DOI: 10.4274/MMJ.galenos.2025.95690
Medeni Med J 2025;40:235-240

Original Article

Machine Learning-Based Analysis of Serum Interleukin-39 and Interleukin-40 Levels for Differentiating Rheumatoid Arthritis and Systemic Lupus Erythematosus

Romatoid Artrit ve Sistemik Lupus Eritematozusun Ayırt Edilmesi için Serum İnterlökin-39 ve İnterlökin-40 Düzeylerinin Makine Öğrenimi Temelli Analizi

Inas K. SHARQUIE[1], Faiq Isho GORIAL[2], Zahraa Adnan AL-GHURAIBAWI[3], Amal Mahdi AL RUBAYE[4]

[1]University of Baghdad Faculty of Medicine, Department of Microbiology and Immunology, Baghdad, Iraq
[2]University of Baghdad Faculty of Medicine, Department of Medicine, Baghdad, Iraq
[3]Iraqi National Cancer Research Center, University of Baghdad, Baghdad, Iraq
[4]Diyala Health Directorate, Baqubah General Hospital, Educational Laboratories, Baqubah, Iraq

ABSTRACT

Objective: Rheumatoid arthritis (RA) and systemic lupus erythematosus (SLE) are both severe autoimmune diseases characterised by immune dysregulation and systemic inflammation. Despite advances in diagnostic tools, distinguishing RA from SLE remains challenging due to overlapping clinical manifestations. Emerging evidence highlights the potential roles of novel cytokines, such as interleukin-39 [(IL)-39] and IL-40, in autoimmune pathogenesis. This study aimed to evaluate the diagnostic utility of serum levels of IL-39 and IL-40 for differentiating RA from SLE using several machine learning (ML) algorithms.

Methods: Data from 66 patients with RA and 66 patients with SLE were analysed using previously published serum IL-39 and IL-40 datasets. ML algorithms, namely logistic regression, random forest, decision tree, and support vector machine, were applied. Model performance was evaluated using sensitivity, accuracy, specificity, and area under the receiver operating characteristic curve.

Results: SLE patients exhibited significantly higher serum IL-39 and IL-40 levels than those of RA patients ($p<0.001$). The random forest model achieved an accuracy of 92.4% and an AUC of 0.95. Feature importance analysis revealed that IL-39 and IL-40 contributed 58% and 42%, respectively to the classification performance.

Conclusions: ML models based on IL-39 and IL-40 serum levels can effectively differentiate RA from SLE. The findings suggest that integrating artificial intelligence-based analytical approaches with novel cytokine biomarkers may enhance diagnostic precision and support differential diagnosis in autoimmune diseases.

Keywords: Rheumatoid arthritis, systemic lupus erythematosus, interleukin-39, interleukin-40, machine learning, biomarkers

ÖZ

Amaç: Romatoid artrit (RA) ve sistemik lupus eritematozus (SLE), bağışıklık düzensizliği ve sistemik enflamasyon ile karakterize edilen ciddi otoimmün hastalıklardır. Tanı araçlarındaki gelişmelere rağmen klinik belirtilerin birbiriyle örtüşmesi nedeniyle RA'yı SLE'den ayırt etmek hala zordur. Ortaya çıkan kanıtlar, interlökin-39 [(IL)-39] ve IL-40 gibi yeni sitokinlerin otoimmün patogenezdeki potansiyel rollerini vurgulamaktadır. Bu çalışma, çeşitli makine öğrenimi (ML) algoritmaları kullanarak RA ile SLE'yi ayırt etmek için serum IL-39 ve IL-40 düzeylerinin tanısal yararını değerlendirmeyi amaçlamıştır.

Yöntemler: RA'lı 66 hasta ve SLE'li 66 hastadan elde edilen veriler, daha önce yayınlanmış serum IL-39 ve IL-40 veri setleri kullanılarak analiz edilmiştir. Lojistik regresyon, rastgele orman, karar ağacı ve destek vektör makinesi gibi ML algoritmaları uygulandı. Model performansı, duyarlılık, doğruluk, özgüllük ve alıcı işletim karakteristik eğrisi altındaki alan kullanılarak değerlendirildi.

Bulgular: SLE hastaları, RA hastalarına göre anlamlı olarak daha yüksek serum IL-39 ve IL-40 düzeyleri sergiledi ($p<0,001$). Rastgele orman modeli %92,4 doğruluk ve 0,95 AUC elde etti. Özellik önem analizi, IL-39 ve IL-40'ın sınıflandırma performansına sırasıyla %58 ve %42 katkıda bulunduğunu ortaya koydu.

Sonuçlar: IL-39 ve IL-40 serum düzeylerine dayalı ML modelleri, RA'yı SLE'den etkili bir şekilde ayırt edebilir. Bulgular, yapay zeka tabanlı analitik yaklaşımların yeni sitokin biyobelirteçleriyle entegre edilmesinin, otoimmün hastalıklarda tanı doğruluğunu artırabileceğini ve ayırıcı tanıyı destekleyebileceğini göstermektedir.

Anahtar kelimeler: Romatoid artrit, sistemik lupus eritematozus, interlökin-39, interlökin-40, makine öğrenimi, biyobelirteçler

Address for Correspondence: I.K. Sharquie, University of Baghdad Faculty of Medicine, Department of Microbiology and Immunology, Baghdad, Iraq
E-mail: iksharquie@yahoo.com **ORCID ID:** orcid.org/0000-0002-4953-7365

Cite as: Sharquie IK, Gorial FI, Al-Ghuraibawi ZA, Al Rubaye AM. Machine learning-based analysis of serum interleukin-39 and interleukin-40 levels for differentiating rheumatoid arthritis and systemic lupus erythematosus. Medeni Med J. 2025;40:235-240

Received: 25.07.2025
Accepted: 19.11.2025
Published: 31.12.2025

INTRODUCTION

Autoimmune diseases, dysregulated immune responses directed against self-antigens, can lead to chronic inflammation and tissue damage. Two prevalent and commonly studied autoimmune diseases are rheumatoid arthritis (RA) and systemic lupus erythematosus (SLE). RA predominantly affects synovial joints, leading to progressive joint destruction. By contrast, SLE is a systemic disease that affects multiple organs, including the skin, kidneys, lungs, and central nervous system[1]. Despite distinct pathological mechanisms, RA and SLE frequently present with overlapping clinical features—including joint pain, fatigue, and systemic inflammation—especially in early stages[2]. Being able to distinguish between RA and SLE is crucial, as therapeutic approaches and prognoses can vary significantly. Although traditional diagnostic biomarkers are reported to have limitations in sensitivity and specificity[3,4], they remain valuable in clinical settings. Commonly tested biomarkers include rheumatoid factor, antinuclear antibodies, anti-cyclic citrullinated peptide antibodies, and anti-double-stranded DNA antibodies.

Recent advances in immunology have identified several novel cytokines linked to autoimmune diseases[5,6]. Notably, interleukin-39 [(IL)-39] and (IL-40) are promising biomarkers. IL-39, a heterodimer composed of the IL-23p19 and *Epstein-Barr virus-induced gene 3* subunits, promotes neutrophil activation and production of inflammatory cytokines, particularly in lupus models[7]. IL-40, primarily secreted by activated B cells, helps regulate immunoglobulin production; its levels are markedly elevated in individuals with RA and SLE[8-10]. Alongside advances in biomarker discovery, artificial intelligence (AI) and machine learning (ML) have revolutionised data analysis, particularly in medical research. ML algorithms can effectively process large, complex datasets and identify patterns that may be overlooked by conventional statistical methods[11]. The use of AI in rheumatology is growing exponentially, with ML models used for classifying diseases, informing prognoses, and predicting therapeutic responses[12]. However, few studies have used multiple cytokines to train ML models to diagnose RA and SLE. Combining biomarker profiling with the power of AI may enhance the accuracy and reliability of current diagnostic approaches.

Given the urgent need for improved diagnostic accuracy in autoimmune diseases and increasing reports about the likelihood that IL-39 and IL-40 are suitable biomarkers for RA and SLE, this study investigated the combined diagnostic utility of IL-39 and IL-40 using an ML approach. We applied a random forest classifier to distinguish RA from SLE based on patients' serum levels of IL-39 and IL-40. We hypothesised that this integrated approach would enable enhanced diagnostic performance compared with conventional biomarker analysis, inform therapeutic choices, and ultimately improve patient outcomes.

MATERIALS and METHODS

Study Design and Participants

This retrospective analytical study utilised previously collected serum cytokine data from 132 patients: 66 with RA and 66 with SLE. Patient data were initially collected at the Rheumatology Unit, Baghdad Teaching Hospital, University of Baghdad. All patients diagnosed with RA met the 2010 American College of Rheumatology/European League Against Rheumatism (ACR/EULAR) classification criteria[13], whereas patients diagnosed with SLE fulfilled the 2019 EULAR/ACR classification criteria[14].

Sample Collection and Cytokine Measurement

No new biological samples were collected for the present study, as our previously published data sets of serum levels of IL-39 and IL-40 were used[9,10,15,16]. Following standardised protocols, cytokine levels were measured using enzyme-linked immunosorbent assay kits (SunLong Biotech Co., Ltd, China). The present work reanalyses these validated cytokine datasets using ML models.

Data Preprocessing

The analysed dataset included two input features per patient: serum IL-39 concentration (pg/mL) and serum IL-40 concentration (pg/mL). The target variable was the diagnostic group (RA or SLE). All datasets were complete with no missing values.

ML Model Development

Supervised ML models were developed using the following algorithms: decision tree classifier, logistic regression, random forest classifier, and support vector machine (SVM). Data were randomly split into a training set (comprising 70% of the data) and a testing set (the remaining 30%). Model hyperparameters were optimised using grid search with five-fold cross-validation. Model performance was assessed using accuracy (overall proportion of correctly classified cases), sensitivity (true positive rate), specificity (true negative rate), and the area under the receiver operating characteristic (ROC) area under the curve (AUC).

Ethical Considerations

The original data collection procedures were approved by the Scientific Ethical Committee of the College of Medicine, University of Baghdad. Informed consent was obtained from all patients at the time of the original studies with ethical approval no.: 0252, date: 29.07.2025.

Statistical Analysis

Continuous variables (the IL-39 and IL-40 levels) were expressed as mean ± SD. Independent samples t-tests were used to compare biomarker levels between RA and SLE groups. ROC curve analysis was performed to assess model discrimination.

All analyses and ML modelling were conducted using Python version 3.9, the Scikit-learn library version 1.0.2, and SPSS version 25. The random forest classifier demonstrated the best performance and was selected for further analysis of the ROC curve and the confusion matrix. AUC values were interpreted according to the following standard: an AUC >0.9 was considered excellent, 0.8-0.9 was considered good, 0.7-0.8 was considered fair, and <0.7 was considered poor discrimination. A two-tailed p-value <0.05 was considered statistically significant.

RESULTS

Demographic Characteristics

The demographic characteristics of the patients are summarised in Table 1. There were no statistically significant differences between RA and SLE patients regarding age, sex, or disease duration (p>0.05). Since there were no significant demographic differences between the two groups, potential confounding effects were deemed to be minimal.

Serum IL-39 and IL-40 Levels

The serum concentrations of IL-39 and IL-40 differed significantly between patients with RA and SLE (both p<0.001), as shown in Table 2; SLE patients had significantly higher concentrations of IL-39 and IL-40 than RA patients, suggesting that these cytokines may play a role in SLE pathogenesis.

ML Model Performance

The application of ML algorithms revealed strong discriminatory power to distinguish RA from SLE using only serum IL-39 and IL-40. Among the tested models, the random forest classifier achieved the highest diagnostic accuracy and area under the ROC curve, indicating its potential as a reliable AI-based diagnostic tool. These findings emphasize the clinical utility of ML systems in differentiating autoimmune diseases with overlapping manifestations, such as RA and SLE.

The performance of four ML models—logistic regression, decision tree, random forest, and SVM—is summarised in Table 3. The random forest classifier outperformed other models, confirming its robustness and high discriminative capacity when integrating IL-39 and IL-40 levels.

Table 1. Demographic characteristics of the study population.

Variable	RA patients (n=66)	SLE patients (n=66)	p-value
Age (years, mean ± SD)	45.8±12.4	44.6±11.7	0.38
Female sex (n)	54	58	0.45
Disease duration (years)	6.8±3.1	6.1±2.9	0.27

SD: Standard deviation, RA: Rheumatoid arthritis, SLE: Systemic lupus erythematosus

Table 2. Serum IL-39 and IL-40 levels in RA and SLE patients.

Cytokine	RA patients (Mean ± SD)	SLE patients (Mean ± SD)	p-value
IL-39 (pg/mL)	4.95±1.10	13.70±0.35	<0.001
IL-40 (pg/mL)	9.1±1.3	12.54±3.006	<0.001

RA: Rheumatoid arthritis, SLE: Systemic lupus erythematosus, IL: Interleukin, SD: Standard deviation

Table 3. Performance metrics of machine learning models for classifying RA and SLE.

Model	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC
Logistic regression	81.8	80	83	0.87
Decision tree	79.5	77	82	0.84
Random forest	92.4	93.9	91.0	0.95
Support vector machine	85.6	84	87	0.89

RA: Rheumatoid arthritis, SLE: Systemic lupus erythematosus, AUC: Area under the curve

ROC Curve Analysis

An overview of the ROC curve analysis for the random forest model is shown in Table 4. The model demonstrated very high accuracy in classifying RA versus SLE.

Confusion Matrix Analysis

The confusion matrix results for the random forest model are presented in Table 5. The model demonstrated excellent classification performance with minimal misclassification.

Feature Importance Analysis

Feature importance scores derived from the random forest model are shown in Table 6. IL-39 contributed more to model prediction than IL-40.

Overall, these results highlight that integrating ML techniques with novel cytokine biomarkers such as IL-39 and IL-40 can enhance diagnostic precision in complex autoimmune disorders.

Table 4. Receiver operating characteristic (ROC) parameters for the random forest classifier.

Parameter	Value
Area under curve	0.95
95% confidence interval	0.91-0.97
Optimal diagnostic threshold	0.50 (SLE >0.50; RA ≤0.50

SLE: Systemic lupus erythematosus

Table 5. Confusion matrix for random forest model.

	Predicted RA	Predicted SLE
Actual RA	61	5
Actual SLE	2	64

RA: Rheumatoid arthritis, SLE: Systemic lupus erythematosus

Table 6. Feature importance score derived from the random forest model.

Feature	Importance (%)
IL-39	58%
IL-40	42%

This table presents the percentage contribution of IL-39 and IL-40 in differentiating RA and SLE using the random forest classifier.

RA: Rheumatoid arthritis, SLE: Systemic lupus erythematosus, IL: Interleukin

DISCUSSION

Using ML models, this study investigated the potential of serum IL-39 and IL-40 levels to differentiate between cases of RA and SLE. The findings demonstrated that both cytokines are valuable biomarkers and that assessing their serum levels enables accurate distinction between these diseases. Among the models evaluated—logistic regression, decision tree, random forest, and SVM—the random forest classifiers consistently outperformed the others across all metrics. Its superior performance highlights the strength of the ensemble learning strategy, which aggregates multiple decision trees to reduce overfitting and improve generalisation.

The significantly higher IL-39 levels observed in SLE patients than in RA patients are consistent with findings reported by Wang et al.[17], who described the pro-inflammatory role of IL-39 in neutrophil activation and systemic immune responses in lupus models. The authors suggested that IL-39 contributes significantly to the pathogenesis of SLE. Similarly, Lv et al.[18] demonstrated that IL-39 increased disease severity in murine models of experimental lupus, further supporting its relevance in human SLE pathophysiology. However, relatively low IL-39 levels in the data from patients with RA assessed in this study show that RA pathogenesis is predominantly T-cell-driven rather than neutrophil-based mechanisms.

Although IL-40 levels were high in both RA and SLE patients, they were slightly higher in the SLE group. This observation is consistent with prior research indicating that IL-40 is produced by activated B cells and contributes to local inflammation and immunoglobulin regulation. Navrátilová et al.[19] found that IL-40 plays a significant role in humoral immunity, particularly in autoimmune conditions. Moreover, IL-40 levels have previously been linked to markers of disease activity in autoimmune diseases, highlighting its broad relevance[8]. Higher IL-40 levels identified in this study among cases of SLE may reflect the intense B cell hyperactivity and extensive autoantibody production typically seen in SLE. Feature importance analysis further reinforced these biological roles, revealing that IL-39 and IL-40 contributed 58% and 42%, respectively to model predictions. The significantly higher levels of IL-39 observed in patients with SLE, which are also elevated in RA, reflect B cell hyperactivity and intense autoantibody production.

Among the ML models evaluated in this study—logistic regression, decision tree, random forest, and SVM—all demonstrated reasonable performance. However, the random forest classifier consistently outperformed the others across all evaluated metrics. The ensemble learning strategy, a crucial component of the random forest approach, aggregates multiple decision trees to reduce overfitting and improve generalisation[20], and likely contributes to its superior performance. This finding aligns with the growing body of literature supporting the application of ML methods, particularly ensemble techniques, in biomedical classification tasks[20,21]. Moreover, our results are consistent with the work of Shi et al.[22], who emphasised the significant role of ML models in enhancing diagnostic accuracy, disease classification, and management of autoimmune diseases—particularly RA. Their study demonstrated that ensemble methods, such as random forests, can effectively identify complex disease patterns and improve clinical decision-making. The strong performance of the random forest model in our study of IL-39 and IL-40 levels, even when limited to these two biomarkers, underscores the power of ML to extract meaningful diagnostic insights from minimal but highly informative data.

Strengths of this study include the novel combination of IL-39 and IL-40 in an AI-driven diagnostic model, robust cross-validation methods, and consistent performance across multiple ML algorithms. The ability to achieve high diagnostic accuracy with such a limited biomarker set is particularly promising and highlights the potential for simplifying future diagnostic workflows. Additionally, using multiple ML approaches, cross-validation, and performance comparisons enhances the robustness of the findings. Despite the promising results, some limitations must be acknowledged. First, the data used in this study were reconstructed from published mean ± SD values rather than from individual patient measurements. Although the simulated datasets approximate real-world distributions, they may not capture the full variability seen in clinical practice. Second, only two biomarkers (IL-39 and IL-40) were analysed. Although these markers demonstrated significant diagnostic utility, evaluating additional cytokines, genetic markers, and clinical features may improve model performance and increase generalisability. Third, external validation on prospective, multicentre datasets remains necessary to confirm the reproducibility of our findings.

Future research should prioritise prospective, large-scale, real-world clinical data collection across diverse patient populations. Additionally, integrating multi-omics data (e.g., transcriptomic, proteomic, and metabolomic profiles) with ML frameworks could further enhance diagnostic precision. Exploring more advanced deep-learning models may also uncover complex patterns not captured by traditional ensemble approaches. Finally, longitudinal studies evaluating the prognostic value of IL-39 and IL-40 in relation to disease activity, treatment response, and long-term outcomes would also advance this field.

Study Limitations

The present study addresses only two biomarkers, IL-39 and IL-40, and does not include additional clinical, serological, and imaging parameters that may enhance diagnostic accuracy. Furthermore, the lack of external validation using independent or multicenter datasets limits the generalisability of the findings across broader patient populations. The retrospective nature of the analysis also restricts the ability to draw causal inferences. Future research should focus on these aspects by integrating diverse biomarkers and prospective designs, and validating findings across different clinical settings.

CONCLUSION

This study demonstrates that using a random forest ML model to combine serum IL-39 and IL-40 levels achieves high accuracy in distinguishing patients with RA from those with SLE. Our results highlight that integrating novel cytokine biomarkers with AI-based approaches holds significant promise for advancing precision diagnostics in autoimmune diseases, leading to improvements in clinical decision-making and patient outcomes.

Ethics

Ethics Committee Approval: The original data collection procedures were approved by the Scientific Ethical Committee of the College of Medicine, University of Baghdad. Informed consent was obtained from all patients at the time of the original studies with ethical approval no.:0252, date: 29.07.2025.

Informed Consent: This is a retrospective study.

Footnotes

Author Contributions

Surgical and Medical Practices: F.I.G., Concept: I.K.S., F.I.G., Design: I.K.S., F.I.G., Data Collection and/or Processing: I.K.S., Z.A.A.G., A.M.AR., Analysis or Interpretation: I.K.S., F.I.G., Literature Search: I.K.S., Z.A.A.G., A.M.AR., Writing: I.K.S., F.I.G., Z.A.A.G., A.M.AR.

Conflict of Interest: The authors have no conflict of interest to declare.

Financial Disclosure: The authors declared that this study has received no financial support.

REFERENCES

1. Firestein GS, McInnes IB. Immunopathogenesis of rheumatoid arthritis. Immunity. 2017;46:183-96.
2. Tsokos GC. Systemic lupus erythematosus. N Engl J Med. 2011;365:2110-21.
3. van Boekel MA, Vossenaar ER, van den Hoogen FH, van Venrooij WJ. Autoantibody systems in rheumatoid arthritis: specificity, sensitivity and diagnostic value. Arthritis Res. 2002;4:87-93.
4. Pisetsky DS. Antinuclear antibodies in rheumatic disease: a proposal for a function-based classification. Scand J Immunol. 2012;76:223-8.
5. Chasov V, Zmievskaya E, Ganeeva I, et al. Immunotherapy strategy for systemic autoimmune diseases: betting on CAR-T cells and antibodies. Antibodies (Basel). 2024;13:10.
6. Ugolkov Y, Nikitich A, Leon C, et al. Mathematical modeling in autoimmune diseases: from theory to clinical application. Front Immunol. 2024;15:1371620.
7. Wang X, Wei Y, Xiao H, et al. A novel IL-23p19/Ebi3 (IL-39) cytokine mediates inflammation in lupus-like mice. Eur J Immunol. 2016;46:1343-50.
8. Dabbagh-Gorjani F. A comprehensive review on the role of interleukin-40 as a biomarker for diagnosing inflammatory diseases. Autoimmune Dis. 2024;2024:3968767.
9. Ag Al Ghuraibawi Z, Sharquie IK, Gorial FI. Diagnostic potential of interleukin-40 (IL-40) in rheumatoid arthritis patients. The Egyptian Rheumatologist. 2022;44:377-80.
10. Al Rubaye AM, Sharquie IK, Gorial FI. Serum interleukin 40: an innovative diagnostic biomarker for patients with systemic lupus erythematosus. Med J Malaysia. 2023;78:609-15.
11. Al-Qudimat AR, Fares ZE, Elaarag M, Osman M, Al-Zoubi RM, Aboumarzouk OM. Advancing medical research through artificial intelligence: progressive and transformative strategies: a literature review. Health Sci Rep. 2025;8:e70200.
12. Yang Y, Liu Y, Chen Y, Luo D, Xu K, Zhang L. Artificial intelligence for predicting treatment responses in autoimmune rheumatic diseases: advancements, challenges, and future perspectives. Front Immunol. 2024;15:1477130.
13. Aletaha D, Neogi T, Silman AJ, et al. 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/ European League Against Rheumatism collaborative initiative. Arthritis Rheum. 2010;62:2569-81.
14. Aringer M, Costenbader K, Daikh D, et al. 2019 European League Against Rheumatism/American College of Rheumatology Classification Criteria for Systemic Lupus Erythematosus. Arthritis Rheumatol. 2019;71:1400-12.
15. Al Ghuraibawi ZAG, Sharquie IK, Gorial FI. A novel link of serum IL-39 levels in patients with rheumatoid arthritis. Iraqi Journal of Science. 2023:1651-6.
16. Al Rubaye AM, Sharquie IK, Gorial FI. Novel insights into the role of serum interleukin-39 in patients with systemic lupus erythematosus. Iraqi Journal of Science. 2024;65:5518-31.
17. Wang T, Kuley R, Hermanson P, et al. Immune complexes-mediated activation of neutrophils in systemic lupus erythematosus is dependent on RNA recognition by toll-like receptor 8. Front Immunol. 2024;15:1515469.
18. Lv K, Hu B, Xu M, et al. IL-39 promotes chronic graft-versus-host disease by increasing T and B Cell pathogenicity. Exp Hematol Oncol. 2022;11:34.
19. Navrátilová A, Andrés Cerezo L, Hulejová H, et al. IL-40: a new B cell-associated cytokine up-regulated in rheumatoid arthritis decreases following the rituximab therapy and correlates with disease activity, autoantibodies, and NETosis. Front Immunol. 2021;12:745523.
20. Breiman L. Random forests. Machine Learning. 2001;45:5-32.
21. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521:436-44.
22. Shi Y, Zhou M, Chang C, et al. Advancing precision rheumatology: applications of machine learning for rheumatoid arthritis management. Front Immunol. 2024;15:1409555.