Immunological signatures in patients with Sjögren's disease: association with systemic disease activity at diagnosis

L. Delgado[1,2], A. Flores-Chávez[2], A. Szántó[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], C. Fugmann[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], H. Nakamura[26], M. López-Dupla[27], M. Vazquez[28], M. Akasbi[29], G. Policarpo Torres[30], B. de Miguel Campo[31], M.R. Suru[1], C. Vericat Queralt[1], I.F. Horváth[3], I. Fischetti[4], A. Vissink[32], L. Kilic[7], V. Manfrè[8], K. Perdan Pirkmajer[9,33], Y. Fujisawa[10], M. Bandeira[12], K. Proc[13], R. Gerli[14], C. Kirana[15], N. Nardi[2], M. Ramos-Casals[2,34], P. Brito-Zerón[1] on behalf of the Sjögren Big Data Consortium

Abstract

Objective

This study aimed to analyse the relationship between distinct autoantibody combinations (immunological signatures) and systemic disease activity in patients with Sjögren's disease (SjD). The hypothesis was that specific multi-autoantibody signatures would be associated with higher systemic disease activity at diagnosis, serving as predictors of a more severe disease course.

Methods

A retrospective observational study was conducted using data from the Big Data Sjögren Project Consortium, an international multicentre registry. The serological status (positive/negative) at diagnosis for ANA, RF, anti-Ro, and anti-La was recorded for each patient. Systemic disease activity was assessed using the EULAR Sjögren's Syndrome Disease Activity Index (ESSDAI) and a simplified Disease Activity Score (DAS) categorised as low, moderate, or high. Statistical analyses included pairwise comparisons, a sensitivity analysis grouping signatures by the number of positive antibodies, and demographic-adjusted ordinal models.

Results

Serum autoantibodies were highly prevalent, with over 94% of patients having at least one autoantibody. The mean ESSDAI values varied significantly across signatures. The fully seronegative group had the lowest mean ESSDAI at 3.61, while the fully seropositive group (ANA+/Ro+/La+/RF+) had the highest among common phenotypes, with a mean of 7.93. A strong dose-response relationship was observed, with each additional positive autoantibody associated with a 1.11-point mean increase in ESSDAI and a 35% increase in the odds of being in a higher DAS category. The rarest signatures, such as ANA−/Ro−/La+/RF+, exhibited the highest mean systemic activity (mean 13.20).

Conclusion

The number and combination of SjD-related autoantibodies at diagnosis are robustly associated with systemic disease activity. Multi-positive profiles, particularly those combining RF with anti-Ro, identify patients at higher risk of systemic activity. Interpreting combined serological patterns offers an immediate, low-cost method for patient stratification and can help guide clinical management.

Key words

Sjögren's disease, Sjögren's syndrome, ESSDAI, EULAR, autoantibodies, immunological signatures, rheumatoid factor, anti-Ro, anti-La, ANA, systemic disease activity

Affiliations: see page 2131.

Lucía Delgado, MD, PhD
Alejandra Flores-Chávez, MD, PhD
Antónia Szántó, 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
Elena Bartoloni, MD
Maureen Rischmueller, MD
Roser Solans, MD, PhD
Sandra G. Pasoto, MD, PhD
Cecilia Fugmann, MD
Isabel Sánchez Berná, MD
Francesco Carubbi, MD, PhD
Virginia Fernandes Moça Trevisani, MD
Valeria Valim, MD
Sheila Melchor, MD
Brenda Maure Noia, PhD
Eva Fonseca-Aizpuru, MD
Hideki Nakamura, MD, PhD
Miguel López-Dupla, MD
Marcos Vazquez, MD
Miriam Akasbi, MD
Guillem Policarpo Torres, MD, PhD
Borja de Miguel Campo, MD, PhD
Mihaela Roxana Suru, MD, MBA
Carmen Vericat Queralt, MD
Ildiko Fanny Horváth, MD
Ilenia Fischetti, MD
Arjan Vissink, DDS, MD, PhD
Levent Kilic, MD
Valeria Manfrè, MD
Katja Perdan Pirkmajer, MD, PhD
Yuhei Fujisawa MD
Matilde Bandeira, MD
Krzysztof Proc, MD
Roberto Gerli, MD
Chandra Kirana, PhD
Norma Nardi, MD
Manuel Ramos-Casals, MD
Pilar Brito-Zerón, 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 2131.
Competing interests: none declared.

Introduction

Sjögren's disease (SjD) is a systemic autoimmune disease characterised by lymphocytic infiltration of exocrine glands, leading to dry eyes and dry mouth (1, 2), along with a spectrum of extra-glandular manifestations (3). Beyond glandular symptoms, up to half of patients develop organ-specific systemic features such as arthritis, vasculitis, neuropathy, and renal or pulmonary involvement, which confer significant morbidity (4). Clinical heterogeneity in SjD poses challenges for prognosis and management, prompting the need for reliable biomarkers to predict disease course and severity (5).

Circulating autoantibodies are a hallmark of SjD and are included in classification criteria (6, 7). The most frequent are antinuclear antibodies (ANA) and anti-Ro, present in the majority of patients, followed by rheumatoid factor (RF) and anti-La (8). Notably, over 95% of patients have at least one of these antibodies, and about 15–20% carry three or more in combination. Serological status has long been observed to correlate with disease phenotype: patients who are seronegative for Ro/La often have milder, glandular-limited disease (9, 10), whereas those who are positive for anti-Ro are more likely to exhibit extra-glandular manifestations (11). In particular, RF, a marker of polyclonal B-cell activation, has been associated with more severe systemic involvement in SjD (12, 13). Previous smaller studies suggested that seropositive patients tend to have higher disease activity and worse outcomes than seronegative patients (5). However, the combined prognostic impact of multiple concurrent autoantibodies has not been quantified on a large scale. It remains unclear whether certain combinations of autoantibodies act synergistically to drive a more aggressive SjD phenotype.

To address this gap, we leveraged the Sjögren Big Data Consortium, the largest international registry of SjD, to analyse how distinct autoantibody combinations ('immunological signatures' relate to systemic disease activity. We focused on the four classic serological markers (ANA, RF, anti-Ro, and anti-La), which represent the core immunological profile in SjD. We hypothesised that specific multi-autoantibody signatures would be associated with higher systemic disease activity at diagnosis, serving as predictors of a more severe disease course. By identifying these high-risk immunological phenotypes, this study aims to enhance risk stratification in SjD and place the basis for personalised management approaches based on serological profiling.

Methods

Study design

We conducted a cross-sectional observational study using data from the Big Data Sjögren Project Consortium, an international multicentre registry of patients fulfilling current international classification criteria (the 2002 American-European Consensus Group criteria and/or the 2016 ACR-EULAR criteria) (6, 7). For each patient, we recorded the serological status (positive/negative) at diagnosis for ANA, RF, anti-Ro, and anti-La. Sixteen possible serological combinations (patterns) can be formed by the presence/absence of the four markers, ranging from completely seronegative (ANA- RF- Ro- La-) to quadruply positive (ANA+ RF+ Ro+ La+). Systemic disease activity at diagnosis was evaluated using the EULAR Sjögren's Syndrome Disease Activity Index (ESSDAI) and the simplified Disease Activity Score (DAS), categorised into three levels corresponding to ESSDAI scores of <5 (low disease activity), 5-13 (moderate activity), and ≥14 (high disease activity) (14-16).

Statistical analysis

We assessed the predictive value of immunological signatures for systemic disease activity by building prediction models for both the mean ESSDAI score and the three-level DAS at diagnosis. Pairwise comparisons against the fully seronegative reference group were performed using two-sided Mann-Whitney U-tests, with Holm adjustment for multiple testing. Cliff's delta was calculated as a non-parametric effect size. To explore whether systemic activity varied according to the cumulative number of positive autoantibodies, a sensitivity analysis was carried out

grouping signatures into five ordinal groups (all four antibodies negative, one positive, two positives, three positives, and four positives). To quantify a predictive trend, we modelled ESSDAI as a function of the 0–4 positivity count using ordinary least squares regression with heteroscedasticity-robust (HC3) standard errors. Monotonicity was evaluated using the Spearman rank correlation.

We finally evaluated the predictive value of immunological signatures for systemic disease activity using demographic-adjusted ordinal models for the three-level DAS outcome (Low, Moderate, High). Three cumulative-logit proportional-odds models were fitted: model A (demographic variables only -age, sex, ethnicity-), model B (model A plus main effects for ANA, Ro, La, RF), and model C (model B plus two-way autoantibody interactions, ANA×Ro, Ro×RF, Ro×La-) to reflect Ro-related context and common serological clustering. Model performance was evaluated using stratified three-fold cross-validation to preserve class balance. For each fold, we calculated accuracy, multiclass log loss, multiclass Brier score, and AUCs for two binary endpoints: Moderate/high *vs*. low and high *vs*. low/moderate. Metrics were averaged across folds with corresponding standard deviations. Ordered models were fitted using BFGS optimisation in *statsmodels* with default convergence settings. Receiver operating characteristic (ROC) curves were generated to compare the discrimination performance of three proportional-odds ordinal logistic regression models for the Disease Activity Score (DAS) endpoints. Two binary endpoints were derived from the original ordinal DAS categories: moderate/high *vs*. low, and high *vs*. low/moderate. Area under the ROC curve (AUC) values were obtained from cross-validated model predictions.

All statistical analyses were developed in collaboration with a large language model (GPT-5, OpenAI, San Francisco, CA, USA) via the ChatGPT Pro platform (August 2025). Analyses were performed in a secure, offline Python 3.10.x environment with no internet access, using only de-identified data in compliance with the EU General Data Protection Regulation. 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 following open-source libraries were employed: *pandas* (v. 1.5.x) and *numpy* (v. 1.23.x) for data handling; *scipy* (v. 1.10.x) for non-parametric tests and correlations; *statsmodels* (v. 0.14.x) for generalised linear and proportional-odds models; *scikit-learn* (v. 1.2.x) for cross-validation and performance metrics; *matplotlib* (v. 3.7.x) and *seaborn* (v. 0.12.x) for data visualisation; *pingouin* (v. 0.5.x) for effect size statistics; and *openpyxl* (v. 3.1.x) for Excel export. Optimisation was performed using BFGS routines in *statsmodels*, with random seeds fixed for reproducibility.

Table I. Summary of the prevalence of each immunological profile.

Immunological signature	n	%	Mean ESSDAI score	SD	Median	IQR
ANA+/Ro+/La+/RF+	3419	25.34	7.93	8.29	6	9
ANA+/Ro+/La-/RF-	2514	18.63	5.35	6.09	4	6
ANA+/Ro+/La-/RF+	1693	12.55	7.06	7.29	5	8
ANA+/Ro+/La+/RF-	1665	12.34	7.3	8.74	5	8
ANA+/Ro-/La-/RF-	1338	9.92	4.69	5.7	3	6
ANA-/Ro-/La-/RF-	967	7.17	3.61	4.89	2	5
ANA-/Ro+/La-/RF-	509	3.77	4.75	5.59	3	7
ANA+/Ro-/La-/RF+	491	3.64	7.26	7.81	5	8.5
ANA-/Ro-/La-/RF+	199	1.48	4.82	5.31	3	6.5
ANA-/Ro+/La-/RF+	157	1.16	6.74	6.25	5	8
ANA-/Ro+/La+/RF-	151	1.12	5.61	6.06	4	8
ANA+/Ro-/La+/RF-	136	1.01	6.03	7	4	7
ANA-/Ro+/La+/RF+	110	0.82	6.92	7.01	5	7
ANA-/Ro-/La+/RF-	79	0.59	4.99	5.1	4	8
ANA+/Ro-/La+/RF+	58	0.43	9.19	9	6.5	9.75
ANA-/Ro-/La+/RF+	5	0.04	13.2	7.82	13	6

Results

Autoantibody prevalence and immunological signatures

Serum autoantibodies were prevalent in this cohort: ANA was positive in 83.8% of tested patients, and anti-Ro in 75.9% (48). RF and anti-La were each positive in roughly 42% of patients (49). Notably, <6% of patients had none of these antibodies (*i.e*. were seronegative for ANA, RF, Ro and La) (51). In contrast, 94% had at least one autoantibody: approximately 73% had two or more positive autoantibodies, about 15–20% of patients had *triple* positivity and ~5% had *quadruple* positivity (all four markers).

Among the 13,491 patients with complete data on the four antibodies and ESSDAI, the most prevalent profile was ANA+/Ro+/La+/RF+ with 3,419 patients (25.34%), followed by ANA+/Ro+/La-/RF- with 2,514 (18.63%), ANA+/Ro+/La-/RF+ with 1,693 (12.55%), and ANA+/ Ro+/La+/RF- with 1,665 (12.34%). When grouping patients by the number of positive antibodies, the cohort contained 3,419 individuals with four positives (25.34%), 3,526 with three (26.14%), 3,454 with two (25.60%), 2,125 with one (15.75%), and 967 with none (7.17%) (Table I).

Association of autoantibody signatures with ESSDAI score

Mean ESSDAI values varied markedly across signatures (Kruskal-Wallis H=698.25, $p<1\times10^{-138}$). Among common phenotypes, the fully seropositive group ANA+/Ro+/La+/RF+ had a mean ESSDAI of 7.93 (n=3,419; 25.34%), followed by ANA+/Ro+/La+/RF– at 7.30 (n=1,665; 12.34%), ANA+/Ro–/La–/RF+ at 7.26 (n=491; 3.64%), and ANA+/Ro+/La–/RF+ at 7.06 (n=1,693; 12.55%). The fully seronegative reference group had the lowest mean ESSDAI at 3.61 (n=967; 7.17%) (Fig. 1).

The Poisson model showed pronounced overdispersion (deviance/df=6.73), confirming the appropriateness of the negative binomial specification

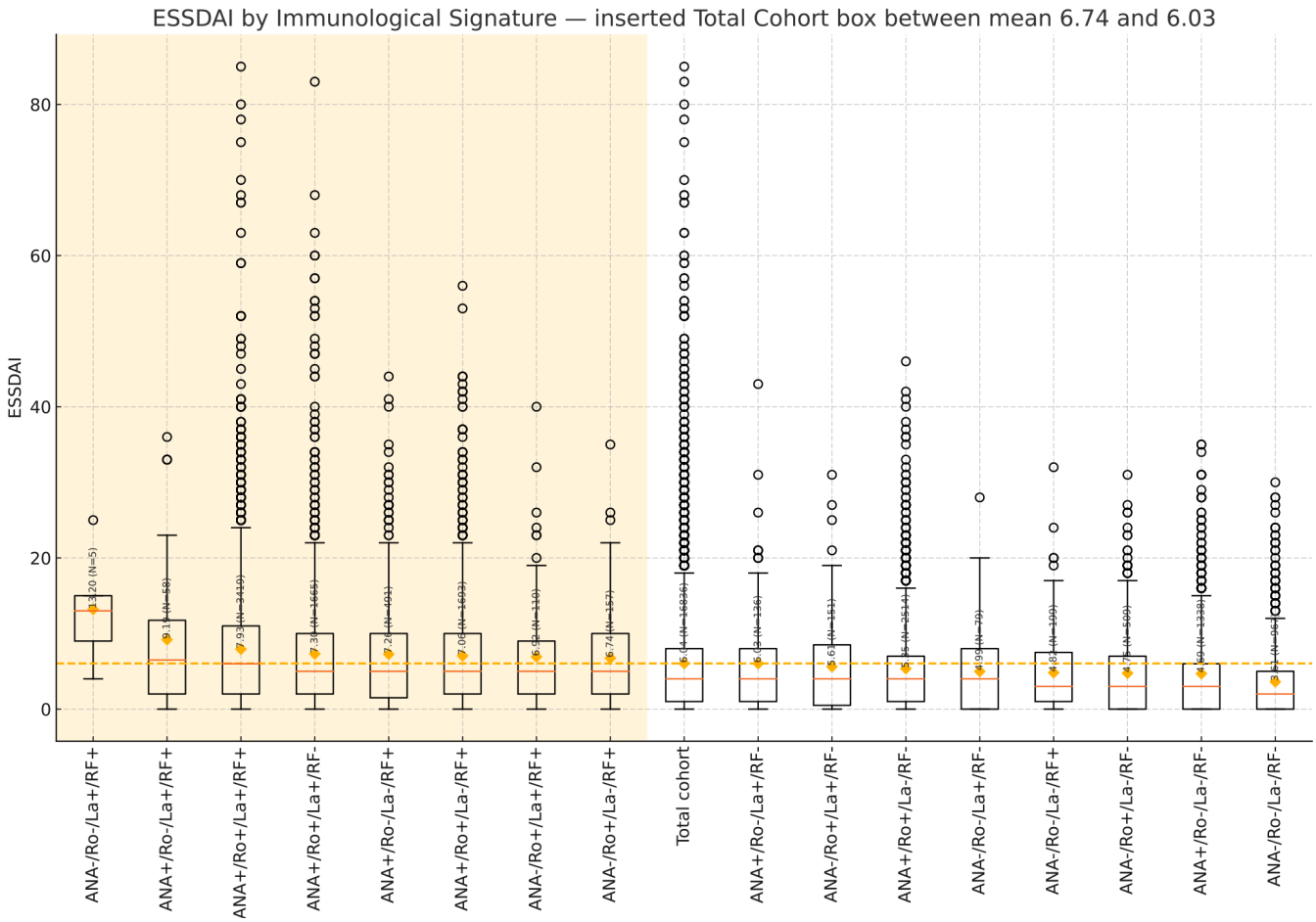

Fig. 1. Boxplots of ESSDAI by immunological signature. Order signatures by median ESSDAI including reference line for cohort median.

(AIC=78,137.02). The rarest signatures exhibited the highest mean systemic activity, with ANA–/Ro–/La+/RF+ reaching a mean of 13.20 (n=5; 0.04%), and ANA+/Ro–/La+/RF+ at 9.19 (n=58; 0.43%).

Using the fully seronegative group (median ESSDAI=2, n=967) as reference, the negative binomial model identified the following signatures most significantly associated with higher systemic activity: ANA+/Ro+/La+/RF+: IRR 2.21 (95% CI 2.04-2.39, p<0.001; n=3,419), ANA+/Ro+/La+/RF–: IRR 2.04 (95% CI 1.87-2.23, p<0.001; n=1,665) and ANA+/Ro–/La–/RF+: IRR 2.03 (95% CI 1.80-2.28, p<0.001; n=491). Holm-adjusted pairwise Mann-Whitney tests confirmed significantly higher ESSDAI scores, with Cliff's delta value up to 0.42 for ANA+/Ro+/La+/RF+. Quantile regression analyses reproduced the direction and relative ranking of signatures, supporting robustness of the findings. The negative binomial model achieved a MAE of 5.06 and RMSE of 7.17 in cross-validation. The rarest signatures exhibited the highest mean systemic activity, with ANA–/Ro–/La+/RF+ reaching a mean of 13.20 (n=5; 0.04%) and IRR 3.77 (95% CI 1.51-9.38, p=0.004; n=5), and ANA+/Ro–/La+/RF+ at 9.19 (n=58; 0.43%) and IRR 2.57 (95% CI 1.94-3.40, p<0.001; n=58).

When autoantibody signatures were split according to the number of concomitant positive markers, in the all-negative group (n=967), ESSDAI averaged 3.61 (SD=4.89) with a median of 2. With one positive marker (n=2,125) the mean was 4.73 (SD=5.61; median 3). With two positives (n=3,454) the mean reached 5.74 (SD=6.45; median 4). With three positives (n=3,526) the mean was 7.20 (SD=8.03; median 5). Patients with four positives (n=3,419) showed the highest activity, mean 7.93 (SD=8.29; median 6) (Fig. 2). The global non-parametric comparison confirmed a statistically significant difference across groups (Kruskal-Wallis H=651.81, p=9.47×10^{-140}), with an effect size of ε^2=0.048, indicating a small-to-moderate magnitude at the population level. The predictive gradient was corroborated by a positive monotonic association between the positivity count and ESSDAI (Spearman ϱ=0.216, p=1.26×10^{-141}). In a simple robust linear model, each additional positive autoantibody was associated with a mean increase of 1.11 points in ESSDAI (p=1.79×10^{-119}; R^2=0.035).

Association of autoantibody signatures with DAS

The distribution of DAS categories differed significantly across immunological signatures (χ^2=490.34, df=30, p=1.15×10^{-84}) (Fig. 3). Among signatures with adequate sample size (≥100 patients), the highest proportions of high activity were observed in ANA+/Ro+/La+/RF+: 608 of 3,419

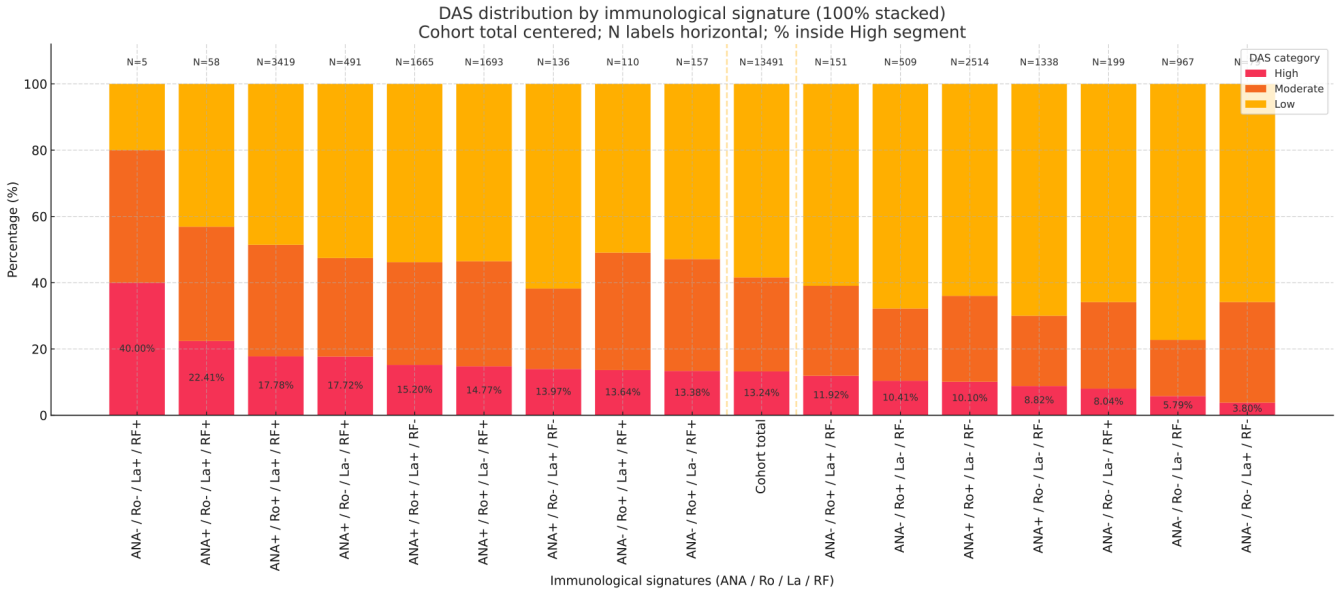

Fig. 2. ESSDAI by number of concomitant positive autoantibodies.

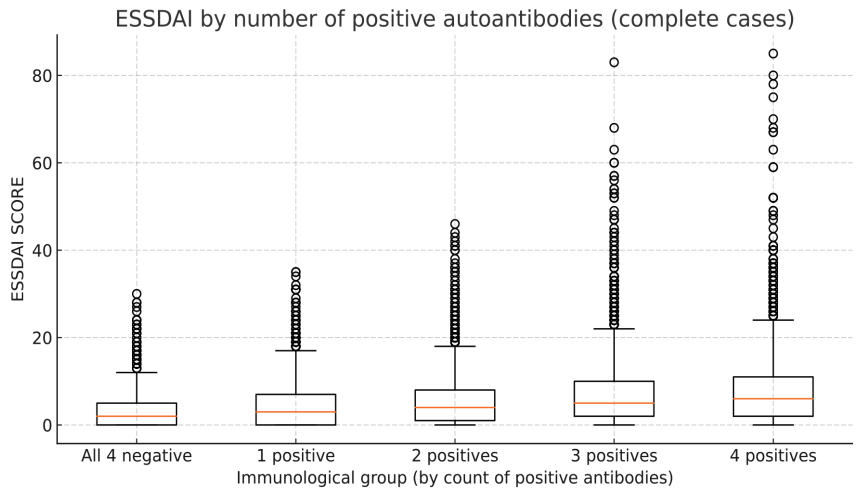

Fig. 3. 100% stacked bar chart of DAS distribution by immunological signature.

(17.78%), ANA+/Ro–/La–/RF+: 87 of 491 (17.72%) and ANA+/Ro+/La+/RF–: 253 of 1,665 (15.20%), while the lowest proportion of high activity was found in ANA–/Ro–/La–/RF–: 56 of 967 (5.79%) (Table II). At the individual marker level, proportional-odds regression demonstrated that positivity for each antibody independently increased the odds of being in a higher DAS category, with adjusted ORs being ANA: 1.32 (95% CI 1.19-1.46; $p=1.60\times10^{-7}$), Ro: 1.25 (95% CI 1.14-1.37; $p=2.50\times10^{-6}$), La: 1.33 (95% CI 1.23-1.43; $p=1.05\times10^{-13}$) and RF: 1.48 (95% CI 1.37-1.58; $p=1.90\times10^{-26}$). There was a strong dose-response relationship when modelling the number of positive antibodies: each additional positive marker increased the odds of higher DAS by 35% (OR 1.35; 95% CI 1.31-1.39; $p=2.30\times10^{-94}$). This monotonic pattern was also evident in the raw proportions of high DAS by the number of positive antibodies: 0 (5.79%), 1 (8.94%), 2 (11.61%), 3 (15.06%), and 4 (17.78%).

Model A (demographics only) had limited discriminatory ability: accuracy 51.56% (SD=0.24%), log loss 0.97 (SD=0.00), Brier score 0.59 (SD=0.00), AUC=0.54 (SD=0.01) for moderate/high *vs*. low, and 0.53 (SD=0.01) for high *vs*. low/moderate.

Adding the four antibody main effects (Model B) improved all metrics: accuracy 51.82% (SD=0.34%), log loss 0.96 (SD=0.00), Brier score 0.58 (SD=0.00), AUC=0.61 (SD=0.01) for moderate/high *vs*. low and 0.60 (SD=0.01) for high *vs*. low/moderate, absolute gains of 0.07 in both AUCs compared with Model A.

Including the interaction terms (Model C) yielded only marginal additional improvement: accuracy 52.03% (SD=0.27%), log loss 0.96 (SD=0.00), Brier score 0.58 (SD=0.00), with unchanged AUCs (0.61 and 0.60).

With respect to decision-curve analysis, for moderate/high *vs*. low, Models B and C had similar net benefits, both surpassing treat-all from mid-range thresholds upward. At a 0.40 threshold, net benefit was 0.16 for Models B and C *vs*. 0.14 for treat-all (absolute gain ≈ 0.02). At a 0.30 threshold, gains were negligible, reflecting the high baseline prevalence. For high *vs*. low/moderate, treat-all yielded negative net benefit across the plausible range, while Model C maintained small positive net benefit at low-to-intermediate thresholds (*e.g*. at 0.20: 0.00 *vs*. -0.08 for treat-all). Across both endpoints, Model C did not materially outperform Model B, confirming that antibody main effects drive most of the predictive signal (Fig. 4).

Table II. Distribution of DAS categories by immunological signature: Signature, n, Low DAS (n, %), Moderate DAS (n, %), High DAS (n, %).

Immunological signature	n	Low DAS	low_%	Moderate DAS	moderate_%	High DAS	high_%
ANA+ / Ro+ / La+ / RF+	3419	1660	48.55	1151	33.66	608	17.78
ANA+ / Ro+ / La- / RF-	2514	1608	63.96	652	25.93	254	10.1
ANA+ / Ro+ / La- / RF+	1693	906	53.51	537	31.72	250	14.77
ANA+ / Ro+ / La+ / RF-	1665	896	53.81	516	30.99	253	15.2
ANA+ / Ro- / La- / RF-	1338	936	69.96	284	21.23	118	8.82
ANA- / Ro- / La- / RF-	967	747	77.25	164	16.96	56	5.79
ANA- / Ro+ / La- / RF-	509	345	67.78	111	21.81	53	10.41
ANA+ / Ro- / La- / RF+	491	258	52.55	146	29.74	87	17.72
ANA- / Ro- / La- / RF+	199	131	65.83	52	26.13	16	8.04
ANA- / Ro+ / La- / RF+	157	83	52.87	53	33.76	21	13.38
ANA- / Ro+ / La+ / RF-	151	92	60.93	41	27.15	18	11.92
ANA+ / Ro- / La+ / RF-	136	84	61.76	33	24.26	19	13.97
ANA- / Ro+ / La+ / RF+	110	56	50.91	39	35.45	15	13.64
ANA- / Ro- / La+ / RF-	79	52	65.82	24	30.38	3	3.8
ANA+ / Ro- / La+ / RF+	58	25	43.1	20	34.48	13	22.41
ANA- / Ro- / La+ / RF+	5	1	20	2	40	2	40

Discussion

In this international multicentre cohort of patients with primary SjD, combinatorial serological profiles composed of ANA, RF, anti-Ro/SSA and anti-La/SSB were robustly associated with systemic disease activity at diagnosis. Systemic activity, assessed both continuously with ESSDAI and categorically using DAS thresholds, displayed a consistent and monotonic relationship with the number and composition of positive autoantibodies. Patients who were negative for all four immunological markers had the lowest mean ESSDAI, whereas those with three or four positive markers had progressively higher means; RF-containing signatures and patterns that included anti-Ro repeatedly ranked among the highest-activity groups. In multivariable ordinal models each antibody retained an independent association with higher activity, and a simple antibody-count metric provided an interpretable correlation: each additional positive marker increased the odds of a higher DAS category by a clinically meaningful increment. These associations were reproduced across non-parametric comparisons, negative-binomial regression and cross-validated proportional-odds models, supporting the biological coherence and robustness of the observations.

These results extend and quantify prior findings from large international series that immunological status strongly shapes the presenting phenotype in SjD. Analyses from the Big Data Sjögren Project and other multinational cohorts previously documented that anti-Ro-positive patients have higher frequencies of extra-glandular manifestations and that seronegative individuals more often present with gland-predominant disease (17-19). Our study complements those reports by systematically evaluating all 16 exact four-marker combinations and by demonstrating that not all seropositive patients are equivalent: specific combinations, particularly those pairing RF with anti-Ro (with or without anti-La), delineate subgroups with substantially greater systemic activity. This pattern merges with mecha-

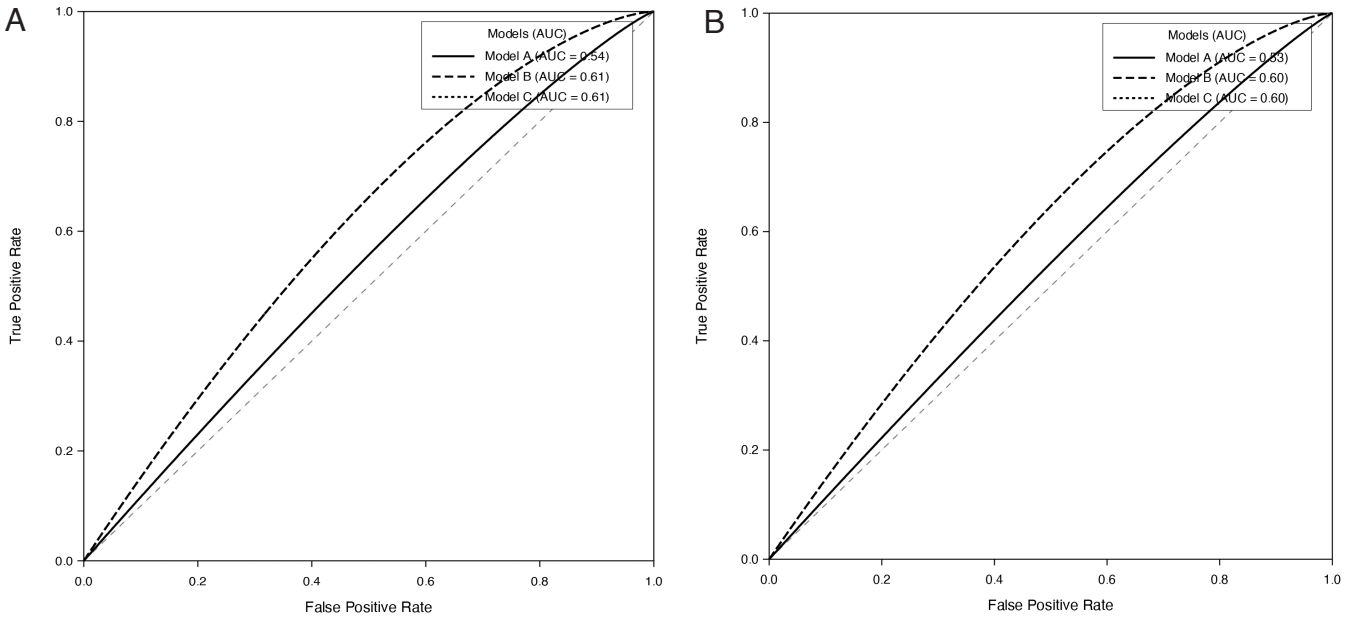

Fig. 4. ROC curves for ordinal DAS prediction models. ROC curves showing the discrimination performance of three proportional-odds models for (**A**) moderate/high *vs*. low DAS (panel a) and (**B**) high *vs*. low/moderate DAS (panel b). Model A included demographics only (age, sex, ethnicity). Model B added main effects for ANA, Ro, La, and RF. Model C further included predefined two-way interaction terms (ANA×Ro, Ro×RF, Ro×La). Curves were generated from the analytical form $y = 1 - (1 - x)^k$ to match the observed AUC values exactly (Model A: 0.54 and 0.53; Models B and C: 0.61 and 0.60 for the two endpoints, respectively). The diagonal dashed line represents chance discrimination.

nistic models in which B-cell hyperactivity, immune-complex formation and interferon pathway activation converge to amplify systemic inflammation (20). Indeed, prior evidence linking cryoglobulinemia and hypocomplementaemia to aggressive clinical phenotypes lends biological plausibility to why RF-containing signatures are associated with a higher likelihood of systemic activity at diagnosis (21-24), even though those additional markers were not uniformly available in our dataset.

Although our analyses are cross-sectional and therefore associative, the observed gradients should be viewed as hypothesis-generating with respect to prognosis and require confirmation in prospective cohorts before any predictive claims can be made. The monotonic increase in ESSDAI with antibody burden, the consistency of effect estimates across analytical approaches, and the improvement in cross-sectional discrimination when serology is added to demographic models (AUC gains from ~0.54 to ~0.61 for moderate/high *vs*. low DAS) suggest that serology contributes independent, reproducible information beyond basic clinical variables. Nonetheless, individual-level discrimination remains modest: serology is a valuable tool for population-level stratification at diagnosis and triage, but, on its own, cannot support precise individual prognostication; prospective validation is needed to establish predictive performance over time. Crucially, 'low risk' does not mean 'no risk': a minority of seronegative patients did present with high systemic activity, underscoring the need to interpret serology within the full clinical context (10).

From a clinical standpoint, the implications are immediately applicable to baseline assessment and monitoring, while therapeutic decisions require individualised clinical judgment and prospective validation. Since ANA, RF, anti-Ro and anti-La are routinely measured at diagnosis, interpreting their combined pattern rather than each marker in isolation offers an immediate, low-cost means to stratify patients. Patients with multi-positive profiles, especially those with RF plus anti-Ro, should prompt clinicians to adopt a lower threshold for targeted screening for extra-glandular involvement and for laboratory evaluation of complements and cryoglobulins when available (25-27). Such patients warrant closer follow-up and a lower threshold for initiating systemic therapies if organ-threatening features arise. Conversely, completely seronegative patients typically present with a gland-predominant course at diagnosis and may often be managed initially with symptomatic measures and standard surveillance; nonetheless, clinicians should remain vigilant because a minority present with, or subsequently develop, systemic manifestations (28). The antibody-count gradient provides a simple, pragmatic rubric for triage, communication with patients about expected risk, and planning the intensity of monitoring.

The study's strengths include the very large multicentre sample, standardised assessment of systemic activity with ESSDAI and explicit DAS categories, and analytic triangulation across multiple statistical frameworks that handle different data properties (non-parametric testing, negative-binomial models for over-dispersed counts, quantile regression and cross-validated proportional-odds models). Evaluating all exact four-marker signatures and the cumulative antibody count allowed a comprehensive mapping of serological phenotypes at diagnosis. Nevertheless, some limitations must be acknowledged. The retrospective, cross-sectional design precludes causal inference and cannot replace prospective validation of longitudinal outcomes such as damage accrual or lymphoma risk. Accordingly, any prognostic interpretation of baseline serological signatures should be considered hypothesis-generating until confirmed in longitudinal cohorts. These design features limit causal inference and restrict any directionality claims to mechanistic plausibility rather than proof. The requirement for complete four-marker serology and ESSDAI may induce selection bias by under-representing milder community-managed cases or centres with incomplete testing. We analysed antibodies as binary variables; quantitative titres, epitope specificities or subclass information, potentially prognostic, were not uniformly available. Moreover, data for complementary prognostic markers (complement components, cryoglobulins, detailed B-cell biomarkers) were incomplete, and omitting these may limit the maximal predictive performance and leave residual confounding. Collinearity among antibodies (for example, frequent co-occurrence of anti-Ro and anti-La) makes it difficult to attribute causal weight to any single marker and emphasises the utility of considering combinatorial signatures rather than isolated tests. Finally, while the cohort was international, external validation across under-represented healthcare settings and ancestries is needed before routine guideline adoption. External validation in geographically and ethnically diverse cohorts, alongside calibration of absolute risk estimates, will be necessary before serology-based algorithms can be considered for routine use.

Looking ahead, the serology-based signatures identified here can be directly leveraged in prospective research. Longitudinal cohorts should pre-specify strata based on baseline antibody burden (*e.g*. 0-1 *vs*. 2 *vs*. 3-4 positive markers) and the presence of RF+anti-Ro combinations to evaluate trajectories of ESSDAI, flare rates, treatment escalation, and damage accrual. In interventional trials, these strata can be used a priori as randomisation factors or enrichment criteria to reduce outcome heterogeneity and increase power, while enabling prespecified tests of effect modification for therapies targeting B-cell activation or interferon pathways (32). Event-driven endpoints (*e.g*. time to systemic escalation or organ-specific events) and repeated measures ESSDAI trajectories can be tailored to baseline risk within each stratum, anchoring sample size calculations to stratum-specific event rates. Embedding biospecimen collection aligned with these strata will facilitate validation of mechanistic surrogates and the development of pragmatic composite scores that combine antibody count with a small set of high-value biomarkers. Importantly, serology should be evaluated as a stratification factor rather than a surrogate endpoint;

head-to-head comparisons of serology enriched *versus* unselected designs will clarify its value for trial efficiency and generalisability. Such prospective work will determine whether serology-based stratification can move from risk communication at diagnosis to guiding trial design and, ultimately, individualised treatment selection.

The results obtained in this study provide a foundation for developing explainable predictive models; emerging AI methods could integrate these serological signatures with multi-omic data, but prospective and transparent evaluation will be essential. Integrating immunophenotypes with molecular signatures, particularly type I interferon gene-expression signatures and B-cell activation markers such as BAFF, free light chains or β2-microglobulin, may delineate pathobiological endotypes (interferon-dominant *vs*. B-cell-dominant) and improve organ-specific risk prediction (29-31). AI-related approaches that combine serology, routine laboratory tests, histopathology and multi-omic biomarkers could reveal novel clusters and produce individualised risk projections, although simpler scores (*e.g*. antibody count plus a few high-value SjD-related biomarkers) may offer the best balance between interpretability and clinical value for everyday use (33).

In conclusion, the arrangement and number of SjD-related autoantibodies at diagnosis delineate immunological signatures that map onto systemic disease burden. Multi-positive profiles, particularly those combining RF with anti-Ro, identify patients at higher risk of systemic activity and thus warrant intensified surveillance and a lower threshold for diagnostic and therapeutic escalation when organ-threatening features are suspected. Integrating serology with molecular biomarkers and prospective outcome data would support personalised management strategies for patients with SjD.

Affiliations

[1]Autoimmune Diseases Unit, Department of Internal Medicine, Hospital CIMA Sanitas, Barcelona, Spain; [2]Postgraduate Degree 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 São 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 São 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*.: 2023 International Rome consensus for the nomenclature of Sjögren disease. *Nat Rev Rheumatol* 2025; 21: 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 Prim* 2016; 2: 16047. https://doi.org/10.1038/nrdp.2016.47
3. MARIETTE X, CRISWELL LA: Primary Sjögren's syndrome. *N Engl J Med* 2018; 378: 931-9. https://doi.org/10.1056/NEJMcp1702514
4. BALDINI C, PEPE P, QUARTUCCIO L *et al*.: Primary Sjögren's syndrome as a multi-organ disease: impact of the serological profile on the clinical presentation of the disease in a large cohort of Italian patients. *Rheumatol-*

ogy 2013; 53: 839-44. https://doi.org/10.1093/rheumatology/ket427
5. BRITO-ZERON 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.
6. 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: 554-8. https://doi.org/10.1136/ard.61.6.554
7. 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: 35-45. https://doi.org/10.1002/art.39859
8. BRITO-ZERÓN P, ACAR-DENIZLI N, NG W-F *et al*.: Epidemiological profile and north-south gradient driving baseline systemic involvement of primary Sjögren's syndrome. *Rheumatology* (Oxford) 2020; 59: 2350-59. https://doi.org/10.1093/rheumatology/kez578
9. LAN J, DENG C, HUANG H *et al*.: Association of immunological features with clinical manifestations in primary Sjögren's syndrome: a single-center cross-sectional study. *Clin Exp Med* 2024; 22: 613-20. https://doi.org/10.1007/s10238-021-00772-2
10. CHATZIS LG, PEZOULAS V, VOULGARI PV *et al*.: Combined seronegativity in Sjögren's syndrome. *Clin Exp Rheumatol* 2021; 39 (Suppl. 133): S80-84. https://doi.org/10.55563/clinexprheumatol/47a4kr
11. FLORES-CHAVEZ A, KOSTOV B, SOLANS R *et al*.: Severe, life-threatening phenotype of primary Sjögren's syndrome: clinical characterisation and outcomes in 1580 patients (GEAS-SS Registry). *Clin Exp Rheumatol* 2018; 36 (Suppl. 112): S121-29.
12. MAŚLIŃSKA M, MAŃCZAK M, KWIATKOWSKA B: Usefulness of rheumatoid factor as an immunological and prognostic marker in PSS patients. *Clin Rheumatol* 2019; 38: 1301-7. https://doi.org/10.1007/s10067-019-04438-z
13. ARADI Z, BÓI B, NAGY G *et al*.: Additive value of rheumatoid factor isotypes in Sjögren's syndrome patients with joint complaints of different etiologies – can rheumatoid factor IgA serve as an early, poor prognostic biomarker candidate? *Int J Mol Sci* 2025; 26(10): 4797. https://doi.org/10.3390/ijms26104797
14. SEROR R, BOWMAN SJ, BRITO-ZERON P *et al*.: EULAR Sjögren's syndrome disease activity index (ESSDAI): a user guide. *RMD Open* 2015; 20: e000022. https://doi.org/10.1136/rmdopen-2014-000022
15. SEROR R, THEANDER E, BRUN JG *et al*.: Validation of EULAR primary Sjögren's syndrome disease activity (ESSDAI) and patient indexes (ESSPRI). *Ann Rheum Dis* 2015; 74: 859-66. https://doi.org/10.1136/annrheumdis-2013-204615
16. SEROR R, BOOTSMA H, SARAUX A *et al*.: Defining disease activity states and clinically meaningful improvement in primary Sjögren's syndrome with EULAR primary Sjögren's syndrome disease activity (ESSDAI) and patient-reported indexes (ESSPRI). *Ann Rheum Dis* 2016; 75: 382-89. https://doi.org/10.1136/annrheumdis-2014-206008
17. NEGRINI S, EMMI G, GRECO M *et al*.: Sjögren's syndrome: a systemic autoimmune disease. *Clin Exp Med* 2022; 22: 9-25. https://doi.org/10.1007/s10238-021-00728-6
18. BRITO-ZERON P, ACAR-DENIZLI N, NG W-F *et al*.: How immunological profile drives clinical phenotype of primary Sjögren's syndrome at diagnosis: analysis of 10,500 patients (Sjögren Big Data Project). *Clin Exp Rheumatol* 2018; 36 (Suppl. 112): S102-12.
19. BRITO-ZERÓN P, KOSTOV B, SOLANS R *et al*.: Systemic activity and mortality in primary Sjögren syndrome: predicting survival using the EULAR-SS Disease Activity Index (ESSDAI) in 1045 patients. *Ann Rheum Dis* 2016; 75: 348-55. https://doi.org/10.1136/annrheumdis-2014-206418
20. BODEWES ILA, AL-ALI S, VAN HELDEN-MEEUWSEN CG *et al*.: Systemic interferon type I and type II signatures in primary Sjögren's syndrome reveal differences in biological disease activity. *Rheumatology* (Oxford) 2018; 57: 921-30. https://doi.org/10.1093/rheumatology/kex490
21. BRITO-ZERÓN P, FLORES-CHÁVEZ A, HORVÁTH IF *et al*.: Mortality risk factors in primary Sjögren syndrome: a real-world, retrospective, cohort study. *EClinicalMedicine* 2023; 61: 102062. https://doi.org/10.1016/j.eclinm.2023.102062
22. RETAMOZO S, GHEITASI H, QUARTUCCIO L *et al*.: Cryoglobulinaemic vasculitis at diagnosis predicts mortality in primary Sjögren syndrome: analysis of 515 patients. *Rheumatology* (Oxford) 2016; 55(8): 1443-51. https://doi.org/10.1093/rheumatology/kew194
23. 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: 1042-50. https://doi.org/10.1136/annrheumdis-2016-209952
24. NOCTURNE G, VIRONE A, NG W-F *et al*.: Rheumatoid factor and disease activity are independent predictors of lymphoma in primary Sjögren's syndrome. *Arthritis Rheumatol* 2016; 68(4): 977-85. https://doi.org/10.1002/art.39518
25. NGUYEN Y, NOCTURNE G, HENRY J *et al*.: Identification of distinct subgroups of Sjögren's disease by cluster analysis based on clinical and biological manifestations: data from the cross-sectional Paris-Saclay and the prospective ASSESS cohorts. *Lancet Rheumatol* 2024; 6: e216-25. https://doi.org/10.1016/S2665-9913(23)00340-5
26. ARGYROPOULOU OD, PEZOULAS V, CHATZIS L *et al*.: Cryoglobulinemic vasculitis in primary Sjögren's syndrome: Clinical presentation, association with lymphoma and comparison with Hepatitis C-related disease. *Semin Arthritis Rheum* 2020; 50: 846-53. https://doi.org/10.1016/j.semarthrit.2020.07.013
27. BRITO-ZERÓN P, KOSTOV B, FRAILE G *et al*.: Characterization and risk estimate of cancer in patients with primary Sjögren syndrome. *J Hematol Oncol* 2017; 10: 90. https://doi.org/10.1186/s13045-017-0464-5
28. DE WOLFF L, VERSTAPPEN G, STEL A *et al*.: Systemic disease activity measured with ESSDAI varies largely over 5 years in a prospective, longitudinal cohort of patients with Sjögren's disease. *RMD Open* 2025; 11(1): e004753. https://doi.org/10.1136/rmdopen-2024-004753
29. BETTACCHIOLI E, SARAUX A, TISON A *et al*.: Association of combined anti-Ro52/TRIM21 and anti-Ro60/SSA antibodies with increased Sjögren disease severity through interferon pathway activation. *Arthritis Rheumatol* 2024; 76: 751-62. https://doi.org/10.1002/art.42789
30. MALEKI-FISCHBACH M, ANDERSON K, FERNÁNDEZ PÉREZ ER: Transcriptomic profiling of peripheral B cells in antibody positive Sjögren's patients reveals interferon signature. *Genes* (Basel) 2024; 15(5): 628. https://doi.org/10.3390/genes15050628
31. NORDMARK G, WANG C, VASAITIS L *et al*.: Association of genes in the NF-κB pathway with antibody-positive primary Sjögren's syndrome. *Scand J Immunol* 2013; 78: 447-54. https://doi.org/10.1111/sji.12101
32. BALDINI C, CHATZIS LG, FULVIO G, LA ROCCA G, PONTARINI E, BOMBARDIERI M: Pathogenesis of Sjögren's disease: one year in review 2024. *Clin Exp Rheumatol* 202; 42(12): 2336-43. https://doi.org/10.55563/clinexprheumatol/i8iszc
33. LIU J, FENG J, ZHU F *et al*.: Analysis of the relationships between interferon-stimulated genes and anti-SSA/Ro 60 antibodies in primary Sjögren's syndrome patients via multiomics and machine learning methods. *Int Immunopharmacol* 2025; 144: 113652. https://doi.org/10.1016/j.intimp.2024.113652