Original Article

A Predictive Model for Estimation Risk of Proliferative Lupus Nephritis

Dong-Ni Chen[1,2], Li Fan[1,2], Yu-Xi Wu[1,2], Qian Zhou[3], Wei Chen[1,2], Xue-Qing Yu[1,2,4]

[1]Department of Nephrology, The First Affiliated Hospital, Sun Yat-sen University, Guangzhou, Guangdong 510080, China
[2]Key Laboratory of Nephrology, Ministry of Health and Guangdong Province, Guangzhou, Guangdong 510080, China
[3]Clinical Trials Unit, The First Affiliated Hospital, Sun Yat-sen University, Guangzhou, Guangdong 510080, China
[4]Guangdong Medical University, Zhanjiang, Guangdong 524023, China

Abstract

Background: Lupus nephritis (LN) is classified by renal biopsy into proliferative and nonproliferative forms, with distinct prognoses, but renal biopsy is not available for every LN patient. The present study aimed to establish an alternate tool by building a predictive model to evaluate the probability of proliferative LN.

Methods: In this retrospective cohort with biopsy-proven LN, 382 patients in development cohort, 193 in internal validation cohort, and 164 newly diagnosed patients in external validation cohort were selected. Logistic regression model was established, and the concordance statistics (C-statistics), Akaike information criterion (AIC), integrated discrimination improvement, Hosmer-Lemeshow test, and net reclassification improvement were calculated to evaluate the performance and validation of models.

Results: The prevalence of proliferative LN was 77.7% in the whole cohort. A model, including age, gender, systolic blood pressure, hemoglobin, proteinuria, hematuria, and serum C3, performed well on good-of-fit and discrimination in the development chohort to predict the risk of proliferative LN (291 for AIC and 0.84 for C-statistics). In the internal and external validation cohorts, this model showed good capability for discrimination and calibration (0.84 and 0.82 for C-statistics, and 0.99 and 0.75 for *P* values, respectively).

Conclusion: This study developed and validated a model including demographic and clinical indices to evaluate the probability of presenting proliferative LN to guide therapeutic decisions and outcomes.

Key words: Biopsy; Lupus Nephritis; Nomogram; Predictive Value of Tests; Risk Factors

Introduction

Systemic lupus erythematosus (SLE) is a chronic autoimmune disease that affects multiple organs and tissues, of which the development of kidney disease is the most important predictor of morbidity and mortality.[1-3] Lupus nephritis (LN) is often associated with a poor long-term prognosis; up to 70% of SLE patients are affected by LN and approximately 10–20% of which will progress to end-stage renal disease (ESRD) within 5 years after diagnosis.[4] LN can be pathologically classified into six classes by renal biopsy, of which Class III, Class IV, and mixed Class V with proliferative lesions, known as proliferative LN manifesting severe symptoms, require intensive therapy and have worse outcomes, compared with nonproliferative LN, including Class I, Class II, and purely Class V.[5] The Kidney Disease Improving Global Outcome suggested that SLE patients presenting with renal function or urine disorders should be subjected to renal biopsy, which is critical for the diagnosis and therapy in LN.[6] However, biopsy was not available for every LN patient because of inadequate medical resources, especially in remote regions, and contraindications.[7,8] Thus far, there has been a lack of alternate evaluation methods to discriminate proliferative LN from nonproliferative LN. Therefore, the aim of this study was to develop and validate a model, including demographic and clinical indices to evaluate the probability of presenting

Address for correspondence: Dr. Wei Chen, Department of Nephrology, The First Affiliated Hospital, Sun Yat-sen University, No. 58, Zhongshan Road II, Guangzhou, Guangdong 510080, China
E-Mail: vvchen66@qq.com

Access this article online

Quick Response Code:

Website: www.cmj.org

DOI: 10.4103/0366-6999.232809

Received: 28-02-2018 **Edited by:** Xin Chen
How to cite this article: 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 2018;131:1275-81.

proliferative LN, which might help the clinical practice and therapeutic decisions.

Methods

Ethical approval

The study was conducted in accordance with the *Declaration of Helsinki* and was approved by the Ethics Committee of the First Affiliated Hospital, Sun Yat-sen University (No. [2016] 215). Informed written consent was obtained from all participants before their enrollment in this study.

Patients

This study population was derived from a retrospectively observational cohort of LN (http://ln.medidata.cn) at the First Affiliated Hospital of Sun Yat-sen University. In this cohort, demographic, clinical, and pathological features, outcomes of ESRD, and all-cause mortality were collected from LN patients older than 14 years who met the 1982 American College of Rheumatology revised criteria for SLE and were confirmed by renal biopsy. Patients who had reached ESRD, which was defined as estimated glomerular filtration rate (eGFR) <15 $ml \cdot min^{-1} \cdot 1.73\ m^{-2}$, dialysis, or renal transplantation at the time of diagnosis; no biopsy information or number of glomeruli <10; drug-induced lupus-like syndrome; complicated with malignant tumor; and inadequate contact information, were excluded from this cohort. Because of extremely distinct outcomes and clinical characteristics, patients with Class VI LN were also excluded from this study. The 4-variable Modification of Diet in Renal Disease formula was used to estimate eGFR by four variables: serum creatinine, age, ethnicity, and gender. Participants were selected between January 1, 1996, and December 31, 2011, and randomly resampled into development and internal validation sets of 2/3 and 1/3, respectively. Newly diagnosed individuals from January 1, 2012, to October 1, 2015, were applied to external validation.

Clinical variables

This study initially considered 97 clinically relevant variables, including demographic information (age and gender), clinical features (blood pressure [BP], symptoms, and comorbidities), and laboratory tests (plasma, serum, and urine). All samples and information were collected before renal biopsy. Missing data for categorical variables were updated from original records as far as possible. Variables with more than 30% missing values were excluded from the analysis. Other data for continuous variables missing <30% were imputed using regression-based maximum-likelihood methods, followed by sensitivity analysis. Variables with skewed distribution were log transformed, and the others were evaluated as linear predictors.

Renal biopsy

All of our patients have been proven by renal biopsy. According to the International Society of Nephrology/Renal Pathology Society 2003 revised criteria,[5] proliferative LN was defined as focal or diffuse proliferative LN, including Class III, Class IV, and mixed Class V; in contrast, nonproliferative LN was defined as normal glomeruli, purely mesangial disease, or membranous glomerulonephritis, including Class I, Class II, and purely Class V. The dependent variable of interest is proliferative LN.

Statistical analysis

For model development, continuous variables were shown as mean ± standard deviation (SD) or median (Q1, Q3) according to the distribution, and categorical variables were shown as frequency. Univariate logistic regression was applied to each variable for the initial model selection with classification of proliferative LN as the dependent variable. Variables with $P < 0.10$ on univariate analysis or that were clinically relevant were included in the multivariable logistic regression model. We entered variables in the multivariable model using a combination strategy of clinical guidance and forward selection with $P < 0.05$. Several multiplicative interactions were evaluated based on clinical grounds. The presence of collinearity was examined using a correlation matrix, followed by evaluation of variance inflation factors and magnitude of standard errors. A sequential series of models was developed and we compared those with more variables (i.e., greater complexity) to simpler ones. Improvement in model performance through addition of new candidate variables in multivariable logistic regression models was tested using metrics for concordance statistics (C-statistics) calculated as measures of discrimination and Akaike information criterion (AIC) calculated as measures of goodness of fit. Sensitivity, specificity, Youden's index, positive predictive value, and negative predictive value were calculated, and the cutoff point was set to maximum the Youden's index.

For model validation, both the internal validation and external validation were performed in this respective dataset, and model performance was assessed by discrimination, reclassification, and calibration. Discrimination, which refers to the ability of a model to correctly distinguish between two classes of outcomes (proliferative LN vs. nonproliferative LN), was measured by calculating concordance statistics (C-statistics) and integrated discrimination improvement (IDI).[9] Reclassification, which refers to the movement of patients from one class to another based on changes to assignment to risk categories, was quantified using the net reclassification improvement (NRI) statistic.[10] Calibration, which describes how closely the calculated probabilities agree numerically with the observed outcomes numerically, was measured by calculating the Hosmer-Lemeshow Chi-square statistic that compared the observed and predicted probability of proliferative LN for each quartile of predicted probability and determined the magnitude of the deviation. We also drew a calibration plot using the bootstrap method (500 draws with replacement of 80% sample once a time).[11]

For model presentation, nomograms are a pictorial representation of a complex mathematical formula.[12]

Statistical analyses were performed using STATA/SE version 14.0 (Stata Corp, College Station, TX, USA). A $P < 0.05$ was defined as statistical significance for a two-tailed test.

Results

Cohort description

The development, internal validation, and external validation cohorts included 382, 193, and 164 patients, respectively. Continuous variables missing <30% were imputed using regression-based maximum-likelihood methods, followed by sensitivity analysis, as shown in Supplementary Table 1. The demographic, clinical, and laboratory characteristics in the development, internal validation, and external validation cohorts were similar, except serum calcium, as shown in Table 1. The prevalence of proliferative LN in the development, internal validation, and external validation cohorts were 304/382 (79.6%), 149/193 (77.2%), and 121/164 (73.8%), respectively. The total prevalence of proliferative LN in this study was 77.7%.

Performance in the development cohort

The univariate logistic regression was performed for every eligible variable in the development cohort, shown in Supplementary Table 2. The odds ratios for the variables, including age, gender, systolic BP, eGFR, serum hemoglobin, proteinuria, hematuria, serum C3, and statistics for discrimination, and goodness of fit for a sequential series of models in the development cohort are shown in Table 2. Model 1 performed poor with C-statistics of 0.54 (0.47–0.61) and AIC of 390. The addition of systolic BP and eGFR successively into models 2 and 3 improved the C-statistics (0.66 and 0.75, respectively) and AIC gradually (373 and 346, respectively). We entered hemoglobin, proteinuria, hematuria, and serum C3 in models 4, 5, and 6. Systolic BP and eGFR were added successively in models 5 and 6, resulting in the greatest improvements in C-statistics and AIC in model 5 with 0.85 (0.80–0.90) for C-statistics and 291 for AIC. Compared with model 5, full model 6 including all variables showed no improvement ($P = 0.55$). Receiver operating characteristic (ROC) curves and area under the ROC curve for sequential models are shown in Supplementary Figure 1. Given these results, models 1 and 6 were not considered in further evaluation steps.

Performance in the internal validation cohort

Discrimination, reclassification, and calibration performance in internal validation cohort are shown in Table 3.

For discrimination, both of the C-statistics and IDI improved gradually in the models, resulting in the best C-statistics of 0.84 (0.80–0.92) in model 5, although the P value was marginal.

For reclassification, NRI, as a measure of reclassification, improved gradually in models.

For calibration, the P value of the Hosmer-Lemeshow Chi-square statistics also indicated an improvement of fit with model 5, compared with models 2, 3, and 4. Given the above results and clinical feasibility, model 5 was chosen as the recommended model. A predictive cutoff point was set to maximum the Youden's index of 0.75, with Youden's index of 0.56, sensitivity of 81.7%, specificity of 74.0%, positive predictive value of 92.5%, and negative predictive value of 50.9% [Supplementary Table 3]. In addition, a calibration plot by the bootstrap method repeating 500 times is shown in Figure 1. In model 5, the observed probability compared with the estimated probability for each quartile of the predicted probability within the range of the vertical line indicated confidence intervals.

For predictive model expression, based on the results of the above analysis, the equation for the predictive model was calculated as follows:

Probability = exponential (−1.065 + [−0.006 × age] + [0.910 × gender] + [0.028 × systolic BP] + [−0.032 × hemoglobin] + [0.458 × proteinuria] + [0.425 × hematuria] + [−1.235 × serum C3])/(1 + exp [−1.065+ [−0.006 × age] + [0.910 × gender] + [0.028 × systolic BP] + [−0.032 × hemoglobin] + [0.458 × proteinuria] + [0.425 × hematuria] + [−1.235 × serum C3]).

Performance in the external validation cohort

We performed external validation with fixed coefficients in 164 newly diagnostic patients from our center. The performance of discrimination and calibration was good, with 0.82 (0.74–0.89) for C-statistics, 6.76 for Hosmer-Lemeshow Chi-square statistics, and 0.75 for P value.

Nomogram

Given the above results, a nomogram was built to present the predictive model, as shown in Figure 2. The nomogram was used to calculate the predictive probability by mapping the values of variables with score and then summing all scores

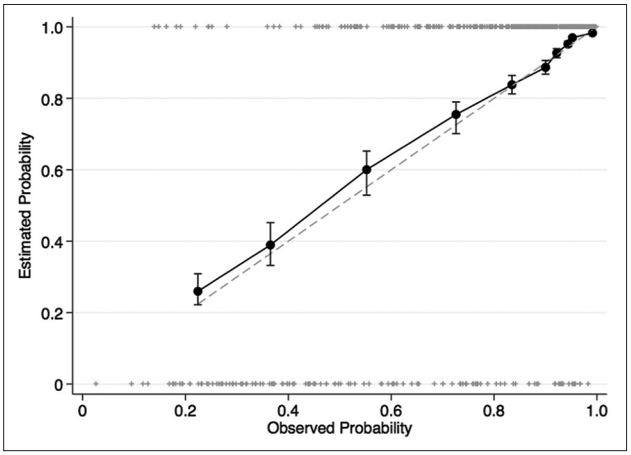


Figure 1: Calibration plot for recommended model 5 by bootstrap method. By bootstrap method repeating 500 times, solid line stands for actual performance and dot line for reference. Plus marks on the top and button indicate the distribution of proliferative lupus nephritis.

Table 1: Baseline demographic and clinical characteristics in development and validation cohorts

Variables	Development cohort (n = 382)	Internal validation cohort (n = 193)	External validation cohort (n = 164)	Development cohort versus internal validation cohort	
				χ^2	P
Demographics					
Age (years)	26 (20, 37)	28 (22, 39)	28 (20, 36)	0.311	0.577
Male gender	66 (17.3)	33 (17.1)	25 (15.2)	0.003	0.957
Physical examination					
Systolic BP (mmHg)	126 (114, 140)	126 (114, 140)	126 (111, 139)	0.071	0.789
Diastolic BP (mmHg)	80 (70, 89)	80 (73, 90)	81 (71, 92)	1.412	0.235
Fever	117 (30.6)	59 (30.6)	36 (22.4)	0.000	0.989
Malar rash	176 (46.7)	68 (35.2)	61 (37.2)	2.512	0.113
Photosensitivity	41 (10.7)	18 (9.3)	19 (11.6)	0.276	0.600
Arthritis	131 (34.3)	58 (30.1)	54 (37.5)	1.045	0.307
Oral ulcer	30 (7.9)	11 (5.7)	8 (4.9)	0.898	0.343
Alopecia	72 (18.9)	25 (13.0)	34 (20.7)	3.177	0.075
Edema	252 (66.3)	135 (70.3)	112 (68.3)	0.931	0.335
Complications					
Hypertension	130 (34.0)	74 (38.3)	56 (34.2)	1.041	0.308
Acute kidney injury	58 (15.2)	38 (19.7)	11 (6.7)	1.872	0.171
Laboratory data					
Baseline eGFR (ml·min^{-1}·1.73 m^{-2})	110 (67, 133)	103 (56, 129)	91 (59, 125)	3.237	0.072
Hemoglobin (g/L)	101 (84, 117)	99 (80, 116)	106 (90, 120)	1.394	0.238
Serum albumin (g/L)	27 (22, 33)	26 (21, 32)	25 (20, 30)	1.707	0.191
Triglyceride (mmol/L)	2.0 (1.4, 3.0)	2.1 (1.5, 3.1)	2.2 (1.6, 2.9)	2.715	0.099
Cholesterol (mmol/L)	5.6 (4.4, 7.2)	5.9 (4.5, 7.5)	6.0 (4.8, 7.2)	1.494	0.222
Uric acid (umol/L)	405 (317, 512)	399 (309, 520)	394 (289, 487)	0.002	0.963
Serum calcium (mmol/L)	2.0 (1.9, 2.2)	2.0 (1.9, 2.1)	2.0 (1.9, 2.1)	4.593	0.032
Serum phosphate (mmol/L)	1.4 (1.2, 1.6)	1.4 (1.2, 1.6)	1.2 (1.1, 1.4)	0.885	0.347
Serum C3 (g/L)	0.4 (0.3, 0.6)	0.4 (0.3, 0.7)	0.4 (0.3, 0.6)	0.104	0.747
Serum C4 (g/L)	0.1 (0.1, 0.2)	0.1 (0.1, 0.2)	0.1 (0.1, 0.2)	0.742	0.389
Haematuria*	1+ (± , 2+)	1+ (± , 2+)	1+ (−, 2+)	0.155	0.694
Proteinuria*	3+ (2+, 3+)	3+ (2+, 3+)	2+ (2+, 3+)	0.551	0.458
Leukocyturia*	− (−, 1+)	− (−, 1+)	− (−, 1+)	0.778	0.375
Urine protein (g/24 h)	1.6 (0.8, 3.0)	1.6 (0.6, 3.3)	3.2 (1.5, 5.9)	0.016	0.900
Positive ANA	369 (96.6)	185 (95.9)	158 (96.9)	0.201	0.654
Positive anti-ds-DNA	313 (81.9)	158 (81.9)	133 (81.6)	0.000	0.983
Pathological data					
Proliferative forms	304 (79.6)	149 (77.2)	121 (73.8)	0.434	0.510
Nonproliferative forms	78 (20.4)	44 (22.8)	43 (26.2)	0.434	0.510
SLEDAI	16 (12, 19)	14 (12, 18)	14 (10, 18)	2.478	0.115
Activity index	6 (5, 8)	7 (5, 8)	7 (4, 9)	0.121	0.728
Chronic index	3 (2, 4)	3 (2, 4)	3 (2, 4)	0.041	0.839

The data were shown as median (Q1, Q3) or n (%). *Hematuria, proteinuria and leukocyturia were divided into six degrees: −, ± , 1+, 2+, 3+, 4+. BP: Blood pressure; eGFR: Estimated glomerular filtration rate; ANA: Antinuclear antibody; SLEDAI: Systemic lupus erythematosus disease activity index.

and reading probabilities on the total score line. An excel calculator based on this predictive model.

Discussion

In this study, by applying a retrospective, observational cohort database, four predictive models were developed and validated to evaluate the probability of presenting proliferative LN in biopsy-proven LN patients. Candidate models 2, 3, 4, and 5, based on demographic characteristics, clinical features, and laboratory data, performed well in discrimination, reclassification, and calibration in both the development and internal validation cohorts. Model 5, including 7 routinely evaluated variables as age, gender, systolic BP, hemoglobin, proteinuria, hematuria, and serum C3, achieved a C-statistics of 0.84 for the development cohort and 0.86 for the internal validation cohort, with a gradual improvement in NRI and a Chi-square value of 2.18 (P = 0.99) for the Hosmer-Lemeshow test. Using the bootstrap method,[11] the magnitude of the deviation between the observed and predicted probability for each quartile of predicted probability was acceptable, which demonstrated that, despite sampling bias, model 5 was able to appropriately predict the probability of proliferative LN

Table 2: Odds ratios and goodness of fit for sequential models in the development cohort*

Variables	Models					
	1	2	3	4	5	6
Age (per 1 year)	1.00†	0.99†	0.96	1.01†	0.99†	0.99†
Male gender	1.77†	1.61†	0.91†	2.79†	2.44†	1.99†
Systolic BP (per 1 mmHg)		1.03	1.02		1.03	1.03
eGFR (per 1 ml·min^{-1}·1.73 m^{-2})			0.97			0.99†
Hemoglobin (per 1 g/L)				0.97	0.97	0.97
Proteinuria (per 1 degree)‡				1.65	1.57	1.53
Hematuria (per 1 degree)‡				1.54	1.53	1.50
Serum C3 (per 1 g/L)				0.34†	0.33†	0.35†
Pseudo R^2	0.01	0.06	0.13	0.25	0.28	0.28
AIC§	390	373	346	299	291	293
C-statistics§	0.54 (0.47–0.61)	0.66 (0.59–0.72)	0.75 (0.69–0.80)	0.82 (0.77–0.88)	0.85 (0.80–0.90)	0.85 (0.80–0.90)
P‖		<0.010	<0.010	0.010	<0.010	0.550

*Data are presented as odds ratios unless otherwise specified; †Odds ratios with $P \geq 0.05$; all other odds ratios are significant (i.e., $P<0.05$); ‡Proteinuria and hematuria were divided into six degrees: −, ±, 1+, 2+, 3+, and 4+; §Null values for C-statistic and AIC are 0.50 and 389, respectively. Higher values for C-statistic and lower values for AIC indicate better models; ‖First-line P values are for comparison of C-statistics between successive models. BP: Blood pressure; eGFR: Estimated glomerular filtration rate; C-statistics: Concordance statistics; AIC: Akaike information criterion.

Table 3: Performance of sequential models in the validation dataset

Metrics	Models			
	2	3	4	5
Discrimination				
C-statistic	0.67 (0.58–0.76)	0.79 (0.71–0.87)	0.84 (0.78–0.91)	0.86 (0.80–0.92)
P	<0.010	<0.010	0.170	0.080
IDI*	0.09	0.11	0.15	0.08
P	<0.010	<0.010	<0.010	<0.010
Calibration				
Hosmer-Lemeshow Chi-square statistic	9.61	8.60	8.68	2.18
P	0.480	0.570	0.560	0.990
Reclassification				
NRI† (%)	77.0	87.0	23.0	21.0
P	<0.010	<0.010	<0.010	<0.010

*IDI is calculated by mean predictive probability of proliferative group minus mean predictive probability of nonproliferative group, and higher values of IDI indicate that the latter model is better; †NRI depends on any change with sign after adding a new variable, and higher values of NRI indicate that the latter model is better. IDI: Integrated discrimination improvement; NRI: Net reclassification improvement.

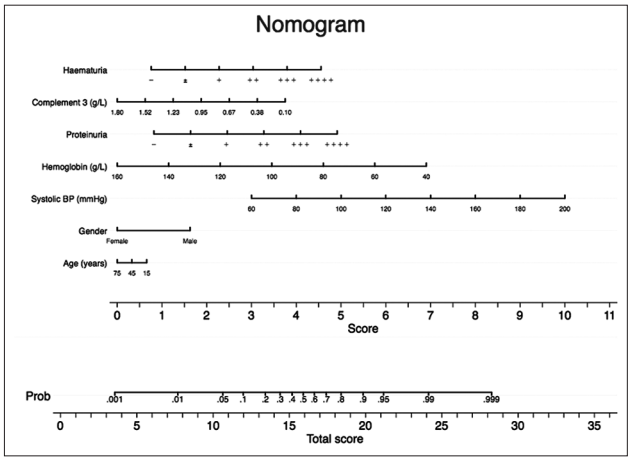


Figure 2: Nomogram graph of predictive model. To provide a quantitative method to better stratify patients with different classes, a nomogram of lupus nephritis was constructed integrating significant independent factors identified in the multivariate analysis.

appropriately. Patients in each decile had a mean predicted probability that was close to the mean observed probability. The performance of discrimination and calibration in the newly diagnostic external validation was good, with a C-statistics of 0.82 and P value of 0.75. Therefore, this predictive model had the potential capability for extended application.

Nomograms have been established and used to estimate the predictive probability of proliferative LN by collecting routine variables: age, gender, systolic BP, hemoglobin, proteinuria, hematuria, and serum C3, with appropriate units. For example, a 35-year-old female LN patient with systolic BP of 110 mmHg, hemoglobin of 120 g/L, proteinuria of (1+), hematuria of (–), and serum C3 of 0.83 g/L, could be mapped with the following scores: age = 0.4, gender = 0, systolic BP = 5.5, hemoglobin = 2.2, proteinuria = 2.4, hematuria = 0.6, serum C3 = 2.3; then, all scores were added to give a sum of 13.4. Finally, we could read the predictive probability corresponding to a score of 13.4 on the total

score line, which was appropriately 0.23, indicating the nonproliferative form with 0.75 as the cutoff point.

Previous studies in our center identified clinical features of distinct histopathological classes of LN. Class I and Class II are associated with mild hypoalbuminemia, hematuria, and proteinuria with normal renal function. Purely Class V is associated with a large amount of proteinuria and severe hypoalbuminemia with relatively preserved renal function.[7,13-15] Class III, Class IV, and mixed classes, the proliferative forms of LN, are associated with the typically systemic presentation of LN, including hypertension, fever, rash, and severe nephrotic syndrome with impaired renal function.[16-18] However, studies that have compared features of Class III with Class IV and those of Class IV-S with Class IV-G have failed to show statistically significant differences in renal outcomes between the two subclasses.[19-21] Meanwhile, this study found that both of the renal and survival outcomes were distinctively different between proliferative LN and nonproliferative LN. In addition, these findings were similar to those in previous studies, which might not be especially designed to identify the predictors but to give clinical information about distinct pathological forms. The experience of previous studies has been considered in the development cohort of this predictive model.

To our knowledge, only one previous study has aimed to build a predictive model based on clinical and routine laboratory parameters to estimate the risk of distinctly histological classes of LN.[22] This study identified three independent models to estimate the risk for each patient separating into different pathological forms. It would be confusing if the subject shared similar risks in distinct models. Nevertheless, this study still provided a useful tool.

Renal biopsy is mandatory for diagnosis and therapy in LN, but absolute contraindications (including uncontrolled severe hypertension, prolonged blood clotting time, active urinary tract infection, uncooperative patients and single kidneys), relative contraindications (including small hyperechoic kidneys, cysts and local skin infections), and particular conditions related to SLE (including anticoagulation because of secondary antiphospholipid syndrome or drug and thrombocytopenia) limited the performance of renal biopsy in LN patients.[1,4,6] Therefore, before making a decision, clinicians need to consider both of the benefit of biopsy and the risk of life-threatening bleeding. Our predictive model may be an ideal solution for this problem. Using seven routinely collected variables to predict the probability of presenting proliferative LN can be applied in biopsy-contraindicated patients to guide further therapy.

The performance of an external validation was good in our predictive model; however, before being applied to other populations, this model needs to be adjusted and carefully validated again. A predictive model with good performance may someday be an alternative to renal biopsy.

Potential limitations should be noticed in this study when we tried to interpret the present findings. The clinical feature is not equal to the histological damage. A study conducted rebiopsy after induction in proliferative LN patients found that one-third of clinical remission patients had persistently high histologic activity and 62% of histologic remission patients were still clinically active.[17] Therefore, the model of this study should be applied carefully and considered as limited information in the patients who cannot receive biopsy. The study population was limited to a single center, which meant that care must be taken when extending results to other populations. Because of nonconformity between clinical and histological findings in SLE, it is difficult to distinguish the pathology class for each case. In addition, this study focused on predicting the probability of proliferative LN, without distinct classification, which made our model simple to apply in clinical practice. Other pathological features, such as tubulointerstitial lesions or vascular involvement,[23-25] were important in LN, but not evaluated in this study. Moreover, biomarkers, which are considered to predict outcome or therapy responses, were not involved in this study.[26-28] A later study would try to combine the traditional model and biomarkers together. Finally, evaluation of the probability of proliferative LN is not the ultimate goal, and further models are highly desired to predict remission or relapsing probability of LN and to evaluate long-term renal outcomes and mortality.

In conclusion, this study developed and validated a model including demographic and clinical indices to evaluate the probability of presenting proliferative LN to guide therapeutic decisions and outcomes. Therefore, these predictors might provide a useful tool to help physicians make decisions regarding treatment, particularly in patients who have contraindication of renal biopsy.

Supplementary information is linked to the online version of the paper on the Chinese Medical Journal website.

Financial support and sponsorship

Study was supported by grants from State Key Program of National Natural Science of China (No. 8113002), Natural Science Foundation of Guangdong (No. 2014B030301023), Guangzhou Committee of Science and Technology (No. 2014YZ-00102), National Natural Science Foundation of China (No. 81470952), and the Chinese National Key Technology R and D Program, Ministry of Science and Technology (No. 2016YFC0906100, No. 2016YFC0906101, No. 2017YFC0907601, No. 2017YFC0907602, and No. 2017YFC0907603).

Conflicts of interest

There are no conflicts of interest.

References

1. D'Cruz DP, Khamashta MA, Hughes GR. Systemic lupus erythematosus. Lancet 2007;369:587-96. doi: 10.1016/S0140-6736(07)60279-7.
2. Wang HP, Wang CY, Pan ZL, Zhao JY, Zhao B. Relationship between clinical and immunological features with magnetic resonance imaging abnormalities in female patients with neuropsychiatric

systemic lupus erythematosus. Chin Med J 2016;129:542-8. doi: 10.4103/0366-6999.176996.
3. Wu CY, Li CF, Wu QJ, Xu JH, Jiang LD, Gong L, *et al.* Chinese systemic lupus erythematosus treatment and research group registry IX: Clinical features and survival of childhood-onset systemic lupus erythematosus in China. Chin Med J 2017;130:1276-82. doi: 10.4103/0366-6999.206346.
4. Rahman A, Isenberg DA. Systemic lupus erythematosus. N Engl J Med 2008;358:929-39. doi: 10.1056/NEJMra071297.
5. Weening JJ, D'Agati VD, Schwartz MM, Seshan SV, Alpers CE, Appel GB, *et al*. The classification of glomerulonephritis in systemic lupus erythematosus revisited. J Am Soc Nephrol 2004;15:241-50. doi: 10.1111/j.1523-1755.2004.00443.x.
6. Members KB. KDIGO 2012 clinical practice guideline for the evaluation and management of chronic kidney disease. Kidney Int Suppl 2013;3:163. doi: 10.1038/kisup.2012.30.
7. Kojo S, Sada KE, Kobayashi M, Maruyama M, Maeshima Y, Sugiyama H, *et al.* Clinical usefulness of a prognostic score in histological analysis of renal biopsy in patients with lupus nephritis. J Rheumatol 2009;36:2218-23. doi: 10.3899/jrheum.080793.
8. Parikh SV, Alvarado A, Malvar A, Rovin BH. The kidney biopsy in lupus nephritis: Past, present, and future. Semin Nephrol 2015;35:465-77. doi: 10.1016/j.semnephrol.2015.08.008.
9. Pencina MJ, D'Agostino RB, Vasan RS. Statistical methods for assessment of added usefulness of new biomarkers. Clin Chem Lab Med 2010;48:1703-11. doi: 10.1515/CCLM.2010.340.
10. Pencina MJ, D'Agostino RB Sr., D'Agostino RB Jr., Vasan RS. Evaluating the added predictive ability of a new marker: From area under the ROC curve to reclassification and beyond. Stat Med 2008;27:157-72. doi: 10.1002/sim.2929.
11. Francq BG, Cartiaux O. Delta method and bootstrap in linear mixed models to estimate a proportion when no event is observed: Application to intralesional resection in bone tumor surgery. Stat Med 2016;35:3563-82. doi: 10.1002/sim.6939.
12. Balachandran VP, Gonen M, Smith JJ, DeMatteo RP. Nomograms in oncology: More than meets the eye. Lancet Oncol 2015;16:e173-80. doi: 10.1016/S1470-2045(14)71116-7.
13. Chan TM. Histological reclassification of lupus nephritis. Curr Opin Nephrol Hypertens 2005;14:561-6. doi: 10.1097/01.mnh.0000168934.18399.97.
14. Yokoyama H, Wada T, Hara A, Yamahana J, Nakaya I, Kobayashi M, *et al.* The outcome and a new ISN/RPS 2003 classification of lupus nephritis in Japanese. Kidney Int 2004;66:2382-8. doi: 10.1111/j.1523-1755.2004.66027.x.
15. Wilhelmus S, Alpers CE, Cook HT, Ferrario F, Fogo AB, Haas M, *et al.* The revisited classification of GN in SLE at 10 years: Time to re-evaluate histopathologic lesions. J Am Soc Nephrol 2015;26:2938-46. doi: 10.1681/ASN.2015040384.
16. Zhang L, Lee G, Liu X, Pascoe EM, Badve SV, Boudville NC, *et al.* Long-term outcomes of end-stage kidney disease for patients with lupus nephritis. Kidney Int 2016;89:1337-45. doi: 10.1016/j.kint.2016.02.014.
17. 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;32:1338-44. doi: 10.1093/ndt/gfv296.
18. Rovin BH, Parikh SV, Alvarado A. The kidney biopsy in lupus nephritis: Is it still relevant? Rheum Dis Clin North Am 2014;40:537-52, ix. doi: 10.1016/j.rdc.2014.04.004.
19. Mittal B, Hurwitz S, Rennke H, Singh AK. New subcategories of class IV lupus nephritis: Are there clinical, histologic, and outcome differences? Am J Kidney Dis 2004;44:1050-9. doi: 10.1053/j.ajkd.2004.08.027.
20. Hill GS, Delahousse M, Nochy D, Bariéty J. Class IV-S versus class IV-G lupus nephritis: Clinical and morphologic differences suggesting different pathogenesis. Kidney Int 2005;68:2288-97. doi: 10.1111/j.1523-1755.2005.00688.x.
21. Yu F, Tan Y, Wu LH, Zhu SN, Liu G, Zhao MH, *et al.* Class IV-G and IV-S lupus nephritis in Chinese patients: A large cohort study from a single center. Lupus 2009;18:1073-81. doi: 10.1177/0961203309106795.
22. Mavragani CP, Fragoulis GE, Somarakis G, Drosos A, Tzioufas AG, Moutsopoulos HM, *et al.* Clinical and laboratory predictors of distinct histopathogical features of lupus nephritis. Medicine (Baltimore) 2015;94:e829. doi: 10.1097/MD.0000000000000829.
23. Alsuwaida AO. Interstitial inflammation and long-term renal outcomes in lupus nephritis. Lupus 2013;22:1446-54. doi: 10.1177/0961203313507986.
24. Wu LH, Yu F, Tan Y, Qu Z, Chen MH, Wang SX, *et al.* Inclusion of renal vascular lesions in the 2003 ISN/RPS system for classifying lupus nephritis improves renal outcome predictions. Kidney Int 2013;83:715-23. doi: 10.1038/ki.2012.409.
25. Song D, Wu LH, Wang FM, Yang XW, Zhu D, Chen M, *et al.* The spectrum of renal thrombotic microangiopathy in lupus nephritis. Arthritis Res Ther 2013;15:R12. doi: 10.1186/ar4142.
26. Anders HJ, Rovin B. A pathophysiology-based approach to the diagnosis and treatment of lupus nephritis. Kidney Int 2016;90:493-501. doi: 10.1016/j.kint.2016.05.017.
27. Zhang X, Nagaraja HN, Nadasdy T, Song H, McKinley A, Prosek J, *et al.* A composite urine biomarker reflects interstitial inflammation in lupus nephritis kidney biopsies. Kidney Int 2012;81:401-6. doi: 10.1038/ki.2011.354.
28. Banchereau R, Hong S, Cantarel B, Baldwin N, Baisch J, Edens M, *et al.* Personalized immunomonitoring uncovers molecular networks that stratify lupus patients. Cell 2016;165:1548-50. doi: 10.1016/j.cell.2016.03.008.

评价增殖型狼疮性肾炎风险的预测模型研究

摘要

背景：狼疮性肾炎通过肾活检可分为增殖型狼疮肾炎和非增殖型狼疮肾炎，两者预后截然不同。然而并不是每位患者都能够进行肾活检。本研究旨在通过建立一个评价增殖型狼疮肾炎风险的预测模型，为不能进行肾活检的狼疮患者提供评估肾脏损伤的方法。

方法：本研究数据来自肾活检证实的狼疮肾炎回顾性队列，随机选取382例患者作为建模队列，193例患者作为内部验证队列，164例新诊断的狼疮肾炎患者作为外部验证队列。构建Logistic模型，并计算C统计量、AIC检验模型拟合优度，计算IDI、NRI检验模型再分类与预测能力。并在外部验证队列中验证最佳模型。

结果：本队列增殖型狼疮肾炎患病率为77.7%。一个包含年龄、性别、收缩压、血红蛋白、蛋白尿半定量、血尿半定量及血清C3水平的模型获得最佳拟合优度与区分度（AIC为291，C统计量为0.84）。经过内部验证和外部验证，该模型的区分度及准确性良好（*C*统计量分别为0.84，0.82；*P*值分别为0.99，0.75）。

结论：我们成功构建了一个模型能通过肾活检前临床及人口学指标评价狼疮肾炎患者为增殖型狼疮肾炎的风险。

Supplementary Table 1: Univariate logistic regression of candidate variables in development cohort

Variables	Odds ratio (*P*)	Variables	Odds ratio (*P*)
Demographics		Laboratory data	
Age (per 1 year)	1.00 (0.920)	Baseline eGFR (per 1 ml·min^{-1}·1.73 m^{-2})	0.98 (<0.010)
Male gender	1.77 (0.140)	Hemoglobin (per 1 g/L)	0.96 (<0.010)
Physical examination		Serum albumin (per 1 g/L)	0.92 (<0.010)
Systolic BP (per 1 mmHg)	1.03 (<0.010)	Triglyceride (per 1 mmol/L)	1.07 (0.400)
Diastolic BP (per 1 mmHg)	1.03 (<0.010)	Cholesterol (per 1 mmol/L)	1.01 (0.780)
Fever	1.25 (0.430)	Uric acid (per 1 mmol/L)	1.00 (<0.010)
Malar rash	0.82 (0.440)	Serum calcium (per 1 mmol/L)	0.26 (0.050)
Photosensitivity	0.45 (0.020)	Serum phosphate (per 1 mmol/L)	2.83 (0.010)
Arthritis	1.42 (0.210)	Serum C3 (per 1 g/L)	0.11 (<0.010)
Oral ulcer	1.03 (0.950)	Serum C4 (per 1 g/L)	0.64 (0.390)
Alopecia	0.79 (0.460)	Hematuria (per 1 degree)*	1.83 (<0.010)
Edema	1.96 (0.010)	Proteinuria (per 1 degree)	1.80 (<0.010)
Complication		Urine protein (per 1 g/24 h)	1.31 (<0.010)
Hypertension	3.13 (<0.010)	Positive ANA	1.02 (0.950)
Acute kidney injury	4.00 (0.010)	Positive anti-ds-DNA	1.42 (0.210)

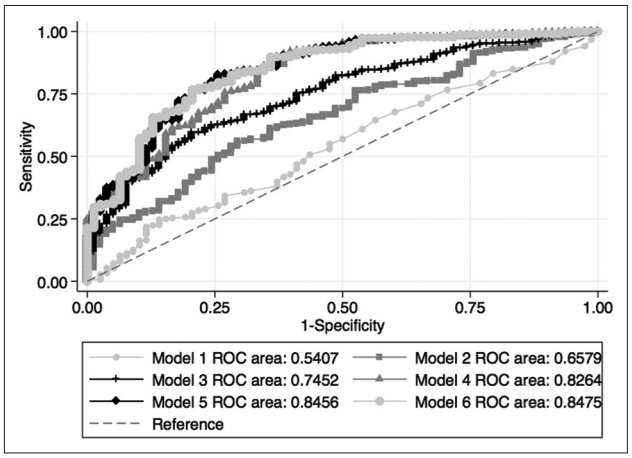
*Hematuria and proteinuria were divided into six degrees: −, ±, 1+, 2+, 3+, and 4+. BP: Blood pressure; eGFR: Estimated glomerular filtration rate; IQR: Interquartile range; ANA: Antinuclear antibody.

Supplementary Table 2: Cutoff points of sequential models in the development cohort

Cutoff point	Model 1*					Model 2				
	Sensitivity (%)	Specificity (%)	YI	PPV (%)	NPV (%)	Sensitivity (%)	Specificity (%)	YI	PPV (%)	NPV (%)
0.79	18.8	88.5	0.07	86.4	21.8	61.2	64.	0.25	86.9	29.8
0.77	NA	NA	NA	NA	NA	65.8	56.4	0.22	85.5	19.7
0.75	NA	NA	NA	NA	NA	74.3	47.4	0.22	84.6	32.2
0.73	NA	NA	NA	NA	NA	78.6	39.7	0.18	83.6	32.3
0.71	NA	NA	NA	NA	NA	83.6	26.9	0.10	81.7	29.6
Cutoff point	**Model 3**					**Model 4**				
	Sensitivity (%)	Specificity (%)	YI	PPV (%)	NPV (%)	Sensitivity (%)	Specificity (%)	YI	PPV (%)	NPV (%)
0.79	61.2	76.9	0.38	91.2	33.7	75.0	72.7	0.48	91.5	42.8
0.77	64.1	71.8	0.36	89.9	33.9	77.0	68.8	0.46	90.6	43.4
0.75	69.1	64.0	0.33	88.2	34.7	80.0	66.2	0.46	90.2	46.0
0.73	73.7	59.0	0.33	87.5	36.5	81.7	66.2	0.48	90.4	48.1
0.71	78.0	53.9	0.32	86.8	38.5	84.7	66.2	0.51	90.7	52.6
Cutoff point	**Model 5†**					**Model 6**				
	Sensitivity (%)	Specificity (%)	YI	PPV (%)	NPV (%)	Sensitivity (%)	Specificity (%)	YI	PPV (%)	NPV (%)
0.79	77.7	76.6	0.54	92.8	46.8	76.7	76.6	0.53	92.7	45.7
0.77	79.3	74.0	0.53	92.3	47.9	78.7	74.0	0.53	92.2	47.1
0.75	81.7	74.0	0.56	92.5	50.9	81.0	71.4	0.52	91.7	49.1
0.73	83.0	70.1	0.53	91.5	51.4	82.7	70.1	0.53	91.5	50.9
0.71	84.3	64.9	0.49	90.4	51.6	83.7	66.2	0.50	99.6	51.0

*Least risk point for Model 1 is 0.78, so lower point is not applicable; †Maximum Youden's index appears in recommended model 5 at 0.75 cutoff point with correctly classified rate of 0.80. YI: Youden's index; PPV: Positive predictive value; NPV: Negative predictive value; NA: Not applicable.

Supplementary Table 3: Sensitivity analysis of original dataset and imputation dataset, median (Q1, Q3)				
Variables	**Original dataset**	**Imputation dataset**	χ^2	***P***
Urine specific gravity	*n* = 571, 1.02 (1.01, 1.02)	*n* = 575, 1.02 (1.01, 1.02)	0.356	0.551
Urine pH	*n* = 571, 6.0 (6.0, 6.5)	*n* = 575, 6.0 (6.0, 6.5)	0.194	0.660
Serum glucose	*n* = 566, 4.5 (4.1, 5.1)	*n* = 575, 4.5 (4.2, 5.1)	1.822	0.177
Serum uric acid	*n* = 560, 402 (315, 515)	*n* = 575, 404 (316, 517)	0.013	0.910
Alanine aminotransferase	*n* = 570, 17 (12, 27)	*n* = 575, 17 (13, 27)	1.842	0.175
Aspartate aminotransferase	*n* = 570, 20 (16, 27)	*n* = 575, 20 (16, 28)	1.845	0.174



Supplementary Figure 1: ROC curves and AUC for sequential models in the development cohort. Model 5 and Model 6 shares the best ROC performance in the development cohort, with AUC 0.846 and 0.848, respectively, $P < 0.05$. AUC: Area under the ROC curve; ROC: Receiver operating characteristic.