Polygenic Score: A Tool for Evaluating the Genetic Background of Sporadic Hidradenitis Suppurativa

JIDOpen

Chiara Moltrasio[1,18], Ronald Moura[2,18], Andrea Conti[3], Luca Fania[4], Wolfram Jaschke[5], Raffaele Dante Caposiena Caro[6], Karin Chersi[6], Flavia Manzo Margiotta[7], Antonella Di Cesare[8], Elia Rosi[8], Florian Regensberger[5], Barbara Boeckle[5], Nina Frischhut[5], Stefania Cappellani[2], Cecilia Del Vecchio[9], Elena Maria Nardacchione[9], Iris Zalaudek[6,9], Esther von Stebut[10], Irene Berti[11], Michele Boniotto[12], Adamo Pio d'Adamo[2,9], Matthias Schmuth[5], Valentina Dini[7], Francesca Prignano[8], Damiano Abeni[13], Andrea Chiricozzi[14,15], Angelo Valerio Marzano[1,16], Sergio Crovella[17,18] and Paola Maura Tricarico[11,18]

Journal of Investigative Dermatology (2025) ■, ■—■; doi:10.1016/j.jid.2024.11.019

Sporadic hidradenitis suppurativa (spHS) is a multifactorial disease in which genetic predisposition is intertwined with environmental factors. Owing to the still-to-date limited knowledge of spHS genetics, we calculated polygenic scores (PGSs) to study the genetic underpinnings that contribute to spHS within European demographic. A total of 256 patients with spHS and 1686 healthy controls were analyzed across 6 European clinical centers. PGSs were calculated using a clumping and thresholding technique on 70% of the total sample, with the remaining 30% used for testing. The PANTHER tool was used to identify overrepresented genes. We generated a PGS characterized by 923 SNPs with a statistically significant association with spHS ($P = 2 \times 10^{-2}$). The statistically significant age-, sex-, and ancestry-adjusted association of our developed PGSs in spHS allows us to attribute a genetic contribution to the susceptibility of spHS (pseudo-R2 = 0.0053). Variants enriched for developing PGSs show a statistically significant preference for mapping to genes that encode primarily for cell adhesion proteins. Although this study developed a polygenic model associated with spHS, the low number of patients enrolled is a limitation. However, we believe that with larger experimental datasets, our model has the potential to serve as a valuable tool for predicting spHS states in future studies.

Keywords: Disease association, Genetics, Polygenic score, Sporadic hidradenitis suppurativa, Susceptibility

INTRODUCTION

Sporadic hidradenitis suppurativa (spHS) is a chronic, autoinflammatory, and debilitating skin disorder clinically characterized by recurrent painful nodules, abscesses, and pus-draining tunnels often accompanied by extensive scarring in apocrine gland—bearing skin (van Straalen et al, 2022). It is characterized by a multifactorial etiology that intertwines genetic predisposition with environmental factors, thereby modulating disease susceptibility, onset, endotype/phenotype, and response to treatment. The prevalence in Europe stands at 0.8%, with a female—male ratio of 3:1 and with patients experiencing a substantial diagnostic delay averaging between 7 and 12 years from the first symptoms to correct diagnosis (Saunte et al, 2015). This delay significantly hinders timely intervention, which is essential for curbing disease progression and for initiating appropriate therapeutic

[1]*Dermatology Unit, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy;* [2]*Department of Advanced Diagnostics, Institute for Maternal and Child Health - IRCCS Burlo Garofolo, Trieste, Italy;* [3]*SS Endocrinologia, auxologia e diabetologia, Institute for Maternal and Child Health - IRCCS Burlo Garofolo, Trieste, Italy;* [4]*Dermatology Unit, IDI-IRCSS, Rome, Italy;* [5]*Department of Dermatology, Venereology and Allergy, Medical University of Innsbruck, Innsbruck, Austria;* [6]*Dermatology Clinic, Hospital Clinics Giuliano Isontino (ASUGI), University of Trieste, Trieste, Italy;* [7]*Department of Dermatology, University of Pisa, Pisa, Italy;* [8]*Section of Dermatology, Department of Health Science, University of Florence, Florence, Italy;* [9]*Department of Medical Surgical and Health Sciences, University of Trieste, Trieste, Italy;* [10]*Department of Dermatology, University of Cologne, Cologne, Germany;* [11]*Pediatric Department, Institute of Maternal and Child Health - IRCCS Burlo Garofolo, Trieste, Italy;* [12]*INSERM, IMRB, Translational Neuropsychiatry, University Paris Est Créteil, Créteil, France;* [13]*Clinical Epidemiology Unit, Istituto Dermopatico dell'Immacolata, IDI-IRCCS, Rome, Italy;* [14]*Dermatologia, Dipartimento Universitario di Medicina e Chirurgia Traslazionale, Università Cattolica Del Sacro Cuore, Rome, Italy;* [15]*Dermatologia, Dipartimento di Scienze Mediche e Chirurgiche, Fondazione Policlinico Universitario A. Gemelli IRCCS, Rome, Italy;* [16]*Department of Pathophysiology and Transplantation, Università degli Studi di Milano, Milan, Italy; and* [17]*Laboratory of Animal Research Center (LARC), Qatar University, Doha, Qatar*

[18]*These authors contributed equally to this work.*

Correspondence: Paola Maura Tricarico, Pediatric Department, Institute of Maternal and Child Health - IRCCS Burlo Garofolo, Via dell'Istria 65/1, Trieste 34137, Italy. E-mail: tricaricopa@gmail.com

Abbreviations: HS, hidradenitis suppurativa; IQR, interquartile range; PGS, polygenic score; PRS, polygenic risk score; QC, quality control; spHS, sporadic hidradenitis suppurativa

Received 25 June 2024; revised 20 November 2024; accepted 25 November 2024; accepted manuscript published online XXX; corrected proof published online XXX

measures. Hidradenitis suppurativa (HS) predominantly manifests as a sporadic skin disorder; however, in around 8% of cases, it occurs as a familial disease with an autosomal dominant inheritance and low penetrance—mostly related to gamma secretase complex genes—and more rarely, it is part of a broader syndrome with different immune/inflammatory-associated conditions, further complicating the disease spectrum (Maronese et al, 2024; Moltrasio et al, 2022).

The multifactorial nature of spHS suggests that its variability arises from a combination of genetic factors—both independent and epistatic—along with environmental influences and gene–environment interactions (van Straalen et al, 2022). A polygenic model could offer a compelling explanation for the observed differences in disease susceptibility, severity, age of onset, and treatment response, reflecting the intricate genetic landscape underlying spHS.

Several GWASs on HS have been conducted, described in the article by Jamec et al (2022), which highlights the importance of registries and multicenter GWAS efforts in advancing our understanding of HS genetics. Notably, the study by Sun et al (2023) stands out as the only one listed in the GWAS catalog that is directly relevant to polygenic risk score (PRS) calculations. In their multicohort study, Sun et al (2023) identified 10 loci associated with HS across diverse populations, making it a reference for our analysis.

Polygenic scores (PGSs) have a different objective and strategies with respect to GWASs, aimed at identifying novel genes associated with the disease, because it combines the effects of SNPs in a single number that predicts the genetic predisposition for a disease. Considering that in a polygenic disorder, a single variant is not informative for assessing disease risk, PGS has been widely applied in polygenic diseases; in fact, PGS is an important tool with the potential to enhance disease risk prediction and predict progression and recurrence of disease (Polygenic Risk Score Task Force of the International Common Disease Alliance, 2021).

To date, owing to the still-limited knowledge of spHS genetics, the disease falls into the low polygenic diseases class (Pace et al, 2022); moreover, a PGS that can contribute to a more accurate disease risk prediction is still not available. The early disease detection is crucial in managing spHS because it could allow for the implementation of targeted treatments or even measures that could potentially prevent the condition from progressing into more severe stages.

All this considered, the primary objective of our research is to investigate the genetic factors that contribute to the occurrence and risk of developing spHS within the European population. To analyze genetic predisposition and assess the combined influence of multiple common variants on spHS susceptibility, we calculated a PGS for our spHS cohort, using data from a large GWAS study (Sun et al, 2023).

This approach, which offers a quantitative measure of an individual's genetic risk modulation on the basis of the specific set of genetic variants they carry, could help elucidate the complex genetic basis of this disease.

RESULTS

Sample characteristics

Among the 256 patients with spHS who passed the quality control (QC) in our experimental dataset, there were 147 females and 109 males, with a female–male ratio of 1.3:1 (Figure 1a).

The median age of disease onset was 20 (interquartile range [IQR] = 14) years, whereas the median age at diagnosis was 29 (IQR = 19) years (Figure 1b) with a diagnostic delay of 4 (IQR = 10) years. In more detail, we found that 36.7% of all patients experienced early disease onset (n = 94 patients), defined as an onset that occurs before age 18 years (15; IQR = 3 years), whereas the remaining 162 patients experienced late disease onset (26; IQR = 15 years) (Figure 1c). In particular, females presented a median age of onset of 18 years (IQR = 11.5 years), significantly lower than the males' onset age of 20 years (IQR = 13 years) (P = .0012) (Figure 1d).

Forty-one patients (16%) were classified as Hurley stage I with a median International Hidradenitis Suppurativa Severity Score System = 2; 133 patients (52%) were classified as Hurley II (median International Hidradenitis Suppurativa Severity Score System = 7), whereas 82 patients (32%) were classified as Hurley III (median International Hidradenitis Suppurativa Severity Score System = 14). Associated conditions were reported in 142 (55.5%) patients, the most common being acne (8.1%) and class I obesity (3.6%). Body mass index was 26.9 (±5.31) for all patients, being 27.1 (±5.51) among the males and 26.8 (±5.17) among females. Finally, 95 patients (37.1%) were smokers, 20 (7.8%) were exsmokers, and 70 (27.3%) never smoked. For 71 patients (27.7%), these data were not available.

The control group consisted of 682 female and 712 male clinically healthy individuals without known history of spHS. The total median age was 39 years (q1 = 34, q3 = 44 years), being the median 38 (q1 = 33, q3 = 42) years for the females and 41 (q1 = 35; q3 = 42) years for the males.

PGSs

The clumping and thresholding technique, implemented using the command-line PLINK software, version 1.9 (Chang et al, 2015), generated multiple PGSs on the basis of 8 different P-value thresholds, which reflect the strength of association between single nucleotide variants and the HS phenotype in selected GWASs. After validation, the PGS with the lowest Akaike Information Criterion value in a logistic regression model—accounting for the effects of sex, age, and ancestry—was selected. This PGS, consisting of 923 SNPs and a P-value threshold of 1.0×10^{-3}, showed a statistically significant association with spHS, adjusted for age, sex, and ancestry ($P = 2.0 \times 10^{-2}$) (Table 1). To avoid biased estimation of phenotypic variance, we used pseudo-R2 on the basis of the area under the curve (https://doi.org/10.1002/gepi.21614) as a metric. The pseudo-R2 for our best PGS model was 0.5%, whereas the null model (which included all covariates except the PGS) had a pseudo-R2 of 3.8%. There was no sign of overfitting in the PGS model, as indicated by the small SD (0.3) from the mean accuracy (84.14%) obtained through 10-fold cross-validation. The receivers' operating characteristic curve analysis of the raw clumping and thresholding PGS, based on a P-value threshold of 1.0×10^{-3}, showed an area under the curve of 53.9%. After adjusting for covariates, the area under the curve increased to 82.0% (Figure 2b), whereas the corresponding null model

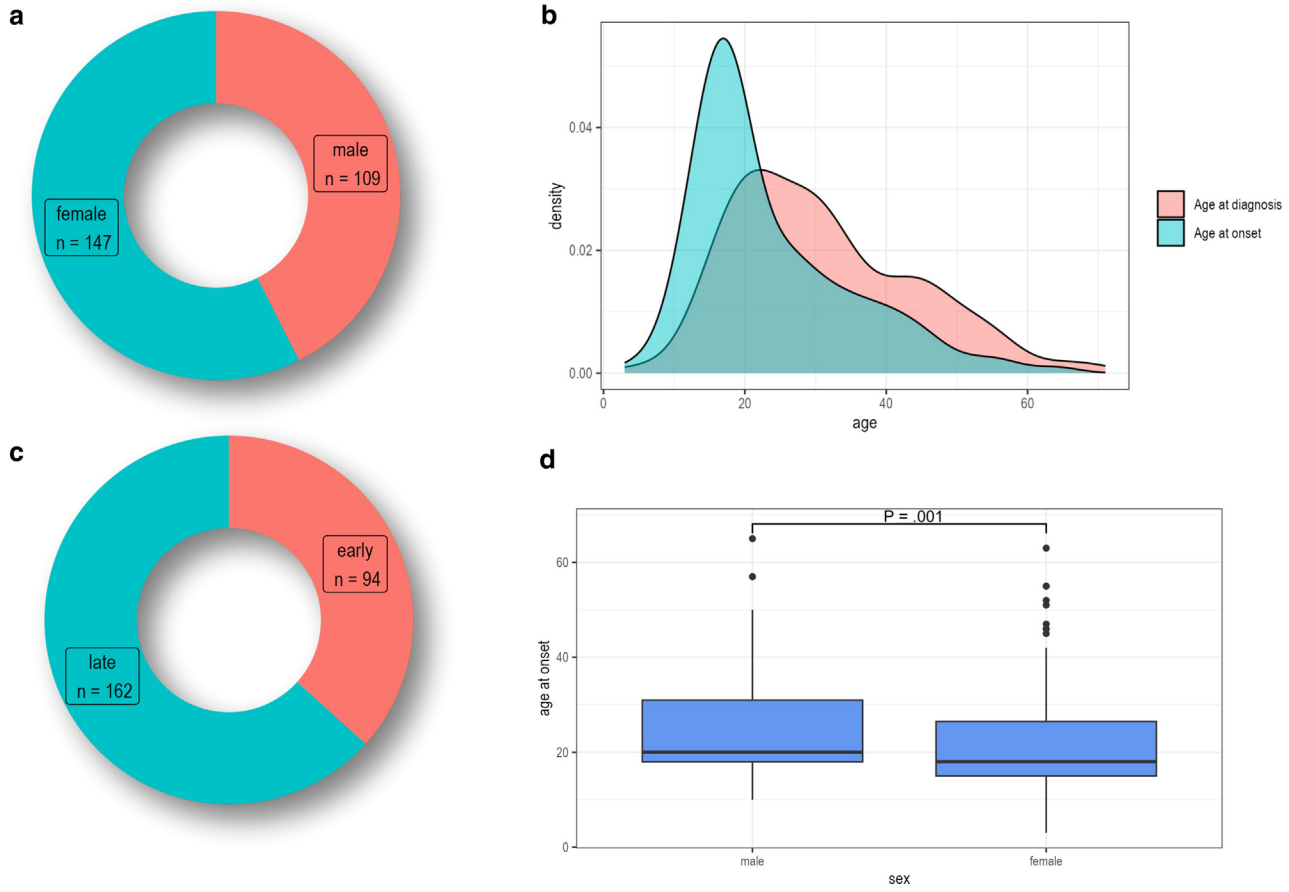

Figure 1. Descriptive analysis of patient data. (**a**) A donut chart of the count of patients by sex. (**b**) A density plot with the distribution of age of diagnosis and age at onset for each patient. (**c**) A donut chart of the count of patients by onset category. (**d**) A boxplot with the *P*-value of the Wilcoxon test to evaluate the significance of age at onset according to sex.

(which included all covariates except the PGS) show an area under the curve value of 81.6% (Figure 2a). Our data showed a difference, although not significant, between patients with early-onset spHS and those with late-onset spHS (early onset: 0.187 ± 0.140, late onset: 0.314 ± 0.204; $P = .0129$) (Figure 3a); moreover, the genetic susceptibility of female patients with spHS was not significantly higher than that of male patients (female: 0.296 ± 0.191, male: 0.195 ± 0.165; $P = .0842$) (Figure 3b).

Table 1. Logistic Regression Model Used to Evaluate the Association between PGS and spHS Disease State

Coefficient Name	β Value	*P*-Value
Intercept	−9.632 1.986	.13690 $5.2\ 10^{-1}$
Polygenic score	3.698e+02 3.28×10^{-1}	.02101 <.0001
Female sex	5.485e-01 4.55×10^{-1}	.05794 3.2×10^{-3}
Age, y	−3.217e-04 -1.99×10^{-2}	.97951 3.98×10^{-3}
PC1	1.888e+03 $1.29 \times 10^{+3}$	<.0001
PC2	1.270e+03 $2.84 \times 10^{+2}$	<.00011.99 × 10^{-2}
PC3	−3.853e+02 $-2.64 \times 10^{+2}$	.00131 <.0001

Abbreviations: PC, principal component; PGS, polygenic score; spHS, sporadic hidradenitis suppurativa.

In order, 1st column reports the coefficient name, 2nd column reports the estimated β value, and 3rd column reports the *P*-value derived from Z-test on β estimated value.

Protein classes–enriched analysis

By analyzing the PGSs using the statistical overrepresentation test provided by the PANTHER software, we retrieved 2 significantly differentially represented protein classes inside the genes where the PGS's variants map (Table 2).

The protein classes were ordered from highest to lowest fold enrichment. Cell adhesion proteins (PC00069) were overrepresented with a positive fold change of 3.06, whereas C2H2 zinc finger transcription factors (PC00248) were underrepresented, showing a fold change of 0.09.

DISCUSSION

In our spHS cohort, female–male ratio is 1.3:1, a very low one compared with literature data from other European cohorts (Morss et al, 2020), and with females having an earlier HS onset than males (18 vs 20 years). Despite that the number of patients enrolled is limited, our findings are consistent with wider epidemiological studies that indicate sex-specific differences in disease manifestation and progression (Rosi et al, 2022; Sabat et al, 2022). The importance

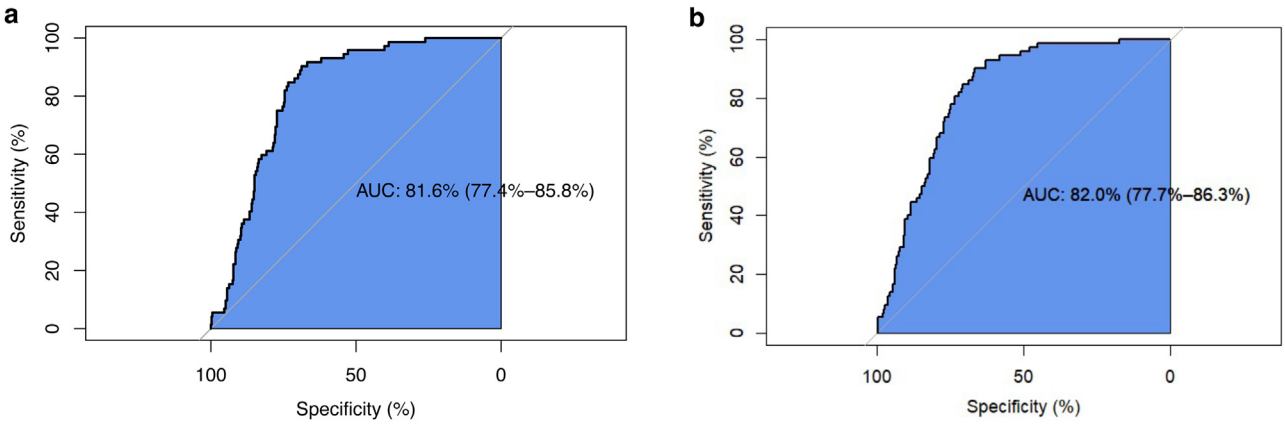

Figure 2. ROC and PGS. ROC curve for (**a**) the base model that included all covariates (ie, sex, age, and first three principal component) except the PGS and (**b**) the one that included PGS. PGS, polygenic score; ROC, receiver operating characteristic.

of considering sex as a key factor in understanding spHS's clinical landscape (Mintoff and Pace, 2024) could be partially attributed to hormonal influences, which are known to differently affect the immune response and skin physiology between males and females. Furthermore, although not statistically significant, we observed a difference between patients with early-onset spHS and those with late-onset spHS; thus, the proportion of patients experiencing early HS onset highlights the potential for genetic predisposition to play a significant role in disease manifestation and progression.

Notably, the multifactorial nature of sporadic HS is a widely accepted concept in dermatology because the clinical phenotypes result from both genetics, polygenic traits, and environmental factors (Visan et al, 2024). In our study, focused on genetic factors, we aimed to evaluate and quantify the extent to which disease susceptibility in spHS can be explained by common associated genetic variants and, as a consequence, the possibility of creating a PGS that can be used as a tool that summarizes the contribution of common genetic variants to spHS disease state. The statistically significant age-, sex-, and ancestry-adjusted association of our developed PGS in spHS allows us to hypothesize a genetic contribution to the susceptibility of spHS.

The proportion of phenotypic variance explained by a PGS (ie, R^2 coefficient of determination) varies greatly depending on how polygenic the disease is. Usually, highly polygenic diseases, that is, diseases that have a major genetic additive influence (Kachuri et al, 2024) (eg, coronary atherosclerosis and type 2 diabetes), are characterized by an R^2 between 5 and 10%, whereas low polygenic diseases, as in our case, are identified by an R^2 value <5%, pointing that gene interaction, epigenetics, and other factors such as environmental and immunology play a role to determine the onset of spHS (Farooqi et al, 2023; Khera et al, 2018).

In addition, in our study, we used the same single nucleotide variants data to investigate the biological pathways involved in the disease. Using a statistical overrepresentation test, we have been able to infer the most informative PANTHER's protein classes according to the PGS we developed. Our findings have elucidated the biological pathways aligned with the known etiopathology of the disease, such as cell adhesion molecules and zinc finger transcription factors. Considering the pathways results, cell adhesion proteins (PC00069) were overrepresented with a positive fold change of 3.06, whereas C2H2 zinc finger transcription factors (PC00248) were underrepresented. Cell adhesion proteins

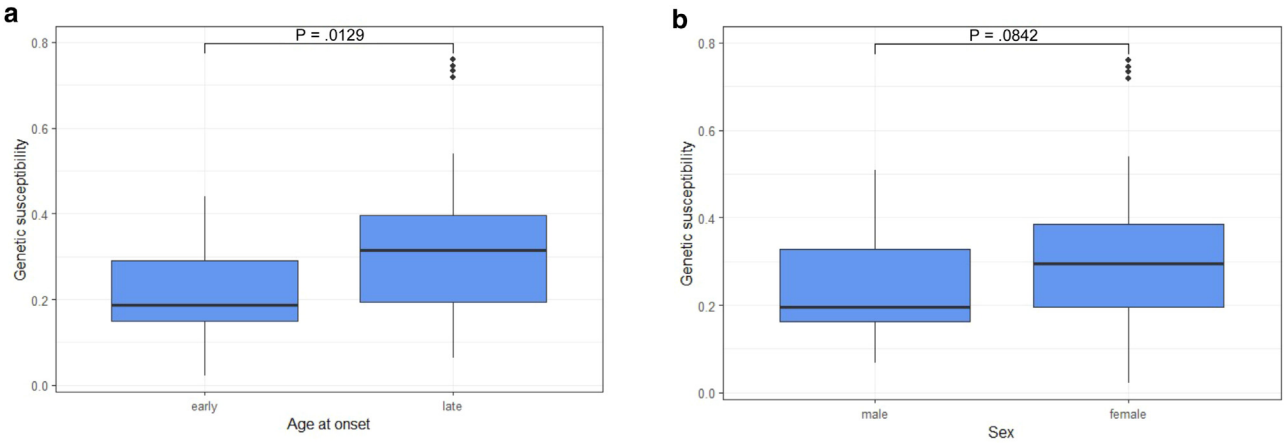

Figure 3. Genetic susceptibility grouped by age of onset and sex. Boxplot and *P*-value of the Wilcoxon test to verify statistical differences of the genetic susceptibility, that is, the polygenic risk, and (**a**) age of onset and (**b**) sex.

Table 2. Protein Classes Enriched by a Statistical Overrepresentation Test

Protein Class ID	PANTHER Protein Class	Observed Genes	Expected Genes	Fold Enrichment	Log2 Fold Enrichment	Adjusted *P*-Value
PC00069	Cell adhesion molecule	17	5.56	3.06	1.61	1.14×10^{-2}
PC00248	C2H2 zinc finger transcription factor	1	10.82	0.09	−3.47	3.91×10^{-2}

Abbreviations: ID, identification; PGS, polygenic score.

In order, 1st and 2nd columns report the protein class unique identifier and protein class name of PANTHER protein class database, respectively, 3rd and 4th columns report respectively the number of observed and expected genes inside PGS gene list, 5th and 6th columns report respectively the fold change and log2 fold change of how many times the number of observed genes is higher than the expected one, and 7th column report the Bonferroni adjusted *P*-value obtained by binomial distribution test.

are essential for the formation and maintenance of a healthy skin; in fact, they are involved in cell−cell contacts and in cellular interactions with the extracellular matrix (D'Arcy and Kiel, 2021). Recently, Jin et al (2023), studying HS epithelial cells, identified changes in gene signatures that revolve around cell−cell adhesion. In addition to this, Sanchez et al (2019) demonstrated that the structure and composition of the extracellular matrix are altered in biopsies of HS lesions. It is well-established that cell adhesion proteins and the extracellular matrix influence biological processes in keratinocytes, such as migration and proliferation, both of which are impaired in HS (Kashyap et al, 2022). The overrepresentation of cell adhesion proteins further supports the idea that their alteration may play a significant role in HS pathology. HS is a chronic inflammatory skin disease, where inflammation and associated immune responses are key drivers of its onset and progression (van Straalen et al, 2022). C2H2 zinc finger transcription factors (PC00248) are a large family of proteins involved in gene regulation, but they are not typically characterized as major players in skin inflammation or immunity. However, some specific members of this family may influence processes related to skin inflammation and immunity indirectly through their role in regulating gene expression. In the context of skin, some C2H2 zinc finger proteins play a role in skin development, keratinocyte differentiation, wound healing, and repair processes, all of which are related to skin health and may indirectly influence immune responses and inflammation. Moreover, although C2H2 zinc finger transcription factors are not directly associated with classical immune pathways, certain family members may regulate the expression of genes that influence immune responses, including cytokines or chemokines, which can impact inflammation. However, their role in immune modulation is not as prominent as those of other well-known transcription factors (eg, NF-κB, signal transducer and activator of transcriptions) that are directly involved in immune and inflammatory pathways. Besides, some C2H2 zinc finger proteins have been implicated in skin diseases, including psoriasis (Liu et al, 2022) and possibly other inflammatory conditions, through their role in epidermal differentiation and skin barrier function. For example, mutations in certain zinc finger proteins can disrupt normal skin homeostasis, potentially leading to inflammation. These are, of course, hypotheses that we have formulated in an attempt to interpret whether the results obtained from the pathway analysis align with the known etiology and pathology of the disease. Our goal was to explore how the identified pathways and molecular mechanisms may contribute to the disease process and whether they provide additional insights into the underlying biological factors. Although the associations that we observed are intriguing, further experimental validation is needed to confirm the relevance of these pathways in disease progression and to establish a more direct link between our findings and the pathophysiology of the condition. Nonetheless, this analysis provides a useful framework for guiding future research and deepening our understanding of the disease mechanisms at play.

In addition to our pathway analysis findings, recent studies have explored the broader genetic landscape of HS through the application of PRSs.

A notable study by Nielsen et al (2024) investigated the genetic correlation between HS and cardiometabolic diseases, such as coronary artery disease and diabetes, using a PRS for HS. Their analysis, conducted in a large cohort from the UK Biobank, revealed that a high PRS for HS (≥75th percentile) was significantly associated with increased odds of both coronary artery disease (OR = 1.09, 95% confidence interval = 1.06−1.12) and diabetes (OR = 1.13, 95% confidence interval = 1.10−1.17). These findings suggest that the genetic susceptibility to HS may also predispose individuals to cardiometabolic conditions, highlighting the pleiotropic effects of genetic variants involved in both inflammatory skin diseases and systemic metabolic disorders.

In the context of our study, these findings underscore the importance of considering genetic factors and their systemic implications in HS. Although our pathway analysis revealed key molecular disruptions in the extracellular matrix and immune responses, the application of PRSs offers a complementary approach to understanding the genetic basis of HS and its comorbidities. Future studies should aim to further dissect the role of specific genetic variants and proteins identified in both our work and others to explore their potential as biomarkers or therapeutic targets for managing not only HS but also its associated comorbid conditions.

Looking forward to a potential clinical application, enrolling a higher number of patients with spHS, as in the study of Nielsen et al (2024), PGS could be a supplementary tool for risk assessment and early diagnosis of HS, moving toward a more tailored approach based on individual genetic risk profiles. Although this approach has not yet been applied to patients with spHS, we believe that our results can contribute to this field as similar studies have done for other dermatological conditions such as psoriasis and atopic dermatitis (Arehart et al, 2022; Yang et al, 2024).

We are aware that our study is limited by the low number of patients enrolled. Future research needs to address this limitation and explore the challenges of integrating genetic testing into routine clinical practice, including issues related to cost, accessibility, and interpretation of genetic risk scores. Despite these limitations, our findings support the role of genetics in spHS and open new possibilities for understanding this complex disease. Incorporating PGS into clinical practice could potentially reduce the delay in disease diagnosis and provide better treatment options for individuals with higher scores, unclear symptoms, or a family history of HS.

MATERIALS AND METHODS

Patient's recruitment

Initially, 263 patients with sporadic HS were enrolled in 6 European clinical centers, and 1686 healthy controls were enrolled at the Institute for Maternal and Child Health IRCCS Burlo Garofolo (Trieste, Italy). The inclusion criteria for recruited patients required compliance to the diagnostic criteria for HS (Revuz and Jemec, 2016). From patient charts, we collected demographic data (age, sex, weight, height, body mass index, smoking status) and clinical data, including age at disease onset, age at disease diagnosis, and comorbidities. Disease and patients' characteristics were stratified according to Hurley staging and International Hidradenitis Suppurativa Severity Score System (van der Zee and Jemec, 2015; Zouboulis et al, 2017).

All participating individuals signed a written informed consent previously approved in agreement with the Helsinki Declaration and local legislations. The study was approved by the institutional review boards of Comitato Etico Unico Regionale of Friuli Venezia Giulia, Institute for Maternal and Child Health IRCCS Burlo Garofolo and Hospital Clinics Giuliano Isontino (ASUGI) (Trieste, Italy), Comitato Etico Milano Area 2 Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico (Milan, Italy), the Ethics committees of Medical University of Innsbruck (Innsbruck, Austria), the Ethics committees of Medical University of Cologne (Cologne, Germany), and Comitato Etico Lazio Area 3 of Policlinico Universitario Agostino Gemelli IRCCS (Rome, Italy).

Genomic DNA extraction and genotyping

Genomic DNA was extracted from saliva samples using the Oragene-DNA (Oragene) kit, following the manufacturer's instructions; Qubit instrument (Invitrogen) was used to evaluate DNA quantity. DNAs from patients with HS were subjected to Illumina Infinium Global Screening array, version 3.0 genotyping, whereas DNAs from healthy controls were subjected to Illumina Infinium OmniExpressExome genotyping, version 1.6. This happened because healthy controls and patients were not genotyped at the same time. However, after QC and genotype imputation (described in the next section), it was possible to merge the 2 arrays using in-house scripts.

QC and data analysis

The patients and controls in the study passed a QC procedure on genotypes files in PLINK binary format, performed separately for each cohort. This was necessary because the groups were genotyped using different arrays. QC was performed by using PLINK, version 1.9 (Purcell et al, 2007); PLINK, version 2.0 (Chang et al, 2015); and R (R Core Team, 2023) to filter out the following:

- Unevaluated variants in at least 5% of each cohort;
- Variants with a minor allele frequency <1%;
- Variants with a Hardy–Weinberg equilibrium $P < 1 \times 10^{-6}$;
- Samples with at least 5% of nongenotyped single nucleotide variants;
- Samples with a rate of heterozygosity outside the range of 6 SDs from the cohort average;
- Second-degree or higher kinship samples; and
- Samples with a non-European ancestry detected by principal component analysis.

After QC procedure, the genotyped data were subjected to a genotyping imputation procedure. The imputation was performed according to standard procedures (Naj, 2019; Zhao et al, 2018), with modifications, using an in-house pipeline and the Italian Genome Reference Panel (IGRP1.0) (Cocca et al, 2020) as reference panel of human haplotypes. Eagle2 software (Loh et al, 2016) was used to perform phasing, whereas impute5 (Rubinacci et al, 2020) was used for phased genotypes imputation.

Our custom genotype imputation consisted of 2 main steps performed to prevent a possible imputation bias given by the different Illumina beadchip used to genotype our 2 cohort of individuals (ie, spHS and healthy controls). In the first step, we imputed both case and control datasets separately and then merged the results by the intersection of variants with an imputation score > 0.99 in both cohorts; in the second step, the merged dataset was imputed again, and the result was filtered to keep only variants with imputation score > 0.80 and minor allele frequency < 0.001. The number of variants and individuals remaining after each step is shown in Figure 4; the final experimental dataset was 256 patients with spHS and 1397 controls.

To compute and test PGSs, we randomly assign our samples to training or testing datasets with a 70–30 proportion. All PGSs were computed on a training dataset (180 spHS and 978 controls) using summary statistics from GWASs of HS (Sun et al, 2023) and PLINK software (Purcell et al, 2007) to perform a polygenic computation through clumping and thresholding technique. The selected GWAS for our PGS computation was created using a cohort of 758,033 individuals (1640 cases and 756,393 controls) fully independent from ours.

To evaluate the performance metrics of obtained PGS, all biostatistical analyses were performed on testing dataset (76 spHS and 419 controls).

In conducting the biostatistical analysis of our PGS, a logistic regression model was employed, adjusted for critical demographic and genetic variables, including sex, age, and population structure. The latter was quantitatively represented by the first 3 principal components derived from data from patients with spHS, ensuring a robust adjustment for potential confounding factors that could influence the association between PGS and HS disease state. This analytical approach not only enhances the accuracy of our PGS association to HS but also underscores the importance of considering a comprehensive array of variables to accurately elucidate the genetic underpinnings of this complex disease (Choi et al, 2020). The best-performed PGS was evaluated using the Akaike Information Criterion value given by the performed logistic regression model; all subsequent statistical analysis were performed on the best-performed PGS.

To qualitatively evaluate our PGS from an etiopathological point of view, a statistical overrepresentation test was performed using

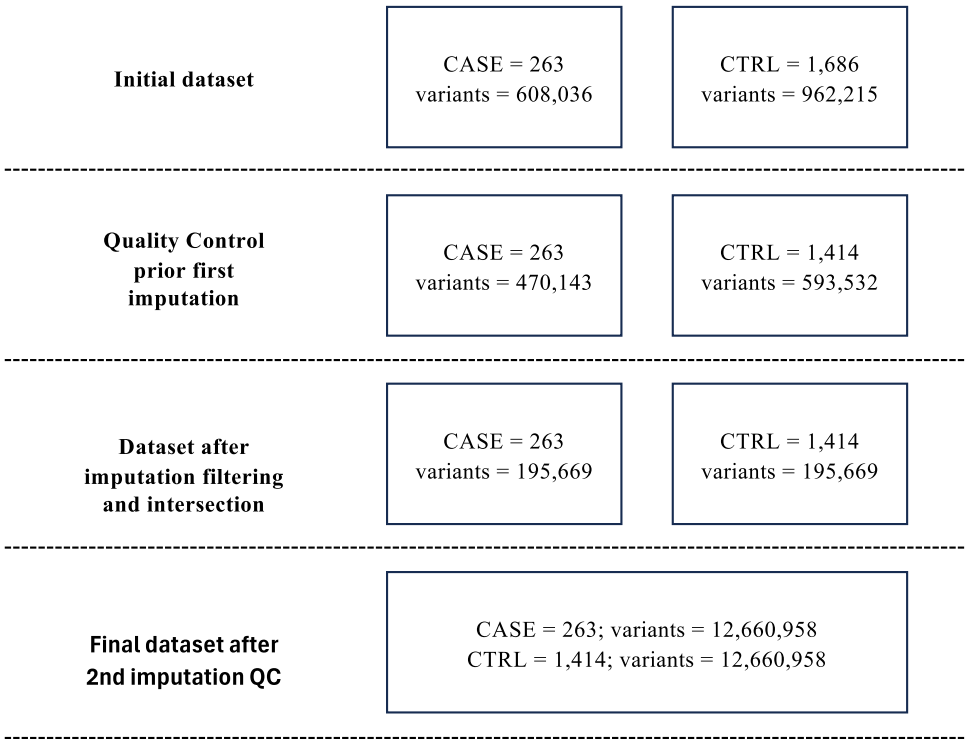

Figure 4. Fluxogram. Fluxogram with the number of samples and variants remaining after each step of preprocessing and QC. QC, quality control.

PANTHER online tool (Mi et al, 2019). The test was performed using as input file the list of genes where the SNPs of PGS mapped, whereas the binomial distribution test was selected as statistical test for overrepresentation against the PANTHER protein class database.

DATA AVAILABILITY STATEMENT

Datasets related to this article can be found at GWAS catalog (https://www.ebi.ac.uk/gwas/), under the accession identification GCST90487900.

ORCIDs

Chiara Moltrasio: http://orcid.org/0000-0003-0060-5870
Ronald Moura: http://orcid.org/0000-0001-8730-9948
Crovella Sergio: http://orcid.org/0000-0001-8493-1168
Tricarico Paola Maura: http://orcid.org/0000-0002-1954-4126

CONFLICT OF INTEREST

The authors state no conflict of interest.

ACKNOWLEDGMENTS

This work was supported by a Biomolecular Analyses for Tailored Medicine in AcneiNversa project funded by ERA PerMed (JTC_2018) through the Italian Ministry of Health, by a grant from the Institute for Maternal and Child Health IRCCS 'Burlo Garofolo/Italian Ministry of Health (RC16/2018), and by a Starting Grant (SG-2019- 12369421) founded by the Italian Ministry of Health and by the Collaborative Grant QU-CG_LARC24/25-423.

AUTHOR CONTRIBUTIONS

Conception: RM, SC, PMT; Data Curation: CM, LF, WJ, RDCC, KC, FMM, ADC, ER, FR, BB, NF, SC, IZ, EvS, IB, MS, VD, FP, AC, AVM; Formal Analysis: RM, PMT, AC, CDV, EMN, MB, SC; Investigation: RM, PMT, AC, CDV, EMN, MB, SC; Writing - Original Draft Preparation: CM, RM, AC, SC, PMT; Writing – Review and Editing: AC, MB, DA, SC

REFERENCES

Arehart CH, Daya M, Campbell M, Boorgula MP, Rafaels N, Chavan S, et al. Polygenic prediction of atopic dermatitis improves with atopic training and filaggrin factors. J Allergy Clin Immunol 2022;149:145–55.

Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaScience 2015;4:7.

Choi SW, Mak TS, O'Reilly PF. Tutorial: a guide to performing polygenic risk score analyses. Nat Protoc 2020;15:2759–72.

Cocca M, Barbieri C, Concas MP, Robino A, Brumat M, Gandin I, et al. A bird's-eye view of Italian genomic variation through whole-genome sequencing. Eur J Hum Genet 2020;28:435–44.

D'Arcy C, Kiel C. Cell adhesion molecules in normal skin and melanoma. Biomolecules 2021;11:1213.

Farooqi R, Kooner JS, Zhang W. Associations between polygenic risk score and Covid-19 susceptibility and severity across ethnic groups: UK Biobank analysis. BMC Med Genomics 2023;16:150.

Jemec GBE, Del Marmol V, Bettoli V, Augustin M, Prens EP, Zouboulis CC. Registries, multicentre and genome-wide association studies in hidradenitis suppurativa. Exp Dermatol 2022;31:22–8.

Jin L, Chen Y, Muzaffar S, Li C, Mier-Aguilar CA, Khan J, et al. Epigenetic switch reshapes epithelial progenitor cell signatures and drives inflammatory pathogenesis in hidradenitis suppurativa. Proc Natl Acad Sci USA 2023;120:e2315096120.

Kachuri L, Chatterjee N, Hirbo J, Schaid DJ, Martin I, Kullo IJ, et al. Principles and methods for transferring polygenic risk scores across global populations. Nat Rev Genet 2024;25:8–25.

Kashyap MP, Khan J, Sinha R, Jin L, Atigadda V, Deshane JS, et al. Advances in molecular pathogenesis of hidradenitis suppurativa: dysregulated keratins and ECM signaling. Semin Cell Dev Biol 2022;128:120–9.

Khera AV, Chaffin M, Aragam KG, Haas ME, Roselli C, Choi SH, et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. Nat Genet 2018;50:1219–24.

Liu S, Yuan X, Su H, Liu F, Zhuang Z, Chen Y. ZNF384: a potential therapeutic target for psoriasis and Alzheimer's disease through inflammation and metabolism. Front Immunol 2022;13:892368.

Loh PR, Danecek P, Palamara PF, Fuchsberger C, A Reshef Y, K Finucane H, et al. Reference-based phasing using the Haplotype Reference Consortium panel. Nat Genet 2016;48:1443–8.

Maronese CA, Moltrasio C, Marzano AV. Hidradenitis suppurativa-related autoinflammatory syndromes: an updated review on the clinics, genetics, and treatment of Pyoderma gangrenosum, Acne and Suppurative Hidradenitis (PASH), Pyogenic Arthritis, Pyoderma gangrenosum, Acne and Suppurative Hidradenitis (PAPASH), synovitis, Acne, Pustulosis, Hyperostosis and osteitis (SAPHO), and Rarer Forms. Dermatol Clin 2024;42:247–65.

Mi H, Muruganujan A, Huang X, Ebert D, Mills C, Guo X, et al. Protocol Update for large-scale genome and gene function analysis with the PANTHER classification system (v.14.0). Nat Protoc 2019;14:703–21.

Mintoff D, Pace NP. Differences in hidradenitis suppurativa patterns of cutaneous involvement between sexes: insights from a cross-sectional study. Hum Immunol 2024;85:110764.

Moltrasio C, Tricarico PM, Romagnuolo M, Marzano AV, Crovella S. Hidradenitis suppurativa: a perspective on genetic factors involved in the disease. Biomedicines 2022;10:2039.

Morss PC, Porter ML, Savage KT, Rosales Santillan M, Giannotti N, Kimball AB. Investigating race and gender in age at onset of hidradenitis suppurativa. J Eur Acad Dermatol Venereol 2020;34:e139–41.

Naj AC. Genotype imputation in genome-wide association studies. Curr Protoc Hum Genet 2019;102:e84.

Nielsen VW, Bundgaard Vad O, Holgersen N, Paludan-Müller C, Meseguer Monfort L, Beyer AF, et al. Genetic susceptibility to hidradenitis suppurativa and predisposition to cardiometabolic disease [published correction appears in JAMA Dermatol. 2024; doi:10.1001/jamadermatol.2024.5294] [e-pub ahead of print] JAMA Dermatol 2024; https://doi.org/10.1001/jamadermatol.2024.3779 (accessed 10 October 2024).

Pace NP, Mintoff D, Borg I. The genomic architecture of hidradenitis suppurativa-A systematic review. Front Genet 2022;13:861241.

Polygenic Risk Score Task Force of the International Common Disease Alliance. Responsible use of polygenic risk scores in the clinic: potential benefits, risks and gaps. Nat Med 2021;27:1876–84.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81:559–75.

R Core Team. The R project for statistical computing. https://www.R-project.org/; 2023 (accessed 10 October 2024).

Revuz JE, Jemec GBE. Diagnosing hidradenitis suppurativa. Dermatol Clin 2016;34:1–5.

Rosi E, Fastame MT, Silvi G, Guerra P, Nunziati G, Di Cesare A, et al. Hidradenitis suppurativa: the influence of gender, the importance of trigger factors and the implications for patient habits. Biomedicines 2022;10:2973.

Rubinacci S, Delaneau O, Marchini J. Genotype imputation using the Positional Burrows Wheeler Transform. PLoS Genet 2020;16:e1009049.

Sabat R, Tsaousi A, Ghoreschi K, Wolk K, Schneider-Burrus S. Sex-disaggregated population analysis in patients with hidradenitis suppurativa. Front Med (Lausanne) 2022;9:1028943.

Sanchez J, Le Jan S, Muller C, François C, Renard Y, Durlach A, et al. Matrix remodelling and MMP expression/activation are associated with hidradenitis suppurativa skin inflammation. Exp Dermatol 2019;28:593–600.

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

Sun Q, Broadaway KA, Edmiston SN, Fajgenbaum K, Miller-Fleming T, Westerkam LL, et al. Genetic variants associated with hidradenitis suppurativa. JAMA Dermatol 2023;159:930–8.

van der Zee HH, Jemec GB. New insights into the diagnosis of hidradenitis suppurativa: clinical presentations and phenotypes. J Am Acad Dermatol 2015;73:S23–6.

van Straalen KR, Prens EP, Gudjonsson JE. Insights into hidradenitis suppurativa. J Allergy Clin Immunol 2022;149:1150–61.

Vișan MA, Căruntu C, Costache RS, Tiplica G, Costache DO. Hidradenitis suppurativa: detangling phenotypes and identifying common denominators. J Eur Acad Dermatol Venereol 2024;38:62–76.

Yang JS, Liu TY, Lu HF, Tsai SC, Liao WL, Chiu YJ, et al. Genome-wide association study and polygenic risk scores predict psoriasis and its shared phenotypes in Taiwan. Mol Med Rep 2024;30:115.

Zhao S, Jing W, Samuels DC, Sheng Q, Shyr Y, Guo Y. Strategies for processing and quality control of Illumina genotyping arrays. Brief Bioinform 2018;19:765–75.

Zouboulis CC, Tzellos T, Kyrgidis A, Jemec GBE, Bechara FG, Giamarellos-Bourboulis EJ, et al. Development and validation of the International Hidradenitis suppurativa Severity Score System (IHS4), a novel dynamic scoring system to assess HS severity. Br J Dermatol 2017;177:1401–9.