Characterization of Hidradenitis Suppurativa Phenotypes: A Multidimensional Latent Class Analysis of the National Italian Registry IRHIS.

Simone Cazzaniga, Elena Pezzolo, Vincenzo Bettoli, Damiano Abeni, Angelo V. Marzano, Annalisa Patrizi, Laura Atzori, Gabriella Fabbrocini, Annamaria Offidani, Claudia Lasagni, Valentina Dini, Marzia Caproni, Serafinella Patrizia Cannavò, Luigi Naldi

PII: S0022-202X(20)32184-9

DOI: https://doi.org/10.1016/j.jid.2020.08.032

Reference: JID 2666

To appear in: *The Journal of Investigative Dermatology*

Received Date: 30 June 2020

Revised Date: 29 July 2020

Accepted Date: 21 August 2020

Please cite this article as: Cazzaniga S, Pezzolo E, Bettoli V, Abeni D, Marzano AV, Patrizi A, Atzori L, Fabbrocini G, Offidani A, Lasagni C, Dini V, Caproni M, Cannavò SP, Naldi L, Characterization of Hidradenitis Suppurativa Phenotypes: A Multidimensional Latent Class Analysis of the National Italian Registry IRHIS., *The Journal of Investigative Dermatology* (2020), doi: https://doi.org/10.1016/j.jid.2020.08.032.

Authors' ORCID:

Simone Cazzaniga: https://orcid.org/0000-0001-8161-6138

Elena Pezzolo: https://orcid.org/0000-0003-0121-4812

Vincenzo Bettoli: https://orcid.org/0000-0002-2760-4600

Damiano Abeni: https://orcid.org/0000-0002-0167-7617

Angelo V. Marzano: https://orcid.org/0000-0002-8160-4169

Annalisa Patrizi: https://orcid.org/0000-0002-8398-0483

Laura Atzori: https://orcid.org/0000-0002-8105-1402

Gabriella Fabbrocini: https://orcid.org/0000-0002-0064-1874

Annamaria Offidani: https://orcid.org/0000-0001-5445-1200

Claudia Lasagni: https://orcid.org/0000-0002-0370-6602

Valentina Dini: https://orcid.org/0000-0002-8537-1999

Marzia Caproni: https://orcid.org/0000-0002-7471-3471

Serafinella Patrizia Cannavò: https://orcid.org/0000-0001-5459-3129

Luigi Naldi: https://orcid.org/0000-0002-3160-2835

Characterization of Hidradenitis Suppurativa Phenotypes: A Multidimensional Latent Class Analysis of the National Italian Registry IRHIS.

Simone Cazzaniga[1,2], Elena Pezzolo[1,3], Vincenzo Bettoli[4], Damiano Abeni[5], Angelo V. Marzano[6,7], Annalisa Patrizi[8], Laura Atzori[9], Gabriella Fabbrocini[10], Annamaria Offidani[11], Claudia Lasagni[12], Valentina Dini[13], Marzia Caproni[14], Serafinella Patrizia Cannavò[15], Luigi Naldi[1,3]

Author affiliations:

[1]Study Center of the Italian Group for Epidemiologic Research in Dermatology (GISED), Bergamo, Italy

[2]Department of Dermatology, Inselspital University Hospital of Bern, Bern, Switzerland

[3]Department of Dermatology, San Bortolo Hospital, Vicenza, Italy

[4]Dermatology Unit, Department of Oncology and Specialistic Medicines, Azienda Ospedaliera University of Ferrara, Italy

[5]Clinical Epidemiology Unit, IDI-IRCCS, Rome, Italy

[6]Dermatology Unit, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy

[7]Department of Physiopathology and Transplantation, Università degli Studi di Milano, Milan, Italy

[8]Department of Experimental, Diagnostic and Specialty Medicine, University of Bologna, Bologna, Italy

[9]Dermatology Clinic, Department of Medical Sciences and Public Health, University of Cagliari, Cagliari, Italy

[10]Dermatology Unit, Department of Clinical Medicine and Surgery, University of Federico II, Naples, Italy

[11]Dermatology Unit, Department of Clinical and Molecular Sciences, United Hospital of Ancona, Polytechnic University of Marche Region, Ancona, Italy

[12]Dermatology Unit, Surgical, Medical and Dental Department of Morphological Sciences related to Transplant, Oncology and Regenerative Medicine, University of Modena and Reggio Emilia, Modena, Italy

[13]Department of Dermatology, University of Pisa, Pisa, Italy

[14]Section of Dermatology, Department of Health Sciences, University of Florence, Florence, Italy

[15]Dermatology Unit, Department of Clinical and Experimental Medicine, Dermatology, University of Messina, Messina, Italy

Word count: Abstract: 202; Text: 2,714; Tables and/or Figures: 2 (2 in the Supplementary material); References: 30

Corresponding author:

Professor Luigi Naldi

Study Center of the Italian Group for Epidemiological Research in Dermatology (GISED)

Via Clara Maffei 4, 24121 Bergamo (Italy),

Tel. +39035223753

luiginaldibg@gmail.com

Abbreviations: HS, hidradenitis suppurativa; LCA, Latent Class Analysis; BIC, Bayesian Information Criterion; AIC, Akaike Information Criterion; IRHIS, Italian Registry of Hidradenitis Suppurativa; VAS, visual analogue scale; DLQI, Dermatology Life Quality Index; OR, odds ratios; CI, confidence intervals.

ABSTRACT

(Current word count is 201 words; max word count is 200 words)

In spite of the large heterogeneity, limited data exist on hidradenitis suppurativa (HS) phenotypes. To identify the HS phenotypes that best explain the disease heterogeneity, a cross-sectional study using Latent Class Analysis (LCA) was conducted on a cohort of patients examined at 17 dermatological centers participating in the Italian Registry of HS (IRHIS), and being enrolled between January 2015 and January 2020. Overall 965 patients aged 32.0±12.4 years (mean±SD) were evaluated. A three-class model in LCA best fitted the data. Patients in latent class 1 (LC1) (20.1%) were females, mostly obese, with a high probability of axillary-groin (0.85) and mammary (0.59) lesions, and the highest HS severity. LC2 patients (29.6%) were nonobese males, with moderate disease severity, a high probability of gluteal (0.50) and genital (0.17) lesions, besides axillary-groin involvement, and with acne and pilonidal cysts. LC3 patients (50%) were nonobese females with a milder disease mostly limited to axillary (0.52) and groin areas (0.66). The stratification of HS patients into a severe “axillary-mammary-groin” phenotype with predominantly anterior-body involvement in females, an “axillary-gluteal-groin” phenotype of intermediate severity mainly affecting males in the posterior-body areas, and an “axillary-groin” phenotype with mildest clinical symptoms and limited skin involvement may help optimizing HS management.

(Current word count is 2,714 words; max word count is 3,500 words)

INTRODUCTION

Hidradenitis Suppurativa (HS) is a chronic inflammatory skin disease characterized by suppurating painful lesions such as nodules and abscesses, sinus tracts, and scars in the intertriginous areas, which significantly impact patient's quality of life (QoL) (Jemec, 2012; Kontis *et al.*, 2017). The condition occurs more frequently in women, with a disease onset in young and middle-age adults (Garg *et al.*, 2017). The prevalence of HS has been estimated as less than 1% in the general population (Bettoli *et al.*, 2016). There is a broad spectrum of clinical presentation and of severity of HS (Jemec, 2012). The lack of knowledge on the heterogeneity of lesion appearance and sites of involvement may delay the diagnosis and the proper management of the condition (Sartorius *et al.*, 2010).

Latent class analysis (LCA) is a statistical way of grouping together individuals of a population of interest into sets of "clusters". These clusters are subgroups of individuals sharing similar characteristics with a specific probability of occurrence (Goodman, 1974). Compared to other methods of data segmentation used when dealing with categorical variables, LCA has the advantage of helping create a probabilistic multivariate model to estimate parameters. In other words, taking a number of pre-selected multifactorial variables, such as demographics, medical history, or other characteristics of the disease of interest, LCA considers the optimal number of clusters as the one that minimizes the degree of relationship among cases belonging to the different clusters. To decide on the optimal number of clusters two methods dealing with the goodness of fit of a statistical model are usually adopted, the Bayesian Information Criterion (BIC) and the Akaike Information Criterion (AIC).

So far, the use of LCA in dermatology has been limited compared to other areas of application (Canoui-Poitrine *et al.*, 2013; Silverberg *et al.*, 2015).

Insights into the phenotypes of HS may provide novel data on its etiology, class changes over time, prognosis and effective treatments.

The aim of this study was to identify, by using LCA, the underlying HS phenotypes that best explain the observed heterogeneity within the national Italian Registry of HS (IRHIS).

RESULTS

Study population

Overall, we enrolled 1,064 consecutive newly diagnosed HS patients. 99 patients (9.3%) with missing data in any of the variables considered for LCA were excluded from the final analysis, which comprised 965 patients.

The main characteristics of the patients included in the analysis are reported in the Table. Almost two third (62.3%) were females, the mean age was 32.0 ± 12.4, the mean body mass index (BMI) was 27.0 ± 6.0; 60.8% were current smokers and 7% were past smokers. The mean age at HS onset was 21.8 ± 10.4 years, and the mean HS duration was 10.2 ± 9.2 years; 20.8% of patients reported a history of HS in first-degree relatives. The mean Sartorius score was 52.8 ± 40.8. About 48.0% of patients were Hurley stage II and 17.8% Hurley stage III. The mean Dermatology Life Quality Index (DLQI) was 13.1 ± 7.8, with 58.5% patients reporting a large impact of HS on their QoL. The most frequent HS-associated lesions were pilonidal cyst (11.4%), acne conglobata (10.1%) and other types of acne (3.1%), while the most frequent HS-related comorbidities included psoriasis (1.8%) and thyroid diseases (1.6%).

Latent Class Analysis classification

The model selection procedure with one to five classes LCA is represented in Table 1 of the Supplementary materials. A three-class model best fitted the data in our LCA, showing the lowest BIC. The estimated class-conditional probabilities of the 11 pre-selected indicators of these three

LCs, including sex, age, BMI, smoking habits, educational attainment, age at HS onset, HS duration, diagnostic delay, Sartorius score, Hurley stage, associated lesion types and related comorbidities, are presented in Table 2 of Supplementary materials. The three corresponding phenotypes of LCs are shown in Figure 1. Based on LCA model estimates, 20.1% of patients belonged to LC1, 29.6% of patients to LC2, and 50.3% of patients to LC3. Compared to other classes, LC1 patients had higher probabilities for axillary and groin (0.85), and mammary involvement (0.59), a positive family history (0.30), and excessive sweating (0.64). LC2 patients had high probabilities for gluteal (0.50) and genital lesions (0.17), besides involvement of axillary (0.63) and groin (0.69) areas, and also involvement of neck, ears, chest, back and legs. These patients had especially high probabilities for follicular lesions, such as pilonidal cysts (0.18), and history of acne conglobata (0.21). In LC3 patients there was an almost exclusive localization on axillae (0.52) and groin areas (0.66) with an overall milder disease. We named the three phenotypes of HS as follows: LC1, “axillary-mammary-groin”; LC2, “axillary-gluteal-groin”; and LC3, “regular axillary-groin”

Patients’ and disease characteristics across the latent classes

We found significant differences with regard to sex, BMI, smoking status, age at disease onset, HS duration, diagnostic delay, and severity scores, when comparing patients and HS characteristics between the three LCs classified based on posterior predicted probabilities (Table).

We used the “regular axillary-groin” phenotype (LC3) as the reference category. Compared to LC3, the “axillary-mammary-groin” phenotype (LC1) was characterized by a higher proportion of females and smokers, a more severe disease (according to both Sartorius and Hurley), a larger proportion of obese patients, earlier disease onset, and longer disease duration. Patients in the “axillary-gluteal-groin” phenotype (LC2) were more often males and smokers, normal/over-weight patients, and had a moderate-to-high disease severity, with quite early disease onset, and long disease duration. The “regular axillary-groin” phenotype (LC3) had a higher proportion of females,

normal/over-weight patients, and had the lowest severity scores compared to the other two phenotypes, with a later disease onset and a shorter disease duration.

DISCUSSION

By relying on LCA, we identified three clinical subtypes of HS at first clinical presentation. One subtype (LC1), was prevalent in obese women and was characterized by a high prevalence of axillary-mammary and groin involvement in a severe form, a second subtype (LC3) was also more prevalent in women and was characterised by an involvement mostly limited to the axillary-groin areas in a milder form. These subtypes correspond to the typical varieties of HS. The remaining clinical subtype (LC2) was characterized by "gluteal and genital" involvement, besides axillary-groin involvement, and accounted for approximately 30% of patients, corresponding to the atypical variety of HS (Jemec, 2012).

A limited number of studies have tried to define clinical-pathological subtypes of HS. To the best of our knowledge, only one cross-sectional French study used LCA to identify these subtypes (Canoui-Poitrine *et al.*, 2013). In a sample of 618 patients, the authors found three LCs subtypes: the "axillary-mammary", the "follicular", and the "gluteal" LC. The "axillary-mammary" LC accounted for about half study population, and was consistent with typical HS. The other half of the study population was divided into two atypical phenotypes: the "follicular" LC, associated with greater disease severity, and the "gluteal" LC, associated with decreased disease severity. Similarly to these results, we identified three LCs, but with a larger proportion of the typical HS varieties (corresponding to the "axillary-mammary-groin" and "regular axillary-groin" subtypes) and a lower proportion of the atypical varieties (corresponding to the "axillary-gluteal-groin" subtype). Compared with the patients enrolled in the above-mentioned study, our patients had more severe disease, with a higher Sartorius score and more advanced Hurley stage, and had an impaired QoL due to HS. These patients were more frequently current smokers, overweight or obese, and a greater proportion of them presented a family history of HS. Notably, compared with the French paper, we

collected a larger number of variables in our study, including for example educational attainment, delay of diagnosis, DLQI. We also considered as a location the genital areas, which were not mentioned in the French study. It is of interest that the proportion of involvement of the axillary-mammary areas in our study was similar to the one observed in the French study, where the probability of such an involvement was 0.74 in their so-called "axillary-mammary" and 0.96 in their "follicular" subtypes. Additional HS-related characteristics, which might influence phenotype classification, such as the presence of hypertrophic scars not considered as a component of the active inflammatory process or folliculitis, which are a rather unspecific feature, were not collected in our registry. It is of interest that follicular lesions such as "acne conglobata" and "pilonidal cysts" were associated with the axillary-gluteal-groin subtype in our study population, and that "papules and folliculitis" were observed with a similar probability (0.71) in the atypical HS varieties identified in the Canoui-Poitrine's study, namely the follicular and gluteal subtypes. Multiple clinically distinct HS phenotypes have been proposed based on less stringent criteria. Usually, five subtypes are recognised (Micheletti, 2014; Heid *et al.*, 2011; Shalom, 2017; Goldburg *et al.*, 2020). The most common is the "regular" subtype, which includes patients who fully comply with the diagnostic criteria for HS and lack any specific feature (van der Zee *et al.*, 2015). Accordingly, in half of our patients we identified the "regular axillary-groin" subtype, which was characterized by an unspecific localization of lesions, and a more benign course (late onset, shorter duration, mild symptoms). Less frequent are the "frictional furuncle" subtype, which features lesions in flexural sites especially in overweight patients, and the "conglobata" subtype, which consists of cysts and acne conglobata mostly on the face and trunk in normal-weight men (van der Zee *et al.*, 2015). In our study, these two subtypes are likely to coexist in a single "axillary-gluteal-groin" subtype with patients who are more often males and smokers, normal or over-weight, and present acne conglobata and pilonidal cysts in frictional sites as well in other areas, such as the face or trunk.

Results of our study are consistent with other previously reported findings. Females, who are usually affected more than twice as often as males (Garg A et al., 2017a; Garg A et al., 2017b),

manifest lesions in the anterior part of the body, including breast and axillae (Canoui-Poitrine et al., 2013; Jemec, 1988; Jansen, 2001). By contrast, the back of the body (Canoui-Poitrine *et al*., 2009), and the gluteal area ((Jemec, 2012; Canoui-Poitrine *et al*., 2013; Poli *et al*., 2010) are more frequently involved in males, who may also present pilonidal cyst (Canoui-Poitrine et al., 2009). In our study, we were able to separate these two distinct subsets represented by our LC1 or LC3 and LC2 respectively. In agreement with other studies (Saunte *et al*., 2015), we found a significant delay in the diagnosis of HS (mean >5 years) particularly in the severe "axillary-mammary-groin" subtype. In such a subtype, with a dramatically decreased QoL, obesity appears to increase the pro-inflammatory response (Delany *et al*., 2018; Kjaersgaard Andersen *et al*., 2018), so that people with the highest BMI had the worse disease severity (confirmed by the highest Sartorius score) (Theut Riis *et al.*, 2018).

Our study has various strengths. By using LCA we recognized phenotypes without *a priori* hypotheses. Indeed, compared with other methods of data segmentation, LCA represents a robust method for stratifying patients sharing similar characteristics, using posterior membership probabilities (Magidson *et al*., 2002). It also has the advantage of helping create a probabilistic multivariate model to estimate multiple parameters, including demographics and nonclinical variables (Magidson *et al.*, 2002). Moreover, our study sample was collected prospectively and only few patients had missing data. Given the large sample size we used, we were able to include multiple indicators and covariates in the analysis. Together with the high number of centers participating to the IRHIS registry, this may make our results more representative of the HS population.

There are some limitations to consider. The cross-sectional study design did not allow to make a validation of our classification scheme based on outcomes. Additional HS-related characteristics and comorbidities such as the presence of hypertrophic scars and folliculitis, comedones, papules, epidermal cysts and/or macrocysts, which might influence phenotypes classification, were not collected in the registry. Moreover, we did not perform a formal validation on an independent

dataset and we did not have enough statistical power to perform split validation in order to test clusters stability; therefore our results should be interpreted with caution. Finally, our registry is limited to tertiary-care outpatient HS cases seen in the Italian population, so our results may not be generalizable to other populations.

By using LCA we found three distinct phenotypes of HS. The "axillary-mammary-groin" phenotype has an anterior body area involvement, mainly in obese women, and carrying the most severe course. The "axillary-gluteal-groin" phenotype has a posterior body area involvement. It mainly affects men, current smokers and non-obese patients, and may present involvement of other sites, such as genitalia, face, trunk, with acne and pilonidal cysts. A third phenotype, the "regular axillary-groin" phenotype has the mildest clinical course and an almost exclusive involvement of the axillae and groin. Although these three phenotypes do not currently guide therapeutic decisions, consideration of them may help optimize disease management (Vekic *et al*., 2018). Future studies should investigate the validity of our findings and the significance of our phenotypes in terms of disease characteristics, genetic background and prospective outcome of treatments.

METHODS

This was a cross-sectional analysis of baseline data from a cohort of consecutive patients with a first clinical diagnosis of HS at a network of 17 Italian dermatological outpatient clinics, participating in the IRHIS registry (Bettoli *et al*, 2019). All consecutive patients with a first ever diagnosis of HS at the participating centers, were included in the study. Patients who were not able to comply with the registry procedures and follow-up requirements were excluded. All patients gave written informed consent before inclusion into the registry. The study was approved by the ethics committee of each participating center.

Data collection

Data were collected in the registry between January 2015 and January 2020 using a centralized electronic data collection form and included: demographics (age, sex, occupation, educational attainment), anthropometric measures (weight and height), smoking habits, clinical history of the disease including age at onset and at first HS diagnosis, family history of HS in first degree relatives, localizations at onset, clinical characteristics at entry into the registry and at regular follow-up intervals, including severity, QoL measure, localizations, presence of worsening factors, presence of comorbidities, previous and current therapies for HS prescribed for at least 1 month. The severity of HS was mainly assessed by using the Sartorius score (Sartorius *et al*., 2003). Harmonization exercises of clinical assessment were conducted among assessors and an online calculator was adopted in order to standardize Sartorius score reporting. Hurley severity stage (Hurley, 1996), and patient's pain assessment based on visual analogue scale (VAS) were also collected. Patients' QoL was evaluated by using the DLQI. DLQI values <6, between 6-10 and ≥11 were considered as no/limited, moderate and large impact on QoL, respectively (Hongbo *et al*., 2005).

Statistical analysis

For descriptive purposes continuous data were presented as means with standard deviations (SD), while categorical data as numbers with percentages. Continuous variables were also categorized by using clinically relevant cut-offs or tertiles of their distribution. LCA was conducted in order to find latent phenotypes in HS presentation. For LCA, the following clinical variables were selected: HS localization and type, main reported HS-related comorbidities (acne conglobata and pilonidal cyst), family history of HS, and presence of excessive sweating. Different LCA models were fitted by varying the number of possible latent classes (LCs). Log-likelihood, AIC and BIC were estimated for each model. The model with lowest BIC was considered as the most reliable statistically. In addition, the following covariates were included in the best-selected model: sex, BMI, Sartorius score, age at HS onset and HS duration since onset. Patients with missing data in any of the included variables were excluded from LCA. Outliers were also removed for better model fitting.

Estimated class population shares as well as predicted class membership based on modal posterior probabilities were also reported. Univariate differences among predicted LCA classes were assessed based on Pearson's X^2 test and Kruskal-Wallis test for categorical and continuous variables, respectively. Multivariate assessment was performed by using multinomial logistic regression including covariates. The strength of variables on each class membership was expressed in terms of odds ratios (OR) along with their 95% confidence intervals (CI) and p-values. All tests were considered statistically significant at P-value <0.05. Analyses were performed with SPSS v.26 (IBM Corp, Armonk, NY, US) and for LCA, R software v.3.6.2 (R Foundation for Statistical Computing, Vienna, Austria) with package poLCA v. 1.4.1.

DATA AVAILABILITY: Datasets related to this article can be found at https://data.mendeley.com/datasets/bv9smj9w99/1 hosted at Mendeley (Cazzaniga, Simone; Naldi, Luigi (2020), "Characterization of Hidradenitis Suppurativa Phenotypes", Mendeley Data, v1 http://dx.doi.org/10.17632/bv9smj9w99.1)

CONFLICT OF INTEREST: Dr. Offidani has been principal investigator in clinical trials and has been paid as a consultant by Abbvie, Almirall, Amgen, Celgene, Eli Lilly, Leo Pharma, Novartis, Regeneron, and Sanofi. None disclosed by the other authors.

ACKNOWLEDGMENTS

We thank Vito Di Lernia, Sandro Pasquinucci, Francesco Cusano, Calogero Pagliarello, Claudio Marasca, Andrea Sechi, Giulia Ruina, Roberta Giuffrida, Luca Fania, Roberta Fusari, Simona Mastroeni, and Davide Ciccone for their invaluable support to this study.

CRediT statement

Conceptualization: LN, VB, SC

Data curation: LN, SC

Formal analysis: SC, LN

Funding acquisition: No funding

Investigation: SC, LN

Methodology: SC, LN

Project administration: LN

Resources: LN, VB, DA, AVM, AP, LA, GF, AO, CL, VD, MC, SPC

Software: SC, LN

Supervision: LN

Validation: SC, LN

Visualization: EP, SC, LN, VB, DA, AVM, AP, LA, GF, AO, CL, VD, MC, SPC

Writing (original draft): EP, SC, LN

Writing (review and editing): EP, SC, LN

Figure 1. Estimated class conditional probability of HS subtype features identified by LCA.

Table. Uni- and multivariate comparisons of patients and disease characteristic among the three HS subtypes identified by LCA

			Predicted class membership**			P***	Multivariate analysis (LC3 ref.)^			
			LC1(axillary-mammary-groin)	LC2 (axillary-gluteal-groin)	LC3 (regular axillary-groin)		LC1		LC2	
		N*=965 (%)	N*=176 (%)	N*=246 (%)	N*=543 (%)		OR (95% CI)	P	OR (95% CI)	P
Sex	Male	364 (37.7)	6 (3.4)170 (96.6)	235 (95.5)	123 (22.7)	<0.001	1		1	
	Female	601 (62.3)		11 (4.5)	420 (77.3)		8.93 (2.26 - 35.26)	0.002	0.001 (0 - 0.003)	<0.001
Age (yrs)	*Mean (SD)*	*32.0 (12.4)*	*33.9 (12.7)*	*33.4 (11.9)*	*30.8 (12.4)*	0.001				
	< 25	*343 (35.5)*	53 (30.1)	66 (26.8)	224 (41.3)		1		1	
	25 - 39	*363 (37.6)*	68 (38.6)	109 (44.3)	186 (34.3)		1.33 (0.45 - 3.92)	0.61	1.56 (0.55 - 4.37)	0.40
	40+	*259 (26.8)*	55 (31.3)	71 (28.9)	133 (24.5)		1.09 (0.14 - 8.72)	0.93	1.59 (0.21 - 11.98)	0.65
BMI	*Mean (SD)*	*27.0 (6.0)*	*31.3 (6.7)*	*26.0 (4.4)* 113	*26.1 (5.8)*	<0.001				
(kg/m^2)	< 25.025.0 - 29.930.0+	*411 (42.6)*	28 (15.9)	(45.9)	270 (49.7)		1		1	
		301 (31.2)	51 (29.0)	99 (40.2)	151 (27.8)		2.55 (1.17 - 5.52)	0.02	0.62 (0.32 - 1.20)	0.15
		253 (26.2)	97 (55 .1)	34 (13.8)	122 (22.5)		13.62 (6.12 - 30.31)	<0.001	0.08 (0.03 - 0.20)	<0.001
Smoking habits	Never smoked	318 (33.0)	54 (30.7)	49 (19.9)	215 (39.6)	<0.001	1		1	
	Smoker	582 (60.3)	111 (63.1)	174 (70.7)	297 (54.7)		2.06 (1.04 - 4.08)	0.04	1.02 (0.50 - 2.08)	0.96
	Ex-smoker	65 (6.7)	11 (6.3)	23 (9.3)	31 (5.7)		2.45 (0.52 - 11.49)	0.25	1.86 (0.53 - 6.56)	0.33
Educational	Primary/Lower secondary	270 (28)	49 (27.8)	72 (29.3)	149 (27.4)	0.78	1		1	

			Predicted class membership**			P***	Multivariate analysis (LC3 ref.)^			
			LC1(axillary-mammary-groin)	LC2 (axillary-gluteal-groin)	LC3 (regular axillary-groin)		LC1		LC2	
		N*=965 (%)	N*=176 (%)	N*=246 (%)	N*=543 (%)		OR (95% CI)	P	OR (95% CI)	P
Attainment (school)	Higher secondary	515 (53.4)	93 (52.8)	135 (54.9)	287 (52.9)		1.42 (0.65 - 3.09)	0.37	0.80 (0.40 - 1.58)	0.52
	University	180 (18.7)	34 (19.3)	39 (15.9)	107 (19.7)		1.41 (0.54 - 3.63)	0.48	1.29 (0.52 - 3.17)	0.58
Age at HS onset (yrs)	*Mean (SD)*	*21.8 (10.4)*	*18.7 (9.2)*	*21.8 (9.4)*	*22.8 (11.0)*	<0.001				
	< 15.0	110 (11.4)	71 (40.3)	43 (17.5)	108 (19.9)		1		1	
	15.0 - 24.0	574 (59.4)	71 (40.3)	128 (52.0)	263 (48.4)		0.44 (0.22 - 0.88)	0.02	0.52 (0.22 - 1.21)	0.13
	25.0+	281 (29.1)	34 (19.3)	75 (30.5)	172 (31.7)		0.06 (0.02 - 0.16)	<0.001	0.44 (0.18 - 1.08)	0.07
HS duration since onset (yrs)	*Mean (SD)*	*10.2 (9.2)*	*15.2 (10.3)*	*11.6 (9.1)* 56	*8.0 (8.1)*	<0.001				
	< 5	316 (32.7)	14 (8.0)	(22.8)	246 (45.3)		1		1	
	5-9	257 (26.6)	52 (29.5)	73 (29.7)	132 (24.3)		1.98 (0.78 - 5.05)	0.15	1.58 (0.73 - 3.41)	0.24
	10+	342 (40.6)	110 (62.5)	117 (47.6)	165 (30.4)		4.15 (1.76 - 9.79)	0.001	3.88 (1.87 - 8.05)	<0.001
Diagnostic delay (yrs)	*Mean (SD)*	*6.7 (7.7)*	*9.1 (9.1)*	*7.5 (7.7)*	*5.6 (7.0)*	<0.001				
	< 2	257 (26.6)	29 (16.5)	57 (23.2)	171 (31.5)		1		1	
	2 - 4	246 (25.5)	28 (15.9)	55 (22.4)	163 (30.0)		0.60 (0.23 - 1.56)	0.30	0.58 (0.26 - 1.30)	0.18
	5+	462 (47.9)	119 (67.6)	134 (54.5)	209 (38.5)		0.61 (0.25 - 1.45)	0.26	0.69 (0.29 - 1.63)	0.40
Sartorius score	*Mean(SD)*	*52.8 (40.8)*	*93.0 (38.6)*	*77.0 (43.4)*	*28.9 (16.1)*	<0.001				
	< 30.0	325 (33.7)	2 (1.1)	11 (4.5)	312 (57.5)		1		1	
	30.0 - 59.0	331 (34.3)	29 (16.5)	99 (40.2)	203 (37.4)		26.63 (12.42 - 57.09)	<0.001	9.81 (3.39 - 28.37)	<0.001

			Predicted class membership**			P***	Multivariate analysis (LC3 ref.)^			
			LC1(axillary-mammary-groin)	LC2 (axillary-gluteal-groin)	LC3 (regular axillary-groin)		LC1		LC2	
		N*=965 (%)	N*=176 (%)	N*=246 (%)	N*=543 (%)		OR (95% CI)	P	OR (95% CI)	P
	60.0+	309 (32.0)	145 (82.4)	136 (55.3)	28 (5.2)		818.9 (253.3 - 2647.7)	<0.001	545.7 (168.4 - 1767.9)	<0.001
Hurley stage	I	330 (34.2)	19 (10.8)	45 (18.3)	266 (49.0)	<0.001	1°		1°	
	II	463 (48)	103 (58.5)	113 (45.9)	247 (45.5)		5.39 (3.06 - 9.49)	<0.001	3.07 (1.87 - 5.06)	<0.001
	III	172 (17.8)	54 (30.7)	88 (35.8)	30 (5.5)		23.33 (11.27 - 48.30)	<0.001	22.02 (10.26 - 47.25)	<0.001
DLQI	*Mean (SD)*	*13.1 (7.8)*	*15.2 (7.2)*	*13.4 (7.8)*	*12.4 (7.8)*	0.001				
	< 6 (no/small)	133 (19.7)	12 (9.8)	28 (17.6)	93 (23.7)		1°		1°	
	6 - 10 (moderate)	147 (21.8)	24 (19.7)	36 (22.6)	87 (22.1)		2.98 (1.22 - 7.27)	0.02	1.95 (0.94 - 4.06)	0.07
	>10 (large)	394 (58.5)	86 (70.5)	95 (59.7)	213 (54.2)		3.78 (1.73 - 8.25)	0.001	2.38 (1.28 - 4.41)	0.006

Abbreviations: BMI: body mass index; DLQI: dermatology life quality index; HS: hidradenitis suppurativa; LC: latent classes; SD: standard deviation

* Numbers may not add up to the total due to missing data

** Based on modal posterior probabilities

*** Pearson's X^2 test and Kruskal-Wallis test were used for categorical and continuous variables respectively

^ Multinomial logistic regression including the following variables: sex, BMI, Sartorius score, age at HS onset, HS duration

° Sartorius score was not included in the model due to collinearity issu

SUPPLEMENTARY MATERIALS

Table 1. Model selection with one to five classes of HS based on LCA

N of classes	Log-likelihood	AIC	BIC	N of estimated Parameters
1	-5415.0	10852.0	10906.6	11
2	-5363.9	10773.7	10888.0	23
3	**-5319.1**	**10708.1**	**10882.1**	**35**
4	-5285.9	10665.7	10899.3	47
5	-5262.0	10642.1	10935.3	59

Abbreviations: AIC: Akaike information criterion; BIC: Bayesian information criterion

Table 2. Results of LCA based on 3 classes of HS

	Prevalence of Indicators	LC1 (axillary-mammary-groin)	LC2 (axillary-gluteal-groin)	LC3 (regular axillary -groin)
Probability of class membership		0.20	0.29	0.50
Conditional probabilities of*				
Localizations				
Neck	4.7%	0.04	0.11	0.01
Axillae	61.8%	0.85	0.63	0.52
Mammary area	21.3%	0.59	0.06	0.15
Groin	70.7%	0.85	0.69	0.66
Genital area	10.5%	0.14	0.17	0.05
Buttocks	29.7%	0.42	0.50	0.13
Other sites	16.8%	0.20	0.28	0.09
Lesion characteristics				
Acne conglobate	10.1%	0.06	0.21	0.05
Pilonidal cysts	11.4%	0.09	0.18	0.08
Family history of HS	20.8%	0.30	0.23	0.16
Excessive sweating	56.8%	0.64	0.59	0.52

Abbreviations: HS: hidradenitis suppurativa; LC: latent classes.
*The following covariates were used in the model: Sartorius score, sex, body max index, age at disease onset, disease duration. Latent class analysis was estimated on 965 patients with complete data.

REFERENCES

Bettoli V, Pasquinucci S, Caracciolo S, Piccolo D, Cazzaniga S, Fantini F, et al. The Hidradenitis suppurativa patient journey in Italy: current status, unmet needs and opportunities. J Eur Acad Dermatol Venereol 2016;30:1965-705.

Bettoli V, Cazzaniga S, Scuderi V, Zedde P, Di Landro A, Naldi L, et al. Hidradenitis suppurativa epidemiology: from the first Italian registry in 2009 to the most recent epidemiology updates - Italian Registry Hidradenitis Suppurativa project 2. J Eur Acad Dermatol Venereol 2019;33:4-6.

Canoui-Poitrine F, Revuz JE, Wolkenstein P, Viallette C, Gabison G, Pouget F, et al. Clinical characteristics of a series of 302 French patients with hidradenitis suppurativa, with an analysis of factors associated with disease severity. J Am Acad Dermatol 2009;61:51-7.

Canoui-Poitrine F, Le Thuaut A, Revuz JE, Viallette C, Gabison G, Poli F, et al. Identification of three hidradenitis suppurativa phenotypes: latent class analysis of a cross-sectional study. J Invest Dermatol 2013;133:1506-11.

Delany E, Gormley G, Hughes R, McCarthy S, Kirthi S, Markham T, et al. A cross-sectional epidemiological study of hidradenitis suppurativa in an Irish population (SHIP). J Eur Acad Dermatol Venereol 2018;3:467-73.

Garg A, Kirby JS, Lavian J, Lin G, Strunk A. Sex- and Age-Adjusted Population Analysis of Prevalence Estimates for Hidradenitis Suppurativa in the United States. JAMA Dermatol. 2017a;153:760-4.

Garg A, Lavian J, Lin G, Strunk A, Alloo A. Incidence of hidradenitis suppurativa in the United States: A sex- and age-adjusted population analysis. J Am Acad Dermatol 2017b;77:118-22.

Goldburg SR, Strober BE, J Payette MJ. Hidradenitis Suppurativa: Epidemiology, clinical presentation, and pathogenesis. J Am Acad Dermatol 2020;82:1045-58.

Goodman LA. Exploratory Latent Structure Analysis Using Both Identifiable and Unidentifiable Models. Biometrika 1974;61:215-31.

Heid E, Chartier C. Hidradenitis suppurativa (Verneuil's disease). Ann Dermatol Venereol 2011;128:158-60.

Hongbo Y, Thomas CL, Harrison MA, Salek MS, Finlay AY. Translating the science of quality of life into practice: what do Dermatology Life Quality Index scores mean? J Invest Dermatol 2005;125:659-64.

Hurley HJ. Axillary hyperhidrosis, apocrine bromhidrosis, hidradenitis suppurativa, and familial benign pemphigus: surgical approach. In: Roenigk RK, Roenigk HH Jr, eds. Roenigk and Roenigk's Dermatologic Surgery: Principles and Practice. 2nd ed. New York, NY: Marcel Dekker. 1996;623-645.

Jansen I, Altmeyer P, Plewig G. Acne inversa (alias hidradenitis suppurativa). J Eur Acad Dermatol Venereol 2001;15:532-40.

Jemec GB. The symptomatology of hidradenitis suppurativa in women. Br J Dermatol 1988;119:345-30.

Jemec GB. Clinical Practice. Hidradenitis suppurativa. N Engl J Med 2012;366:158-64.

Kjaersgaard Andersen R, Theut Riis P, Jemec GBE. Factors predicting the self-evaluated health of hidradenitis suppurativa patients recruited from an outpatient clinic. J Eur Acad Dermatol Venereol 2018;32:313-7.

Kontis V, Bennett JE, Mathers CD, Li G, Foreman K, Ezzati M. Future Life Expectancy in 35 Industrialised Countries: Projections With a Bayesian Model Ensemble. Lancet 2017;389:1323–35.

Magidson J, Vermunt JK. Latent class models for clustering: a comparison with K-means. Can J Mark Re 2002;20: 37-44.

Micheletti RG. Natural history, presentation, and diagnosis of hidradenitis suppurativa. Semin Cutan Med Surg 2014;33:S51-3.

Poli F, Wolkenstein P, Revuz J. (2010) Back and face involvement in hidradenitis suppurativa. Dermatology 221: 137-41.

Sartorius K, Lapins J, Emtestam L, Jemec GBE. (2003) Suggestions for uniform outcome variables when reporting treatment effects in hidradenitis suppurativa. Br J Dermatol 149: 211–13.

Sartorius K, Killasli H, Heilborn J, Jemec GBE, Lapins J, Emtestam L. (2010) Interobserver Variability of Clinical Scores in Hidradenitis Suppurativa Is Low. Br J Dermatol 162:1261-8.

Saunte DM, Boer J, Stratigos A, Szepietowski JC, Hamzavi I, Kim KH, et al. (2015) Diagnostic delay in hidradenitis suppurativa is a global problem. Br J Dermatol 173:1546-9.

Shalom G. (2017) Hidradenitis suppurativa: epidemiology, clinical features, associated comorbidities and treatment. G Ital Dermatol Venereol 152:46-57.

Silverberg JI, Garg NK, Paller AS, Fishbein AB, Zee PC. (2015) Sleep Disturbances in Adults With Eczema Are Associated With Impaired Overall Health: A US Population-based Study. J Invest Dermatol 135:56-66.

Theut Riis P, Saunte DM, Benhadou F, Del Marmol V, Guillem P, El-Domyati M et al. (2018) Low and high body mass index in hidradenitis suppurativa patients-different subtypes? J Eur Acad Dermatol Venereol 32:307-12.

van der Zee HH, Jemec GB. (2015) New insights into the diagnosis of hidradenitis suppurativa: Clinical presentations and phenotypes. J Am Acad Dermatol 73:S23-26.

Vekic DA, Frew J, Cains GD. (2018) Hidradenitis suppurativa, a review of pathogenesis, associations and management. Part 1. Australas J Dermatol 59:267-77.

Table S1. Model selection with one to five classes of HS based on LCA

N of classes	Log-likelihood	AIC	BIC	N of estimated Parameters
1	-5415.0	10852.0	10906.6	11
2	-5363.9	10773.7	10888.0	23
3	**-5319.1**	**10708.1**	**10882.1**	**35**
4	-5285.9	10665.7	10899.3	47
5	-5262.0	10642.1	10935.3	59

Abbreviations: AIC: Akaike information criterion; BIC: Bayesian information criterion

Table S2. Results of LCA based on 3 classes of HS

	Prevalence of Indicators	**LC1 (axillary-mammary-groin)**	**LC2 (axillary-gluteal-groin)**	**LC3 (regular axillary -groin)**
Probability of class membership		0.20	0.29	0.50
Conditional probabilities of*				
Localizations				
Neck	4.7%	0.04	0.11	0.01
Axillae	61.8%	0.85	0.63	0.52
Mammary area	21.3%	0.59	0.06	0.15
Groin	70.7%	0.85	0.69	0.66
Genital area	10.5%	0.14	0.17	0.05
Buttocks	29.7%	0.42	0.50	0.13
Other sites	16.8%	0.20	0.28	0.09
Lesion characteristics				
Acne conglobate	10.1%	0.06	0.21	0.05
Pilonidal cysts	11.4%	0.09	0.18	0.08
Family history of HS	20.8%	0.30	0.23	0.16
Excessive sweating	56.8%	0.64	0.59	0.52

Abbreviations: HS: hidradenitis suppurativa; LC: latent classes.

*The following covariates were used in the model: Sartorius score, sex, body max index, age at disease onset, disease duration. Latent class analysis was estimated on 965 patients with complete data.