

Clinical characteristics and prognosis of pure and mixed proliferative LN: a random survival forest model integrating histopathological and clinical features

Anjing Wang

Xijing 986 Hospital Department, Fourth Military Medical University.

Yunlong Qin

Department of Nephrology, Bethune International Peace Hospital,

Xiaorong Zou

Xijing 986 Hospital Department, Fourth Military Medical University

Jin Zhao

Xijing Hospital, The Fourth Military Medical University

Shiren Sun

sunrsun@163.com

Xijing Hospital <https://orcid.org/0000-0001-8580-9746>

Research Article

Keywords: proliferative lupus nephritis, risk factors, prognostic model, machine learning

Posted Date: August 22nd, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10514337/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Objective

This retrospective cohort study aimed to characterize clinical and prognostic differences between pure and mixed proliferative lupus nephritis (LN), identify outcome predictors, and develop a predictive model.

Methods

Biopsy-proven proliferative LN patients (2013–2023) were enrolled. Primary outcome was $\geq 50\%$ eGFR decline, end-stage kidney disease, or death. Machine learning and Cox regression were used for model development, assessed by C-index, integrated Brier score (IBS), net reclassification index (NRI), and integrated discrimination improvement (IDI).

Results

Among 408 patients (192 pure, 216 mixed), median follow-up was 47.8 months. Composite endpoints occurred in 24.0% (pure) and 25.5% (mixed). Mixed LN had higher chronic index (CI) ($P = 0.016$) and laboratory differences, but similar therapies and prognosis. Seven variables (CI, eGFR, age, blood urea nitrogen, red blood cell count, basophil percentage, uric acid) formed a random survival forest (RSF) model, achieving C-index 0.842 and IBS 0.154, with superior reclassification and discrimination.

Conclusion

Pure and mixed proliferative LN show heterogeneous features but similar outcomes, highlighting limits of pathology-based risk stratification. The RSF model integrating seven factors offers improved prediction, enabling earlier precise intervention.

1. Introduction

Systemic Lupus Erythematosus (SLE) is a prototypical systemic autoimmune disorder characterized by multisystem involvement and highly variable clinical manifestations (1, 2). Lupus nephritis (LN) represents one of the most severe organ manifestations of SLE, occurring in an estimated 30–60% of adult patients, and serving as a major driver of SLE-related morbidity and mortality (3, 4). Renal histopathology had been considered as the gold standard for diagnosing LN, providing critical guidance for therapeutic decision-making and prognostic assessment (5). Notably, up to 90% of patients with SLE exhibit pathological manifestations of LN when they undergo renal biopsy (6). Among the six histological classes defined by the International Society of Nephrology/Renal Pathology Society (ISN/RPS) classification system, proliferative LN is the most prevalent subtype and is strongly associated with severe clinical manifestations, the necessity for aggressive immunosuppression, and a poorer renal

prognosis (7). Therefore, it is imperative to conduct further research on proliferative LN to advance long-term renal outcomes.

Proliferative LN, encompassing both pure and mixed proliferative forms, is recommended to receive the same initial treatment regimen according to current guidelines (8). However, whether there are differences in clinical features and prognosis between these two pathological subtypes remains controversial in clinical studies (9–11). More critically, although the level of creatinine, estimated glomerular filtration rate (eGFR), and chronicity index (CI) have been determined that was linked to adverse outcomes in pure proliferative LN, patients with the mixed proliferative form have not been included from these analyses (12, 13). Therefore, the prognostic determinants specific to proliferative LN remain incompletely defined, which hindered the development of stratified management strategies. Therefore, comparing the clinical features and long-term outcomes of between pure and mixed proliferative LN and identifying prognostic determinants of proliferative LN are crucial for enabling individualized risk assessment and improving patient prognosis.

Recent years have witnessed growing application of machine learning (ML) in medicine, particularly for disease prognosis prediction (14). The Random Survival Forest (RSF), a ML method specialized for survival analysis that demonstrates robust performance in predicting patient survival (15). eXtreme Gradient boosting (XGboost) has excelled in numerous ML and data mining (16). However, ML-based approaches have yet to be systematically employed in predicting the long-term prognosis of patients with proliferative LN. Therefore, the present study seeks to perform a comparative analysis of clinical characteristics and prognostic trajectories, identify key factors influencing the long-term outcomes of proliferative LN patients, and develop a precise predictive model to facilitate early therapeutic intervention and support clinical decision-making by providing data-driven prognostic insights.

2. Materials and methods

2.1 Study population

The clinical and histopathological profiles of patients diagnosed with proliferative LN following renal biopsy at Xijing Hospital, Air Force Medical University, were collected for the period spanning December 2013 to December 2023. Patient inclusion required: (1) age ≥ 18 years; (2) renal biopsy-confirmed proliferative LN with complete follow-up data. Exclusion criteria were: (1) presence of other concomitant secondary nephropathy (e.g., diabetic nephropathy); (2) inadequate biopsy specimen (< 8 glomeruli) or missing original pathology reports; (3) suffering other life-limiting disease with an expected survival < 1 year; or (4) missing outcome data. We followed the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) guidelines throughout this study (17) (Table S1). Due to the retrospective nature of this study, informed consent was not required, as approved by the Institutional Review Board of Xijing Hospital (ethical number: KY20213027-1).

2.2 Outcomes

The primary outcome was a composite measure incorporating any of the following events, whichever occurred first, including a sustained eGFR decline of 50% or more, development of ESKD, or death. The eGFR was calculated using the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) Eq. (18). ESKD was defined as an eGFR < 15 ml/min/1.73 m², requiring initiation of long-term renal replacement therapy, including hemodialysis, peritoneal dialysis, or kidney transplantation (19). The follow-up duration was calculated from the date of renal biopsy until the final outpatient visit or telephone interview.

2.3 Data collection

Two investigators (Wang and Zhao) collected baseline demographic, clinical, and pathological data from the electronic medical record system (Table S2). Baseline characteristics were compared between the pure and mixed proliferative LN groups. Therapeutic approaches were stratified into four types: (1) corticosteroids alone, (2) corticosteroids plus immunosuppressive agents, (3) combination therapy with corticosteroids, immunosuppressants, and novel biological agents, and (4) other treatments. Comorbidity or other diseases were documented based on the diagnostic results of patients. A minimum follow-up period of six months was required for inclusion, barring earlier attainment of study endpoints. During follow-up, we assessed survival outcomes, ESKD progression, dialysis initiation, and laboratory data. Stringent confidentiality protocols were maintained throughout data collection and all subsequent research phases.

2.4 Renal pathological assessment

The pathological assessment of LN cases at this institution was performed by senior pathologists from Shaanxi Provincial Renal Pathology Research Center. The assessment applied the AI and CI scoring system, which was established by the National Institutes of Health (NIH) (20). Detailed information is provided in Table S3.

2.5 Model Development and Performance

R software (Version 4.3.2) was employed for model development and internal validation. Variable selection was predominantly guided by variable importance (VIMP), an intrinsic metric derived from the RSF algorithm (21). To assess the stability of VIMP estimates, bootstrap resampling with 1000 iterations was performed to generate 95% confidence intervals. Predictors achieving a mean VIMP > 0.03 were retained as significant prognostic factors. The cohort was randomly partitioned into training (70%) and testing (30%) sets. Grid search was utilized to identify optimal hyperparameters. The final RSF model was configured with 150 trees, an mtry of 3, and a nodesize of 6 (Figure S1). For the XGboost model, For the XGboost model, the parameters were: learning rate at 0.01, tree depth at 1, subsample at 1, colsample_bytree at 1, and gamma at 0.5. The specific R packages utilized were “survival”, “MASS” “randomForestSRC”, “survivalsvm”, and “XGboost”.

We evaluated model discrimination using the area under the receiver operating characteristic curve (AUC) and the concordance index (C-index). Time-dependent AUC (tAUC) was computed to capture discriminative performance across the entire study timeframe. Calibration ability was assessed through

the integrated Brier score (IBS) and visual inspection of calibration plots. Decision curve analysis (DCA) was performed to determine clinical net benefit. Additionally, we calculated the integrated discrimination improvement (IDI) and net reclassification improvement (NRI) to quantify improvements in risk stratification. Superior NRI values signified better reclassification capability of the new models compared to the conventional Cox model, whereas elevated IDI values denoted enhanced predictive accuracy.

2.6 Statistical analysis

We excluded variables with missing data rates exceeding 20% or correlation coefficients surpassing 0.75. Random forest imputation was employed to address missingness. Categorical data were summarized as counts and percentages, with comparisons conducted using the χ^2 test or Fisher's exact test as indicated. Continuous variables following a normal distribution were expressed as mean $\pm$ standard deviation, while those with skewed distributions were reported as median with interquartile range. We employed the Mann-Whitney U test or Student's t-test to compare continuous variables between groups. To determine optimal cut-off points for continuous predictors associated with survival, we utilized the Jsuvival-Survival Module in the ClinicoPath software suite within Jamovi (22). Kaplan-Meier curves were constructed to visualize cumulative survival, and the log-rank test was applied to compare survival distributions. Statistical significance was defined as a two-tailed P value < 0.05 . All analyses were carried out using R (Version 4.3.2), SPSS (IBM Version 27.0), and Jamovi (Version 2.3, Sydney, Australia).

3. Results

3.1 Patient characteristics

We enrolled 408 patients with renal biopsy-confirmed proliferative LN, comprising 192 with pure proliferative LN and 216 with mixed proliferative LN (Fig. 1). During a median observation period of 47.8 (19.2, 87.7) months, combined events occurred in 46 (24.0%) pure and 55 (25.5%) mixed proliferative LN patients, with no statistically significant difference ($P = 0.725$). The systemic lupus erythematosus disease activity index (SLEDAI) was 10 ± 4 . The platelet was $152 (104,205) \times 10^9/L$, the white blood cell was $5.10 (3.49, 7.10) \times 10^9/L$, the hypersensitive C-reactive protein was $1.98 (0.81, 7.38) \text{ mg/L}$, the MAP was $100 (87.0, 112) \text{ mmHg}$, the urinary total protein (UTP) was $1916 (924, 3600) \text{ mg/24h}$, and the eGFR was $72.0 (47.8, 95.7) \text{ mL/min/1.73 m}^2$. The AI score was $6.00 (4.00, 8.00)$ and CI score was $2.00 (1.00, 3.00)$ (Table 1).

Table 1
Baseline and follow-up characteristics of the study subjects.

Characteristics	Total N = 408	Training cohort N = 288	Testing cohort N = 120	P-value
Baseline (at renal biopsy)				
Age (years)	36.1 ± 13.7	36.1 ± 13.6	36.1 ± 14.0	0.302
Female	351(86.0%)	249 (86.5%)	102 (85.0%)	0.818
LN classification				0.153
Class III	36 (8.8%)	28 (9.7%)	8(6.7%)	
Class IV	156 (38.2%)	118 (41.0%)	38 (31.7%)	
Cass III + V	59 (14.5%)	39 (13.5%)	20 (16.7%)	
Class IV + V	157 (38.5%)	103 (35.8%)	54 (45.0%)	
Duration of SLE (months)	8.0 (2.0, 48.0)	8.00 (2.0, 48.0)	10.0 (2.0, 36.0)	0.839
^a BMI (kg/m ²)	22.9 ± 3.93	22.7 ± 3.73	23.5 ± 4.33	0.071
^b MAP (mmHg)	100 (87 ,112)	100 (87,113)	99 (89,110)	0.914
UTP (mg/24h)	1916 (924,3600)	1856 (878,3508)	2084 (1086,3944)	0.152
WBC (×10 ⁹ /L)	5.10 (3.49,7.10)	5.00 (3.41,6.85)	5.40 (3.91,7.76)	0.122
BASO% (%)	0.221 ± 0.271	0.207 ± 0.220	0.254 ± 0.363	0.190
RBC (× 10 ⁹ /L)	3.61 ± 0.76	3.59 ± 0.76	3.64 ± 0.78	0.589
Hb (g/L)	105 ± 22.6	105 ± 22.6	106 ± 22.7	0.486
PLT (×10 ⁹ /L)	152 (104,205)	150 (103,196)	154 (104,226)	0.311
Urinary RBC (μL)	131 (36.3,430)	126 (38.7,467)	144 (35.5,418)	0.935
Urinary CAST (uL)	1.44 (1.00,8.54)	1.27 (1.00,8.35)	1.79 (1.02,8.76)	0.994
BUN (mmol/l)	11.2 ± 29.4	10.8 ± 28.5	12.0 ± 31.7	0.721
eGFR (ml/min/1.73 m ²)	72.0 (47.8,95.7)	72.7 (46.1,97.0)	68.6 (50.1,95.2)	0.844
UA (μmol/l)	363 (281,448)	355 (276,447)	386 (291,460)	0.243
CysC (mg/L)	1.60 (1.23,2.30)	1.60 (1.21,2.28)	1.63 (1.26,2.45)	0.672
Serum albumin (g/L)	28.0 ± 7.61	28.2 ± 7.42	27.4 ± 8.04	0.322

Characteristics	Total N = 408	Training cohort N = 288	Testing cohort N = 120	P-value
Baseline (at renal biopsy)				
C3 (mg/dl)	0.42 (0.29,0.61)	0.41 (0.28,0.59)	0.43 (0.32,0.62)	0.317
C4 (mg/dl)	0.07 (0.04,0.13)	0.07 (0.04,0.13)	0.07 (0.04,0.12)	0.790
hs.CRP (mg/L)	1.98 (0.81,7.38)	2.04 (0.81,7.07)	1.93 (0.81,8.14)	0.694
SLEDAI	9.80 ± 4.11	9.87 ± 4.20	9.62 ± 3.91	0.564
Positive ANA	403 (98.8%)	284 (98.6%)	119 (99.2%)	1.000
Positive anti-dsDNA	162 (39.7%)	118 (41.0%)	44 (36.7%)	0.485
Positive anti-RNP	171 (41.9%)	123 (42.7%)	48 (40.0%)	0.693
Positive anti-Sm	127 (31.1%)	93 (32.3%)	34 (28.3%)	0.503
Positive anti-SSA	225 (55.1%)	153 (53.1%)	72 (60.0%)	0.245
Positive anti-SSB	53 (13.0%)	38 (13.2%)	15 (12.5%)	0.977
AI	6.00 (4.00,8.00)	6.00 (4.00,8.00)	6.00 (4.00,8.00)	0.427
CI	2.00 (1.00,3.00)	2.00 (1.00,3.00)	2.00 (1.00,3.00)	0.805
Extrarenal manifestations				
Fever	85 (20.8%)	63 (21.9%)	22 (18.3%)	0.504
Arthralgia	127 (31.2%)	88 (31.0%)	39 (32.5%)	0.805
Photosensitivity	47 (11.5%)	35 (12.2%)	12 (10.0%)	0.652
Rash	48 (11.8%)	35 (12.2%)	13 (10.8%)	0.835
Hypertension	95 (23.3%)	71 (24.7%)	24 (20.0%)	0.376
Alopecia	139 (34.1%)	100 (34.7%)	39 (32.5%)	0.751
Oral ulcers	41 (10.0%)	31 (10.8%)	10 (8.33%)	0.573
First-line treatment				
Induction therapy				0.195
Glucocorticoid only	23 (5.64%)	20 (6.94%)	3 (2.50%)	
Immunosuppressive agent pluse	371 (90.9%)	258 (89.6%)	113 (94.2%)	
Novel biological agents pluse	11 (2.70%)	7 (2.43%)	4 (3.33%)	

Characteristics	Total N = 408	Training cohort N = 288	Testing cohort N = 120	P-value
Baseline (at renal biopsy)				
Other	3 (0.74%)	3 (1.04%)	0 (0.00%)	
Follow-up parameters				
Follow-up (months)	47.8 (19.2,87.7)	47.5 (19.3,87.7)	47.9 (18.9,87.70	0.809
Endpoint	101 (20.7%)	71 (24.7%)	30 (10.4%)	0.941
Note: Values are presented as the mean $\pm$ standard deviation, median (interquartile range) or n (%). LN, lupus nephritis. SLE, systemic lupus erythematosus. MAP, mean arterial pressure. BMI, body mass index. UTP, urinary total protein. WBC, white blood cell. BASO%, basophil percentage. RBC, red blood cell. Hb, hemoglobin. HCT, haematocrit. PLT, platelet. UA, uric acid. eGFR, estimated glomerular filtration rate. hs.CRP, hypersensitive C-reactive protein. AI, activity index. CI, chronicity index. SLEDAI, systemic lupus erythematosus Disease Activity Index. ANA, antinuclear antibodies. dsDNA, anti-double stranded deoxyribonucleic antibody. anti-Sm, anti-Smith antibody. anti-RNP, anti-ribosomal P antibody. anti-SSA, anti-SSA antibody. anti-SSB, anti-SSB antibody. ^a BMI was calculated as weight/height²; ^b MAP was calculated by diastolic blood pressure + systolic blood pressure/3 based on initial blood pressure measurements upon admission; ^c eGFR was calculated by the CKD-EPI (Chronic Kidney Disease Epidemiology Collaboration) equation.				

3.2 Comparisons of the baseline data and prognosis of pure and mixed proliferative LN patients

In clinical characteristics, we observed that patients with mixed proliferative LN presented with higher MAP ($P = 0.033$) and UTP ($P = 0.041$), coupled with lower serum albumin ($P = 0.003$), reduced urinary RBC ($P = 0.026$), and SLEDAI scores ($P < 0.001$). Patients in both cohorts primarily received a regimen consisting of glucocorticoids combined with either mycophenolate mofetil ($P = 0.221$) or cyclophosphamide ($P = 0.724$). Histopathological evaluation demonstrated comparable AI score between the two groups ($P = 0.916$). But the CI score was markedly elevated in the cohort with mixed proliferative LN ($P = 0.016$) (Table S4). Nevertheless, Kaplan-Meier analysis revealed that long-term renal survival rates were comparable between patients with pure and mixed proliferative LN ($P = 0.770$) (Fig. 2).

3.3 Screening of variables

Following the application of the RSF algorithm for missing data imputation and the removal of collinear variables, a total of 99 covariates were retained from an initial set of 114 for subsequent analysis. Based on the VIMP scores, seven key prognostic indicators including CI, eGFR, age, BUN, RBC, UA, BASO% were

identified as significant predictors (Table S5). These selected variables were subsequently incorporated into the final prognostic mode (Fig. 3).

3.4 Model Development and Performance

3.4.1 Comparison of the discriminatory capabilities of RSF, XGboost, and Cox model

We randomly partitioned the cohort into 288 patients for training and 120 for testing (Table 1). The two groups were well-matched in terms of clinical, pathological, and follow-up data, indicative of a robust study design. At the 12-month follow-up, the tAUC for the RSF model remained consistently higher than that of both the Cox and XGBoost models across the training and testing cohorts (Fig. 4).

In the training set, the RSF model achieved 12-, 36-, and 60-month AUCs of 0.941 (0.907, 0.975), 0.948 (0.920, 0.975) and 0.950 (0.920, 0.980), respectively. By contrast, the corresponding AUCs for the XGBoost model were 0.781 (0.666, 0.896), 0.805 (0.798, 0.936) and 0.837 (0.767, 0.907), and those for the Cox model were 0.833 (0.732, 0.934), 0.779 (0.727, 0.883), 0.770 (0.689, 0.850). In the testing set, the RSF model maintained superior discrimination with AUCs of 0.942 (0.895, 0.989), 0.896 (0.824, 0.968), and 0.890 (0.813, 0.967). In contrast, the XGBoost model yielded AUCs of 0.801 (0.704, 0.897), 0.769 (0.653, 0.885), and 0.731 (0.599, 0.863), while the Cox model exhibited AUCs of 0.852 (0.761, 0.943), 0.787 (0.670, 0.904), and 0.796 (0.688, 0.905), respectively (Fig. 5). Additionally, the C-index for the RSF model was significantly superior to that of the XGBoost and Cox models in both the training [0.917 (0.895, 0.939) vs 0.743 (0.659, 0.805) vs 0.753 (0.694, 0.812), $P < 0.001$] and testing cohorts 0.842 (0.764, 0.901) vs 0.683 (0.431, 0.765) vs 0.760 (0.668, 0.830), $P < 0.001$] (Table 2).

Table 2
The C-index and IBS score in the training and testing cohort.

Prediction models	Training cohort		Testing cohort	
	C-index (95%CI)	IBS (95%CI)	C-index (95%CI)	IBS (95%CI)
RSF model	0.917 (0.895, 0.939)	0.094 (0.087, 0.100)	0.842 (0.764, 0.901)	0.154 (0.125, 0.155)
XGboost model	0.743 (0.659, 0.805)	0.134 (0.128, 0.142)	0.683 (0.431, 0.765)	0.178 (0.155, 0.178)
Cox model	0.753 (0.694, 0.812)	0.140 (0.138, 0.142)	0.760 (0.668, 0.830)	0.171 (0.135, 0.176)
P	< 0.001	-	< 0.001	-

Note: IBS, integrated brier score. RSF, random survival forests. Cox, Cox regression. XGboost, eXtreme Gradient Boosting. C-index, concordance index. IBS, integrated brier score.

3.4.2 Comparison of the Calibration accuracy of the RSF, XGboost, and Cox model

The IBS was employed to evaluate the calibration of the RSF, XGBoost, and Cox models. In the testing cohort, the RSF model exhibited the lowest IBS at 0.154 (0.125–0.155), outperforming both the XGBoost model [IBS: 0.178 (0.155–0.178)] and the Cox model [IBS: 0.171 (0.135–0.176)] (Table 2).

3.4.3 Comparison of prediction efficiency of RSF, XGboost, and Cox model

Assessment of model improvement using categorical NRI and IDI revealed that the RSF model consistently outperformed the Cox model. The RSF model yielded positive NRI values at 12, 36, and 60 months were 0.411 (0.136, 0.630), 0.520 (0.297, 0.661) and 0.425 (0.210, 0.586), respectively, along with positive IDI values over the same intervals [0.103 (0.019–0.191), 0.171 (0.095–0.234), and 0.152 (0.077–0.216)]. By contrast, the XGBoost model showed negative NRI and IDI values across all time points, with all confidence intervals encompassing zero [NRI range: -0.145 to -0.071; IDI range: -0.047 to -0.003]. These findings confirm the superior predictive capacity of the RSF model relative to the Cox model (Table 3).

Table 3
NRI and IDI of prediction models compared with reference model.

Models		Cox model	RSF model	XGboost model	<i>P</i> ^a	<i>P</i> ^b
NRI	12-month	reference	0.411 (0.136, 0.630)	-0.095 (-0.068, 0.070)	0.004	0.837
	36-month		0.520 (0.297, 0.661)	-0.071 (-0.276, 0.056)	< 0.001	0.210
	60-month		0.425 (0.210, 0.586)	-0.145 (-0.315, 0.026)	< 0.001	0.008
IDI	12-month	reference	0.103 (0.019, 0.191)	-0.003 (-0.068, 0.007)	0.016	0.837
	36-month		0.171 (0.095, 0.234)	-0.027(-0.090, 0.033)	< 0.001	0.350
	60-month		0.152 (0.077, 0.216)	-0.047(-0.114, 0.016)	< 0.001	0.144

Note: NRI, net reclassification improvement. IDI, integrated discrimination improvement. RSF random survival forest. XGboost, eXtreme Gradient boosting.

P^a was calculated by comparing RSF model with the Cox model.

P^b was calculated by comparing XGboost model with Cox mod

3.4.4 Decision curve analysis of RSF, XGboost, and Cox model

Decision curve analysis (DCA) was performed to assess the clinical applicability of each model. As illustrated in Fig. 6, the RSF model demonstrated a superior net benefit compared to both the XGBoost and Cox models across all threshold probabilities at 12, 36, and 60 months. These findings indicate that the RSF model may provide greater clinical utility than the other two models.

3.5 Renal outcomes of proliferative LN patients under different risk factors

The variables including CI, eGFR, age, BUN, RBC, UA, and BASO% were identified as predictors of renal outcomes, ranked in descending order of the value of VIMP. Using the maximally selected rank statistics method (MSRSM), we determined optimal prognostic thresholds for each variable in patients with proliferative LN patients (Figure S2). Subsequent stratification into high- and low-risk groups—based on a $CI \geq 4$, $eGFR < 45.2 \text{ ml/min/1.73m}^2$, $\text{age} \geq 52 \text{ years}$, $BUN \geq 15.0 \text{ mg/dL}$, $RBC < 2.98 \times 10^{12}/L$, $UA \geq 510 \text{ } \mu\text{mol/L}$, or $BASO\% \leq 0.003\%$ —revealed that all high-risk subgroups experienced significantly poorer renal survival on Kaplan-Meier analysis (all $P < 0.05$) (Figure S3).

4. Discussion

In this study, we performed a comparative analysis of clinical features between patients with pure and mixed proliferative LN and subsequently conducted survival analyses to evaluate prognosis. We identified key risk factors influencing the prognosis of proliferative LN patients who underwent renal biopsy and developed novel prognostic prediction models. Although pure and mixed proliferative LN exhibit distinct pathological and clinical characteristics, they share comparable treatment strategies and long-term outcomes, suggesting that the traditional ISN/RPS classification alone is insufficient for accurate risk stratification. Seven variables including CI, eGFR, age, BUN, RBC, UA, and BASO% were identified as independent risk factors affecting long-term outcomes in proliferative LN. Regarding predictive performance and clinical applicability, the RSF model demonstrated superior prognostic capability compared to the XGBoost and Cox regression models.

The long-term renal prognosis of mixed versus pure proliferative LN remains contentious, with studies reporting inferior either outcomes in mixed LN (9, 10) or comparable risks between the two classes (23, 24). This discrepancy may stem from population heterogeneity, limited sample sizes, and varying class proportions across cohorts. In our study, the number of patients with pure proliferative LN was comparable to that of the mixed form, and no significant difference in long-term prognosis was observed between the two groups. Notably, patients with mixed proliferative LN presented a distinct profile characterized by elevated proteinuria, hypoalbuminemia, and higher chronicity index, yet lower SLEDAI scores, indicating a greater burden of chronic renal damage relative to systemic inflammatory activity. This aligns with the reports of Najafi et al. and Ikeuchi et al that highlights severe renal injury in mixed LN (9, 10), yet contrasts with those reporting higher systemic activity in pure forms (23, 24). Collectively, mixed proliferative LN exhibits a distinct high-risk phenotype, yet its long-term prognosis is comparable to pure proliferative LN under standardized therapy. This highlights the limitation of ISN/RPS

classification alone in prognostic prediction. Therefore, multidimensional models integrating clinical and pathological features are essential for precise risk stratification and optimizing outcomes in proliferative LN.

In the present study, the RSF model demonstrated superior overall performance in predicting disease progression among patients with proliferative LN. Compared to the Cox model and the XGBoost model, RSF exhibited enhanced discriminative ability, reflected by higher C-index and tAUC values, along with improved calibration accuracy indicated by a lower IBS. Furthermore, its significant improvements in clinical reclassification metrics, such as the NRI and IDI, further validated its enhanced predictive utility. The superior performance of RSF can be primarily attributed to its distinctive algorithmic advantages. On one hand, it does not rely on the proportional hazards assumption or log-linear relationship constraints required by the Cox model, allowing it to more flexibly capture potential complex nonlinear relationships and time-varying effects often present in clinical data (25–27). On the other hand, compared to XGBoost, which is also proficient in capturing nonlinear patterns, RSF employs dual randomization (bootstrap sampling and random feature selection) to effectively control overfitting, thereby enhancing the robustness of the model and generalizability, particularly when working with limited clinical datasets (28–30). These findings suggest that the individualized risk assessment provided by the RSF model can be better leveraged to identify high-risk proliferative LN patients who might otherwise be overlooked, thereby offering a valuable tool for precision management.

CI, a complementary tool to the 2003 ISN/RPS classification, quantifies chronic renal damage via semi-quantitative scoring of glomerulosclerosis and tubulointerstitial fibrosis (3, 31–34). This study confirms the CI as the foremost predictor of renal outcomes in proliferative LN, underscoring the decisive prognostic impact of baseline chronic irreversible damage. Current standard treatment plan, glucocorticoids combined with CYC or MMF, effectively suppresses active inflammation and induces remission(35, 36). However, evidence shows that chronic renal damage can progress rapidly despite a good initial immunosuppressive response, (37) indicating that that standard regimens poorly reverse established chronic lesions and that conventional remission markers may not reflect the underlying renal progression. Therefore, the initial biopsy CI demands high clinical priority. For high-risk proliferative LN patients ($CI \geq 4$), therapy must integrate the organ protection and the anti-fibrotic strategies early. Additionally, if renal function declines despite improved active markers, chronic progression likely dominates. Here, repeated biopsy is critical to pivot therapy from intensive immunosuppression toward enhanced organ protection.

eGFR, BUN, and UA were established as predictors of prognosis in proliferative LN in this study. eGFR and BUN are widely recognized as core indicators for prognostic assessment in renal diseases, and this is equally true for proliferative LN (38). Notably, our results highlight that patient with $eGFR < 45$ ml/min/1.73m² and/or $BUN \geq 15.0$ mg/dL represents a high-risk subgroup that warrants closer clinical monitoring. UA was also established as a predictor of prognosis in patients with proliferative LN in this study, a finding consistent with previous research (39–41). Elevated serum uric acid levels not only indicate a more severe clinical phenotype but also increase the risk of progression to ESRD or death in

patients with LN (41). In terms of pathogenesis, excessive UA has been shown to form micro-crystals in sites such as renal tubules, activating the NLRP3 inflammasome pathway, thereby amplifying local renal inflammation (42). Concurrently, elevated serum UA levels can damage the microvasculature and promote a pro-fibrotic phenotypic shift in renal resident cells, driving both microangiopathy and tubulointerstitial fibrosis (39, 43). Therefore, for proliferative LN patients with elevated SUA, particularly those with UA ≥ 510 $\mu\text{mol/L}$, proactive management of hyperuricemia should no longer be regarded merely as complication control but rather re-evaluated as a potential adjunctive therapeutic strategy that may confer renal benefits by mitigating the “metabo-inflammatory” pathway.

Age is a major prognostic predictor for proliferative LN, with its VIMP ranking second only to eGFR and CI, and it is a significant independent risk factor for LN mortality (44, 45). Elderly LN patients often present with insidious initial clinical presentation, but more severe renal chronic lesions and poor tolerance to intensive immunosuppression, which leading to diagnostic delays, therapeutic difficulties, and poor long-term renal prognosis (45–48). Pathologically, aging-related mitochondrial dysfunction promotes renal fibrosis through increased renal tubular reactive oxygen species production and impaired mitophagy, thereby exacerbating kidney injury (49). The mechanism likely underlies the more pronounced chronic renal damage and limited response to immunosuppression observed in elderly LN patients. Additionally, elderly LN patients often present with comorbidities such as hypertension and diabetes, which can impair glomerular filtration function and renal microcirculation, thereby exacerbating tubulointerstitial injury and promoting chronic fibrosis (47). In summary, for elderly patients (≥ 52 years) with proliferative LN, an early monitoring system integrating comorbidities and chronic injury biomarkers should be established, and anti-fibrotic therapies (e.g., metformin) should be combined with immunosuppression to improve prognosis.

BASO% was established in this study as a clinical predictor for prognosis in patients with proliferative LN, which was consistent with our previous finding (50). In recent years, a growing body of studies have revealed a close association between BASOs and LN. Clinically, Liang et al. found that patients with active LN typically present with reduced circulating BASOs (51). In terms of pathogenesis, BASO can be activated and recruited to secondary lymphoid organs (SLOs) via prostaglandin D2 or autoreactive IgE (52, 53). Within SLOs, activated BASO can promote the pathological accumulation of follicular helper T cells, which increased production of autoreactive anti-RNP IgG and renal IgG deposits (54). Concurrently, they also can activate both T and B cells to promote a TH2 environment, driving plasma cells to produce autoreactive IgE and IgG (55). These pathogenic immune complexes not only deposit in the kidney, inducing local inflammation and tissue injury, but also can enhance BASO activation, establishing an amplification loop that perpetuates LN progression (55). Thus, BASO% serves not only as a key prognostic predictor for proliferative LN, but also as an indicator of its disease-amplifying loop. For proliferative LN patients with reduced BASO%, particularly $< 0.003\%$, close monitoring and early targeted intervention should be emphasized, which may improve the long-term outcome in proliferative LN.

RBC was confirmed as a risk factor for adverse prognosis in proliferative LN patients in the present study. Immunohistological studies demonstrate that in LN, activated CD4 + lymphocytes and

macrophages infiltrate the renal interstitial regions and secrete cytokines. These cytokines not only impair Epo production but also disrupt the responsiveness of erythroid precursor cells to this hormone, ultimately resulting in decreased RBC counts and the development of anemia (56). Additionally, persistently low RBC counts reduce the oxygen-carrying capacity of the blood, leading to tissue hypoxia that exacerbates renal interstitial fibrosis and further worsens kidney prognosis. This pathological process has been confirmed in studies of anemia associated with kidney disease, indicating that a hypoxic state disrupts normal renal physiological homeostasis and accelerates the decline of renal function (57, 58). Notably, multiple studies have identified anemia as a significant prognostic factor associated with poor outcomes in patients with LN (59–61). As the core laboratory parameter for evaluating anemia, changes in red blood cell count can directly reflect the risk of disease progression. Therefore, for proliferative LN patients with low RBC counts ($\text{RBC} < 2.98 \times 10^{12}/\text{L}$), timely correction of anemia is indicated. This intervention helps alleviate the hypoxic state, protect renal function, delay disease progression, and ultimately improve the prognosis of proliferative LN.

This study compared the baseline characteristics and long-term prognosis between pure and mixed proliferative LN, performed variable selection and subgroup analyses, and constructed a RSF model to predict the prognosis of patients with proliferative LN, providing a reference for clinical risk stratification and individualized management. The main strengths of this study are as follows. First, the use of the RSF ML method effectively handles complex nonlinear relationships and interactions, improving predictive accuracy. Second, through rigorous variable selection and subgroup analyses, we not only identified key prognostic factors but also preliminarily explored their differential effects across various subgroups, enhancing the model's clinical interpretability and utility. At the same time, this study has certain limitations. First, the study population consisted exclusively of Chinese patients, lacking validation in cohorts from different geographical regions or ethnicities, which limits the generalizability of the model. Second, the study lacks external validation. To address these constraints, a large-scale, multicenter prospective study is planned to further validate the proposed model. Despite the aforementioned limitations, the insights derived from this study may contribute to early therapeutic intervention and enhanced long-term prognosis in proliferative LN patients.

Conclusion

Although patients with pure and mixed proliferative LN present with distinct pathological and clinical profiles, they receive comparable treatment and experience similar long-term outcomes, highlighting the limitations of the ISN/RPS classification system in predicting prognosis. To address this gap, we developed an RSF model incorporating seven readily accessible predictors including CI, eGFR, age, BUN, RBC, UA, and BASO%, which demonstrated superior discriminative ability and clinical utility compared to both XGBoost and Cox regression models. These findings establish a new foundation for risk stratification and management paradigms in proliferative LN, with their value lying in translating prognostic insights into actionable clinical strategies, thereby enabling earlier and more precise interventions to ultimately achieve the optimization of long-term outcomes.

Declarations

Conflict of Interest:

The authors report there are no competing interests to declare.

Note

NRI, net reclassification improvement. IDI, integrated discrimination improvement. RSF random survival forest. XGboost, eXtreme Gradient boosting.

Funding:

National Natural Science Foundation of China grants (reference number: 8270735, 82570807), Key Research and Development Plan of Shaanxi Province grant (reference number: No.2023-ZDLSF-15), Medical Science Research Project of Hebei (reference number: 20261407)

, and Key Project of Shaanxi Province (2022ZDLSF03-12) supported this study.

Author Contributions:

Conceptualization, A. Wang, and S. Sun; Methodology, J. Zhao; Software, A. Wang, Y. Qin; Validation, A. Wang, and Y. Qin; Data Collection, A. Wang; Writing-original draft, A. Wang, X. Zou, and Y. Qin, Writing-review and editing, B. Ming and S. Sun; Fund supporting, S. Sun. Each of the authors contributed an important role in drafting the manuscript and ensuring the accuracy or completeness of the overall work is properly investigated and resolved.

Acknowledgment:

The authors would like to acknowledge all statisticians for participating in this study.

Data Availability:

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

1. Kiriakidou M, Ching CL. Systemic Lupus Erythematosus. Ann Intern Med. 2020;172(11):ITC81–96.

2. Owen KA, Grammer AC, Lipsky PE. Deconvoluting the heterogeneity of SLE: The contribution of ancestry. *J Allergy Clin Immunol*. 2022;149(1):12–23.
3. Yu C, Li P, Dang X, Zhang X, Mao Y, Chen X. Lupus nephritis: new progress in diagnosis and treatment. *J Autoimmun*. 2022;132:102871.
4. Mohan C, Zhang T, Putterman C. Pathogenic cellular and molecular mediators in lupus nephritis. *Nat Rev Nephrol*. 2023;19(8):491–508.
5. Yu F, Haas M, Glassock R, Zhao MH. Redefining lupus nephritis: clinical implications of pathophysiologic subtypes. *Nat Rev Nephrol*. 2017;13(8):483–95.
6. Gasparotto M, Gatto M, Binda V, Doria A, Moroni G. Lupus nephritis: clinical presentations and outcomes in the 21st century. *Rheumatology (Oxford)*. 2020;59(Suppl5):v39–51.
7. Chen DN, Fan L, Wu YX, Zhou Q, Chen W, Yu XQ. A Predictive Model for Estimation Risk of Proliferative Lupus Nephritis. *Chin Med J (Engl)*. 2018;131(11):1275–81.
8. Kidney Disease. Improving Global Outcomes Lupus Nephritis Work G. KDIGO 2024 Clinical Practice Guideline for the management of LUPUS NEPHRITIS. *Kidney Int*. 2024;105(1S):S1–69.
9. Najafi CCKS. Significance of histologic patterns of glomerular injury upon long-term prognosis in severe lupus glomerulonephritis. *Kidney Int*. 2001;59(6):2156–63.
10. Ikeuchi H, Hiromura K, Kayakabe K, Tshilela KA, Nojima YJJJR. Renal outcomes in mixed proliferative and membranous lupus nephritis (Class III/IV + V): A long-term observational study. 2016;26(6):908–13.
11. Ilori TO, Enofe N, Oommen A, Cobb J, Navarrete J, Adedinsewo DA, et al. Comparison of Outcomes between Individuals with Pure and Mixed Lupus Nephritis: A Retrospective Study. *PLoS ONE*. 2016;11(6):e0157485.
12. Villalobos Navarro A, Xipell M, Hernandez EG, Valjalo R, Contreras F, Ruiz de Arechavaleta A et al. Prediction model for renal outcomes in Latin American Mestizo patients with pure proliferative lupus nephritis. *Lupus Sci Med*. 2025;12(2).
13. Ayodele OE, Okpechi IG, Swanepoel CR. Long-term renal outcome and complications in South Africans with proliferative lupus nephritis. *Int Urol Nephrol*. 2013;45(5):1289–300.
14. Handelman GS, Kok HK, Chandra RV, Razavi AH, Lee MJ, Asadi H. eDoctor: machine learning and the future of medicine. *J Intern Med*. 2018;284(6):603–19.
15. Tian D, Yan HJ, Huang H, Zuo YJ, Liu MZ, Zhao J, et al. Machine Learning-Based Prognostic Model for Patients After Lung Transplantation. *JAMA Netw Open*. 2023;6(5):e2312022.
16. Hou N, Li M, He L, Xie B, Wang L, Zhang R, et al. Predicting 30-days mortality for MIMIC-III patients with sepsis-3: a machine learning approach using XGboost. *J Transl Med*. 2020;18(1):462.
17. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ*. 2015;350:g7594.

18. Inker LA, Eneanya ND, Coresh J, Tighiouart H, Wang D, Sang YY, et al. New Creatinine- and Cystatin C-Based Equations to Estimate GFR without Race. *New Engl J Med*. 2021;385(19):1737–49.
19. Jia Q, Ma F, Zhao J, Yang X, Sun R, Li R, et al. Effect of corticosteroids combined with cyclophosphamide or mycophenolate mofetil therapy for IgA nephropathy with stage 3 or 4 chronic kidney disease: A retrospective cohort study. *Front Pharmacol*. 2022;13:946165.
20. Bajema IM, Wilhelmus S, Alpers CE, Bruijn JA, Colvin RB, Cook HT, et al. Revision of the International Society of Nephrology/Renal Pathology Society classification for lupus nephritis: clarification of definitions, and modified National Institutes of Health activity and chronicity indices. *Kidney Int*. 2018;93(4):789–96.
21. Ishwaran H, Kogalur UB, Blackstone EH, Lauer MS. Random survival forests. *Kidney Int*. 2008;2(3):18–28.
22. Qin Y, Yu Z, Wu H, Wang A, Wang F, Wang D, et al. Prognostic factors affecting long-term outcomes in patients with concurrent IgA nephropathy and membranous nephropathy. *Heliyon*. 2024;10(1):e23436.
23. Titilayo O, Ilori, Nosayaba, Enofe A et al. Comparison of Outcomes between Individuals with Pure and Mixed Lupus Nephritis: A Retrospective Study. 2016;11(6):e0157485.
24. Ahn SS, Yoo J, Lee SW, Song JJ, Park YB, Jung SMJL. Clinical characteristics and long-term outcomes in patients with mixed Class III/IV + V and pure proliferative lupus nephritis: A single-center experience. 2022;31(5):588–95.
25. Taylor JM. Random Survival Forests. *J Thorac Oncol*. 2011;6(12):1974–5.
26. Zhou Y, McArdle JJ. Rationale and Applications of Survival Tree and Survival Ensemble Methods. *Psychometrika*. 2015;80(3):811–33.
27. Yao W, Frydman H, Larocque D, Simonoff JS. Ensemble methods for survival function estimation with time-varying covariates. *Stat Methods Med Res*. 2022;31(11):2217–36.
28. Luo L, Orlow I, Kanetsky PA, Thomas NE, Fang S, Lee JE, et al. No prognostic value added by vitamin D pathway SNPs to current prognostic system for melanoma survival. *PLoS ONE*. 2017;12(3):e0174234.
29. Yeo HJ, Noh D, Son E, Kwon S, Cho WH. Machine Learning for 1-Year Mortality Prediction in Lung Transplant Recipients: ISHLT Registry. *Transpl Int*. 2025;38:14121.
30. Li SM, Ren MY, Gan J, Zhang SG, Kang MT, Li H, et al. Machine Learning to Determine Risk Factors for Myopia Progression in Primary School Children: The Anyang Childhood Eye Study. *Ophthalmol Ther*. 2022;11(2):573–85.
31. Austin HA 3rd, Boumpas DT, Vaughan EM, Balow JE. Predicting renal outcomes in severe lupus nephritis: contributions of clinical and histologic data. *Kidney Int*. 1994;45(2):544–50.
32. Moroni G, Pasquali S, Quaglini S, Banfi G, Casanova S, Maccario M, et al. Clinical and prognostic value of serial renal biopsies in lupus nephritis. *Am J Kidney Dis*. 1999;34(3):530–9.

33. Moroni G, Vercelloni PG, Quaglini S, Gatto M, Gianfreda D, Sacchi L, et al. Changing patterns in clinical-histological presentation and renal outcome over the last five decades in a cohort of 499 patients with lupus nephritis. *Ann Rheum Dis*. 2018;77(9):1318–25.
34. Parodis I, Adamichou C, Aydin S, Gomez A, Demoulin N, Weinmann-Menke J, et al. Per-protocol repeat kidney biopsy portends relapse and long-term outcome in incident cases of proliferative lupus nephritis. *Rheumatology (Oxford)*. 2020;59(11):3424–34.
35. De Vriese AS, Sethi S, Fervenza FC. Lupus nephritis: redefining the treatment goals. *Kidney Int*. 2025;107(2):198–211.
36. Anders HJ, Rovin B. A pathophysiology-based approach to the diagnosis and treatment of lupus nephritis. *Kidney Int*. 2016;90(3):493–501.
37. Malvar A, Pirruccio P, Alberton V, Lococo B, Recalde C, Fazini B, et al. Histologic versus clinical remission in proliferative lupus nephritis. *Nephrol Dial Transpl*. 2017;32(8):1338–44.
38. Li C, Zhang K, Liu L, Shen J, Wang Y, Tan Y et al. Study of the Mechanism of Astragali Radix in Treating Type 2 Diabetes Mellitus and Its Renal Protection Based on Enzyme Activity, Network Pharmacology, and Experimental Verification. *Molecules*. 2023;28(24).
39. Dos Santos M, Veronese FV, Moresco RN. Uric acid and kidney damage in systemic lupus erythematosus. *Clin Chim Acta*. 2020;508:197–205.
40. Ugolini-Lopes MR, Gavinier SS, Leon E, Viana VT, Borba EF, Bonfa E. Is serum uric acid a predictor of long-term renal outcome in lupus nephritis? *Clin Rheumatol*. 2019;38(10):2777–83.
41. Wang H, Qiu F, Liu J, Luo C, Liu X. Elevated serum uric acid is associated with renal arteriopathy and predict poor outcome in patients with lupus nephritis. *Clin Exp Rheumatol*. 2024;42(1):30–8.
42. Luo Y, Long M, Wu X, Zeng L. Targeting the NLRP3 inflammasome in kidney disease: molecular mechanisms, pathogenic roles, and emerging small-molecule therapeutics. *Front Immunol*. 2025;16:1703560.
43. Wang H, Chen Q, Ye Q, Liu L, Wei L. Development and validation of a nomogram for predicting the interstitial fibrosis and tubular atrophy in patients with lupus nephritis. *Rheumatology (Oxford)*. 2025;64(5):2647–55.
44. Rodríguez-Almaraz E, Gutiérrez-Solís E, Rabadán E, Rodríguez P, Carmona L, Morales E, et al. Something new about prognostic factors for lupus nephritis? A systematic review. *Lupus*. 2021;30(14):2256–67.
45. Calatroni M, Andrulli S, Doti F, Bello F, De Vivo G, Mastrangelo A, et al. Long-term prognosis of lupus nephritis: comparison between pediatric, adult, and advanced age onset. *Front Immunol*. 2025;16:1531675.
46. Yu Y, Jin H, Zhou S, Zhao M, Wu H, Long H, et al. The differential panorama of clinical features of lupus erythematosus patients with different onset ages: a cross-sectional multicenter study from China. *Clin Rheumatol*. 2023;42(9):2353–67.
47. Sun S, Li S, Chang X, Wu J. Age-stratified risk factors and predictive models for progression to lupus nephritis in patients with systemic lupus erythematosus: a review. *Front Immunol*. 2025;16:1682280.

48. Ceasovschih A, Alexa RE, Sorodoc V, Asaftei A, Stoian DC, Morarasu BC et al. Systemic Lupus Erythematosus Features in Elderly Patients: Case-Based Review. *J Clin Med*. 2025;14(8).
49. Cai W, Zhang H, Li Z, Cai M, Yu S, Chen P et al. Tissue Fibrosis Decoded via Cellular Senescence: Mechanisms, Treatments, and Emerging Technologies. *Aging Dis*. 2025.
50. Wang A, Qin Y, Xing Y, Yu Z, Huang L, Yuan J, et al. Clinical characteristics, prognosis, and predictive modeling in class IV +/- V lupus nephritis. *Front Immunol*. 2025;16:1580146.
51. Liang P, Tang Y, Lin L, Zhong H, Yang H, Zeng Y, et al. Low level of circulating basophil counts in biopsy-proven active lupus nephritis. *Clin Rheumatol*. 2018;37(2):459–65.
52. Pellefigues C, Dema B, Lamri Y, Saidoune F, Chavarot N, Loheac C, et al. Prostaglandin D(2) amplifies lupus disease through basophil accumulation in lymphoid organs. *Nat Commun*. 2018;9(1):725.
53. Charles N, Hardwick D, Daugas E, Illei GG, Rivera J. Basophils and the T helper 2 environment can promote the development of lupus nephritis. *Nat Med*. 2010;16(6):701–U107.
54. Tchen J, Simon Q, Chapart L, Thaminy MK, Vibhushan S, Saveanu L, et al. PD-L1- and IL-4-expressing basophils promote pathogenic accumulation of T follicular helper cells in lupus. *Nat Commun*. 2024;15(1):3389.
55. Charles N, Chemouny JM, Daugas E. Basophil involvement in lupus nephritis: a basis for innovation in daily care. *Nephrol Dial Transpl*. 2019;34(5):750–6.
56. Voulgarelis M, Kokori SI, Ioannidis JP, Tzioufas AG, Kyriaki D, Moutsopoulos HM. Anaemia in systemic lupus erythematosus: aetiological profile and the role of erythropoietin. *Ann Rheum Dis*. 2000;59(3):217–22.
57. Velo-Garcia A, Castro SG, Isenberg DA. The diagnosis and management of the haematologic manifestations of lupus. *J Autoimmun*. 2016;74:139–60.
58. Palm F, Nordquist L. Renal tubulointerstitial hypoxia: cause and consequence of kidney dysfunction. *Clin Exp Pharmacol Physiol*. 2011;38(7):474–80.
59. Ikeuchi H, Sugiyama H, Sato H, Yokoyama H, Maruyama S, Mukoyama M, et al. A nationwide analysis of renal and patient outcomes for adults with lupus nephritis in Japan. *Clin Exp Nephrol*. 2022;26(9):898–908.
60. Yang J, Liang D, Zhang H, Liu Z, Le W, Zhou M, et al. Long-term renal outcomes in a cohort of 1814 Chinese patients with biopsy-proven lupus nephritis. *Lupus*. 2015;24(14):1468–78.
61. Shayakul C, Ong-aj-yooth L, Chirawong P, Nimmannit S, Parichatikanond P, Laohapand T, et al. Lupus nephritis in Thailand: clinicopathologic findings and outcome in 569 patients. *Am J Kidney Dis*. 1995;26(2):300–7.

Figures

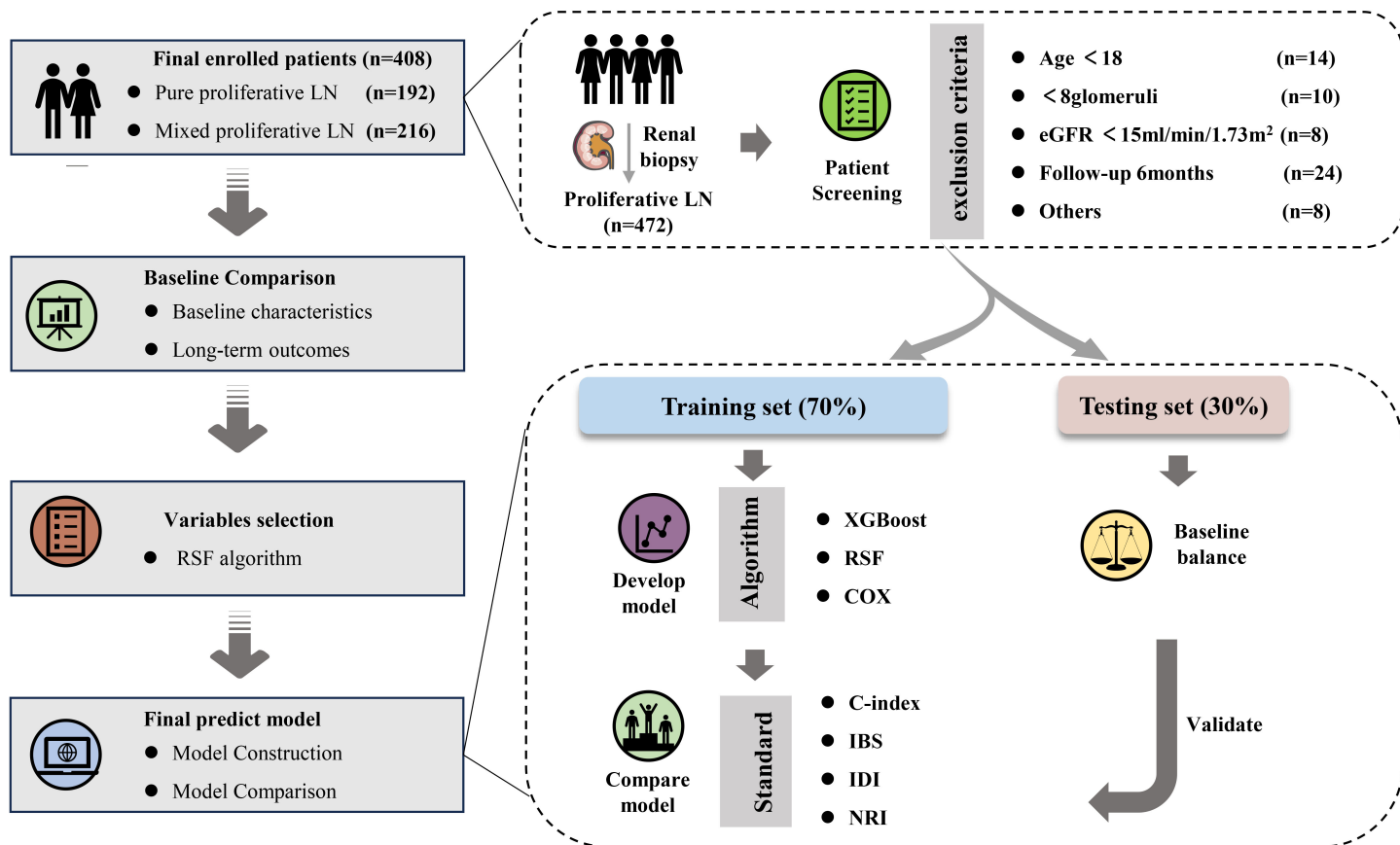


Figure 1

Study flowchart. LN, lupus nephritis. eGFR, estimated glomerular filtration rate. RSF, random survival forests. XGboost, eXtreme Gradient Boosting. C-index, concordance index. IBS, integrated brier score. NRI, net reclassification index. IBS, integrated discrimination improvement.

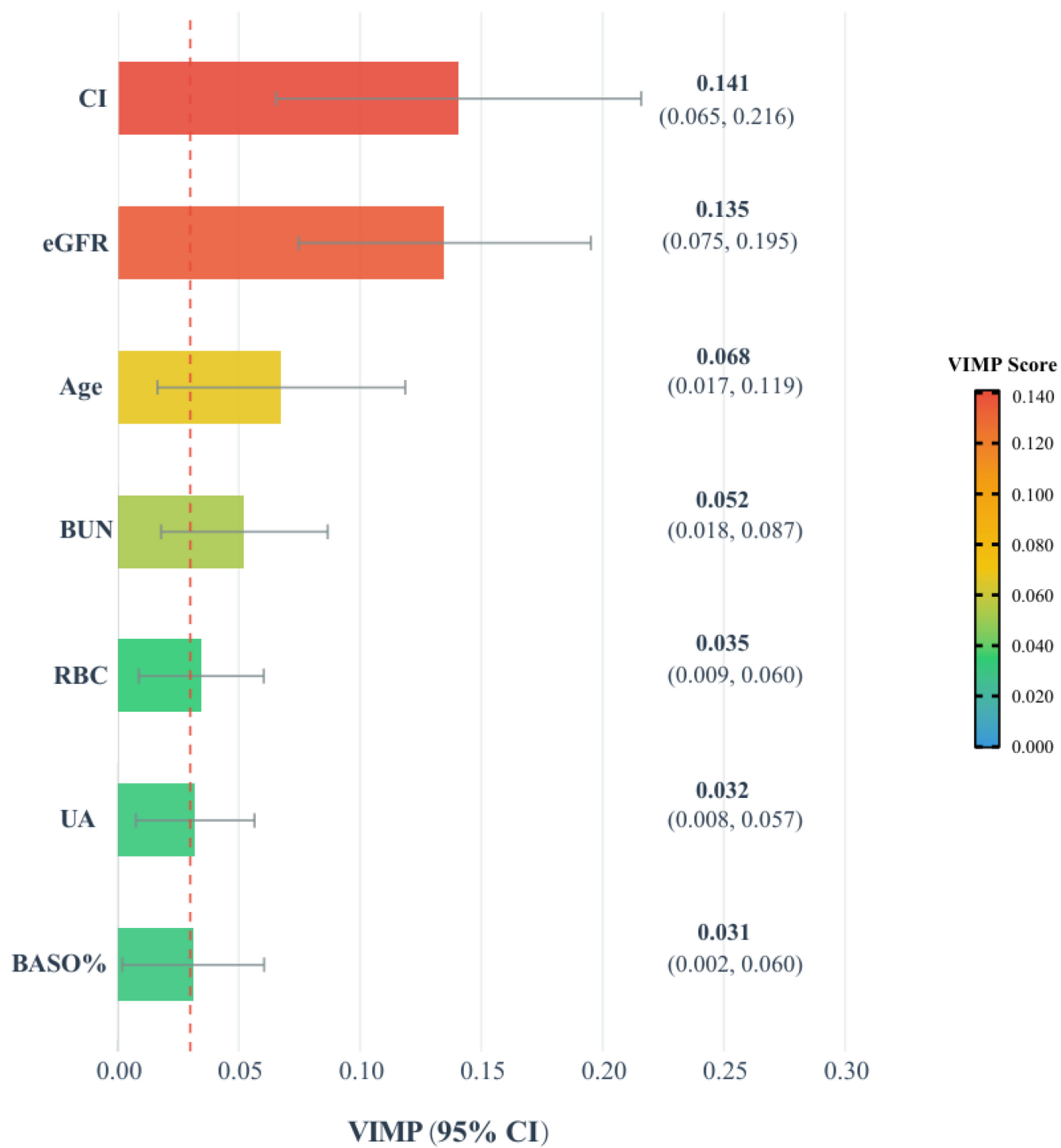


Figure 2

Survival curve of pure and mixed proliferative LN patients. LN, lupus nephritis

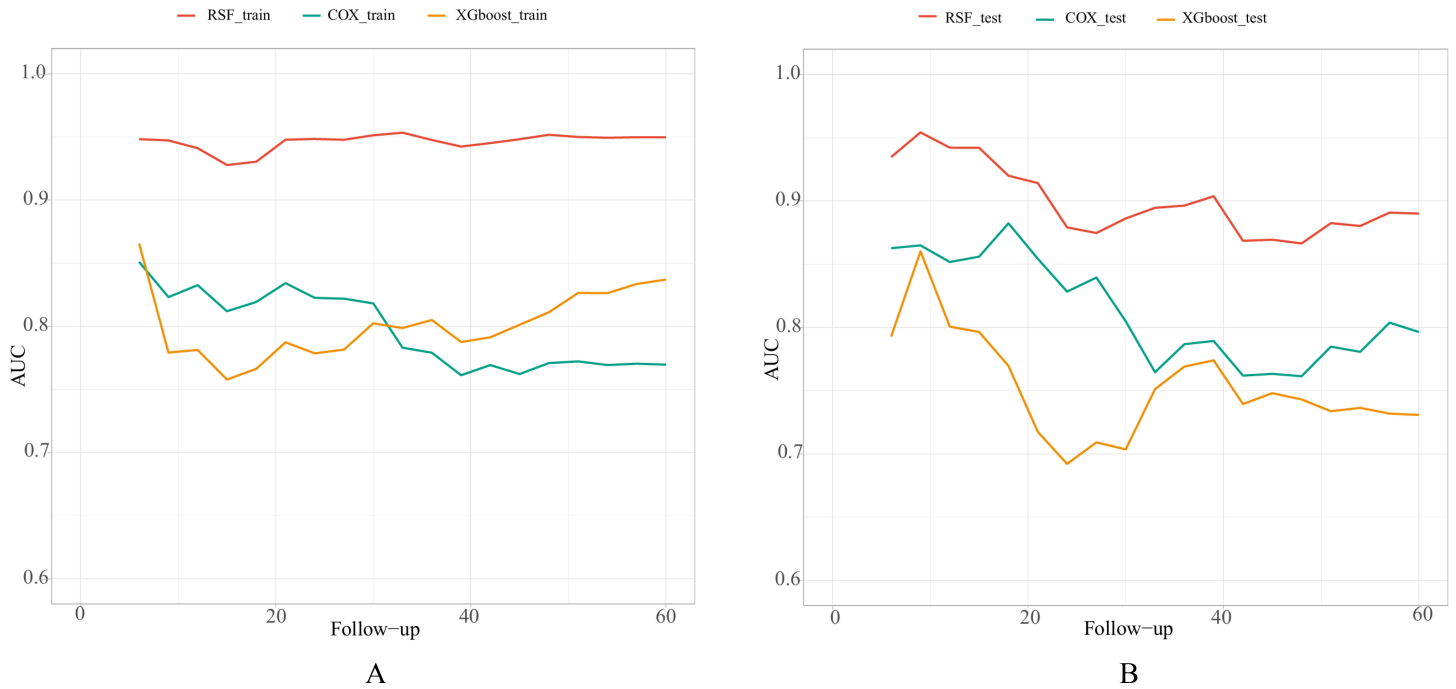


Figure 3

VIMP of indicators for developing model. CI, chronic index. eGFR, estimated glomerular filtration rate. BUN, blood urea nitrogen. RBC, red blood cell. UA, uric acid. BASO%, basophil percentage. VIMP, variable importance. Red dotted line, VIMP equal to 0.03.

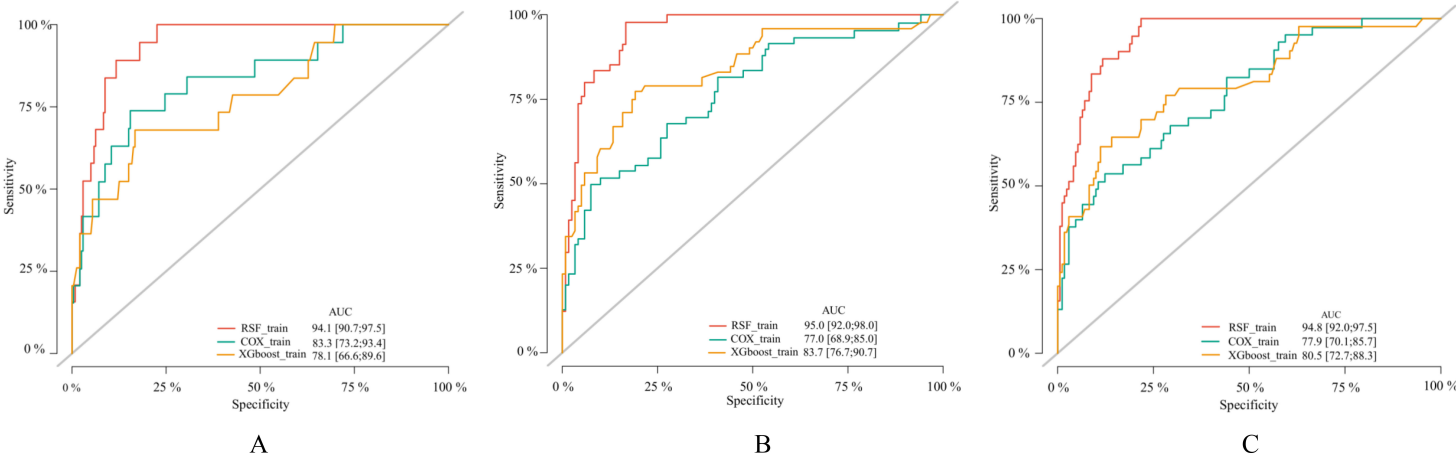


Figure 4

Comparison of tAUC among the RSF, XGboost, and Cox model in the training and testing set. (A) comparison of tAUC among the RSF, XGboost, and Cox model in the training set. (B) comparison of tAUC among the RSF, XGboost, and Cox model in the testing set. RSF, random survival forests. XGboost, eXtreme Gradient Boosting. Cox, cox regression model. tAUC, time-independent area under curve.

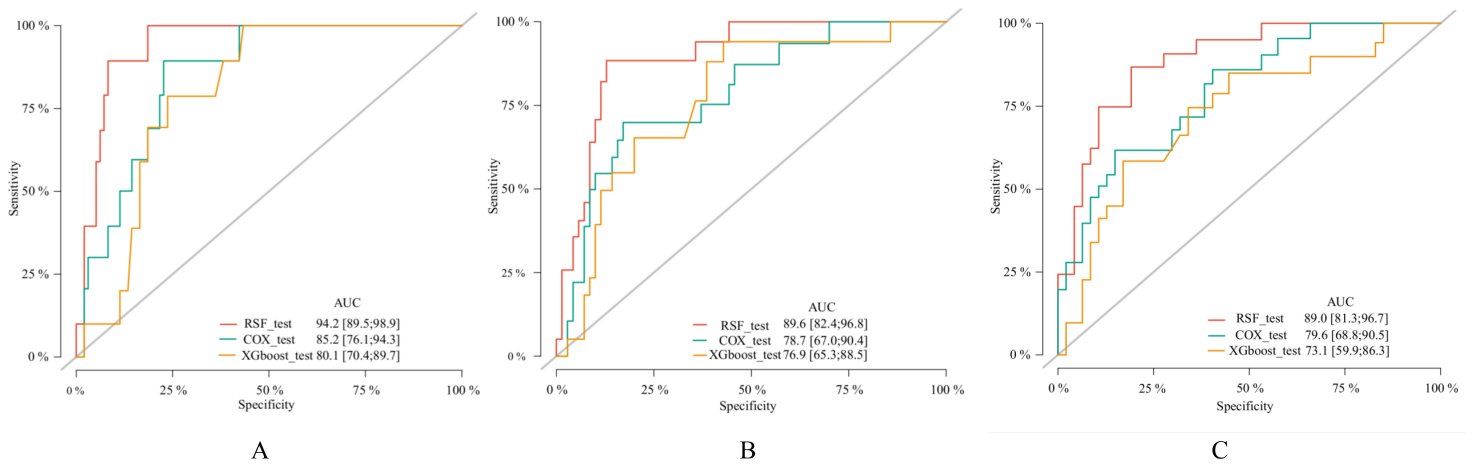


Figure 5

Comparison of AUC among the RSF, XGboost, and Cox model in the training and testing cohort. (A) AUC at 12 months. **(B)** AUC at 36 months. **(C)** AUC at 60 months. RSF, random survival forests. XGboost, eXtreme Gradient Boosting. Cox, cox regression model. AUC, area under curve.

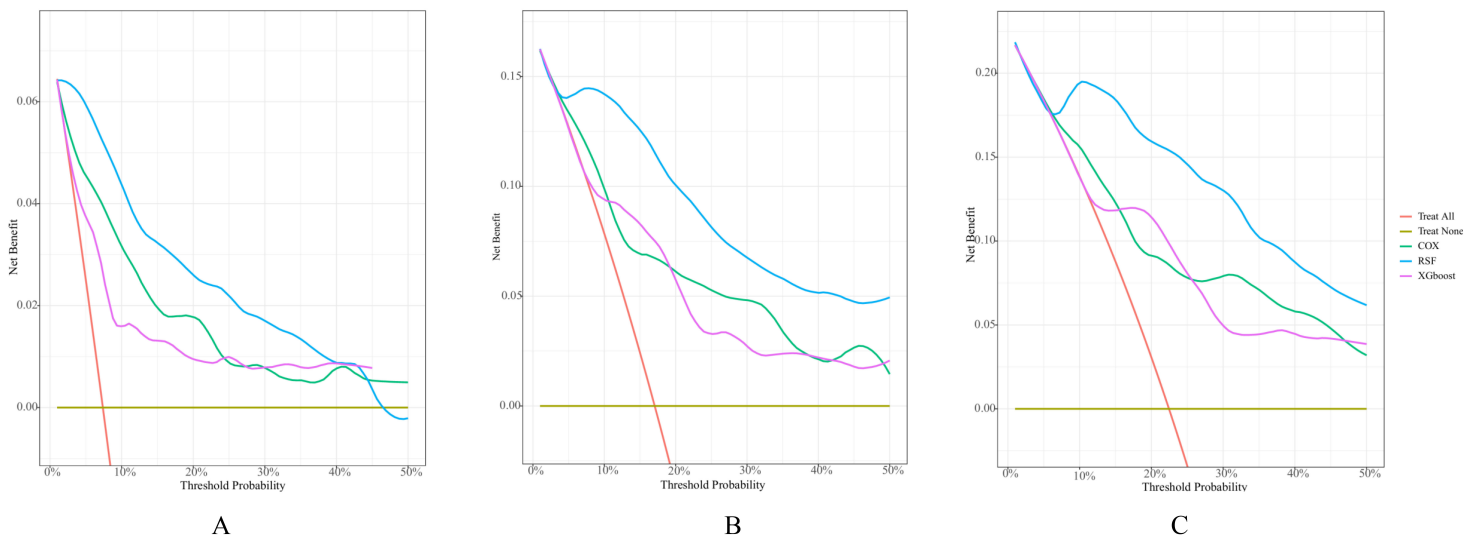


Figure 6

Comparison of DCA plots among the RSF, XGboost, and Cox model. (A) DCA plots at 12 months. **(B)** DCA plots at 36 months. **(C)** DCA plots at 60 months. DCA, decision curve analysis. The X-axis is the threshold probability for clinical intervention (range: 0-50%). Y-axis is the net benefit. The red line represents "treat all" that means the assumption that all patients receive intervention. The yellow line represents "treat none" representing the assumption that no patients receive intervention. The green line represents the RSF model. The blue line represents the COX model. The purple line represents the XGBoost model. RSF, random survival forests. XGboost, eXtreme Gradient Boosting. Cox, cox regression model.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementarymaterials.docx](#)