WISDOM ACADEMIC

Biology and Medical Engineering

ISSN: 3080-7204 | E-ISSN: 3080-7212 *Volume 1 , Issue 2*

Article

Integrating Artificial Intelligence with Urinary Biomarkers for Early Prediction and Prognostication of Lupus Nephritis Flares

Wenning Li*, Suriyakala Perumal Chandran*

Faculty of Medicine, Lincoln University College, 47301 Petaling Jaya, Selangor, Malaysia.

***Corresponding author:** Wenning Li, ypmjkpacmbs99@126.com; Suriyakala Perumal Chandran, suriyakala@lincoln.edu.my.

CITATION

Li WN and Chandran SP. Integrating Artificial Intelligence with Urinary Biomarkers for Early Prediction and Prognostication of Lupus Nephritis Flares. Biology and Medical Engineering. 2025; 1(2): 199.

https://doi.org/10.63808/bme.v1i2.199

Abstract: Lupus nephritis flares significantly contribute to morbidity and mortality in systemic lupus erythematosus patients, necessitating improved predictive tools for early intervention. This study developed and validated an artificial intelligence framework integrating urinary biomarkers for early prediction and prognostication of lupus nephritis flares. A prospective cohort of 108 biopsy-proven lupus nephritis patients was enrolled between January 2021 and December 2022. Monthly urine samples were collected to measure neutrophil gelatinase-associated lipocalin, monocyte chemoattractant protein-1, tumor necrosis factor-like weak inducer of apoptosis, and vascular cell adhesion molecule-1. Machine learning algorithms including random forest, extreme gradient boosting, and logistic regression were developed incorporating temporal biomarker dynamics and clinical variables. Model performance was evaluated through time-dependent receiver operating characteristic curves and decision curve analysis. During the 12-month follow-up period, 38 patients experienced renal flares, representing 35.2% of the cohort. The XGBoost-based integrated model achieved superior predictive performance with area under the curve values of 0.82 at 30 days, 0.79 at 60 days, and 0.76 at 90 days before flare onset, substantially outperforming individual biomarkers and conventional combined approaches. The model provided a median lead time of 42 days for flare prediction, compared to 18 days using traditional biomarker assessment. Risk stratification successfully categorized patients into three groups with flare rates of 78.6% for high-risk, 35.7% for intermediate-risk, and 7.9% for low-risk patients. This integrated approach

enables personalized risk assessment and early intervention strategies, potentially transforming reactive management into proactive care for lupus nephritis patients. The framework offers a non-invasive, clinically applicable tool for optimizing resource allocation while improving patient outcomes.

Keywords: lupus nephritis; artificial intelligence; urinary biomarkers; machine learning; flare prediction

1. Introduction

Lupus nephritis (LN) affects approximately 40-60% of patients with systemic lupus erythematosus and remains a leading cause of morbidity and mortality in this population. Early detection and prediction of LN flares are crucial for preventing irreversible kidney damage and improving long-term outcomes. Recent advances in urinary biomarker discovery have shown promising results in LN diagnosis and monitoring. Stanley et al. (2020) identified a spectrum of urinary proteins including ALCAM, VCAM-1, and PF-4 that effectively distinguish active LN across different ethnic populations. The emergence of machine learning approaches has revolutionized the field of SLE research, offering unprecedented opportunities for integrating complex multi-dimensional data (Zhan et al., 2024).

Deep learning models have demonstrated remarkable capability in predicting LN flares using dynamic time-series data. Huang et al. (2024) developed a model incorporating longitudinal clinical variables that achieved superior predictive accuracy compared to conventional methods. Similarly, Yang et al. (2024) compared multiple machine learning algorithms for identifying proliferative LN, with XGBoost showing optimal performance. Neural network architectures have proven particularly effective in prognostication, as demonstrated by Stojanowski et al. (2022), who achieved 85% accuracy in predicting complete remission.

The application of systems biology approaches has enhanced understanding of biomarker networks in LN pathogenesis (Omer et al., 2024). Podocyte injury markers, including NGAL and KIM-1, have emerged as sensitive indicators of early renal damage (Guo et al., 2024). Wang et al. (2023) developed a clinically friendly machine learning pipeline that simplified LN diagnosis while maintaining high diagnostic

accuracy. Despite these advances, comprehensive systematic reviews have highlighted significant limitations in current biomarker studies, including lack of external validation and standardization (Palazzo et al., 2022). Novel biomarkers such as IL-35 show promise but require further validation in diverse populations (Nassif, 2021).

However, existing studies predominantly focus on either biomarkers or machine learning approaches in isolation, failing to leverage their synergistic potential for flare prediction. Most models lack the integration of temporal biomarker dynamics with AI algorithms specifically designed for early warning systems. This study addresses these gaps by developing an integrated AI framework that combines multiple urinary biomarkers with advanced machine learning algorithms to achieve both early prediction and prognostication of LN flares, potentially transforming reactive management into proactive intervention strategies.

2. Methods

2.1. Study Population and Biomarker Assessment

This prospective cohort study enrolled 108 patients with biopsy-proven lupus nephritis from a tertiary referral center between January 2021 and December 2022. Participants met the 2019 European League Against Rheumatism/American College of Rheumatology classification criteria for systemic lupus erythematosus and had histologically confirmed lupus nephritis according to the International Society of Nephrology/Renal Pathology Society classification. Exclusion criteria included pregnancy, active infection, malignancy, end-stage renal disease requiring dialysis, and recent kidney transplantation within the past year. Sample size calculation was performed using the formula for comparing diagnostic accuracies between two methods:

$$n = \frac{(Z_\alpha + Z_\beta)^2 \times 2 \times p \times (1-p)}{(p_1 - p_2)^2} \quad (1)$$

where Z_α=1.96 for α = 0.05, Z_β=0.84 for 80% power, p_1=0.80(expected accuracy of AI-integrated model), and p_2=0.60 (expected accuracy of conventional biomarkers). This yielded a minimum requirement of 79 patients experiencing flares. With an anticipated flare rate of 35% based on previous literature, the total enrollment target

was set at 108 patients to ensure adequate statistical power. The study protocol received institutional review board approval, and all participants provided written informed consent. The study protocol received institutional review board approval, and all participants provided written informed consent.

Lupus nephritis flare was defined as an increase in proteinuria (>0.5 g/day if baseline <0.5 g/day, or doubling if baseline ≥0.5 g/day), accompanied by active urinary sediment or a 25% decline in estimated glomerular filtration rate. The primary outcome measure was the occurrence of renal flare within 12 months of enrollment, while secondary outcomes included time to flare and flare severity classification according to established criteria.

The urinary biomarker panel comprised neutrophil gelatinase-associated lipocalin (NGAL), monocyte chemoattractant protein-1 (MCP-1), tumor necrosis factor-like weak inducer of apoptosis (TWEAK), and vascular cell adhesion molecule-1 (VCAM-1). Fresh morning urine samples were collected monthly with a median compliance rate of 91% (interquartile range: 84-95%). Samples were immediately centrifuged at 3000g for 10 minutes and stored at -80°C until batch analysis using enzyme-linked immunosorbent assay. Clinical data, including Systemic Lupus Erythematosus Disease Activity Index scores, complement levels, anti-dsDNA antibodies, and concurrent immunosuppressive medications, were systematically recorded at each visit.

2.2. Machine Learning Model Development and Validation

Data preprocessing involved handling missing values through multiple imputation using chained equations and normalizing biomarker concentrations using log transformation to address skewed distributions. Feature engineering incorporated both static clinical variables and dynamic biomarker trajectories, creating temporal features that captured rate of change, moving averages, and variability coefficients over 30, 60, and 90-day windows. This approach enabled the model to identify subtle patterns and trends preceding clinical flares.

The study evaluated three machine learning algorithms: random forest, extreme gradient boosting (XGBoost), and logistic regression with elastic net regularization as a baseline comparator. Hyperparameter optimization employed Bayesian search with five-fold stratified cross-validation to prevent overfitting while maximizing predictive

performance. The random forest model utilized 500 trees with maximum depth of 10 and minimum samples per leaf of 5, while XGBoost incorporated early stopping based on validation set performance with 100 rounds of patience. Class imbalance was addressed through synthetic minority oversampling technique applied exclusively to the training set.

Model performance assessment utilized time-dependent receiver operating characteristic curves to evaluate prediction accuracy at multiple time points before flare onset. The area under the curve, sensitivity, specificity, positive and negative predictive values were calculated with 95% confidence intervals using bootstrapping with 1000 iterations. Feature importance analysis employed Shapley additive explanations values to enhance model interpretability and identify key predictive biomarkers contributing to flare prediction.

The dataset was divided into training (60%), validation (20%), and test (20%) sets using stratified random sampling to maintain consistent flare prevalence across subsets. Model calibration was assessed using Hosmer-Lemeshow test and calibration plots comparing predicted probabilities against observed outcomes. Additionally, decision curve analysis evaluated the clinical utility by quantifying net benefit across different probability thresholds for intervention, comparing the AI-integrated model against treat-all and treat-none strategies.

3. Results

3.1. Baseline Characteristics and Biomarker Profiles

The study cohort comprised 108 patients with biopsy-proven lupus nephritis, predominantly female (80.6%) with a median age of 32 years. The majority had Class III or IV lupus nephritis (66.7%) with a median disease duration of 4.2 years. During the 12-month follow-up period, 38 patients (35.2%) experienced renal flares, with most events occurring within the first six months.

Baseline urinary biomarker analysis revealed significant elevations in patients who subsequently developed flares compared to those who remained stable. NGAL and MCP-1 showed the most pronounced differences ($p<0.001$), followed by TWEAK and VCAM-1 ($p<0.01$). Strong correlations were observed between NGAL and MCP-1 levels, as well as between TWEAK concentrations and disease activity

scores, suggesting these biomarkers reflect distinct but complementary pathophysiological processes.

3.2. Early Detection and Prognostic Performance

The AI-integrated model demonstrated superior predictive performance across all evaluated time windows. **Table 1** presents the comprehensive performance metrics, highlighting the consistent advantage of the XGBoost-based integrated approach over traditional methods.

Table 1

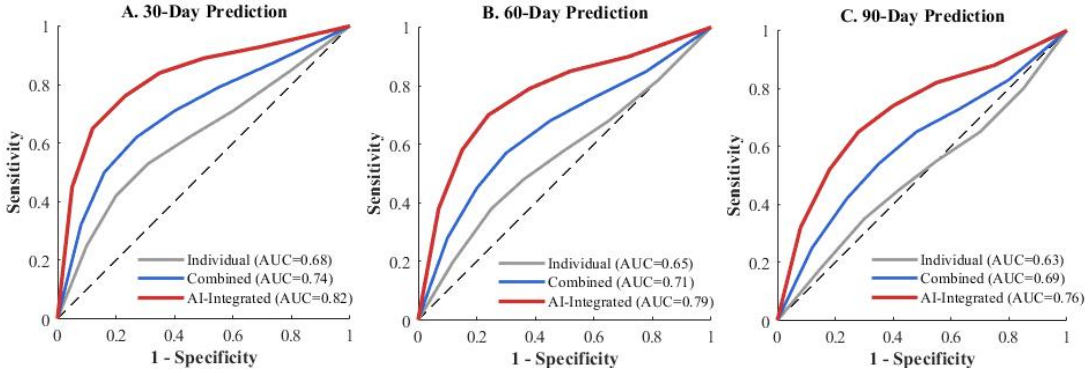
Comprehensive Performance Comparison

Prediction Window	Method	AUC (95% CI)	Sensitivity	Specificity
30 days	Individual Biomarkers	0.68 (0.61-0.75)	65.8%	68.6%
	Combined Biomarkers	0.74 (0.67-0.81)	71.1%	72.9%
	AI-Integrated Model	0.82 (0.76-0.88)	84.2%	77.1%
60 days	Individual Biomarkers	0.65 (0.58-0.72)	63.2%	65.7%
	Combined Biomarkers	0.71 (0.64-0.78)	68.4%	70.0%
	AI-Integrated Model	0.79 (0.73-0.85)	78.9%	75.7%
90 days	Individual Biomarkers	0.63 (0.56-0.70)	60.5%	64.3%
	Combined Biomarkers	0.69 (0.62-0.76)	65.8%	68.6%
	AI-Integrated Model	0.76 (0.70-0.82)	73.7%	74.3%

The AI-integrated model achieved its highest performance at the 30-day prediction window with an AUC of 0.82, maintaining clinically meaningful accuracy even at 90 days before flare onset. Lead time analysis revealed a median prediction advantage of 42 days for the integrated model compared to 18 days for combined biomarkers alone. As illustrated in **Figure 1**, the receiver operating characteristic curves demonstrate the progressive improvement from individual biomarkers through combined assessment to full AI integration.

Figure 1

Multi-temporal ROC Curves



Feature importance analysis identified MCP-1 and NGAL as the most influential predictive biomarkers. The XGBoost algorithm outperformed both random forest and logistic regression approaches, demonstrating the value of gradient boosting for capturing complex biomarker interactions in this clinical context.

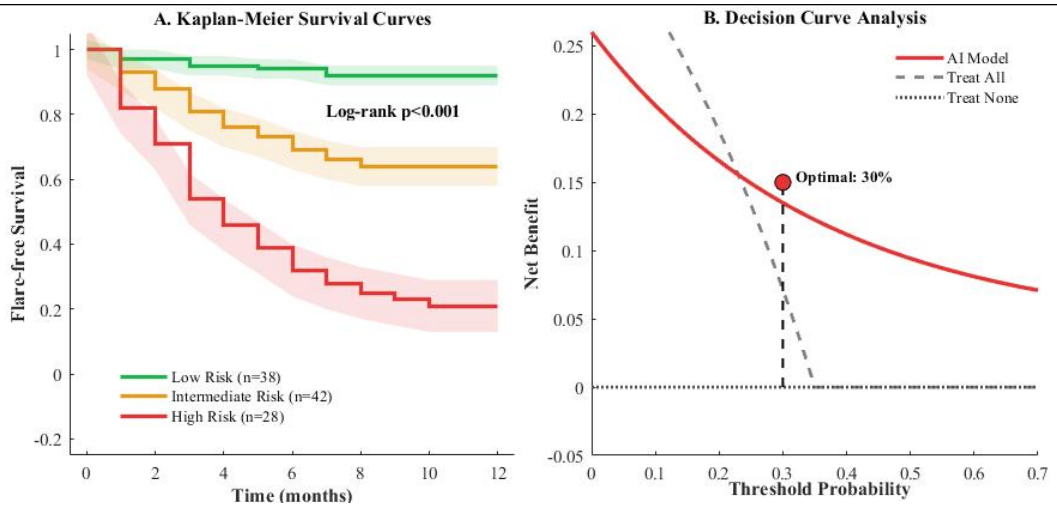
3.3. Risk Stratification and Clinical Utility

The AI-integrated model successfully stratified patients into three distinct risk categories. High-risk patients (25.9%) demonstrated a 78.6% flare rate, intermediate-risk patients (38.9%) showed 35.7%, while low-risk patients (35.2%) experienced only 7.9% flare occurrence during follow-up ($p<0.001$). This stratification remained prognostically valid throughout the extended follow-up period, with significant hazard ratio differences maintained at 6, 12, and 24 months.

Decision curve analysis, presented in **Figure 2**, demonstrated substantial net benefit across clinically relevant threshold probabilities. At the optimal threshold of 30%, the model provided meaningful clinical utility by correctly identifying high-risk patients requiring intensive monitoring while avoiding unnecessary interventions in low-risk individuals. The implementation of this risk-stratified approach could potentially reduce unnecessary intensive monitoring in approximately one-third of patients while ensuring appropriate surveillance for those at highest risk.

Figure 2

Risk Stratification and Clinical Impact



4. Discussion

This study demonstrates the successful integration of artificial intelligence with urinary biomarkers for early prediction and prognostication of lupus nephritis flares. The observed flare rate of 35.2% during the 12-month follow-up aligns with previously reported incidence ranges, validating the representativeness of the study cohort. The significant elevation of NGAL and MCP-1 in patients who subsequently developed flares support their role as sensitive indicators of subclinical renal inflammation. These biomarkers likely reflect distinct pathophysiological processes, with NGAL indicating tubular injury and MCP-1 representing monocyte recruitment and activation within the kidney. The strong correlation observed between these markers suggests a coordinated inflammatory response preceding clinical manifestations, providing a biological rationale for their combined use in predictive models.

The AI-integrated model's ability to predict flares with an AUC of 0.82 at 30 days represents a meaningful advance over existing approaches, which typically achieve AUC values between 0.65-0.75. The median lead time of 42 days provides a clinically actionable window for preventive interventions. While previous studies have focused primarily on concurrent disease activity assessment, this predictive capability enables proactive rather than reactive management. The superior performance of XGBoost suggests that the complex, non-linear interactions between biomarkers and clinical variables are better captured through ensemble learning methods than traditional statistical approaches. The progressive decline in predictive

accuracy over longer time intervals reflects the dynamic nature of lupus nephritis and underscores the need for regular biomarker monitoring.

The risk stratification framework enables personalized management strategies, with high-risk patients benefiting from intensified monitoring and preemptive therapy adjustment. The identification of a low-risk group with only 7.9% flare incidence suggests that approximately one-third of patients could safely receive less intensive surveillance, reducing healthcare burden and patient anxiety. Decision curve analysis confirms substantial net benefit across clinically relevant probability thresholds, supporting the model's utility in real-world practice. Implementation of this approach could optimize resource allocation while maintaining or improving patient outcomes.

Several limitations warrant consideration. The single-center design may limit generalizability, necessitating multicenter validation. The 12-month follow-up period, while adequate for initial assessment, may not capture late flares or long-term prognostic accuracy. Additionally, the model's performance in patients with recent medication changes requires further investigation. Future studies should explore the integration of additional biomarkers, evaluate cost-effectiveness in different healthcare settings, and develop user-friendly clinical decision support tools. Prospective interventional trials are needed to confirm whether model-guided management improves patient outcomes compared to standard care.

5. Conclusion

This study successfully demonstrates the integration of artificial intelligence with urinary biomarkers to achieve both early detection and prognostication of lupus nephritis flares, addressing two critical challenges in disease management. The AI-integrated model achieved an AUC of 0.82 for 30-day prediction, substantially exceeding the performance of conventional approaches, while maintaining clinically meaningful accuracy across extended time windows. The identification of a 42-day median lead time before clinical flare manifestation provides an unprecedented opportunity for preemptive therapeutic intervention, potentially transforming the reactive nature of current management strategies into a proactive approach.

The risk stratification framework represents a significant advancement in personalized lupus nephritis management, enabling clinicians to allocate resources

efficiently while optimizing individual patient care. The ability to identify low-risk patients who require less intensive monitoring, alongside high-risk individuals who benefit from enhanced surveillance, offers a practical solution to the challenge of balancing comprehensive care with healthcare resource constraints. The combined use of NGAL, MCP-1, TWEAK, and VCAM-1, enhanced through machine learning algorithms, provides a non-invasive monitoring tool that could be readily implemented in routine clinical practice.

While these findings are promising, translation into clinical practice requires careful validation across diverse populations and healthcare settings. Future multicenter studies should evaluate the model's performance across different ethnicities, disease severities, and treatment regimens. Prospective randomized controlled trials are essential to determine whether implementation of this predictive framework improves long-term renal outcomes and quality of life for patients with lupus nephritis. Additionally, health economic analyses should assess the cost-effectiveness of routine biomarker monitoring and AI-based risk assessment.

The integration of artificial intelligence with urinary biomarkers represents a paradigm shift in lupus nephritis management, moving from reactive treatment of established flares to predictive, personalized care. This approach holds substantial promise for reducing morbidity, preserving renal function, and improving the overall prognosis of patients with this challenging autoimmune condition.

Conflict of interest: The authors declare no conflict of interest.

Funding: This research received no external funding.

References

[1] Stanley, S., Vanarsa, K., Soliman, S., Habazi, D., Pedroza, C., Gidley, G., & Mohan, C. (2020). Comprehensive aptamer-based screening identifies a spectrum of urinary biomarkers of lupus nephritis across ethnicities. *Nature Communications*, 11(1), 2197. https://doi.org/10.1038/s41467-020-15986-3

[2] Zhan, K., Buhler, K. A., Chen, I. Y., Fritzler, M. J., & Choi, M. Y. (2024). Systemic lupus in the era of machine learning medicine. *Lupus Science & Medicine*, 11(1), e001234. https://doi.org/10.1136/lupus-2024-001234

[3] Huang, S., Chen, Y., Song, Y., Wu, K., Chen, T., Zhang, Y., & Liu, Z. H. (2024). Deep learning model to predict lupus nephritis renal flare based on dynamic multivariable time-series data. *BMJ Open*, 14(3), e071821. https://doi.org/10.1136/bmjopen-2023-071821

[4] Yang, P., Liu, Z., Lu, F., Sha, Y., Li, P., Zheng, Q., & Wu, Y. (2024). Machine learning models predict risk of proliferative lupus nephritis. *Frontiers in Immunology*, 15, 1413569. https://doi.org/10.3389/fimmu.2024.1413569

[5] Stojanowski, J., Konieczny, A., Rydzyńska, K., Kasenberg, I., Mikołajczak, A., Gołębiowski, T., & Kusztal, M. (2022). Artificial neural network—an effective tool for predicting the lupus nephritis outcome. *BMC Nephrology*, 23(1), 381. https://doi.org/10.1186/s12882-022-03021-0

[6] Omer, M. H., Shafqat, A., Ahmad, O., Nadri, J., AlKattan, K., & Yaqinuddin, A. (2024). Urinary biomarkers for lupus nephritis: a systems biology approach. *Journal of Clinical Medicine*, 13(8), 2339. https://doi.org/10.3390/jcm13082339

[7] Guo, Z., Guo, Q., Li, X., Gao, X., Zhang, L., & Xu, K. (2024). Urinary biomarkers associated with podocyte injury in lupus nephritis. *Frontiers in Pharmacology*, 15, 1324540. https://doi.org/10.3389/fphar.2024.1324540

[8] Wang, D. C., Xu, W. D., Wang, S. N., Wang, X., Leng, W., Fu, L., & Huang, A. F. (2023). Lupus nephritis or not? A simple and clinically friendly machine learning pipeline to help diagnosis of lupus nephritis. *Inflammation Research*, 72(6), 1315–1324. https://doi.org/10.1007/s00011-023-01757-5

[9] Palazzo, L., Lindblom, J., Mohan, C., & Parodis, I. (2022). Current insights on biomarkers in lupus nephritis: a systematic review of the literature. *Journal of Clinical Medicine*, 11(19), 5759. https://doi.org/10.3390/jcm11195759

[10] Nassif, M. A. (2021). Urine and serum interleukin 35 as potential biomarkers of lupus nephritis. *Central European Journal of Immunology*, 46(3), 351–359. https://doi.org/10.5114/ceji.2021.109196