Machine Learning for Prediction and Risk Stratification of Lupus Nephritis Renal Flare

Yinghua Chen^a Siwan Huang^b Tiange Chen^b Dandan Liang^a Jing Yang^a
Caihong Zeng^a Xiang Li^b Guotong Xie^{b, c, d} ZhiHong Liu^a

^aNational Clinical Research Centre of Kidney Diseases, Jinling Hospital, Nanjing University School of Medicine, Nanjing, China; ^bPing An Healthcare Technology, Beijing, China; ^cPing An Health Cloud Co. Limited, Beijing, China; ^dPing An International Smart City Technology Co, Beijing, China

Keywords

Lupus nephritis · Renal flare · Machine learning · Prediction · Risk factors

Abstract

Background: Renal flare of lupus nephritis (LN) is strongly associated with poor kidney outcomes, and predicting renal flare and stratifying its risk are important for clinical decision-making and individualized management to reduce LN flare. **Methods:** We randomly divided 1,694 patients with biopsy-proven LN, who had achieved remission after treatment, into a derivation cohort ($n = 1,186$) and an internal validation cohort ($n = 508$), at a ratio of 7:3. The risk of renal flare 5 years after remission was predicted using an eXtreme Gradient Boosting (XGBoost) method model, developed from 59 variables, including demographic, clinical, immunological, pathological, and therapeutic characteristics. A simplified risk score prediction model (SRSPM) was developed from important variables selected by XGBoost model using stepwise Cox regression for practical convenience. **Results:** The 5-year relapse rates were 39.5% and 38.2% in the derivation and internal validation cohorts, respectively. Both the XGBoost model and the SRSPM had good predictive performance,

with a C-index of 0.819 (95% confidence interval [CI]: 0.774–0.857) and 0.746 (95% CI: 0.697–0.795), respectively, in the validation cohort. The SRSPM comprised 6 variables, including partial remission and endocapillary hypercellularity at baseline, age, serum Alb, anti-dsDNA, and serum complement C3 at the point of remission. Using Kaplan-Meier analysis, the SRSPM identified significant risk stratification for renal flares ($p < 0.001$). **Conclusions:** Renal flare of LN can be readily predicted using the XGBoost model and the SRSPM, and the SRSPM can also stratify flare risk. Both models are useful for clinical decision-making and individualized management in LN.

© 2021 S. Karger AG, Basel

Introduction

Lupus nephritis (LN) causes significant morbidity and mortality in patients with systemic lupus erythematosus (SLE) that progress to end-stage kidney disease (ESKD) in 10–20% of patients over 5–10 years [1, 2]. The high

Yinghua Chen, Siwan Huang and Tiange Chen contributed equally to this work.

frequency of renal flares is a crucial contributing factor to poor kidney outcomes in patients with LN [2–8]. Patients suffer from irreversible aggravation of kidney damage after each relapse and even progress to ESKD. Lupus flares often require intensive immunosuppressive therapy, leading to an increase in treatment-related comorbidities [9]. Therefore, it is necessary to develop risk prediction models for renal flares.

Previous studies are limited to evaluate the impact of risk factors on LN flare [4, 6, 8, 10–13]. They are constrained by small derivation of sample sizes, varying populations, and the inclusion of relatively few variables. Patients with LN have high heterogeneity in terms of clinical presentation, histologic lesions, response to treatment, and disease outcome; therefore, it is difficult for conventional statistical methods to cover all of these features. Machine learning involves a range of statistical algorithms used to analyze multidimensional heterogeneous data and discover key features or patterns of relationships among multiple risk factors, which can make accurate data-driven predictions. Prediction models based on machine learning methods have shown good predictive power for disease relapse, such as cancer [14–16]. However, models for prediction and risk stratification of renal flare in kidney disease have yet to be developed, especially concerning LN where relapse is prone to occur. This study aimed to use machine learning methods to analyze data from 1,694 patients with LN who had been followed up long-term, to develop models for prediction and risk stratification of renal flare, and to provide favorable tools for clinical decision-making and individualized management.

Materials and Methods

Study Cohort

A total of 1,694 patients with LN who had been registered in the Nanjing Glomerulonephritis Registry at the National Clinical Research Center of Kidney Diseases, Jinling Hospital, from January 1985 to December 2010 were included in this study. The inclusion criteria comprised all patients who: (a) met the 1997 American College of Rheumatology criteria for SLE, (b) had renal involvement and biopsy-proven LN, and (c) had achieved remission after treatment. All eligible patients were followed up from the time of renal biopsy until LN relapse or any censoring event (loss to follow-up, death, ESKD, or study end date).

Definition of Variables

Baseline was defined as the time of renal biopsy. The relapse time was defined as the time between the remission and the relapse. The renal biopsy specimens were reviewed by the same pathologist, in accordance with the 2018 Revision of the International

Society of Nephrology/Renal Pathology Society classification [17]. Categorical variables involving too many levels were recategorized into binary components using the survival tree method of exploiting an equivalence between a proportional hazard full likelihood model and a Poisson likelihood model. For endocapillary hypercellularity, fibrinoid necrosis, neutrophil infiltration/karyorrhexis, subendothelial immune deposits/hyaline deposits, subepithelial immune deposits, acute tubulointerstitial injury, interstitial inflammation, tubular atrophy, and interstitial fibrosis, their absence was denoted as 0 and presence as 1. The intensity of IgG, IgA, IgM, C3, and C1q staining was described as negative, 1+, 2+, and 3+. There were few patients with an intensity of 3+; therefore, the intensity of 3+ was merged with 2+ for analysis.

All patients were treated with the standard induction therapy for LN at our center [2], including glucocorticoids alone, glucocorticoids plus intravenous cyclophosphamide pulse, mycophenolate mofetil, calcineurin inhibitor, or combination therapy (combining corticosteroids, mycophenolate mofetil, and tacrolimus). Complete remission was defined as urinary protein (UPro) ≤ 0.4 g/24 h, the absence of active urine sediments, serum Alb ≥ 35 g/L, and normal SCr. Partial remission was defined as a reduction of $\geq 50\%$ in proteinuria and UPro < 3.5 g/24 h, serum Alb ≥ 30 g/L, and normal or $\leq 25\%$ increase in SCr from baseline [2].

Clinical Outcomes

We modified the 2012 KDIGO Clinical Practice Guideline of LN [18] and defined renal relapse as the presence of any one of the following: (a) an increase in glomerular hematuria to $> 10 \times 10^4$ /mL and/or recurrence of urine sediment RBC cast, WBC cast (no infection), or both; (b) if the baseline UPro was < 0.5 g/24 h, an increase to ≥ 1 g/24 h; 0.5 – 1.0 g/24 h, an increase to ≥ 2 g/24 h; and (c) if baseline SCr was < 2.0 mg/dL, an increase to ≥ 0.2 mg/dL; ≥ 2.0 mg/dL, an increase to ≥ 0.4 mg/dL.

Statistical Analyses

The entire cohort was randomly divided into derivation ($n = 1,186$) and internal validation ($n = 508$) cohorts at a 7:3 ratio using the R package caret, version 6.0–84 [19].

Descriptive Statistics

Fifty-nine variables with missing rates $< 5\%$ were considered for analysis (details in online suppl. Table 1–3; for all online suppl. material, see www.karger.com/doi/10.1159/000513566). Candidate predictors included demographic characteristics, pathological data at baseline, clinical and immunological data at the point of remission, and treatment information. Continuous variables were presented as the mean $\pm$ standard deviation or median (interquartile range [IQR]) and compared using Mann-Whitney U tests. Categorical variables were presented as frequency (percentages) and compared using χ^2 tests.

Predictive Modeling

Two risk models were developed to predict renal flare in patients with LN at the point of remission. The method flowchart was illustrated in Figure 1. The prediction window was 5 years after remission. We first developed a machine learning-based risk prediction model applying the eXtreme Gradient Boosting (XGBoost) algorithm. A simplified risk score prediction model (SRSPM) with a selected number of variables was then derived from a stepwise Cox regression model for practical convenience [20].

Fig. 1. Flowchart illustrating model derivation and validation.

Prediction Model Using the EXtreme Gradient Boosting Method

We first applied the XGBoost method using the gradient boosting decision tree algorithm to construct a risk prediction model. The model generates the importance score of each feature automatically by calculating the average improvement for each feature of the model after it was introduced to a branch [21–24]. A higher importance score indicated that the variable had a higher predictive value for the model. We trained the model with all candidate predictors in the derivation cohort and reached our final model through a hyper-parameter selection process using the k -fold cross-validation ($k = 5$) (online suppl. Table 4). We started with default parameters, tuned the maximum depth of a tree and minimum sum of the instance weight needed in a child, and tested various numbers of other parameters to avoid overfitting on the derivation cohort. We then validated the final model in the internal validation cohort. To facilitate the use of the XGBoost prediction model, a web-based risk calculator was developed.

We performed the Shapley Additive exPlanation (SHAP) method on the final model to better interpret the nonlinear relationship between variables and outcomes [25]. SHAP computes all possible combinations with and without certain features and returns its average contribution. A higher SHAP absolute value means that the contribution of a specific feature to the risk prediction is larger.

SRSPM-Derived from a Stepwise Cox Regression Model

We constructed a stepwise Cox regression model with all the important variables selected using the XGBoost method (Table 1). A Cox regression model cannot manage missing data automatically; therefore, we generated multiple imputations for incomplete multivariate data using the R package mice, version 3.6.0 [26]. After a stepwise Cox regression model was developed, we categorized continuous variables from the model according to their cutoff values derived from the survival tree method. Risk score was assigned to each variable from the coefficients of the Cox regression model. The score of each variable was added together as a relapse risk score. The Kaplan-Meier method was used to estimate the relapse-free renal survival rate. The comparison of different risk groups was evaluated using a log-rank test.

Model Validation

The C-index (area under receiving operating characteristic curve) was used to measure the models' predictive performance in both the derivation and internal validation cohorts. Calibration of the SRSPM was assessed using a quintile plot of observed versus expected risk and the Hosmer-Lemeshow test [27]. A p value >0.05 confirmed a good calibration by Hosmer-Lemeshow test.

All statistical analyses were performed on Python 3.7.3 and R 3.6.1 (packages: mice, survival, survivalROC, boot, rpart, XGBoost, and resourceSelection). A two-tailed p value <0.05 was considered statistically significant.

Table 1. Variables selected using XGBoost model and their importance score

Feature	Importance score
Partial remission	0.140
Age at the point of remission (years)	0.086
Serum C3 at the point of remission (g/L)	0.079
WBC count at the point of remission (/μL)	0.070
Serum UA at the point of remission (μmol/mL)	0.067
Hb at the point of remission (g/L)	0.060
Duration of LN (months)	0.049
Serum Alb at the point of remission (g/L)	0.048
SCr at the point of remission (mg/dL)	0.045
UPro at the point of remission (g/24 h)	0.045
Serum C4 at the point of remission (g/L)	0.043
Duration of SLE (months)	0.041
Proportion of cellular crescents (%)	0.032
Serum globulin at the point of remission (g/L)	0.030
Time to remission (months)	0.025
Activity index	0.025
Proportion of global sclerosis (%)	0.025
Anti-dsDNA at the point of remission	0.021
Chronicity index	0.011
URBC count ($\times 10^4$ /mL)	0.009
Mesangial deposition of C3	0.008
Proportion of segmental sclerosis (%)	0.007
Endocapillary hypercellularity	0.006
Neutrophils/karyorrhexis	0.006
Treatment with MMF (%)	0.004
Mesangial deposition of C1q	0.004
Intensity of IgG staining	0.004
Tubular basement membrane deposition of C3	0.003
Interstitial fibrosis	0.003126
Intensity of C1q staining	0.002824

The importance score is the relative number of times a variable is used to distribute the data across all trees. Alb, albumin; Hb, hemoglobin; LN, lupus nephritis; MMF, mycophenolate mofetil; SCr, serum creatinine; SLE, systemic lupus erythematosus; UA, uric acid; UPro, urinary protein; URBC, urine sediment RBC; XGBoost, eXtreme Gradient Boosting.

Results

Study Cohort

Of 1,694 patients, 1,305 (77.0%) achieved complete remission, and 1,482 (87.5%) were female. The average age at the point of remission was 32.4 ± 9.9 years. The median follow-up time was 4.1 (IQR, 1.7–6.7 years), the median relapse time was 2.0 (IQR, 1.0–3.8) years, and the mean follow-up frequency was 6.4 ± 4.6 visits per year.

In the derivation cohort, 915 (77.2%) achieved complete remission and 390 (76.8%) in the validation cohort.

During the 5-year follow-up after remission, renal relapse occurred in 663 patients. The 5-year relapse rate was 39.5% in the derivation cohort and 38.2% in the validation cohort. Online suppl. Tables 1–3 summarize the comparison of 16 variables regarding demographic, clinical characteristics, and laboratory data at the point of remission, of 37 variables including the duration of disease and pathologic data at baseline, and of 6 variables on induction treatment between relapse and nonrelapse patients.

Performance and Interpretation of the EXtreme Gradient Boosting Model

The XGBoost model had a good predictive performance, with a C-index of 0.822 (95% confidence interval [CI]: 0.798–0.855) in the derivation cohort and 0.819 (95% CI: 0.774–0.857) in the validation cohort. The 30 important features derived from the final XGBoost model are presented in Table 1, ranging from high to low in order of importance. A web-based calculator was developed from the XGBoost model, allowing clinicians and patients to easily obtain the relapse risk (<http://101.89.95.81:8250>).

The SHAP method was applied to visualize how these features contributed to the risk of LN relapse. The SHAP plots of 3 continuous variables included in the SRSPM are shown in online suppl. Figures 1–3. Patients with aged <21.5 years at the point of remission had the highest risk and that the risk of relapse decreased as patients aged. For patients aged >28 years at the point of remission, the SHAP value was below 0, indicating that the nonrelapse probability was higher than the relapse probability (online suppl. Fig. 1). Patients with lower serum Alb and serum complement C3 levels at the point of remission had a higher risk of LN relapse (online suppl. Fig. 2, 3).

Construction and Evaluation of the Simplified Risk Score Prediction Model

The XGBoost model showed good predictive performance; however, it had several limitations, such as challenges in interpretability and difficulty in calculating due to model complexity. Based on the SHAP method, the relationship between variables and outcomes could be explained; however, developing a ready-to-hand simplified model for clinical practice is necessary.

The SRSPM was constructed from 30 important variables selected by the XGBoost method using stepwise Cox regression model. Six variables were derived comprising partial remission and endocapillary hypercellularity at baseline, age, serum Alb, anti-dsDNA, and serum complement C3 at the point of remission (online suppl. Table

Fig. 2. Kaplan-Meier curves of the scoring model in the derivation (**a**) and validation cohorts. Comparisons of different risk scores were evaluated using the log-rank test (**b**).

Table 2. Derived score of the simplified Cox risk score prediction model

Feature	Beta	HR (95% CI)	<i>p</i> value	Points
Partial remission	1.302	3.676 (2.971–4.549)	<0.001	2
Age at the point of remission (<28 years)	0.578	1.782 (1.483–2.141)	<0.001	1
Serum Alb at the point of remission (<35 g/L)	0.386	1.471 (1.108–1.953)	0.01	1
Anti-dsDNA positive at the point of remission	0.219	1.245 (0.999–1.551)	0.05	1
Serum C3 at the point of remission (<0.58 g/L)	0.490	1.633 (1.287–2.072)	<0.001	1
Endocapillary hypercellularity	0.300	1.350 (1.081–1.684)	0.01	1

Alb, albumin; CI, confidence interval; HR, hazard ratio.

5). The cutoff values of age, serum Alb, and complement C3 at the point of remission were derived from the survival tree method. Risk points were derived from the coefficient of the variables in the Cox regression model (Table 2). Partial remission was scored as 2 points; endocapillary hypercellularity at baseline and age <28 years, serum Alb <35 g/L, positive anti-dsDNA, and serum complement C3 <0.58 g/L at the point of remission were each scored as 1 point. A risk score was obtained by adding all the points of the 6 variables together. Due to only 25 pa-

tients with a risk score of >5, patients with a risk score of ≥5 were combined into one group. The Kaplan-Meier survival curves showed significant differences in the risk of renal flare over 5 years among patients with different risk scores ($p < 0.001$) (Fig. 2). For patients with risk scores of 0, 1, 2, 3, 4, and ≥5, the risk of relapse in the 5 years after remission was 17.8%, 18.6%, 38.6%, 53.7%, 74.1%, and 84.1% in the derivation cohort, respectively, and 20.0%, 28.3%, 38.9%, 53.8%, 82.1%, and 91.4% in the validation cohort, respectively.

Fig. 3. Calibration plot of scoring model in derivation (a) and validation cohorts (b). Patients were divided into quintiles based on their predicted risk. The black bar represents the observed proportion of events, and the grey bar represents the predicted risk. The Hosmer-Lemeshow test indicated no evidence of poor fit in the derivation cohort ($p = 0.35$) and the validation cohort ($p = 0.62$).

The C-index of the SRSPM was 0.747 (95% CI: 0.717–0.772) in the derivation cohort and 0.746 (95% CI: 0.697–0.795) in the validation cohort. The Hosmer-Lemeshow calibration test of the scoring model indicated good fit in the derivation cohort ($p = 0.35$) and validation cohort ($p = 0.62$) (Fig. 3).

Discussion/Conclusion

Renal flares are strongly associated with the prognosis of LN, with a relapse rate of 30–60% [2–6, 8]. They reflect a new immunological and inflammatory attack on the kidney that may lead to the development of further active renal lesions. The recurrence of these lesions exposes patients to aggravation of glomerular sclerosis and interstitial fibrosis with consequent progression to ESKD. LN flares are independently associated with an increased risk of renal function deterioration [4–8]. Therefore, prevention of renal flares is a major challenge in the treatment of LN. In addition, LN flares require intensive immunosuppressive therapy, which makes treatment more difficult and may lead to prolonged treatment duration. All these factors increase treatment-related comorbidities and impair patient survival and quality of life [9]. There-

fore, development of a risk prediction model for renal flares is urgently needed, and preventive measures based on risk assessment of LN flares should be taken.

Our study is the first to apply machine learning methods to develop models for prediction and risk stratification of renal flares in LN, including an XGBoost prediction model and an SRSPM, both of which showed promising performances. Our study also comprised the largest sample size of related studies to date, involving 1,694 patients with biopsy-proven LN and 4.8 (IQR, 1.7–6.7) years of follow-up. Based on comprehensive and long-term follow-up data, machine learning methods were used to process the complex data [21, 23, 24]. The XGBoost model was constructed from 59 variables, combining demographic, clinical, immunological, pathological, and therapeutic characteristics. To maximize the value of variables to the models and ensure best predictive performance, the multi-categorical pathological variables were recategorized using the survival tree method, which overcame the limitation involved with pathological classification and pathological semiquantitative scores included in previous studies. All variables were measured routinely in clinical settings and were collected not only at baseline but also at the point of remission. The point of remission is an opportune time to evaluate the risk of flare and determine

subsequent treatment. To further enhance clinical practicability and operability, the SRSPM was developed comprising 6 variables, namely, partial remission, and endocapillary hypercellularity at baseline, age, serum Alb, anti-dsDNA, and C3 at the point of remission. An online calculator allows for simple implementation of the models.

Partial remission, found to be the most important variable in the XGBoost model and the variable with the highest score in the SRSPM, showed the highest value in renal flare prediction. Patients with partial remission had proteinuria, hematuria, or renal damage, indicating that the disease was active, possibly leading to a higher risk of disease relapse. Previous studies have also shown that partial remission is a risk factor for LN flare [3, 4, 8, 28, 29]. The induction phase has generally been considered to comprise a short term of 6 months for LN treatment in guidelines and clinical trials. Divisions between the induction and maintenance phases of treatment are often blurred. In this study, partial remission had the greatest impact on LN flare. Our findings indicate that induction therapy should strive to achieve complete remission in patients with LN and that the duration should be prolonged in patients with partial remission, rather than simply limited to 6 months.

This study showed that the lower the serum Alb level at the point of remission, the higher the risk of relapse. Several factors may affect serum Alb levels. The predominant factors are albuminuria and systemic inflammation associated with SLE. All of these factors may contribute to the disease activity of LN. A previous study has shown that serum Alb is inversely associated with disease activity and the AI of renal histological lesions in patients with SLE [30, 31]. This explains why patients with low serum Alb levels at the point of remission had a higher risk of LN flare.

This study showed that patients with endocapillary proliferation were also at a higher risk of renal flare. Endocapillary proliferation was an indicator of renal disease activity. It was found to be common in active LN and was the predominant residual active inflammation observed on repeat biopsies after the completion of induction treatment [32]. Residual histologic activity and more specifically the extent of endocapillary proliferation were a risk factor for LN flare [32, 33]; therefore, patients with LN with endocapillary proliferation were more likely to relapse. However, the definition of current complete remission was relay mainly on clinical criteria and did not include a kidney histology component despite a growing evidence suggesting significant discordance between

clinical and histological evidence of ongoing activity. Nearly one-third of patients with complete clinical remission had active inflammatory lesions on repeat biopsy [34]. Achieving a complete renal histological remission may be reasonable to minimize the risk of LN flare.

Positive anti-dsDNA and low complement C3 levels at the point of remission were found to involve a high risk of renal flare in our study. Autoantibodies and complement are crucial to the pathogenesis of LN. It has been reported that an increase in anti-dsDNA or a decrease in complement can predict subsequent flares [6, 35, 36]. Prophylactic treatment in patients with a rise in anti-dsDNA and a fall in serum complement reduced subsequent disease flares [37]. Patients have positive anti-dsDNA and low C3 at the point of remission, suggesting that autoantibodies and complement continue to play a role in kidney damage, which may induce disease relapse at any time. The risk prediction model developed in this study once again confirmed that immunological markers were important for predicting LN flares.

In this study, younger patients with LN had a higher risk of relapse. The incidence of LN is the highest during reproductive years, when women are most hormonally active, in contrast to postmenopause when the incidence of LN is lower and the female-to-male ratio is considerably less marked [38]. A strong correlation between hormone levels and LN has been supported by epidemiologic data. Previous studies have shown a high risk of relapse with younger age [6, 8, 10–12], which is consistent with our study findings.

The SRSPM integrated 6 variables, all of which had a rational link to disease activity, as mentioned above. Therefore, the model could comprehensively predict and stratify the risk of renal flare in LN. The risk of renal flare was significantly different among patients with different risk scores, and the 5-year renal relapse rate significantly increased with an increase in the risk scores. However, even for patients with risk scores of 0, the risk of relapse was up to 20%, which indicated that maintenance therapy was necessary for all LN patients to prevent relapse. Relapse was a characteristic of LN and was influenced by many factors. For this group of patients, XGBoost model could be used to further well predict the risk of relapse.

This study had several limitations. First, this cohort study did not involve a prospective therapeutic trial. It included patients with LN who had been treated in our department for the past 25 years. We also did not consider information concerning maintenance treatment regimens for analysis. Second, the prediction model was developed based on data obtained from a Chinese popu-

lation. Whether it is applicable to other ethnicities and regions remains to be verified because of the ethnic differences in the incidence, response to treatment, and prognosis of LN [39].

This study developed 2 prediction models, including an XGBoost model and an SRSPM, both of which showed good performance in predicting renal flare in LN, and the SRSPM could also stratify the risk of flare. Web-based calculators were easily implemented for clinical practicality and operability. These models provide useful tools for clinicians and patients to predict and stratify the risk of renal flare, as well as for individualized management of LN. Considering some of the inherent limitations, such as patients' heterogeneity, variations in induction regimen, and evolving management over time, the clinical application, accuracy, and advantage of machine learning and prediction models remain to be further validated.

Acknowledgements

We would like to thank Kang Li, Ying Jin, Duqun Chen, and Hui Du for their assistance in collecting data of the patients.

References

- 1 Tektonidou MG, Dasgupta A, Ward MM. Risk of end-stage renal disease in patients with lupus nephritis, 1971–2015: a systematic review and bayesian meta-analysis. *Arthritis Rheumatol*. 2016 Jun;68(6):1432–41.
- 2 Shao SJ, Hou JH, Xie GT, Sun W, Liang DD, Zeng CH, et al. Improvement of outcomes in patients with lupus nephritis: management evolution in Chinese patients from 1994 to 2010. *J Rheumatol*. 2019 Aug;46(8):912–9.
- 3 Illei GG, Takada K, Parkin D, Austin HA, Crane M, Yarboro CH, et al. Renal flares are common in patients with severe proliferative lupus nephritis treated with pulse immunosuppressive therapy: long-term followup of a cohort of 145 patients participating in randomized controlled studies. *Arthritis Rheum*. 2002 Apr;46(4):995–1002.
- 4 Mok CC, Ying KY, Tang S, Leung CY, Lee KW, Ng WL, et al. Predictors and outcome of renal flares after successful cyclophosphamide treatment for diffuse proliferative lupus glomerulonephritis. *Arthritis Rheum*. 2004 Aug;50(8):2559–68.
- 5 Gibson KL, Gipson DS, Massengill SA, Doolley MA, Primack WA, Ferris MA, et al. Predictors of relapse and end stage kidney disease in proliferative lupus nephritis: focus on children, adolescents, and young adults. *Clin J Am Soc Nephrol*. 2009 Dec;4(12):1962–7.
- 6 Moon SJ, Park HS, Kwok SK, Ju J, Choi BS, Park KS, et al. Predictors of renal relapse in Korean patients with lupus nephritis who achieved remission six months following induction therapy. *Lupus*. 2013 Apr;22(5):527–37.
- 7 Parikh SV, Nagaraja HN, Hebert L, Rovin BH. Renal flare as a predictor of incident and progressive CKD in patients with lupus nephritis. *Clin J Am Soc Nephrol*. 2014 Feb;9(2):279–84.
- 8 Mejía-Vilet JM, Córdova-Sánchez BM, Arreola-Guerra JM, Morales-Buenrostro LE, Uribe-Uribe NO, Correa-Rotter R. Renal flare prediction and prognosis in lupus nephritis Hispanic patients. *Lupus*. 2016 Mar;25(3):315–24.
- 9 Parikh SV, Rovin BH. Current and emerging therapies for lupus nephritis. *J Am Soc Nephrol*. 2016 Oct;27(10):2929–39.
- 10 Ioannidis JP, Boki KA, Katsorida ME, Drosos AA, Skopouli FN, Boletis JN, et al. Remission, relapse, and re-remission of proliferative lupus nephritis treated with cyclophosphamide. *Kidney Int*. 2000 Jan;57(1):258–64.
- 11 Mosca M, Neri R, Giannesi S, Pasquariello A, Puccini R, Bencivelli W, et al. Therapy with pulse methylprednisolone and short course pulse cyclophosphamide for diffuse proliferative glomerulonephritis. *Lupus*. 2001;10(4):253–7.
- 12 Mosca M, Bencivelli W, Neri R, Pasquariello A, Batini V, Puccini R, et al. Renal flares in 91 SLE patients with diffuse proliferative glomerulonephritis. *Kidney Int*. 2002 Apr;61(4):1502–9.
- 13 Cortés-Hernández J, Ordi-Ros J, Labrador M, Segarra A, Tovar JL, Balada E, et al. Predictors of poor renal outcome in patients with lupus nephritis treated with combined pulses of cyclophosphamide and methylprednisolone. *Lupus*. 2003;12(4):287–96.
- 14 Wu X, Ye Y, Barcenas CH, Chow WH, Meng QH, Chavez-MacGregor M, et al. Personalized prognostic prediction models for breast cancer recurrence and survival incorporating multidimensional data. *J Natl Cancer Inst*. 2017 Jul 1;109(7):djw314.
- 15 Alabi RO, Elmusrati M, Sawazaki-Calone I, Kowalski LP, Haglund C, Coletta RD, et al. Machine learning application for prediction of locoregional recurrences in early oral tongue cancer: a web-based prognostic tool. *Virchows Arch*. 2019 Oct;475(4):489–97.
- 16 Sala Elarre P, Oyaga-Iriarte E, Yu KH, Baudin V, Arbea Moreno L, Carranza O, et al. Use of machine-learning algorithms in intensified preoperative therapy of pancreatic cancer to predict individual risk of relapse. *Cancers*. 2019 Apr 30;11(5):606.

Statement of Ethics

This retrospective study was in compliance with the Declaration of Helsinki and approved by the Ethics Committee of Jinling Hospital.

Conflict of Interest Statement

The authors have declared no conflicts of interest.

Funding Sources

This study was funded by Key Program of Social Development Project of Jiangsu Province (BE2016747) and Clinical Research Center Program of Jiangsu Province (YXZX2016003).

Author Contributions

Z.H.L., G.T.X., and C.H.Z. designed the study; Y.H.C., S.W.H., and T.G.C. carried out the study; Y.H.C., D.D.L., and J.Y. collected the data; S.W.H., T.G.C., and X.L. analyzed the data; S.W.H. and T.G.C. made the figures; Y.H.C., S.W.H., and T.G.C. drafted and revised the paper; all the authors approved the final version of the manuscript.

- 17 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 Apr; 93(4):789–96.
- 18 Kidney Disease: Improving Global Outcomes (KDIGO) Glomerulonephritis Work Group. KDIGO clinical practice guidelines for glomerulonephritis. Chapter 12: lupus nephritis. *Kidney Int Suppl.* 2012 Jun;2(2):221–32.
- 19 Kuhn M. Building predictive models in R using the caret package. *J Stat Soft.* 2008;28(5): 1–26.
- 20 Chen T, Guestrin C. XGBoost: a scalable tree boosting system: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York, NY ACM; 2016. p. 785–94.
- 21 Kilic A, Goyal A, Miller JK, Gjekmarkaj E, Tam WL, Gleason TG, et al. Predictive utility of a machine learning algorithm in estimating mortality risk in cardiac surgery. *Ann Thorac Surg.* 2020 Jun;109(6):1811–9.
- 22 Shung DL, Au B, Taylor RA, Tay JK, Laursen SB, Stanley AJ, et al. Validation of a machine learning model that outperforms clinical risk scoring systems for upper gastrointestinal bleeding. *Gastroenterology.* 2020 Jan;158(1): 160–7.
- 23 Chen T, Li X, Li Y, Xia E, Qin Y, Liang S, et al. Prediction and risk stratification of kidney outcomes in IgA nephropathy. *Am J Kidney Dis.* 2019 Sep;74(3):300–9.
- 24 Mortazavi BJ, Bucholz EM, Desai NR, Huang C, Curtis JP, Masoudi FA, et al. Comparison of machine learning methods with national cardiovascular data registry models for prediction of risk of bleeding after percutaneous coronary intervention. *JAMA Netw Open.* 2019 Jul 3;2(7):e196835.
- 25 Lundberg S, Lee SI. A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems;* 2017. p. 4765–74.
- 26 Hapfelmeier A, Hothorn T, Riediger C, et al. Mice: multivariate imputation by chained equations in R. *Int J Biostat.* 2014;45:1–67.
- 27 Lemeshow S, Hosmer DW Jr. A review of goodness of fit statistics for use in the development of logistic regression models. *Am J Epidemiol.* 1982 Jan;115(1):92–106.
- 28 So MW, Koo BS, Kim YG, Lee CK, Yoo B. Predictive value of remission status after 6 months induction therapy in patients with proliferative lupus nephritis: a retrospective analysis. *Clin Rheumatol.* 2011 Nov;30(11):1399–405.
- 29 Yap DYH, Tang C, Ma MKM, Mok MMY, Chan GCW, Kwan LPY, et al. Longterm data on disease flares in patients with proliferative lupus nephritis in recent years. *J Rheumatol.* 2017 Sep;44(9):1375–83.
- 30 Yip J, Aghdassi E, Su J, Lou W, Reich H, Bargman J, et al. Serum albumin as a marker for disease activity in patients with systemic lupus erythematosus. *J Rheumatol.* 2010 Aug 1; 37(8):1667–72.
- 31 Sui M, Jia X, Yu C, Guo X, Liu X, Ji Y, et al. Relationship between hypoalbuminemia, hyperlipidemia and renal severity in patients with lupus nephritis: a prospective study. *Cent Eur J Immunol.* 2014;39(2): 243–52.
- 32 De Rosa M, Azzato F, Toblli JE, De Rosa G, Fuentes F, Nagaraja HN, et al. A prospective observational cohort study highlights kidney biopsy findings of lupus nephritis patients in remission who flare following withdrawal of maintenance therapy. *Kidney Int.* 2018 Oct; 94(4):788–94.
- 33 Hill GS, Delahousse M, Nochy D, Thervet E, Vrtovsnik F, Rémy P, et al. Outcome of relapse in lupus nephritis: roles of reversal of renal fibrosis and response of inflammation to therapy. *Kidney Int.* 2002 Jun;61(6):2176–86.
- 34 Malvar A, Pirruccio P, Alberton V, Lococo B, Recalde C, Fazini B, et al. Histologic versus clinical remission in proliferative lupus nephritis. *Nephrol Dial Transplant.* 2017 Aug 1; 32(8):1338–44.
- 35 Ho A, Magder LS, Barr SG, Petri M. Decreases in anti-double-stranded DNA levels are associated with concurrent flares in patients with systemic lupus erythematosus. *Arthritis Rheum.* 2001 Oct;44(10):2342–9.
- 36 Tseng CE, Buyon JP, Kim M, Belmont HM, Mackay M, Diamond B, et al. The effect of moderate-dose corticosteroids in preventing severe flares in patients with serologically active, but clinically stable, systemic lupus erythematosus: findings of a prospective, randomized, double-blind, placebo-controlled trial. *Arthritis Rheum.* 2006 Nov;54(11):3623–32.
- 37 Yap DYH, Kwan LPY, Ma MKM, Mok MMY, Chan GCW, Chan TM. Preemptive immunosuppressive treatment for asymptomatic serological reactivation may reduce renal flares in patients with lupus nephritis: a cohort study. *Nephrol Dial Transplant.* 2019 Mar 1; 34(3):467–73.
- 38 Tsokos GC. Systemic lupus erythematosus. *N Engl J Med.* 2011 Dec 1;365(22):2110–21.
- 39 Hanly JG, O’Keefe AG, Su L, Urowitz MB, Romero-Diaz J, Gordon C, et al. The frequency and outcome of lupus nephritis: results from an international inception cohort study. *Rheumatology.* 2016 Feb;55(2):252–62.