Sex disparities in the phenotype at diagnosis of Sjögren's disease: artificial intelligence-driven characterisation in 17,416 patients

P. Brito-Zerón[1], A. Flores-Chávez[2], I.F. Horváth[3], R. Priori[4,5], H. Bootsma[6], B. Armagan[7], L. Quartuccio[8], S. Praprotnik[9], Y. Suzuki[10], G. Hernandez-Molina[11], V.C. Romão[12], A. Sebastian[13], E. Bartoloni[14], M. Rischmueller[15], R. Solans[16], S.G. Pasoto[17], G. Nordmark[18], I. Sánchez Berná[19], F. Carubbi[20], V. Fernandes Moça Trevisani[21], V. Valim[22], S. Melchor[23], B. Maure Noia[24], E. Fonseca-Aizpuru[25], L. Delgado[1,2], H. Nakamura[26], M. López-Dupla[27], M. Vazquez[28], M. Akasbi[29], G. Policarpo Torres[30], B. de Miguel Campo[31], R. Rouco[1], A. Szántó[3], A. Gattamelata[4], A. Vissink[32], L. Kilic[7], V. Manfrè[8], K. Perdan Pirkmajer[9,33], Y. Fujisawa[10], R. Pereira da Costa[12], P. Wiland[13], R. Gerli[14], C. Kirana[15], N. Nardi[2], M. Ramos-Casals[2,34] on behalf of the Sjögren Big Data Consortium

Abstract

Objective

Sjögren disease (SjD) predominantly affects females, but the early disease presentation in male patients remains poorly characterised due to historically small sample sizes. The aim of this study was to investigate sex-based differences in the clinical phenotype at diagnosis of SjD and identify predictors of patient sex using a large international cohort and AI-enhanced analysis.

Methods

Cross-sectional analysis of an anonymised dataset comprising 17,416 worldwide patients fulfilling the 2002/2016 classification criteria (Sjögren Big Data Registry). We stratified the dataset by sex and conducted a comparative analysis of baseline glandular and systemic involvement, organ-specific ESSDAI domains, and immunological profiles. Multivariate logistic regression models were developed, adjusting for epidemiological confounders (age and ethnicity) to identify predictors of sex classification. We used a generative AI (OpenAI's GPT-4o model) environment with Python (version 3.9) and the pandas (1.4.3), numpy (1.21.5), and matplotlib (3.5.1) libraries. All analyses adhered to GDPR standards, with anonymized patient data and strictly controlled secure environments.

Results

The cohort included 1,161 (6.67%) men and 16,255 (93.33%) women, with a mean age at diagnosis of 51.11 years (SD=14.45). Men showed a higher mean age at diagnosis (54.09 vs. 51.42 years in women; t=6.08, p<0.0001), a higher average ESSDAI score (7.65 vs. 5.93; t=7.91, p<0.0001) and higher frequencies in severe DAS categories (i.e. high activity 20% vs. 12% in women, $\chi^2 = 81.15$, p<0.0001). The epidemiologically-adjusted logistic regression model (pseudo R-squared value of 0.026) identified statistical significance for age (coefficient =0.009, p=0.024; each additional year in age increased the likelihood of being female by 1.4%), ethnicity (coefficient=0.579, HR=1.78, p=0.004), ocular dryness (coefficient=-0.607, HR=0.54, p<0.001), and systemic activity in the glandular (coefficient=0.359, HR=1.43, p=0.006) and pulmonary (coefficient=0.445, HR=1.56, p=0.004) ESSDAI domains.

Conclusion

Male SjD patients present a distinct, more systemic phenotype at diagnosis. Awareness of sex-specific features can improve early recognition and tailored management.

Key words

Sjögren's disease, sex disparities, phenotype at diagnosis, artificial intelligence, big data

Affiliations: see page 2140.

Pilar Brito-Zerón, MD, PhD
Alejandra Flores-Chávez, MD, PhD
Ildiko Fanny Horváth, MD
Roberta Priori, MD
Hendrika Bootsma, MD, PhD
Berkan Armagan, MD,
Luca Quartuccio, MD
Sonja Praprotnik, MD
Yasunori Suzuki, MD
Gabriela Hernandez-Molina, MD
Vasco C. Romão, MD, PhD
Agata Sebastian, MD, PhD
Elena Bartoloni, MD
Maureen Rischmueller, MD
Roser Solans, MD, PhD
Sandra G. Pasoto, MD, PhD
Gunnel Nordmark, MD, PhD
Isabel Sánchez Berná, MD, PhD
Francesco Carubbi, MD, PhD
Virginia Fernandes Moça Trevisani, MD
Valeria Valim, MD
Sheila Melchor, MD, PhD
Brenda Maure Noia, MD, PhD
Eva Fonseca-Aizpuru, MD
Lucía Delgado, MD
Hideki Nakamura, MD
Miguel López-Dupla, MD
Marcos Vazquez, MD, PhD
Miriam Akasbi, MD
Guillem Policarpo Torres, MD, PhD
Borja de Miguel Campo, MD, PhD
Rosana Rouco, MD
Antónia Szántó, MD
Angelica Gattamelata, MD
Arjan Vissink, DDS, MD, PhD
Levent Kilic, MD
Valeria Manfrè, MD
Katja Perdan Pirkmajer, MD, PhD
Yuhei Fujisawa. MD
Roberto Pereira da Costa, MD
Piotr Wiland, MD
Roberto Gerli, MD
Chandra Kirana, PhD
Norma Nardi, MD
Manuel Ramos-Casals, MD, PhD

Other members of the Sjögren Big Data Consortium-Sjögren GEAS-SEMI who contributed to this study are listed in Appendix in the Supplementary material.

Please address correspondence to:
Manuel Ramos-Casals
Servei de Malalties Autoimmunes Sistèmiques,
Hospital Clínic, C/Villarroel 170, 08036 Barcelona, Spain.
E-mail: mramos@clinic.cat

Received on September 18, 2025; accepted in revised form on October 20, 2025.

Funding: see page 2141.
Competing interests: none declared.

Introduction

Sjögren's disease (SjD) (1) is a systemic autoimmune disease that primarily affects people between the fourth and sixth decades of life, but it also can be diagnosed at any age (2). While dryness caused by autoimmune glandular impairment is the key clinical feature of SjD, affecting more than 95% of patients (3), the clinical spectrum of SjD can extend from these symptoms to a range of systemic organ-specific manifestations, clearly influencing the prognosis of the disease (4-6).

SjD disproportionately affects females. According to a recent systematic review by Qin *et al*. (7) the SjD estimated incidence rate ratio (IRR) for females *versus* males was 9.29 (95% 6.61–13.04) (8). The characterisation of disease presentation in males remains rare and understudied since previous reports specifically exploring male patients has been severely constrained by small absolute numbers. The largest single-centre study to date analysed 140 men (9) while most other cohorts enrolled ≤100 males (10-12). In this epidemiological scenario, the use of sizable datasets may offer more conclusive evidence of sex-based differences and a more robust, generalizable picture. Artificial intelligence (AI), which has recently revolutionised the field of statistical analysis in medical research, offers tools for efficient data processing and enhanced interpretability of complex, large clinical datasets. This approach may address a critical gap in the literature, namely, the underrecognition of gender disparities in initial disease severity and immunologic patterns within SjD.

The primary objective of this study was to characterise how male patients with SjD present with a distinct phenotype at diagnosis compared to female patients and to identify potential associations between baseline features and the likelihood of gender classification. By assessing the statistical significance and effect sizes of various predictors using AI, we aimed to understand better the differential impact of these variables between male and female patients, ultimately guiding more personalised clinical management.

Methods

Study design

We conducted a cross-sectional, observational study using data from a large international cohort of patients diagnosed with primary SjD. The Sjögren Big Data Consortium is an international, multicentre registry established in 2014 to provide a real-world understanding of the phenotypes and outcomes of individuals with primary SjD through the cooperative merging of existing individual databases from leading centres in SjD clinical research. Inclusion criteria for the registry consisted of a clinical diagnosis of primary SjD based on fulfilling 2002/2016 classification criteria (13, 14). Diagnostic tests for SjD (ocular tests, oral tests, and salivary gland biopsy) were conducted according to the recommendations of 2002/2016 criteria (13, 14). Exclusion criteria for considering SjD as a primary disease were chronic HCV/HIV infections and associated systemic autoimmune diseases (at diagnosis or during the follow-up). The longitudinal cohort collected cases diagnosed before 2014 and incident cases diagnosed after 2014 (individual databases are updated annually).

This study utilises the cross-sectional of anonymised patient records from multiple participating centres worldwide including baseline information at diagnosis The database consisted of a minimum anonymised basic data set (MBDS) of across 33 variables encompassing demographic, clinical, and laboratory data which are considered essential for characterising the disease phenotype at the time of diagnosis (15, 16). To harmonise databases from each centre, we applied AI-assisted specific preprocessing techniques such as detecting and treating outliers, influential observations, errors in naming standards, and other errors, and excluding observations with missing data for mandatory variables (age, sex, symptoms, criteria fulfilment). The study was approved by the Ethics Committee of the coordinating centre (Hospital Clinic, Barcelona, Spain, registry HCB0869).

Statistics

We stratified the dataset by sex and conducted a comparative analysis of

baseline features. Statistical analyses included descriptive statistics (mean, median, standard deviation) for numerical variables and frequency distributions for categorical variables. Differences between male and female groups were assessed using Student's t-tests or Mann-Whitney U-test when appropriate for numerical data. For categorical variables, we constructed contingency tables between 'Gender' and each variable, applying chi-square or Fisher's exact test based on cell frequency assumptions. Measures of association (*e.g.* Phi coefficient, Cramer's V) were also calculated. An ordinal analysis was conducted to assess potential associations between the disease severity variable 'DAS' and gender, using contingency tables and the chi-square test. A multivariate logistic regression model was constructed to estimate the odds ratios (ORs) for being male at diagnosis, adjusted for age and ethnicity. For each predictor, we calculated the logistic coefficient (B), exponentiated coefficient (HR or OR), 95% confidence intervals (CI), and *p*-values. Statistical significance was set at $p<0.05$. The integration of generative AI into the statistical pipeline not only optimised the efficiency and depth of data interpretation but also supported comprehensive and accessible reporting of findings, an innovative approach that enriched traditional statistical methods. Data processing and analysis workflows adhered to GDPR standards to ensure patient privacy. All patient data were anonymised prior to analysis, and no identifiable information was accessed at any point. The statistical analysis was conducted in a secure computational environment using generative AI (via OpenAI's GPT-4 model) with Python (v. 3.11) and essential libraries including pandas (1.4.3) for data manipulation, numpy (1.21.5) for numerical computations, and matplotlib (3.5.1) and seaborn (0.11.2) for visualisations. Code modularity and reproducibility were prioritised, with all scripts managed in version control systems (*e.g.* Git) to enable transparency.

Table I. Baseline features of 17,416 patients at diagnosis of primary Sjögren's disease.

Logical determinants	Variables	n=17,416	%
Epidemiological features	Sex (women)	16255	93.33
	Mean age (± SD)	51.11 (14.45)	–
	Ethnicity		
	White	12941/17011	76.07
	Asian	2498/17011	14.68
	Hispanic	1113/17011	6.54
	Black-African American (BAA)	248/17011	1.46
	Others	211/17011	1.24
Glandular involvement	Dryness of the mouth	16006	91.9
	Dryness of the eyes	15794	90.69
	Ocular tests (abnormal results)	12134/14404	84.24
	Oral tests (abnormal results)	8584/11093	77.38
	Focal lymphocytic sialadenitis	9484/11009	86.15
Immunological markers	Antinuclear antibodies	13179/15730	83.78
	Rheumatoid factor	6563/15327	42.82
	Ro autoantibodies	12349/16275	75.88
	La autoantibodies	6766/16150	41.89
	Low C3 value	1845/13710	13.46
	Low C4 value	1784/13676	13.04
	Serum cryoglobulins	567/7387	7.68
Systemic activity	Mean ESSDAI	6.03961749	–
	DAS		
	Low activity	9022/16836	53.59
	Moderate activity	5754/16836	34.18
	High activity	2060/16836	12.24
	Individual ESSDAI domains		
	Constitutional	1673/16823	9.95
	Lymphadenopathy	1437/16826	8.54
	Glandular	3137/16822	18.65
	Articular	6200/16821	36.86
	Cutaneous	1488/16825	8.84
	Pulmonary	1700/16820	10.11
	Renal	624/16810	3.72
	Muscular	334/16805	2.02
	PNS	915/16746	5.46
	CNS	304/16827	1.81
	Haematological	3551/16663	21.31
	Biological	7632/16633	45.89

Results

Description

This study utilises a cross-sectional database comprising 17,416 anonymised records included in the anonymised dataset on July 8, 2025 (Table I). In total, 16,255 women (93.3%) and 1,161 men (6.7%) were analysed, with a mean age at diagnosis of 51.11 years (SD=14.45). The sex ratio men: women were 1:15. The mean EULAR Sjögren's Syndrome Disease Activity Index (ESSDAI) score was 6.04 (SD=7.04), with a median of 4. The key features (>75%) at diagnosis included dry mouth (91.9%), dry eyes (90.7%), positive salivary biopsy (86.1%), abnormal ocular tests (84.2%), positive ANA (83.8%), abnormal oral tests (77.4%) and positive anti-Ro antibodies (75.9%).

Univariate analysis

Univariable comparisons by sex are summarised in Table II. Male patients were diagnosed at a significantly older age than females (54.09 *vs.* 51.42 years; t=6.08, $p<0.0001$) (Fig. 1). Ethnic distribution also differed significantly by sex: among women, 73.9% identified as White in comparison with 80.2% among men ($\chi^2=23.67$; $p<0.0001$).

Regarding glandular and systemic manifestations at diagnosis, seven variables showed statistically significant sex differences. Men reported fewer sicca symptoms: ocular dryness was present in 85.4% of men *vs.* 91.1% of women, and oral dryness in 87.3% of men *vs.* 92.2% of women (both $p<0.001$). Similarly, oral tests investigating salivary glandular function were less frequently reported as abnormal in men: positive oral tests in 71.8% *vs.* 77.8% in women ($p<0.001$). In contrast, there were no significant differences in the frequency of positive ocular tests.

Table II. Univariable comparisons of epidemiological, clinical and immunological features at diagnosis by sex.

Epidemiological determinants	Variables	Men	Women	Statistic (T for numeric, Chi-square for categoric)	*p*-value
Epidemiological features	Mean age (± SD)	54.09	51.42	6.08	**<0.001**
	Ethnicity (White)	931	12010	23.67	**<0.001**
Glandular involvement	Dryness of the mouth	1014	14992	34.85	**<0.001**
	Dryness of the eyes	992	14802	40.49	**<0.001**
	Ocular tests (abnormal results)	800	11334	0.32	0.570
	Oral tests (abnormal results)	527	8057	14.00	**<0.001**
	Focal lymphocytic sialadenitis	638	8846	2.10	0.148
Immunological markers	Antinuclear antibodies	865	12314	2.67	0.102
	Rheumatoid factor	480	6083	4.08	**0.043**
	Ro autoantibodies	835	11514	0.56	0.453
	La autoantibodies	495	6271	7.38	**0.007**
	Low C3 value	132	1713	1.06	0.303
	Low C4 value	109	1675	0.74	0.390
	Serum cryoglobulins	52	515	3.85	**0.049**
Systemic activity	DAS			81.15	**<0.001**
	Low activity	501	8521		
	Moderate activity	388	5366		
	High activity	227	1833		
	Individual ESSDAI domains				
	Constitutional	127	1546	2.75	0.097
	Lymphadenopathy	128	1309	13.13	**<0.001**
	Glandular	271	2866	24.79	**<0.001**
	Articular	371	5829	6.60	**0.010**
	Cutaneous	93	1395	0.39	0.534
	Pulmonary	157	1543	20.75	**<0.001**
	Renal	39	585	0.15	0.700
	Muscular	34	300	7.06	**0.008**
	PNS	112	803	50.43	**<0.001**
	CNS	29	275	4.31	**0.038**
	Haematological	244	3307	0.63	0.426
	Biological	519	7113	0.65	**<0.001**

Conversely, men exhibited greater systemic involvement at diagnosis. Males presented a higher average ESSDAI score (M=7.65, SD=8.26) compared to females (M=5.93, SD=6.93). This difference was also statistically significant (t=7.91, $p<0.0001$) (Fig. 2a). The distribution across DAS categories revealed that male patients presented more frequently in severe DAS categories ($\chi^2=81.15$, $p<0.0001$), whereas female patients were more represented in the lower severity categories (Fig. 2b). With respect to the specific phenotype composed by involvement of the different organ-specific ESSDAI domains, men were more likely to show systemic activity in the glandular (23.3% *vs*. 17.6%; $\chi^2=24.8$, $p<0.001$), lymphadenopathy (11% *vs*. 8.1%; $\chi^2=13.1$, $p=0.0003$), pulmonary (13.5% *vs*. 9.5%; $\chi^2=20.7$, $p=0.000005$), muscular (2.9% *vs*. 1.8%; $\chi^2=7.1$, $p=0.008$) and peripheral nervous system (9.6% *vs*. 4.9%; $\chi^2=50.4$, $p<0.0001$) domains, but less likely to show activity in the articular domain (32% *vs*. 35.9%; $\chi^2=6.6$, $p=0.01$).

Regarding the serological profile, men showed a higher positivity rate for anti-La antibodies (42.6% *vs*. 38.6%; $\chi^2=7.4$, $p=0.007$), RF (41.3% *vs*. 37.4%; $\chi^2=4.1$, $p=0.043$) and cryoglobulins (4.5% *vs*. 3.2%; $\chi^2=3.8$, p=0.049), whereas no significant sex differences were found for ANA, hypocomplementemia, or anti-SSA/Ro antibodies.

Multivariate analysis

The logistic regression model included age, ethnicity, and 11 additional variables relevant to the clinical characterization of SjD (Table III). After exclusion of records with missing data, 5,103 cases were entered in the model. The model demonstrated a significant pseudo-R squared value of 0.026, and several variables showed statistically significant associations with gender (Fig. 3). Male gender was positively associated with older age at diagnosis (HR 1.01, 95% CI 1.00, 1.02, $p=0.024$), ethnicity classified as White (HR 1.78, 95% CI 1.21–2.63, $p=0.004$), positive RF (HR 1.35, 95% CI 1.07–1.70, $p=0.011$) and systemic activity in the glandular (coefficient=0.359, HR=1.43, $p=0.006$) and pulmonary (coefficient=0.445, HR=1.56, $p=0.004$) ESSDAI domains. Conversely, the presence of dry eyes at diagnosis was associated with lower odds of being male (HR 0.54, 95% CI 0.38–0.78, $p=0.001$). No other ESSDAI domains or immunological variables showed significant independent associations with sex.

Discussion

SjD is one of the most common systemic autoimmune diseases, with a striking female predominance of 9:1 to 16:1 (17). Males consistently represent 2.9% to 14.6% of primary SjD cohorts

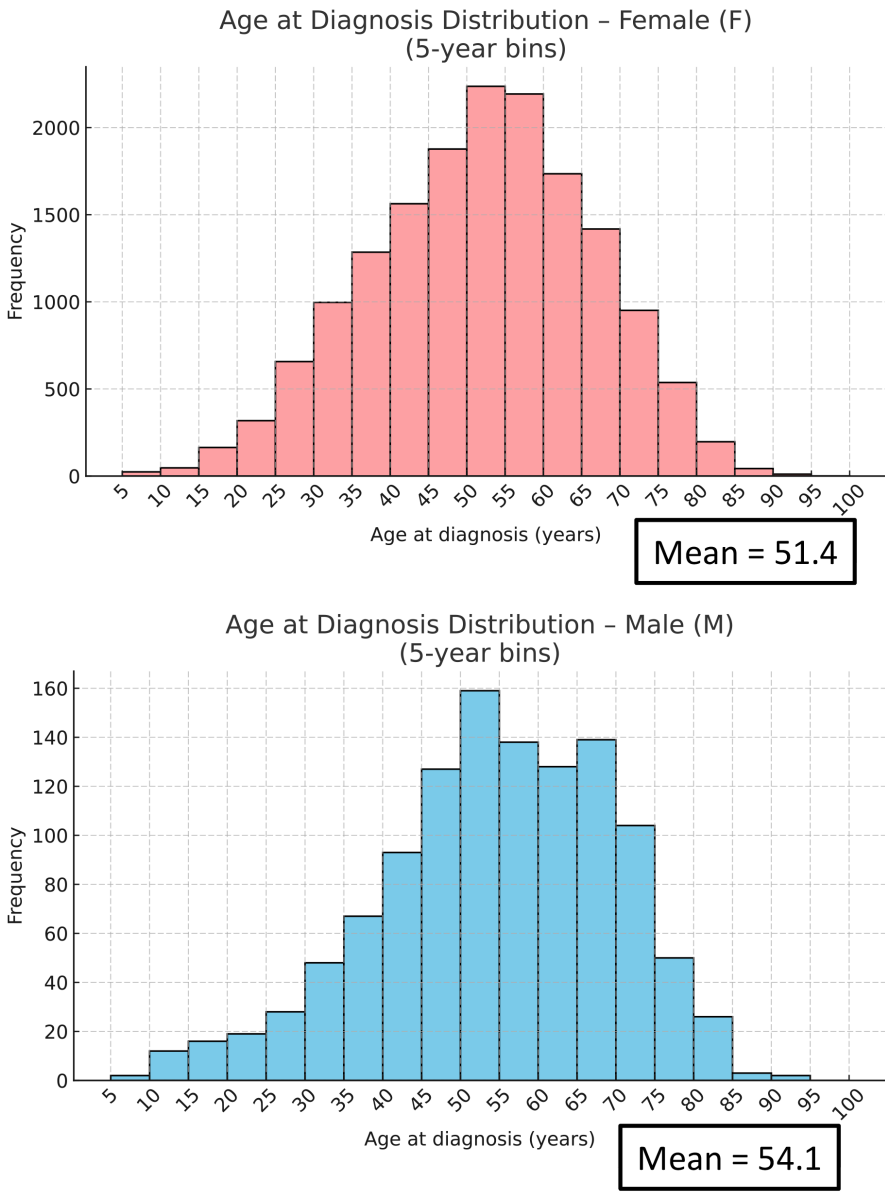

Fig. 1. Histogram of the age at diagnosis of SjD in men and women (5-year bins).

(9, 11, 18, 19). This sex bias has been attributed to hormonal influences, genetic factors, and sex chromosome effects (20). Understanding disease presentation in males is crucial, as it may be underdiagnosed and potentially more severe in this population.

We present the largest study to date characterising SjD in males, including nearly 1,200 affected men, ten times larger than the previous record of 140 men and 25–40 times larger than most published male-specific cohorts (30). This unprecedented sample size provides substantial statistical power to explore sex-related differences. Our analysis evaluated independent associations between clinical and demographic variables with gender in SjD patients, adjusting for age and ethnicity to identify factors associated with differences in disease manifestation between sexes. The multivariable logistic regression was designed to assess associations rather than to build an individual-level predictive classifier.

Males represent 6.7% of the Sjögren Big Data Cohort (ratio male:female 1:15). A key finding was the verification of epidemiological differences in age at diagnosis and ethnicity between males and females with SjD. Previous studies reported that males are typically diagnosed at 55–60 years compared to 50–55 years in females (21, 22). Our cohort confirmed this trend (54 years in males *vs*. 51 years in females), though differences were smaller than previously reported. And for the first time, we demonstrate significant ethnic disparities by gender, with overrepresentation of White individuals among males and Asian individuals among females. This novel finding highlights the importance of considering gender-specific ethnic patterns when addressing health disparities and access to care.

Consistent with previous literature, males with primary SjD exhibit less severe glandular dysfunction than females in both subjective symptoms and objective assessments (19). Males typically report milder xerophthalmia and xerostomia (9, 19), with reduced salivary flow and lower patient-reported disease burden (19, 23). The mechanisms underlying reduced sicca symptoms in males remain unclear but may involve differences in hormone levels, glandular function, and pain perception (20, 24). Our findings suggest a significant inverse association between male sex and the presence of keratoconjunctivitis sicca at initial diagnosis. The odds ratio of 0.54 indicates that male patients exhibited approximately 46% lower odds in this cohort of presenting with dry eye symptoms at the time of diagnosis compared to their female counterparts. Ophthalmologic assessment warrants heightened vigilance in male patients, who indicate increased susceptibility to undetected severe ocular complications (25).

Previous evidence also suggests males experience a more severe disease course, more frequently present with acute organ dysfunction requiring immunosuppressive therapy and increased lymphoma and mortality risk (9, 18, 19, 21, 23, 25). Specific systemic manifestations reported more commonly in males include parotid enlargement (9), pulmonary involvement (9, 19, 23), and vasculitis (11, 23). Our univariate analysis revealed that male patients have a significantly higher prevalence of systemic involvement across several ESSDAI domains. However, following multivariate adjustment for potential confounders, glandular and pulmonary involvements emerged as the systemic

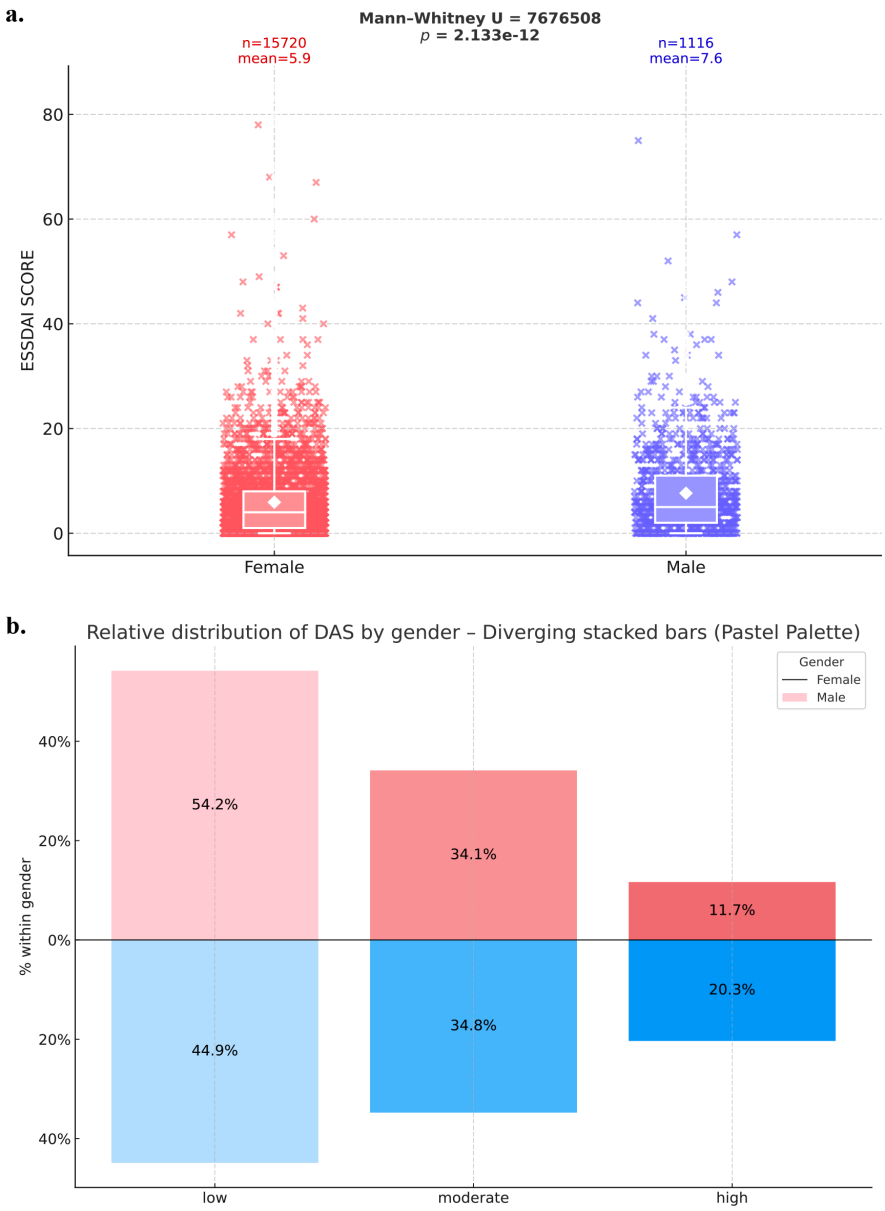

Fig. 2. a. Average ESSDAI score at diagnosis by sex in men (7.65, SD=8.26) compared to women (5.93, SD=6.93). This difference was also statistically significant (t=7.91, p<0.0001).
b. Relative distribution across DAS categories by sex.

manifestations independently associated with male sex, with a 56% increased odds of lung involvement in comparison with women.

Previous investigations have consistently reported heightened immunologic activity in females, characterised by increased erythrocyte sedimentation rates and more pronounced autoantibody profiles encompassing anti-Ro, anti-La, ANA, and RF (9, 17, 18, 23). Conversely, male patients have been documented to exhibit lower autoantibody titres and increased seronegativity rates (9, 19, 21), including anti-Ro and anti-La antibodies (21), although anti-ribonucleoprotein antibodies demonstrated male predominance (19). Our findings diverge from established literature by suggesting modestly elevated immunological parameters in male SjD patients. The substantial sample size disparity in our cohort requires cautious interpretation of these contrasting immunological findings. This 15:1 female-to-male ratio may introduce inherent selection bias, as male patients captured within the registry potentially constitute a more immunologically active subset requiring comprehensive evaluation. Additionally, our expanded male cohort (n=1,161) provides unprecedented statistical power compared to previous smaller-scale studies, potentially revealing immunological patterns that remained undetected due to insufficient sample sizes and broader confidence intervals in earlier investigations.

Hormonal factors likely play a significant role. Sex hormones modulate immune responses through effects on lymphocyte maturation, activation, and cytokine production, while maintaining glandular homeostasis. Hormonal differences may influence clinical presentation and progression, with females particularly affected during menopause (24, 26). These hormones affect immune cell infiltration and glandular apoptosis, suggesting potential therapeutic targets. In addition, genetic factors may also contribute to sex-based differences. X chromosome variants associated with SjD show sex-specific effects, potentially explaining the higher female prevalence and severity (27, 28). The increased frequency of 47,XXY (Klinefelter's syndrome) among males with SjD suggests that an additional X chromosome enhances autoimmune susceptibility (29).

Several limitations merit consideration. The predominant European representation may restrict applicability to other ethnicities. Cultural and behavioural differences in symptom reporting may also influence clinical presentation and diagnosis. Additionally, participating centres are primarily tertiary university hospitals serving as referral centres, potentially selecting for more severe phenotypes, and diagnostic test availability varied across centres, potentially introducing bias and limiting generalisability. The retrospective design resulted in missing data for some diagnostic variables, particularly affecting multivariate models, notably for cryoglobulins, tested in only half the cohort. The modest pseudo-R-squared value (0.02542) indicates that while these predictors explain some gender differences, substantial variance remains unexplained, suggesting additional unmeasured factors contribute to disease manifestation. This low pseudo-R^2 underscores that the model accounts for only a small proportion of variance in sex classification, limiting its predic-

Table III. Multivariate logistic regression model estimating the odds ratios (ORs) for being male at diagnosis, adjusted for age and ethnicity.

Domains	Variables at diagnosis	At risk	Coefficient (B)	HR	95% CI (I)	95% CI (S)	*p*-value
Epidemiological	Age	..	0.0093	1.01	1.00	1.02	**0.024**
	Ethnicity	White	0.5791	1.78	1.21	2.63	**0.004**
Glandular	Dryness of eyes	Presence	0.6071	0.54	0.38	0.78	**0.001**
	Dryness of mouth	Presence	0.1982	0.82	0.55	1.22	0.324
	Oral tests	Normal	0.2184	0.80	0.62	1.05	0.105
Systemic	Lymphadenopathy	Activity	0.1635	1.18	0.84	1.66	0.350
	Glandular	Activity	0.359	1.43	1.11	1.85	**0.006**
	Articular	Activity	0.1497	0.86	0.68	1.09	0.213
	Pulmonary	Activity	0.4457	1.56	1.15	2.12	**0.004**
	Muscular	Activity	0.2614	1.30	0.78	2.18	0.321
	PNS	Activity	0.2117	1.24	0.83	1.84	0.294
	CNS	Activity	0.4497	1.57	0.81	3.05	0.185
Immunological	La antibodies	Positive	0.18	1.19	0.94	1.52	0.143
	RF	Positive	0.2993	1.35	1.07	1.70	**0.011**
	Cryoglobulins	Positive	0.0196	1.02	0.69	1.51	0.922

RF: rheumatoid factor; HR: hazard ratio; CI: confidence interval.

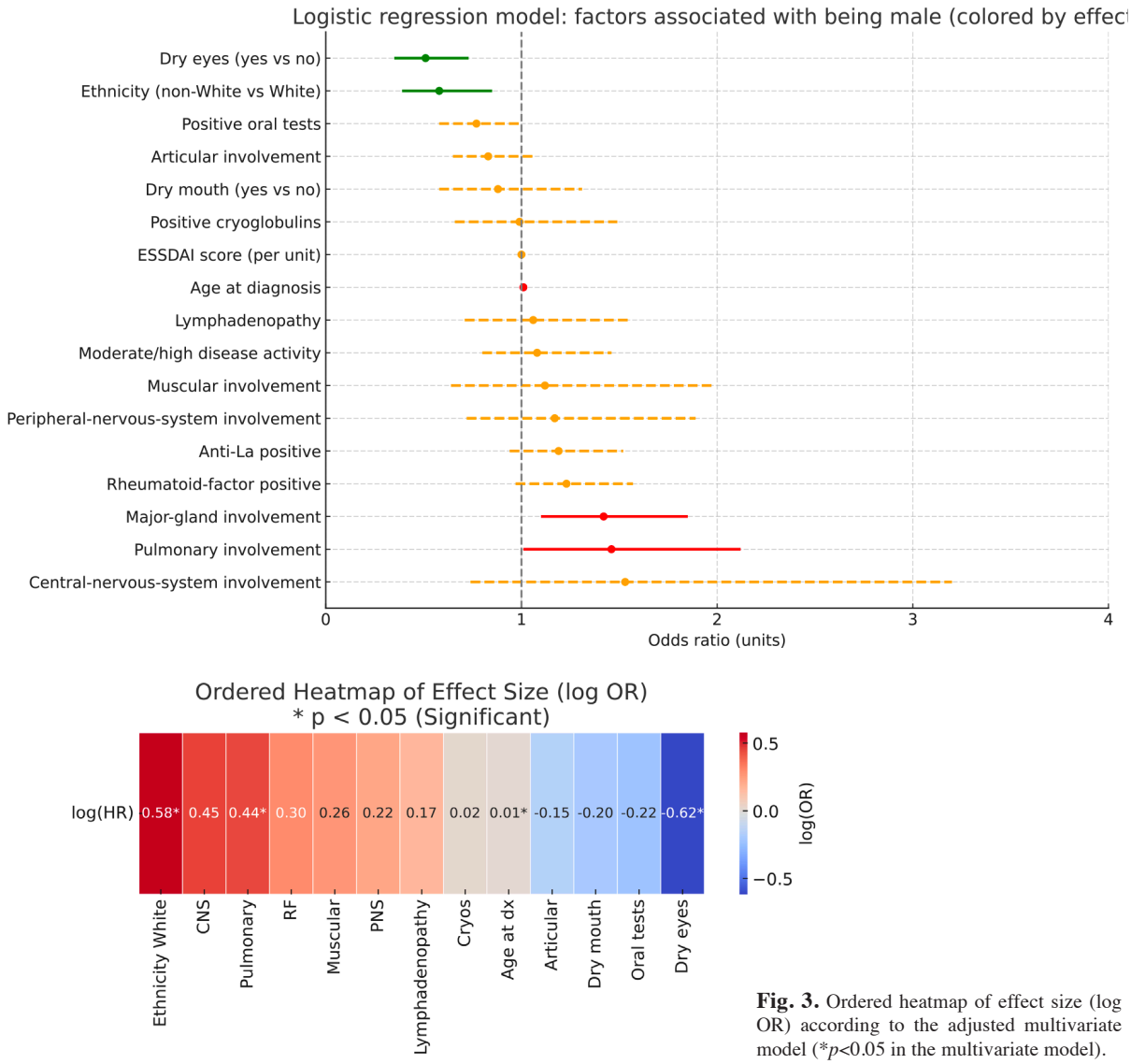

Fig. 3. Ordered heatmap of effect size (log OR) according to the adjusted multivariate model (*p<0.05 in the multivariate model).

tive utility for individual patients. Accordingly, the multivariable model has limited discriminative/predictive value for classifying sex at the individual-patient level, and findings should be interpreted as associations rather than predictive estimates.

This study's primary strength is its unprecedented sample size of nearly 1,200 SjD males, providing exceptional statistical power for detecting gender-specific differences. The international scope, encompassing 27 countries across five continents, offers unique insights into gender disparities across diverse populations. The research may inform actionable insights for sex-driven personalised medicine approaches in SjD. Our findings are consistent with the broader literature describing the clinical spectrum of male patients (30) while refining estimates with greater precision.

In addition, AI techniques facilitated analysis of this large dataset, automating descriptive analysis, multivariable modelling, and clustering. Natural language processing enabled precise report drafting aligned with publication standards. AI-powered visualisations enhanced accessibility of statistical findings. This integration optimised efficiency and depth of data interpretation while supporting comprehensive reporting, demonstrating AI's essential role in managing complex medical research datasets.

Our findings reveal significant gender-specific presentations in SjD with important clinical implications. Males present with distinct phenotypes characterised by more severe systemic manifestations despite less prominent sicca symptoms, and clinicians should maintain high suspicion for SjD in males with compatible symptoms, even without classical sicca features. Gender-specific predispositions to glandular and pulmonary manifestations may warrant consideration of tailored screening protocols. Furthermore, males could warrant consideration of more intensive monitoring, and these findings may inform hypotheses about targeted therapeutic strategies pending confirmation in prospective studies.

These are population-level associations and should not be interpreted as individual-level predictions. In summary, we report that men with SjD are more likely to be older and White at diagnosis and to present with glandular and pulmonary involvement, and positive rheumatoid factor, while being significantly less likely to report dry eye symptoms. This study provides robust population-level evidence for sex-based differences in SjD presentation, emphasising the need for sex-specific considerations in diagnosis and management. These insights may enhance recognition of SjD in males and inform clinical vigilance and hypothesis generation, rather than individual-level prediction.

Affiliations

[1]Autoimmune Diseases Unit, Internal Medicine Department, Hospital CIMA Sanitas, Barcelona, Spain; [2]Postgraduate Degree Program in Artificial Intelligence in Medicine, Department of Medicine, University of Barcelona, Spain; [3]Division of Clinical Immunology, Faculty of Medicine, University of Debrecen, Hungary; [4]Department of Internal Medicine and Medical Specialties, Rheumatology Clinic, Sapienza University of Rome, Italy; [5]Saint Camillus International University of Health Science, UniCamillus, Rome, Italy; [6]Department of Rheumatology and Clinical Immunology, University of Groningen, University Medical Center Groningen, the Netherlands; [7]Division of Rheumatology, Department of Internal Medicine, Hacettepe University, Faculty of Medicine, Ankara, Turkey; [8]Division of Rheumatology, Department of Medicine, University of Udine, Hospital Santa Maria della Misericordia, ASUFC, Udine, Italy; [9]Department of Rheumatology, University Medical Centre, Ljubljana, Slovenia; [10]Division of Rheumatology, Kanazawa University Hospital, Kanazawa, Ishikawa, Japan; [11]Immunology and Rheumatology Department, Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, México City, Mexico. [12]Rheumatology Department, Hospital de Santa Maria, Unidade Local de Saúde Santa Maria and Faculdade de Medicina, Universidade de Lisboa, Lisbon Academic Medical Centre, Lisbon, Portugal; [13]Department and Clinic of Rheumatology and Internal Medicine, Wroclaw Medical University, Poland; [14]Rheumatology Unit, Department of Medicine and Surgery, University of Perugia, Italy; [15]Department of Rheumatology, The Queen Elizabeth Hospital, Discipline of Medicine, University of Adelaide, South Australia, Australia; [16]Department of Internal Medicine, Hospital Vall d'Hebron, Barcelona, Spain; [17]Rheumatology Division, Hospital das Clinicas HCFMUSP, Faculdade de Medicina, Universidade de Sao Paulo, Brazil; [18]Rheumatology, Department of Medical Sciences, Uppsala University, Uppsala, Sweden; [19]Department of Internal Medicine, Virgen de las Nieves University Hospital, Granada; and ibs.GRANADA Biomedical Research Institute, Granada, Spain; [20]Department of Clinical Medicine, Life, Health and Environmental Sciences, Internal Medicine and Nephrology Division, ASL1 Avezzano-Sulmona-L'Aquila, San Salvatore Hospital, University of L'Aquila, Italy; [21]Division of Health Based Evidence, Federal University of Sao Paulo, Brazil; [22]Federal University of Espírito Santo, Brazil; [23]Department of Rheumatology, Hospital 12 de Octubre, Madrid, Spain; [24]Department of Internal Medicine, Complexo Hospitalario Universitario de Vigo, Spain; [25]Department of Internal Medicine, Hospital de Cabueñes, Gijón, Spain; [26]Division of Haematology and Rheumatology, Department of Medicine, Nihon University School of Medicine, Oyaguchi Kami-cho, Itabashi-ku, Tokyo, Japan; [27]Department of Internal Medicine, Hospital Joan XXIII, Tarragona, Spain; [28]Department of Rheumatology, Hospital de Clínicas, San Lorenzo, Paraguay; [29]Department of Internal Medicine, Hospital Infanta Leonor, Madrid, Spain; [30]Department of Internal Medicine, Girona Biomedical Research Institute-IDIBGI, Hospital Universitari Doctor Josep Trueta, Girona, Spain; [31]Department of Internal Medicine, Hospital 12 de Octubre, Madrid, Spain; [32]Department of Oral and Maxillofacial Surgery, University of Groningen, University Medical Center Groningen, The Netherlands; [33]Department of Internal Medicine, Faculty of Medicine, University of Ljubljana, Slovenia; [34]Department of Autoimmune Diseases, ICMiD, Hospital Clínic, Barcelona, Spain.

Funding

S.G. Pasoto has received a research grant from the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (no. 303116/2022-6).

References

1. RAMOS-CASALS M, BAER AN, BRITO-ZERÓN P *et al*.: International Task Force on Nomenclature of Sjögren disease. 2023 International Rome consensus for the nomenclature of Sjögren disease. *Nat Rev Rheumatol* 2025; 21(7): 426-37. https://doi.org/10.1038/s41584-025-01268-z
2. BRITO-ZERÓN P, BALDINI C, BOOTSMA H *et al*.: Sjögren syndrome. *Nat Rev Dis Primers* 2016; 2: 16047. https://doi.org/10.1038/nrdp.2016.47
3. RETAMOZO S, ACAR-DENIZLI N, HORVÁTH IF *et al*.: Influence of the age at diagnosis in the disease expression of primary Sjögren syndrome. Analysis of 12,753 patients from the Sjögren Big Data Consortium. *Clin Exp Rheumatol* 2021; 39 (Suppl. 133): S166-74. https://doi.org/10.55563/clinexprheumatol/egnd1i
4. VITALI C, PALOMBI G, BALDINI C *et al*.: Sjögren Syndrome Disease Damage Index and disease activity index: scoring systems for the assessment of disease damage and disease activity in Sjögren syndrome, derived from an analysis of a cohort of Italian patients. *Arthritis Rheum* 2007; 56(7): 222331. https://doi.org/10.1002/art.22658
5. BALDINI C, PEPE P, QUARTUCCIO L *et al*.: Primary Sjögren's syndrome as a multiorgan disease: impact of the serological profile on the clinical presentation of the disease in a large cohort of Italian patients. *Rheumatology* (Oxford) 2014; 53(5): 839-44. https://doi.org/10.1093/rheumatology/ket427
6. QIN B, WANG J, YANG M *et al*.: Epidemiology of primary Sjögren syndrome: a systematic review and meta-analysis. *Ann Rheum Dis* 2015; 74(11): 1983-89. https://doi.org/10.1136/annrheumdis-2014-205375
7. FULVIO G, LA ROCCA G, CHATZIS LG *et al*.: Impact of gender and age at onset on Sjögren's syndrome presentation and outcome: state of the art. *Clin Exp Rheumatol* 2023; 41(12): 2547-54. https://doi.org/10.55563/clinexprheumatol/lygrzv
8. ZHANG Y, CHEN JQ, YANG JY *et al*.: Sex difference in primary Sjögren syndrome: a medical records review study. *J Clin Rheumatol* 2023; 29(5): e78-e85. https://doi.org/10.1097/rhu.0000000000001962
9. NORTEY J, SHIBOSKI C, ROSE-NUSSBAUMER J *et al*.: How are sicca signs and symptoms associated with depression among men classified with and without Sjögren disease? *Am J Ophthalmol* 2023; 247: 96-102. https://doi.org/10.1016/j.ajo.2022.09.016
10. GONDRAN G, FAUCHAIS A, LAMBERT M *et al*.: Primary Sjögren's syndrome in men. *Scand J Rheumatol* 2008; 37(4): 300-5. https://doi.org/10.1080/03009740802001426
11. RAMOS-CASALS M, SOLANS R, ROSAS J *et al*.: Primary Sjögren syndrome in Spain: clinical and immunologic expression in 1010 patients. *Medicine* (Baltimore) 2008; 87(4): 210-19. https://doi.org/10.1097/MD.0b013e318181e6af
12. VITALI C, BOMBARDIERI S, JONSSON R *et al*.: Classification criteria for Sjögren's syndrome: a revised version of the European criteria proposed by the American-European Consensus Group. *Ann Rheum Dis* 2002; 61(6): 554-58. https://doi.org/10.1136/ard.61.6.554
13. SHIBOSKI CH, SHIBOSKI SC, SEROR R *et al*.: 2016 American College of Rheumatology/European League Against Rheumatism classification criteria for primary Sjögren's syndrome. A consensus and data-driven methodology involving three international patient cohorts. *Arthritis Rheumatol* 2017; 69(1): 35-45. https://doi.org/10.1002/art.39859
14. BRITO-ZERÓN P, ACAR-DENIZLI N, ZEHER M *et al*.: Influence of geolocation and ethnicity on the phenotypic expression of primary Sjögren's syndrome at diagnosis in 8310 patients: a cross-sectional study from the Big Data Sjögren Project Consortium. *Ann Rheum Dis* 2017; 76 (6): 1042-50. https://doi.org/10.1136/annrheumdis-2016-209952
15. BRITO-ZERÓN P, ACAR-DENIZLI N, NG WF *et al*.: How immunological profile drives clinical phenotype of primary Sjögren syndrome at diagnosis: analysis of 10,500 patients (Sjögren Big Data Project). *Clin Exp Rheumatol* 2018; 36 (Suppl. 112): S102-12. https://doi.org/10.55563/clinexprheumatol/255qsx
16. BRANDT JE, PRIORI R, VALESINI G, FAIRWEATHER D: Sex differences in Sjögren's syndrome: a comprehensive review of immune mechanisms. *Biol Sex Differ* 2015; 6: 19. https://doi.org/10.1186/s13293-015-0037-7
17. HORVATH IF, SZODORAY P, ZEHER M: Primary Sjögren's syndrome in men: clinical and immunological characteristic based on a large cohort of Hungarian patients. *Clin Rheumatol* 2008; 27(12): 1479-83. https://doi.org/10.1007/s10067-008-0944-7
18. PARK Y, LEE J, PARK SH, KWOK SK: Male patients with primary Sjögren's syndrome: a distinct clinical subgroup? *Int J Rheum Dis* 2020; 23(10): 1388-95. https://doi.org/10.1111/1756-185x.13940
19. KONTTINEN YT, FUELLEN G, BING Y *et al*.: Sex steroids in Sjögren's syndrome. *J Autoimmun* 2012; 39(1-2): 49-56. https://doi.org/10.1016/j.jaut.2012.01.004
20. MATHEWS PM, HAHN S, HESSEN M *et al*.: Ocular complications of primary Sjögren syndrome in men. *Am J Ophthalmol* 2015; 160(3): 447-52. https://doi.org/10.1016/j.ajo.2015.06.004
21. BRUNO KA, MORALES-LARA AC, BITTENCOURT EB *et al*.: Sex differences in comorbidities associated with Sjögren's disease. *Front Med* (Lausanne) 2022; 9: 958670. https://doi.org/10.3389/fmed.2022.958670
22. RAMÍREZ SEPÚLVEDA JI, KVARNSTRÖM M, BRAUNER S, BALDINI C, WAHREN-HERLENIUS M: Difference in clinical presentation between women and men in incident primary Sjögren's syndrome. *Biol Sex Differ* 2017; 8: 16. https://doi.org/10.1186/s13293-017-0132-x
23. XUAN Y, ZHANG X, WU H: Impact of sex differences on the clinical presentation, pathogenesis, treatment and prognosis of Sjögren's syndrome. *Immunology* 2024; 171(4): 513-24. https://doi.org/10.1111/imm.13740
24. BRITO-ZERÓN P, RETAMOZO S, RAMOS-CASALS M: Phenotyping Sjögren's syndrome: towards a personalised management of the disease. *Clin Exp Rheumatol* 2018; 36 (Suppl. 112): S198-209. https://doi.org/10.55563/clinexprheumatol/263xbc
25. NIKOLOV N, ILLEI G: Pathogenesis of Sjögren's syndrome. *Curr Opin Rheumatol* 2009; 21 (5): 465-70. https://doi.org/10.1097/bor.0b013e32832eba21
26. CHATZIS LG, GOULES AV, TZIOUFAS AG: Searching for the "X factor" in Sjögren's syndrome female predilection. *Clin Exp Rheumatol* 2021; 39 (Suppl. 133): S206-14. https://doi.org/10.55563/clinexprheumatol/88dyrn
27. THORLACIUS G, BJÖRK A, WAHREN-HERLENIUS M: POS0820 Sex-eQTL analysis reveals a significant sex-bias in the effects of Sjögren syndrome associated alleles on gene transcription. *Ann Rheum Dis*. 2023; 82 Suppl. 1: 705. https://doi.org/10.1136/annrheumdis-2023-eular.6230
28. HARRIS VM, SHARMA R, CAVETT J *et al*.: Klinefelter's syndrome (47,XXY) is in excess among men with Sjögren's syndrome. *Clin Immunol* 2016; 168: 25-29. https://doi.org/10.1016/j.clim.2016.04.002
29. CHATZIS LG, PEZOULAS VC, FERRO F *et al*.: Sjögren's syndrome: the clinical spectrum of male patients. *j clin med* 2020; 9(8): 2620. https://doi.org/10.3390/jcm9082620