Arthritis Research & Therapy

RESEARCH

Open Access

Predicting autoimmune thyroiditis in primary Sjogren's syndrome patients using a random forest classifier: a retrospective study

Jia-yun Wu[1†], Jing-yu Zhang[1†], Wen-qi Xia[1], Yue-ning Kang[1], Ru-yi Liao[1], Yu-ling Chen[1], Xiao-min Li[1], Ya Wen[1], Fan-xuan Meng[1], Li-ling Xu[1], Sheng-hui Wen[1], Hui-fen Liu[1], Yuan-qing Li[2,3], Jie-ruo Gu[1,4*], Qing Lv[1*] and Yong Ren[5,6*]

Abstract

Background Primary Sjogren's syndrome (pSS) and autoimmune thyroiditis (AIT) share overlapping genetic and immunological profiles. This retrospective study evaluates the efficacy of machine learning algorithms, with a focus on the Random Forest Classifier, to predict the presence of thyroid-specific autoantibodies (TPOAb and TgAb) in pSS patients.

Methods A total of 96 patients with pSS were included in the retrospective study. All participants underwent a complete clinical and laboratory evaluation. All participants underwent thyroid function tests, including TPOAb and TgAb, and were accordingly divided into positive and negative thyroid autoantibody groups. Four machine learning algorithms were then used to analyze the risk factors affecting patients with pSS with positive and negative for thyroid autoantibodies.

Results The results indicated that the Random Forest Classifier algorithm (AUC = 0.755) outperformed the other three machine learning algorithms. The random forest classifier indicated Age, IgG, C4 and dry mouth were the main factors influencing the prediction of positive thyroid autoantibodies in pSS patients. It is feasible to predict AIT in pSS using machine learning algorithms.

Conclusions Analyzing clinical and laboratory data from 96 pSS patients, the Random Forest model demonstrated superior performance (AUC = 0.755), identifying age, IgG levels, complement component 4 (C4), and absence of dry mouth as primary predictors. This approach offers a promising tool for early identification and management of AIT in pSS patients.

Trial registration This retrospective study was approved and monitored by the Ethics Committee of The Third Affiliated Hospital of Sun Yat-sen University (No.II2023-254-02).

[†]Jia-yun Wu and Jing-yu Zhang contributed equally to this work.

*Correspondence:
Jie-ruo Gu
gujieruo@163.com
Qing Lv
lvqing@sysush.com
Yong Ren
koalary@qq.com

Full list of author information is available at the end of the article

Keywords Primary Sjogren's syndrome, Autoimmune thyroiditis, Predictors, Machine learning algorithms

Background

Primary Sjogren's syndrome(pSS) is a chronic autoimmune systemic inflammatory disease characterized by involvement of exocrine glands such as salivary glands and lacrimal glands. Because of the interaction between both genetics, hormonal and environmental factors in disease's pathogenesis [1], pSS can also manifest systemically, affecting various organs and potentially leading to complications such as interstitial lung disease, renal tubular acidosis, and lymphoma. This systemic nature and the complex interplay of immunological factors make pSS a multifaceted disease with significant diagnostic and management challenges.

Autoimmune thyroiditis(AIT), including Graves' disease and Hashimoto's thyroiditis(HT), represent the most prevalent form of organ-specific autoimmune disorders. They are characterized by the dysfunction of the thyroid gland due to immune-mediated destruction, often accompanied by the presence of circulating thyroid-specific autoantibodies such as anti-thyroid peroxidase (TPOAb) and anti-thyroglobulin (TgAb) [2]. Like pSS, AIT involves a complex pathogenesis where genetic predisposition interplays with environmental triggers. Multiple studies have found AIT was frequently associated with diffuse connective tissue diseases, especially systemic lupus erythematosus(SLE), rheumatoid arthritis(RA) and pSS [3, 4]. Studies have reported the prevalence of primary Sjogren's syndrome with thyroid involvement is estimated to be approximately as high as 30-40% [5, 6]. The coexistence of pSS and AIT has been well-documented, suggesting shared pathological mechanisms and perhaps a common genetic background predisposing to autoimmunity [2, 7]. For example, SS and AIT both exhibit a familial clustering pattern. Additionally, The expression of HLA-B8 and DR3 in both AIT and pSS. Furthermore, some environmental factors, such as virus infection and vitamin D deficiency, might be implicated in the pathogenesis of AIT and pSS [8, 9]. The frequent overlap of these conditions necessitates a deeper understanding of their co-morbidity, which could lead to more effective screening and management strategies. Besides, there was a 67-fold increased risk of thyroid MALT lymphoma in autoimmune thyroiditis and a 44-fold increased risk of parotid lymphoma in pSS [10]. A high risk of lymphoma in pSS and AIT patients was both reported, which further reduce the quality of life of patients.

Emerging evidence suggests that machine learning techniques can provide significant insights into the complex patterns of autoimmune diseases, potentially predicting disease onset and progression with greater accuracy than traditional statistical methods. For instance, machine learning methods are used to assess comorbidity clusters in RA [11]. Similarly, it may be a promising tool for dissecting the intricate interactions between clinical and laboratory features in pSS patients, predicting the likelihood of concurrent AIT.

This study aims to harness the power of advanced machine learning algorithms to predict the presence of AIT in patients with pSS. By identifying key predictive factors, this research seeks to contribute to the personalized management of pSS, facilitating early intervention and tailored treatment strategies to improve patient outcomes.

Methods

Study design and population

This is a retrospective cross-sectional study was conducted at the Seventh Affiliated Hospital of Sun Yat-sen University, Shenzhen, China, from October 2020 to December 2022. The study was approved and monitored by the Ethics Committee of The Third Affiliated Hospital of Sun Yat-sen University (No.II2023-254-02). Because of the retrospective nature of the study, the requirement for informed consent was waived. All 96 pSS patients who presented to rheumatology Immunology department of our center were consecutively included. All patients were aged between 20 and 70 and met the classification established in 2002 by the American and European Consensus Group [12]. All pSS patients were divided into 2 groups based on TPOAb/TgAb positive or negative.

Clinical and Laboratory assessments

All participants underwent a comprehensive clinical evaluation including symptoms assessment for dry mouth and dry eyes. Basic tear secretion was evaluated by a professional ophthalmologist using the Schirmer test. Laboratory investigations included complete thyroid function tests (TPOAb and TgAb levels determined by chemiluminescent immunoassay), complete blood count, immunoglobulin levels (IgG, IgA, IgM), complement components (C3, C4), and vitamin D (VitD) levels. TPOAb and TgAb were also determined using chemiluminescent immunoassay with an optimum cut-off level of 4.11 IU/ml and 5.61 IU/ml, respectively.

Data preprocessing

Our study utilizes a dataset from patients diagnosed with primary Sjogren's syndrome to predict the presence of AIT, specifically measured by the presence of thyroid-specific antibodies (binary outcome: 0 or 1). The dataset includes 14 features, including Age, Gender, Dryness of

eye, Dryness of mouth, Basic tear secretion, ANA titer, SSB, SSA, VitD, IgG, C4, IgA, C3, IgM.

The data preprocessing steps involved the following 4 points:

- Handling Missing Values: Missing values were imputed using the median for continuous variables and the mode for categorical variables. This method ensures that the imputed values do not skew the data distribution.
- Normalization: Continuous variables were standardized to have zero mean and unit variance. Standardization is essential to ensure that features with larger ranges do not dominate the model training process.
- Encoding: Categorical variables, such as gender, were encoded using one-hot encoding to convert them into a numerical format suitable for machine learning algorithms. This encoding avoids ordinal relationships between categories that could mislead the model.
- Binarization: The target variable (thyroid-specific antibodies) was binarized into 0 and 1. This binarization simplifies the problem into a binary classification task.

Model building

Four machine learning algorithms were employed for classification modeling:

1. Logistic Regression (LR): A linear model commonly used for binary classification tasks. It models the probability of the default class (presence of thyroid-specific antibodies) using a logistic function.
2. Support Vector Machine (SVM): A classifier that finds the optimal hyperplane that separates the data into classes. We used a radial basis function (RBF) kernel to handle non-linear relationships.
3. Random Forest Classifier (RFC): An ensemble learning method that builds multiple decision trees and merges them to get a more accurate and stable prediction.
4. Extreme Gradient Boosting (XGB): An advanced implementation of gradient boosting algorithm that optimizes the model's performance by combining the predictions of multiple weak models to form a strong learner.

The models were trained and evaluated using 5-fold cross-validation to ensure robustness. Cross-validation helps to mitigate overfitting and ensures that the model's performance is consistent across different subsets of the data. The evaluation metrics included the area under the receiver operating characteristic curve (AUC-ROC), accuracy (ACC), positive predictive value (PPV), negative predictive value (NPV), sensitivity (SENS), specificity (SEPC), and F1-score, as depicted in the ROC curve plot and the table of metrics.

Statistical analysis

Statistical analyses were conducted using Python's Scikit-learn library. Differences in clinical characteristics between the antibody-positive and antibody-negative groups were analyzed using the Mann-Whitney U test for continuous variables and the chi-square test for categorical variables. A p-value of less than 0.05 was considered statistically significant.

Interpretability analysis

To interpret the feature importance in the models, we used SHapley Additive exPlanations (SHAP). The SHAP summary plot highlights the impact of each feature on the model's output. Importance: The SHAP summary plot demonstrates the importance of each feature in predicting the presence of thyroid-specific antibodies. The features are ranked based on their mean absolute SHAP values, with higher values indicating greater importance.

Results

Participant characteristics

Of the 96 patients enrolled in the study, 48 tested positive for thyroid-specific autoantibodies (TPOAb and/or TgAb), while the other 48 were negative. The average age of the antibody-positive group was 47 ± 11 years, compared to 48 ± 14 years in the antibody-negative group. The majority of participants in both groups were female, with 92% in the antibody-positive group and 90% in the antibody-negative group. 71% (68/96) of all patients had a disease duration of less than five years. 34% (33/96) of the patients had received glucocorticoid treatment, while 46% (44/96) had received disease-modifying antirheumatic drugs(DMARDs). Among patients taking DMARDs, hydroxychloroquine was most commonly used, followed by methotrexate. The results of clinical characteristics of two groups were shown and compared in Table 1.

Clinical and laboratory findings

Comparison between the two groups revealed that patients with positive thyroid autoantibodies reported significantly lower prevalence of dry mouth symptoms. The levels of IgG and C4 were significantly lower in pSS with positive thyroid-specific autoantibodies. Other clinical and laboratory findings, including basic tear secretion and VitD levels, showed no statistically significant difference in Table 1. (Table 1 was shown at the end of the document)

Table 1 Characteristics, Clinical and Laboratory data of 96 participants with or without thyroid specific autoantibodies

Variable	pSS with TPOAb and/or TgAb (*n* = 48)	pSS without TPOAb and TgAb (*n* = 48)
Age (years)	47 ± 11	48 ± 14
Female (n, %)	44 (92%)	43 (90%)
Disease Duration (n, %)		
< 1 year	19 (40%)	16 (33%)
≥1 year, <5 years	18 (38%)	15 (31%)
≥ 5 years	11 (23%)	17 (35%)
Medication history		
Glucocorticoid	15 (31%)	18 (38%)
DMARDs	19 (40%)	25 (52%)
Symptoms(n, %)		
Dryness of mouth	22 (46%)	37 (77%)
Dryness of eye	38 (79%)	37 (77%)
Basic tear secretion (mm)	7.66 ± 4.42	9.55 ± 5.18
IgA (g/L)	2.65 ± 0.98	2.75 ± 0.98
IgG (g/L)	15.27 ± 3.82	16.35 ± 5.47
IgM (g/L)	1.14 ± 0.58	1.25 ± 0.83
C3 (g/L)	0.99 ± 0.17	1.05 ± 0.19
C4 (g/L)	0.21 ± 0.06	0.22 ± 0.08
VitD (ng/ml)	23.86 ± 6.78	24.91 ± 6.25
ANA (n, %)		
<1/320	16 (38%)	18 (33%)
≥ 1/320	30 (63%)	32 (68%)
SSA positive	31 (65%)	35 (73%)
SSB positive	12 (25%)	25 (52%)

Remarks: pSS primary Sjogren's syndrome, TPOAb anti-thyroid peroxidase, TgAb anti-thyroglobulin, VitD vitamin D, DMARDs disease-modifying antirheumatic drugs

Model performance

The performance metrics for the four machine learning algorithms are as follows: Random Forest Classifier (RFC) achieved the highest AUC of 0.755, with an accuracy of 77%, PPV of 72%, NPV of 82%, sensitivity of 87%, specificity of 64%, and an F1 score of 0.75. Support Vector Machine (SVM) recorded an AUC of 0.690, with metrics slightly lower than those of RFC. Extreme Gradient Boosting (XGB) and Logistic Regression (LR) showed moderate performance with AUC values of 0.652 and 0.552, respectively. The AUC values and other performance metrics are summarized in the Table 2.

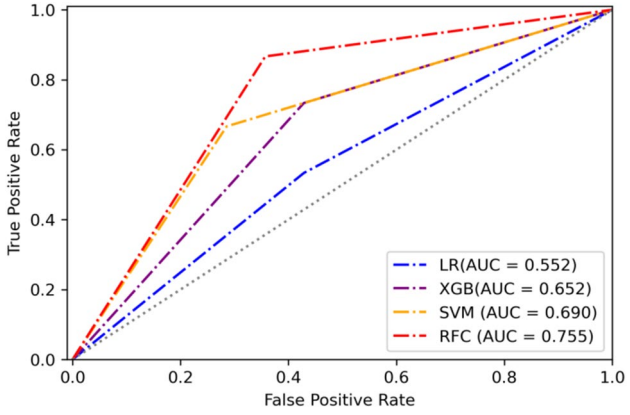

Fig. 1 The ROC curves for the four machine learning algorithms

ROC curve analysis

The ROC curves for each model are depicted in Fig. 1 below. The ROC plot illustrates that RFC not only provided the highest AUC but also maintained a balance between sensitivity and specificity, outperforming the other models in predicting the presence of thyroid-specific autoantibodies in pSS patients.

Feature importance

The Random Forest model indicated that the most influential predictors were age, IgG levels, C4, and symptoms of dry mouth. A detailed SHAP summary plot illustrates the impact of each feature on the model's output in Fig. 2 below, confirming the significant role of these factors in the model's predictions.

Discussion

In this study, 96 patients diagonosed pSS were enrolled and analyzed to identify the predictors for AIT. We collected 14 demographic and clinical characteristics to identify the predictors for AIT. We created a random forest classifier, which is a part of machine learning, providing a more accurate prediction for the risk of AIT (AUC=0.755) than logistic regression model (AUC=0.552). In our study, AUC and ROC curves confirmed that the RFC model, based on predictors per our findings, had greater predictive efficiency than LR model, was widely reported in other disciplines as well [13, 14].

Autoimmune diseases affect approximately one in ten individuals [15]. There may be common pathogenesis or

Table 2 The AUC values and other performance metrics for the four machine learning algorithms

Metric	ROC	ACC	PPV	NPV	SENS	SEPC	F1
LR	0.552	0.55	0.57	0.53	0.53	0.57	0.55
XGB	0.652	0.66	0.65	0.67	0.73	0.57	0.66
SVM	0.690	0.69	0.71	0.67	0.67	0.71	0.69
RFC	0.755	0.77	0.72	0.82	0.87	0.64	0.75

Remarks: LR Logistic Regression, XGB Extreme Gradient Boosting, SVM Support Vector Machine, RFC Random Forest Classifier, ROC receiver operator characteristic curve, ACC accuracy, PPV positive predictive value, NPV negative predictive value, SENS sensitivity, SEPC specificity

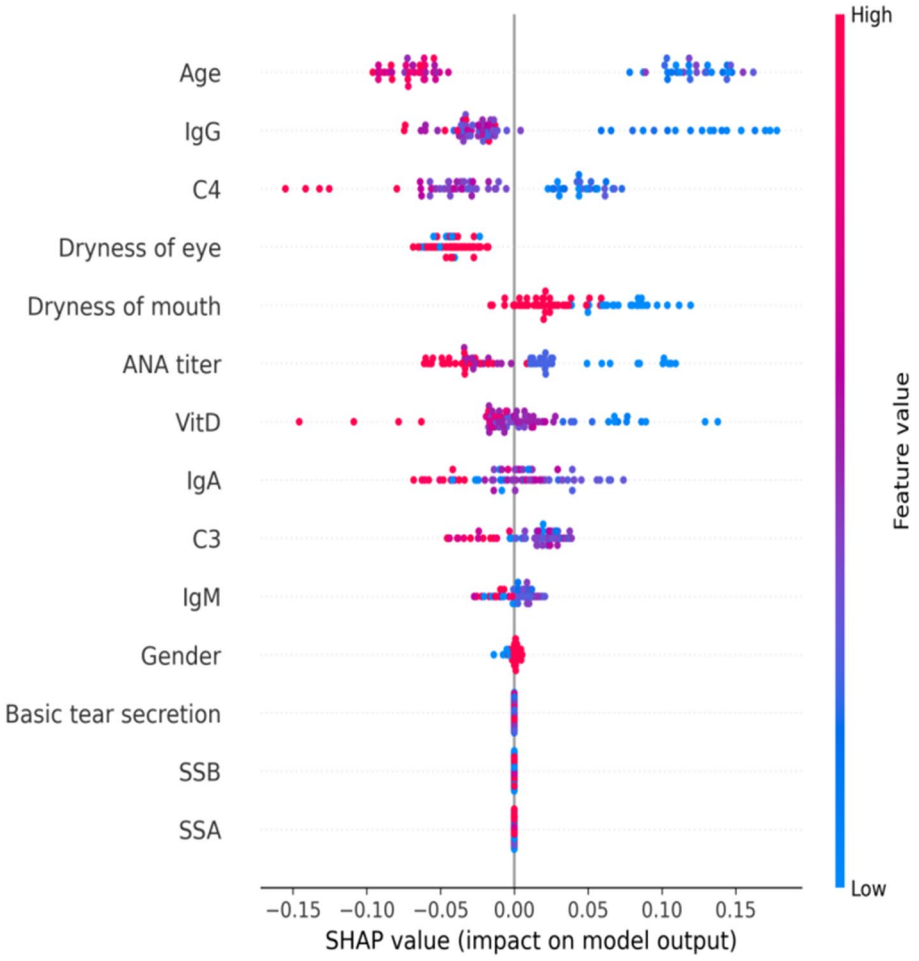

Fig. 2 The impact of 14 features in pSS patients on the model's output

predisposing factors between different autoimmune diseases, especially between connective tissue diseases and endocrine diseases. pSS, one of chronic autoimmune diseases, was characterized by lymphocyte infiltration and specific autoantibodies. The most common thyroid disorders of pSS are autoimmune thyroiditis, mainly caused by TPOAb and TgAb [16]. Therefore, in this paper, the detection of TPOAb and TgAb was used as an objective index to evaluate AIT.

Age, IgG, C4, and dryness of mouth, were the top four most important characteristics. In our series, age accounted for the largest proportion of the feature weight. As known, the peak prevalence of pSS and AIT occurs between 30 and 50 years in both. Previous studies found that the peak age of onset in pSS with AITD patients ranges 45 to 54 years old, consistent with our study [17]. The patients of pSS with AIT are younger than those without AIT, in agreement with others [18]. So younger patients with pSS should pay more attention to the possibility of autoimmune thyroiditis.

B-cell hyperactivity is a hallmark of pSS [19]. Hyperglobulinemia is one of the most common clinical manifestations of B cell hyperactivity. Significant serum immunological features in pSS patients were hyperglobulinemia and positive for multiple autoantibodies, such as anti-SSA, anti-SSB and ANA. Previous researches proved that more than 70% of pSS patients have elevated serum IgG levels [20]. Recent studies showed that hyperimmunoglobulinemia can indeed increase death and important damage to the renal, pulmonary, interstitial, haematological or neurological involvement [21–23]. In our study, elevated level or normal range but close to upper limit IgG levels were observed in both group. However, the group of pSS with TPOAb and TgAb presented lower levels of IgG than those without. IgG is classified into four subclasses, IgG1 to IgG4. Previous studies have indicated that.

IgG1 and IgG4 are mainly elevated in AIT patients, and IgG1 and IgG3 are mainly in pSS [24, 25]. Despite the high degree of homology in the amino acid sequence of IgG subtypes, they have unique characteristics in binding antigens, activating complement, triggering effect cells, and tissue distribution. In several studies recently, immunohistochemistry has indicated numerous IgG4-positive

plasma cells in a fraction of HT [26, 27]. Another study showed that IgG4 HT, grouped based on immunostaining results of thyroid, had higher levels of thyroid autoantibodies compared to those classified as non-IgG4 positive [28]. These findings indicate that the IgG4 subtype of AIT is prone to deposit and destroy thyroid cells in endocrine gland tissue. Therefore, although serum IgG levels in pSS patients with AIT do not appear to be higher, there may be a specific IgG subtype that is more readily distributed to the affected glands. Regrettably, our study did not investigate the distribution of IgG subtypes and IgG deposition in the affected tissues of patients with positive thyroid autoantibodies, which may be an important research direction for the future. We hope future studies can better answer this question.

In addition, we found that C4 levels of pSS with TPOAb or TgAb are lower. The same results was found in pSS with HT [18]. The results of this study showed that pSS-AIT patients had lower C4 levels. In a large cohort of 921 Spanish patients with pSS, reduced C3 and C4 complement fractions at the pSS diagnosis was observed and this finding was associated with disease activity [29]. Hypocomplementaemia was detected in 24% of patients with pSS and low C4 level was also closely associated with the two main adverse outcomes of lymphoma development and death [30]. Another survey of 723 patients with pSS found that low levels of C4 and obvious purpura were the most threatening risk factors for lymphoma after 6 years follow-up [31]. High titer TPOAb is one of the main causes of AIT. The mechanism is that TPOAb destroys thyroid tissue through antibody-dependent cytotoxicity and complement-dependent cytotoxicity, leading to reduction of complement components. This implies that high titer thyroid autoantibodies in pSS patients may activate complement-dependent cytotoxicity leading to high disease activity or multiple organ dysfunction, and even lymphoma.

In our study, no symptom of dry mouth is a risk factor for developing pSS with AIT. From the analysis of 100 Italian patients affected by pSS, it was similarly found that pSS associated with HT, a common subtype of AIT, complained less frequently of xerostomia [18]. A cross-sectional study included 305 subjects with sicca symptoms showed that high prevalence of AIT in subjects with sicca symptoms, especially mouth dryness [4]. In this study, AIT precedes SS in only 13% of patients. In 50% of patients, the diagnosis of HT was secondary to pSS after a mean follow-up of 5.5 years [17]. It means that symptoms of eyes or mouth dryness should be seriously considered as a possible symptom of SS, as well as AIT (sometimes both). Therefore, we need to actively screen thyroid function and thyroid autoantibodies in the initial diagnosis of pSS and subsequent follow-up.

Our study is not without limitations. First, retrospective design limits our ability to draw causal inferences from the associations found. The sample size, while adequate for initial explorations, is relatively small for generalizing the findings across diverse populations. Second, this study lacks the collection of thyroid ultrasound data. The absence of TPOAb and TgAb in some AIT patients may be helpful in the diagnosis based on the ultrasound features of the thyroid gland. Thirdly, our study did not explore the distribution of serum IgG subtypes and the deposition of IgG subtypes in thyroid glands. Future studies should aim to include a larger cohort to validate these predictors and potentially uncover new ones.

Conclusion

In summary, this study demonstrates the utility of a Random Forest Classifier for predicting the presence of autoimmune thyroiditis in patients with primary Sjogren's syndrome, identifying age, IgG levels, C4, and absence of dry mouth as significant predictors. These findings are significant given the high prevalence of pSS-AIT co-occurrence and the potential for early therapeutic intervention.

Abbreviations

pSS	Primary Sjogren's syndrome
AIT	Autoimmune thyroiditis
HT	Hashimoto's thyroiditis
TPOAb	Anti-thyroid peroxidase
TgAb	Anti-thyroglobulin
SLE	Systemic lupus erythematosus
RA	Rheumatoid arthritis
LR	Logistic Regression
SVM	Support Vector Machine
RFC	Random Forest Classifier
XGB	Extreme Gradient Boosting
AUC	ROC-The area under the receiver operating characteristic curve
ACC	Accuracy
PPV	Positive predictive value
NPV	Negative predictive value
SENS	Sensitivity
SEPC	Specificity
DMARDs	Disease-modifying antirheumatic drugs

Acknowledgements

The authors acknowledge funding received from Shenzhen Science and Technology Program (No. JCYJ20220530145001002).

Author contributions

JYW, JYZ, JRG, QL, YR contributed to study conception/design and data analysis. FXM, LLX, SHW, HFL contributed to data collection. YQL contributed to data analysis. JYW, JYZ contributed to data interpretation. WQX, YNK, RYL, YLC, XML, YW collectively interpreted the results. JYW wrote the manuscript. JYZ, QL, YR substantively revised it. All authors had access to the data, and reviewed and approved the final manuscript before submission.

Funding

Shenzhen Science and Technology Program (No. JCYJ20220530145001002).

Data availability

No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate
The study protocol was approved by the Ethics Committee of The Third Affiliated Hospital of Sun Yat-sen University (No.II2023-254-02). Because of the retrospective nature of the study, the requirement for informed consent was waived.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Rheumatology, The Seventh Affiliated Hospital of Sun Yat-Sen University, Shenzhen, China
[2]School of Automation Science and Engineering, South China University of Technology, Guangzhou, China
[3]Research Center for Brain-Computer Interface, Pazhou Lab, Guangzhou, China
[4]Department of Rheumatology, The Third Affiliated Hospital of Sun Yat-Sen University, Guangzhou, China
[5]Scientific Research Project Department, Guangdong Artificial Intelligence and Digital Economy Laboratory (Guangzhou), Pazhou Lab, Guangzhou, China
[6]The Seventh Affiliated Hospital of Sun Yat-Sen University, Shenzhen, China

Received: 8 September 2024 / Accepted: 28 December 2024
Published online: 02 January 2025

References

1. Jisha JN, Tejas PS. Sjogren's syndrome: Review of the aetiology, Pathophysiology & Potential therapeutic interventions. J Clin Exp Dent. 2017;9(4).
2. Antonelli A, Ferrari SM, Corrado A, Di Domenicantonio A, Fallahi P. Autoimmune thyroid disorders. Autoimmun Rev. 2015;14(2):174–80.
3. Tagoe CE. Rheumatic symptoms in Autoimmune Thyroiditis. Curr Rheumatol Rep. 2015;17(2).
4. Vera DM, Goran R, Ivan B, Sanja O, Radmila P, Marija R-P et al. High prevalence of autoimmune thyroid disease in subjects with sicca symptoms without Sjogren's syndrome. Rheumatology (Oxford). 2013;52(4).
5. Lin DF, Yan SM, Zhao Y, Zhang W, Li MT, Zeng XF, et al. Clinical and prognostic characteristics of 573 cases of primary Sjögren's syndrome. Chin Med J (Engl). 2010;123(22):3252–7.
6. Mario G-P, Ivette C-B, Vanessa S-G, Yemil A-F, Gustavo B, Carlos A, A-S et al. Autoimmune Thyroid Disease in Primary Sjögren's Syndrome: Real-life Screening Practice and Clinical Outcomes. Curr Rheumatol Rev. 2022;18(3).
7. Cruz-Tapias P, Rojas-Villarraga A, Maier-Moore S, Anaya JM. HLA and Sjögren's syndrome susceptibility. A meta-analysis of worldwide studies. Autoimmun Rev. 2012;11(4):281–7.
8. Ramagopalan SV, Goldacre R, Disanto G, Giovannoni G, Goldacre MJ. Hospital admissions for vitamin D related conditions and subsequent immune-mediated disease: record-linkage studies. BMC Med. 2013;11:171.
9. Brito-Zerón P, Gheitasi H, Retamozo S, Bové A, Londoño M, Sánchez-Tapias JM, et al. How hepatitis C virus modifies the immunological profile of Sjögren syndrome: analysis of 783 patients. Arthritis Res Ther. 2015;17(1):250.
10. Berthold S, Daniela H, Stefan W, Andreas C, Markus R. Frequency of chromosomal aberrations involving MALT1 in mucosa-associated lymphoid tissue lymphoma in patients with Sjögren's syndrome. Clin Cancer Res. 2004;10(2).
11. Solomon DH, Guan H, Johansson FD, Santacroce L, Malley W, Guo L, et al. Assessing clusters of comorbidities in rheumatoid arthritis: a machine learning approach. Arthritis Res Ther. 2023;25(1):224.
12. Vitali C, Bombardieri S, Jonsson R, Moutsopoulos H, Alexander E, Carsons S, 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):554–8.
13. Rongbo W, Kuo Z, Qihang Z, Leqi Z, Qizhi L, Guanyu Y et al. Machine learning-based random forest predicts anastomotic leakage after anterior resection for rectal cancer. J Gastrointest Oncol. 2021;12(3).
14. Bin C, Li Q, Tang J, Dai C, Jiang T, Xie X, et al. Machine learning models for predicting the risk factor of carotid plaque in cardiovascular disease. Front Cardiovasc Med. 2023;10:1178782.
15. Conrad N, Misra S, Verbakel J, Verbeke G, Molenberghs G, Taylor P, et al. Incidence, prevalence, and co-occurrence of autoimmune disorders over time and by age, sex, and socioeconomic status: a population-based cohort study of 22 million individuals in the UK. Lancet (London England). 2023;401(10391):1878–90.
16. Luis JJ, Carmen N, María del Pilar B-Z, Mario G-C, Ricardo OE, Manuel R-C. Thyroid disease in Sjögren's syndrome. Clin Rheumatol. 2007;26(10).
17. Margit Z, Ildiko Fanny H, Antonia S, Peter S. Autoimmune thyroid diseases in a large group of Hungarian patients with primary Sjögren's syndrome. Thyroid. 2009;19(1).
18. Caramaschi P, Biasi D, Caimmi C, Scambi C, Pieropan S, Barausse G, et al. The co-occurrence of Hashimoto thyroiditis in primary Sjogren's syndrome defines a subset of patients with milder clinical phenotype. Rheumatol Int. 2013;33(5):1271–5.
19. Divi C, Valérie D-P, Gabriel JT, Jacques-Olivier P, Sandrine J-J, Alain S. B cells in Sjögren's syndrome: from pathophysiology to diagnosis and treatment. J Autoimmun. 2012;39(3).
20. Ren H, Wang W, Chen X, Zhang W, Pan X, Wang X, et al. Renal involvement and followup of 130 patients with primary Sjögren's syndrome. J Rhuematol. 2008;35(2):278–84.
21. Hui Z, Yanhong W, Pingting Y, Xinwang D, Yongfu W, Jian X et al. Hyperglobulinemia predicts increased risk of mortality in primary Sjögren's syndrome: Based on a Chinese multicentre registry. Mod Rheumatol. 2023;34(1).
22. López-Morales J, Cortes-Muñoz D, Astudillo-Ángel M, Hernández-Molina G. Persistent serological activity in primary Sjögren's syndrome. Clin Rheumatol. 2020;39(3):919–23.
23. Tabea S, Emelie K, Franz Felix K, Nadine Z, Sonja B, Nils Kristian P et al. Sjögren's syndrome with and without neurological involvement. J Neurol. 2023;270(6).
24. Forouhi N, McLachlan S, Middleton S, Atherton M, Baylis P, Clark F, et al. T cell regulation of thyroglobulin autoantibody IgG subclasses in Hashimoto's thyroiditis. Clin Exp Immunol. 1987;69(2):314–22.
25. Liu Y, Li J. Preferentially immunoglobulin (IgG) subclasses production in primary Sjögren's syndrome patients. Clin Chem Lab Med. 2011;50(2):345–9.
26. Kojima M, Hirokawa M, Kuma H, Nishihara E, Masawa N, Nakamura N, et al. Distribution of IgG4- and/or IgG-positive plasma cells in Hashimoto's thyroiditis: an immunohistochemical study. Pathobiol J ImmunoPathol Mol Cell Biol. 2010;77(5):267–72.
27. Li Y, Bai Y, Liu Z, Ozaki T, Taniguchi E, Mori I, et al. Immunohistochemistry of IgG4 can help subclassify Hashimoto's autoimmune thyroiditis. Pathol Int. 2009;59(9):636–41.
28. Li Y, Zhou G, Ozaki T, Nishihara E, Matsuzuka F, Bai Y, et al. Distinct histopathological features of Hashimoto's thyroiditis with respect to IgG4-related disease. Mod Pathology: Official J United States Can Acad Pathol Inc. 2012;25(8):1086–97.
29. Manuel R-C, Pilar B-Z, Roser S, María-Teresa C, Arnau C, Bernardo S et al. Systemic involvement in primary Sjogren's syndrome evaluated by the EULAR-SS disease activity index: analysis of 921 Spanish patients (GEAS-SS Registry). Rheumatology (Oxford). 2013;53(2).
30. Ramos-Casals M, Brito-Zerón P, Yagüe J, Akasbi M, Bautista R, Ruano M, et al. Hypocomplementaemia as an immunological marker of morbidity and mortality in patients with primary Sjogren's syndrome. Rheumatology (Oxford). 2005;44(1):89–94.
31. Ioannidis J, Vassiliou V, Moutsopoulos H. Long-term risk of mortality and lymphoproliferative disease and predictive classification of primary Sjögren's syndrome. Arthritis Rheum. 2002;46(3):741–7.

Publisher's note