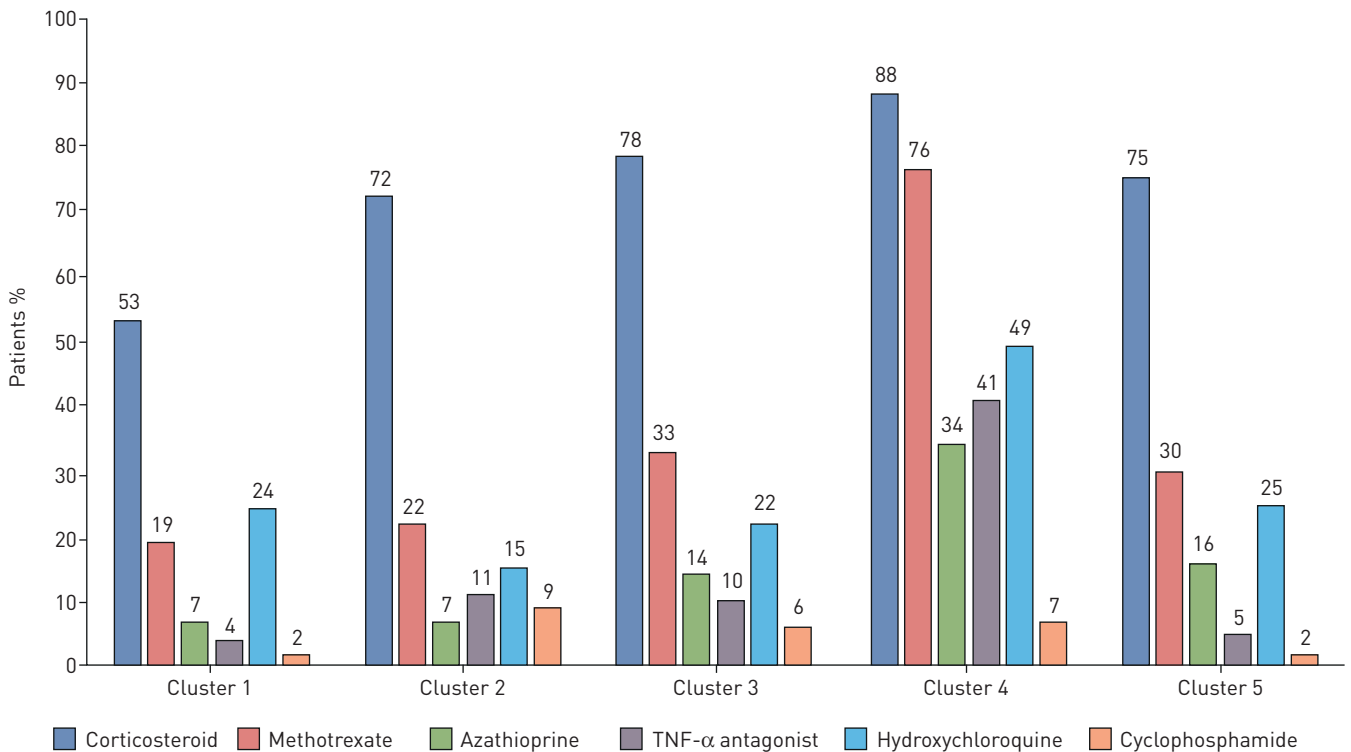

FIGURE 3 Treatments administered in the five phenotypes of sarcoidosis. Percentages of patients receiving corticosteroids, methotrexate, azathioprine, hydroxychloroquine, cyclophosphamide and tumor necrosis factor-α (TNF-α) antagonists in the five phenotypes of sarcoidosis.

Socioeconomic factors related to sarcoidosis phenotypes

The phenotypes varied according to socioeconomic characteristics (table 3). Females were represented differently according to the phenotypes (p<0.002), and the female to male ratio was higher in phenotypes 1 (66% women) and 4 (68% women) than in phenotypes 2 (56% women), 3 (51% women) and 5 (53% women). There was no difference in the age at diagnosis of sarcoidosis between phenotypes.

The occupational repartition differed significantly between the five phenotypes (p=0.04). Labour workers (40.5%) were the most frequent occupational population among the whole cohort. The proportion of labour workers was higher for phenotype 3 (46%), phenotype 2 (44%) and phenotype 5 (35%).

The geographical origin repartition was significantly different regarding the phenotype (p=0.01). Patients of European/Caucasian origin (44%) were the most represented geographical category in the study population. The proportion of European/Caucasian origin was higher for phenotype 1 (51%) and phenotype 2 (62.5%) than for phenotypes 3 (42%), 5 (37.5%) or 4 (20%).

The adjusted-multinomial analysis confirmed the results of the univariate analysis. The female proportion was significantly higher for phenotype 1 than for the other phenotypes (OR 0.6, 95% CI 0.4–0.9, p=0.008); the labour worker proportion was significantly lower for phenotypes 1 (OR 0.6, 95% CI 0.4–0.7, p=0.03) and 5 (OR 0.6, 95% CI 0.4–0.9, p=0.01) than for the other phenotypes. European/Caucasian origin was significantly higher for phenotype 1 (OR 1.7, 95% CI 1.1–2.5, p=0.01) and phenotype 2 (OR 2.2, 95% CI 1.4–3.4, p<0.001) than for the other phenotypes.

Discussion

The EpiSarc study identified five phenotypes of sarcoidosis in a large French multicentre and multi-ethnic cohort of 1237 well-characterised patients. These phenotypes may be useful for better management of the disease [11]. The five phenotypes significantly differed by sex, geographical origin and occupational categories. The five phenotypes were 1) erythema nodosum, joint involvement and hilar lymph nodes, mainly European/Caucasian female; 2) neurological, digestive and/or kidney involvement; 3) parenchymal lung involvement and fibrosis, cardiac and skin involvement, mainly in non-European/Caucasian patients; 4) lupus pernio and severe involvement; and 5) parenchymal pulmonary involvement, peripheral nodes, and hepatic, splenic and bone involvement, mainly in non-European/Caucasian patients. In addition to clustering phenotypes, we identified a preferential association of organ involvements, characterised by the

TABLE 3 Factors related to clusters of sarcoidosis of patients in the EpiSarc study

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Subjects n	180	137	630	41	249
Male sex					
OR (95% CI)	**0.6 (0.4–0.9)**	0.8 (0.5–1.3)	1	0.6 (0.3–1.5)	0.9 (0.6–1.3)
p-value	**0.008**	0.4	NA	0.3	0.5
Labour worker					
OR (95% CI)	**0.6 (0.4–0.7)**	0.9 (0.6–1.5)	1	0.5 (0.2–1.3)	**0.6 (0.4–0.9)**
p-value	**0.03**	0.7	NA	0.1	**0.01**
European/Caucasian					
OR (95% CI)	1.7 (1.1–2.5)	**2.2 (1.4–3.4)**	1	0.4 (0.2–1.1)	0.9 (0.6–1.3)
p-value	0.01	**<0.001**	NA	0.08	0.6
Sub-Saharan African					
OR (95% CI)	0.6 (0.4–1.2)	0.7 (0.4–1.3)	1	1.1 (0.4–2.9)	1.2 (0.8–1.8)
p-value	0.1	0.3	NA	0.9	0.5
Caribbean					
OR (95% CI)	0.5 (0.2–1.2)	0.4 (0.2–1.3)	1	2.5 (0.9–7.1)	1.3 (0.7–2.2)
p-value	0.1	0.1	NA	0.08	0.4
North African					
OR (95% CI)	1 (0.6–1.6)	0.7 (0.4–1.2)	1	1.3 (0.6–3.2)	0.9 (0.6–1.4)
p-value	0.98	0.2	NA	0.5	0.7
Other					
OR (95% CI)	0.5 (0.2–1.4)	0.6 (0.2–1.7)	1	1.2 (0.3–5.3)	0.9 (0.5–1.8)
p-value	0.2	0.4	NA	0.8	0.8

Odds ratios and 95% confidence intervals in each cluster adjusted for sex, geographical origin and socioprofessional activities. Reference is cluster 3. Significant results are represented in bold type. NA: not applicable.

association of erythema nodosum with joint involvement; hepatic, splenic, bone and peripheral node involvement; pulmonary with cardiac involvements; kidney with digestive involvement; lupus pernio with ENT involvement; and uveitis with neurological involvement.

In the GenPhenReSa study, which included 2163 patients, the phenotypes were slightly distinct from those of the EpiSarc study. They included 1) abdominal organ involvement, 2) ocular-cardiac-cutaneous-central nervous system (CNS) disease involvement, 3) musculoskeletal-cutaneous involvement, 4) pulmonary and intrathoracic lymph node involvement and 5) extrapulmonary involvement. However, the GenPhenReSa study only included Caucasian patients. Organ involvement was different in the GenPhenReSa study: we observed more skin (32% *versus* 16%), eye (24% *versus* 8%), CNS (22% *versus* 3%), cardiac (10% *versus* 3%) and gastrointestinal involvement (3% *versus* 0.6%). Several explanations may be proposed: our cohort is a multi-ethnic cohort, and African American patients usually have more organ involvement [7]. Moreover, our patients were mostly enrolled from internal medicine centres, whereas the GenPhenReSa study participants were mostly enrolled from pulmonology centres. Our French centres are highly specialised centres, and it is likely that we recruited patients with more severe disease, particularly for cardiac, eye or CNS involvement. We only recruited hospitalised patients. Finally, we excluded patients with isolated intrathoracic sarcoidosis from our cohort, in order to highlight the extrapulmonary localisations in the clustering analysis. Thus, we think that our results are complementary to those of the GenPhenReSa study. In a Spanish study of 1230 patients, including 91.4% white patients, those with pulmonary involvement had a lower frequency of skin and salivary gland involvement, and a higher frequency of liver involvement [16]. This study used multinomial logistic regression analyses to test the associations, whereas we used unsupervised hierarchical clustering. Thus we better detected phenotypes than this Spanish study, which studied associations. In another study of 195 patients from pulmonology centres, the authors used ^{18}F-FDG PET to identify four phenotypes based on not only organ involvement, but also disease activity [10]. However, the phenotypes were not compared for severity and socioeconomic factors. The impact of this study is also limited because not all patients with sarcoidosis need ^{18}F-FDG PET, which is an expensive examination with radiation exposure.

In the EpiSarc study, unsupervised hierarchical clustering allowed the aggregation of organ involvement corresponding to Löfgren syndrome [17]. We also demonstrated female and European/Caucasian predominance among these patients. Löfgren syndrome was first described in 1946 and is recognised as a distinct phenotype of sarcoidosis. Patients typically have acute disease onset together with bilateral hilar lymphadenopathy, erythema nodosum and/or bilateral ankle arthritis or arthralgia. Patients with Löfgren syndrome typically have favourable outcomes, especially in HLA-DRB1*03+ individuals [18]. We found that Löfgren/phenotype 1 was less treated, with a mean number of treatments of 1.1±1.3 and lower rates for each molecule. It was also associated with better outcomes with no deaths. Extrapulmonary sarcoidosis specificities were recently reported in an American cohort [9]. In this multicentre study, the skin was the most frequently involved extrathoracic organ in patients with or without lung involvement. Our study focused on hospitalised patients, and this could explain the difference in organ involvement frequencies from other cohorts.

As previously reported, in the EpiSarc study, lupus pernio was strongly associated with a severe phenotype [19, 20]. With this phenotype, TNF-α antagonists were more frequently used (41%), and patients received more treatments than with other phenotypes. Lupus pernio was associated with pulmonary fibrosis and ENT, CNS, bone and joint involvement and was more frequently seen in Afro-Caribbean women.

This study has several strengths. The patients were recruited from 11 centres highly specialised in sarcoidosis. The cohort was well characterised: in addition to the verification of inclusion criteria, organ involvement was assessed from individual medical records using the WASOG instrument, leading to a strong validity of organ involvement. The main limitations of the study were the retrospective design and a possible selection bias of particularly severe patients with complicated, multi-organ involvements because we selected only hospitalised patients. We had only one cohort, so we could not obtain external validation. Additionally, the analysis did not take into account the time of appearance of organ involvement from the onset of disease. Both newly diagnosed cases and cases that had chronic disease were included in the cohort: in incident cases, additional organ involvement could have occurred during the disease course and accumulated, even though most of the time organ involvement in sarcoidosis is present at diagnosis. We provided a clustering based only on the basis of organ involvement, not on disease activity. With this study, we did not obtain a complete understanding of why sarcoidosis phenotypes exist: genetic, environmental, epigenetic and immunologic factors are increasingly being evaluated as possible mechanisms for sarcoidosis aetiology, but they are probably also major determinants of disease heterogeneity [21]. This should be evaluated in further studies.

In conclusion, our multicentre study highlights and confirms the existence of distinct phenotypes of sarcoidosis, with a non-random distribution of organ involvement. These phenotypes are associated with

sex, geographical origin and socioprofessional category in a multi-ethnic cohort. These phenotypes should be used to understand sarcoidosis not as a single disease but as various syndromes. The genetic and environmental determinants of such phenotypes have to be elucidated in future studies.

References

1 Valeyre D, Bernaudin JF, Uzunhan Y, et al. Clinical presentation of sarcoidosis and diagnostic work-up. Semin Respir Crit Care Med 2014; 35: 336–351.

2 Lofgren S, Lundback H. The bilateral hilar lymphoma syndrome; a study of the relation to tuberculosis and sarcoidosis in 212 cases. Acta Med Scand 1952; 142: 265–273.

3 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.

4 Baughman RP, Teirstein AS, Judson MA, et al. Clinical characteristics of patients in a case control study of sarcoidosis. Am J Respir Crit Care Med 2001; 164: 1885–1889.

5 Loddenkemper R, Kloppenborg A, Schoenfeld N, et al. Clinical findings in 715 patients with newly detected pulmonary sarcoidosis--results of a cooperative study in former West Germany and Switzerland. WATL Study Group. Wissenschaftliche Arbeitsgemeinschaft fur die Therapie von Lungenkrankheitan. Sarcoidosis Vasc Diffuse Lung Dis 1998; 15: 178–182.

6 Pereira CA, Dornfeld MC, Baughman R, et al. Clinical phenotypes in sarcoidosis. Curr Opin Pulm Med 2014; 20: 496–502.

7 Judson MA, Boan AD, Lackland DT. The clinical course of sarcoidosis: presentation, diagnosis, and treatment in a large white and black cohort in the United States. Sarcoidosis Vasc Diffuse Lung Dis 2012; 29: 119–127.

8 Schupp JC, Freitag-Wolf S, Bargagli E, et al. Phenotypes of organ involvement in sarcoidosis. Eur Respir J 2018; 51: 1700991.

9 James WE, Koutroumpakis E, Saha B, et al. Clinical features of extrapulmonary sarcoidosis without lung involvement. Chest 2018; 154: 349–356.

10 Papiris SA, Georgakopoulos A, Papaioannou AI, et al. Emerging phenotypes of sarcoidosis based on 18F-FDG PET/CT: a hierarchical cluster analysis. Expert Rev Respir Med 2020; 14: 229–238.

11 Baughman RP, Scholand MB, Rahaghi FF. Clinical phenotyping: role in treatment decisions in sarcoidosis. Eur Respir Rev 2020; 29: 190145.

12 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.

13 Judson MA, Costabel U, Drent M, et al. The WASOG Sarcoidosis Organ Assessment Instrument: an update of a previous clinical tool. Sarcoidosis Vasc Diffuse Lung Dis 2014; 31: 19–27.

14 Judson MA, Baughman RP, Costabel U, et al. Efficacy of infliximab in extrapulmonary sarcoidosis: results from a randomised trial. Eur Respir J 2008; 31: 1189–1196.

15 Lê S, Josse J, Husson F. FactoMineR: an R package for multivariate analysis. J Stat Softw 2008; 25: 1–28.

16 Perez-Alvarez R, Brito-Zeron P, Kostov B, et al. Systemic phenotype of sarcoidosis associated with radiological stages. Analysis of 1230 patients. Eur J Intern Med 2019; 69: 77–85.

17 Karakaya B, Kaiser Y, van Moorsel CHM, et al. Lofgren's syndrome: diagnosis, management, and disease pathogenesis. Semin Respir Crit Care Med 2017; 38: 463–476.

18 Grunewald J, Eklund A, Olerup O. Human leukocyte antigen class I alleles and the disease course in sarcoidosis patients. Am J Respir Crit Care Med 2004; 169: 696–702.

19 Aubart FC, Ouayoun M, Brauner M, et al. Sinonasal involvement in sarcoidosis: a case-control study of 20 patients. Medicine 2006; 85: 365–371.

20 Heidelberger V, Ingen-Housz-Oro S, Marquet A, et al. Efficacy and tolerance of anti-tumor necrosis factor alpha agents in cutaneous sarcoidosis: a French study of 46 cases. JAMA Dermatol 2017; 153: 681–685.

21 Moller DR, Rybicki BA, Hamzeh NY, et al. Genetic, immunologic, and environmental basis of sarcoidosis. Ann Am Thorac Soc 2017; 14: Suppl. 6, S429–S436.