Development and Validation of a Nomogram for the Early Prediction of IgA Nephropathy in Primary Glomerulonephritis

Hou-Long Luo

Peking University Shenzhen Hospital

Lixia Cui

Peking University Shenzhen Hospital

Mingyang Li

Peking University Shenzhen Hospital

Chencui Huang

Beijing Deepwise & League of PHD Technology Co., Ltd

Wei Liu

Beijing Deepwise & League of PHD Technology Co., Ltd

Yong Xia (✉ sunmoonrain78@163.com)

Peking University Shenzhen Hospital

Ling Ji

Peking University Shenzhen Hospital

Research Article

Keywords: IgA nephropathy, Primary glomerulonephritis, Prediction, Nomogram

Posted Date: October 11th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-941889/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background: To make early prediction of immunoglobulin A nephropathy (IgAN) before renal biopsy, we developed and validated a new non-invasive nomogram for the early prediction of IgAN in primary glomerulonephritis (GN) in south China.

Methods: A total of 431 patients were included in this study and additional 113 patients were included as the independent test cohorts to validate our results. A stepwise regression model was used for features selection. Multivariate logistic regression analysis with 5-fold cross validation was used to validate the result of the stepwise selection. Performance of the logistic regression model was assessed with respect to its calibration, discrimination, and clinical usefulness. Independent test was assessed.

Results: We developed a model incorporating age of patients and four clinical chemistry signatures, including serum IgA, serum albumin (ALB), serum phosphorus (P) and 24-hour urinary protein (24hUpro) and presented with a nomogram. The area under the receiver operating characteristic (ROC) curve (AUC) reached 0.89 (95% CI: 0.86–0.92) and 0.88 (95%: 0.85-0.92) in the training set and validation set, respectively. The model also had good performance in independent test cohorts (AUC of 0.82, 95% CI: 0.75–0.90). The DCA and calibration plot of the model also shows good performance.

Conclusions: The logistic regression model presented in this study incorporates age of patients, IgA, ALB, P and 24hUpro and can be conveniently used to facilitate the individualized prediction of IgAN.

Background

Immunoglobulin A nephropathy (IgAN) is the most common type of primary glomerulonephritis (GN) globally^[1, 2], especially in the Asia-Pacific region^[2, 3]. IgAN is characterized by the predominant deposition of IgA in the mesangium of glomeruli^[4]. The range of clinical manifestations of IgAN is broad, from asymptomatic microscopic hematuria to rapidly progressive GN^[5]. The typical mode of presentation varies according to age group and biopsy practice patterns. Epidemiological investigation showed that IgAN accounts for ~45.3% of primary GN cases in China^[6]. IgAN reduces life expectancy by more than 10 years and leads to kidney failure in 20–40% of patients within 20 years after the diagnostic renal biopsy^[7]. Therefore, it's vital for the early diagnosis and intervention of IgAN.

Until now, the diagnosis of IgAN relies on revealing IgA as the dominant or codominant immunoglobulin in the glomerular mesangium by renal biopsy^[8]. However, biopsy registry data underestimate disease burden as patients with mild disease may not undergo renal biopsy, and in countries lacking screening programs disease may not be detected^[5]. And renal biopsy is invasive and cannot be repeated frequently to evaluate therapeutic effect. Therefore, it is necessary to develop novel, non-invasive diagnostic model for IgAN.

As a statistical tool, a nomogram can solve the above problems in a more accurate way. A large number of studies showed that a nomogram can predict the prognosis of some kidney disease^[9–11]. Mao Yonghui et al.^[9] proposed a predictive model for the progression of patients with primary membranous nephropathy and nephrotic syndrome based on age, sPLA2R-Ab, proteinuria and Ua1m/Cr. Another study from China developed and assessed a predictive nomogram for the progression of IgA nephropathy^[12]. However, there are few reports about the diagnostic model of IgAN form primary GN. In particular, non-invasive diagnostic models are scarcer. Therefore, the purpose of this study is to develop and validate an accurate, valid and simple non-invasive nomogram model to predict IgAN from primary GN to help clinicians make clinical decision.

Methods

Study Design and Patients

This is a retrospective monocentric analysis of patients who underwent biopsy-proven primary GN at Peking University Shenzhen Hospital between January 2019 and May 2021. Study participants split into two independent set. 431 patients with primary GN enrolled at Peking University Shenzhen Hospital from January 2019 to December 2020 as training cohorts. Another independent

test cohorts included 113 patients from January 2021 to May 2021 were enrolled as validation cohorts to validate our results. All study participants provided their written informed consent.

The inclusion criteria were as follows: with biopsy-proven primary GN; 18 years or older; with eGFRs ≥ 30 mL/min/1.73 m² at the time of renal biopsy. We excluded those with clinical or serologic evidence of systemic lupus erythematosus, antinuclear cytoplasmic antibody–associated vasculitis, Henoch-Schönlein purpura, or liver cirrhosis and those with urinary tract infection or obstruction at the time of biopsy.

Clinical and Laboratory Data

Baseline information were collected at the time of renal biopsy, including age, gender, serum albumin (ALB), serum total protein (TP), serum creatinine (SCr), blood urea nitrogen (BUN), uric acid (UA), blood glucose (GLU), total cholesterol (TC), triglyceride (TG), low-density lipoprotein cholesterol (LDL-C), high-density lipoprotein cholesterol (HDL-C), immunoglobulin G (IgG), IgM, IgA, complement 3 (C3), C4, total complement activity (CH50), serum calcium (Ca), serum magnesium (Mg), serum phosphorus (P), C-reactive protein (CRP), antistreptolysin O (ASO), 24-hour urinary protein (24hUpro), random urinary protein creatinine ratio (uPCR). Estimated glomerular filtration rate (eGFR in mL/min/1.73 m²) was calculated by the equation put forward by the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) and was categorized according to the KDIGO 2012 Clinical Practice Guideline^[13].

Renal Biopsy

For all patients, renal biopsy was performed at the time of diagnosis. Pathological renal examination included light microscopy, immunofluorescence, and electron microscopy. Direct immunofluorescence analysis was performed on frozen renal tissue sections to detect IgG, IgM, IgA, C3, C4d, C1q, Fibrin, HBcAg, HBsAg, Kappa, Lambda, PLA2R, IgG1, IgG2, IgG3, IgG4 and Albumin. Glomerular membranous nephropathy lesions were classified into four stages (I, II, III and IV) based on Ehrenreich and Churg's criteria. IgAN histologic lesions were graded according to MEST-C score^[14].

Statistical Analysis

Statistical analyses were performed using Deepwise & Beckman Coulter DxAI platform, version 1.0.3 (<https://dxonline.deepwise.com>) and SPSS, version 17.0 (IBM Corp., Armonk, NY, USA). In all analyses, statistical significance is implied at two-sided with alpha of 0.05.

Characteristics of patients and clinical and laboratory data are presented as frequency (percentages), mean (SD) for normally distributed variables and median (IQR) for non-normally distributed variables. Stepwise regression analysis using the Akaike information criterion (AIC) was used to identify candidates to be incorporated into the multivariable logistic regression model. The odds ratio (OR), 95% confidence interval (CI) and *p* value were calculated for each candidate to predict possible diagnosis. Selected candidates were incorporated into the multivariate logistic regression model and presented with nomogram to visualize the probability of IgAN diagnosis using Deepwise & Beckman Coulter DxAI platform in the training cohorts.

We applied both internal validation and independent test cohorts to evaluate the performance of logistic regression model. 5-fold cross validation was used as internal validation to develop the model in the training set. A calibration plot was used to assess the consistency between predicted probability and actual probability of the model^[15]. To quantify the discrimination performance of the model, the area under ROC curve (AUC) was calculated. The performance of the model was also tested in the independent test cohorts. The logistic regression formula formed in the training set was applied to patients in the independent test set, with calculation of the total points for each patient. Then, logistic regression analysis was performed using the total points as a factor. Decision curve analysis (DCA) is a method for evaluating the benefits of a diagnostic test across a range of patient preferences for accepting risk of undertreatment and overtreatment to facilitate decisions about test selection and use^[16].

Results

Characteristics of the study cohort

A total of 544 patients with biopsy-proven primary GN were enrolled in this study (shown in Fig. 1). Median age at biopsy was 37 (IQR 30–47) years. The male/female ratio was 1.125: 1 (288/256). The median eGFR was 78.08 (IQR 56.12– 103.23) mL/min/1.73 m², and the 24hUpro was 1.62 (IQR 0.74– 3.72) g/24 h. The study population consisted of two independent datasets, training set (431 patients) and independent test set (113 patients). The clinical characteristics were similar in the 2 sets (shown in Table 1). The percentage of IgAN was 56.4% (243/431) and 54.0% (61/113) in the training and validation set, respectively. Patient characteristics of IgAN and non-IgAN in the 2 groups are shown in the Table 1.

Table 1
Baseline characteristics of training cohorts and independent test cohorts at biopsy

	Training cohorts				Independent test cohorts				<i>p</i> -Value ^b
	Total	IgAN	non-IgAN	<i>p</i> -Value ^a	Total	IgAN	non-IgAN	<i>p</i> -Value ^a	
No. of patients	431	243 (56.4%)	188 (43.6%)	-	113	61 (54.0%)	52 (46.0%)	-	0.648
men(n[%])	229 (53.1%)	119 (49.0%)	110 (58.5%)	0.049	59 (52.2%)	32 (54.2%)	27 (45.8%)	0.955	0.862
Age, years	35 (29-44)	34 (28-40)	40 (31-50)	<0.001	38 (31-49)	35 (31-41)	41 (32.5-54.5)	0.036	0.39
IgG (g/L)	10.24 (7.49-11.91)	10.77 (9.25-12.07)	7.79 (4.46-10.96)	<0.001	9.52 (7.13-11.40)	10.07 (8.58-11.55)	8.46 (5.87-10.87)	0.003	0.359
IgM (g/L)	1.23±0.61	1.22±0.58	1.25±0.64	0.936	1.14±0.53	1.15±0.53	1.13±0.53	0.834	0.176
IgA (g/L)	3.04±1.11	3.44±1.02	2.45±1.00	<0.001	2.59±0.88	2.81±0.84	2.33±0.85	0.003	0.001
C3 (g/L)	1.12(0.97-1.25)	1.06 (0.94-1.19)	1.20 (1.02-1.37)	<0.001	1.07 (0.93-1.17)	1.08 (0.92-1.16)	1.07 (0.94-1.17)	0.786	0.052
C4 (g/L)	0.25(0.21-0.31)	0.24 (0.20-0.30)	0.27 (0.22-0.35)	0.047	0.26 (0.22-0.32)	0.26 (0.22-0.32)	0.26 (0.22-0.33)	0.833	0.05
IgA/C3	2.64(1.89-3.55)	3.11 (2.48-3.90)	1.89 (1.51-2.48)	<0.001	2.33 (1.85-2.94)	2.59 (2.04-3.08)	2.01 (1.62-2.74)	0.017	0.054
CH50 (U/mL)	40.2(34.7-44.8)	40.8 (35.6-44.8)	39.3 (33.6-44.9)	0.089	38.6 (30.4-44.4)	40.5 (33.5-44.7)	33.8 (28.3-43.6)	0.032	0.104
K (mmol/L)	3.97(3.77-4.21)	3.99 (3.79-4.24)	3.95 (3.74-4.17)	0.532	3.92 (3.69-4.13)	3.92 (3.74-4.13)	3.92 (3.68-4.11)	0.802	0.154
Na (mmol/L)	141(140-142)	141 (140-142)	141 (139-143)	0.164	141 (140-143)	141 (139-142)	142 (141-143)	0.025	0.432
Cl (mmol/L)	106.4(104.8-108.1)	106.1 (104.9-107.8)	107.0 (104.7-108.6)	0.001	105.7 (104.4-108.1)	105.2 (104.2-107.5)	106.45 (104.9-109.5)	0.010	0.28
Ca (mmol/L)	2.21(2.10-2.28)	2.24 (2.18-2.31)	2.09 (1.97-2.23)	<0.001	2.21 (2.10-2.30)	2.23 (2.17-2.30)	2.11 (2.01-2.25)	<0.001	0.381
Mg (mmol/L)	0.83(0.77-0.86)	0.83 (0.78-0.86)	0.83 (0.76-0.87)	0.472	0.83 (0.78-0.87)	0.83 (0.79-0.86)	0.82 (0.77-0.87)	0.775	0.922
P (mmol/L)	1.23±0.19	1.217±0.17	1.27±0.20	<0.001	1.26±0.19	1.22±0.18	1.32±0.18	0.004	0.086
ALT (U/L)	15 (10-23)	13 (10-21)	17 (12-24)	0.002	17 (12-26)	15 (11-21)	20 (14-27)	0.011	0.055
TB (μmol/L)	9.6(7.4-12.6)	10.2 (8.2-13.8)	8.6 (5.9-10.9)	<0.001	9.3 (7.1-11.9)	9.7 (7.7-13.6)	8.3 (6.7-11.1)	0.053	0.108
DB (μmol/L)	1.8±1.0	2.0±1.0	1.5±0.9	<0.001	1.6±0.8	1.8±0.7	1.5±0.8	0.022	0.162

	Training cohorts				Independent test cohorts				<i>p</i> -Value ^b
	Total	IgAN	non-IgAN	<i>p</i> -Value ^a	Total	IgAN	non-IgAN	<i>p</i> -Value ^a	
TP (g/L)	63.6 (55.3-68.0)	66.0 (61.6-69.0)	55.3 (44.4-65.5)	<0.001	63.4 (54.2-67)	65.6 (61.8-68.9)	57.5 (48.8-63.9)	<0.001	0.73
ALB (g/L)	35.9 (30.4-39.3)	37.6 (35.0-40.2)	29.8 (22.3-35.9)	<0.001	35.1 (29.9-38.8)	37.5 (34.7-39.6)	31 (24.8-35.3)	<0.001	0.731
TG (mmol/L)	1.49(1.02-2.15)	1.40 (0.92-1.90)	1.63 (1.17-2.40)	<0.001	1.5 (1.05-2.23)	1.38 (1.02-1.99)	1.71 (1.20-2.52)	0.098	0.866
TC (mmol/L)	5.41(4.63-6.44)	5.22 (4.48-5.89)	5.87 (4.86-8.28)	<0.001	5.47 (4.66-6.38)	5.25 (4.72-6.04)	5.92 (4.56-7.62)	0.025	0.941
HDL-C (mmol/L)	1.23(1.01-1.59)	1.20(0.98-1.52)	1.35 (1.02-1.70)	0.003	1.15 (0.98-1.48)	1.08 (0.92-1.35)	1.27 (1.08-1.57)	0.005	0.173
LDL-C (mmol/L)	3.57(2.92-4.40)	3.34 (2.90-3.91)	4.05 (3.08-5.40)	<0.001	3.74 (3.10-4.37)	3.61 (3.21-4.03)	3.87 (2.87-5.33)	0.184	0.665
Urea (mmol/L)	5.21(4.03-7.14)	5.34 (4.32-7.17)	4.77 (3.71-7.04)	0.096	5.43 (4.30-7.03)	5.86 (4.68-7.48)	4.92 (4.00-6.56)	0.021	0.447
Cr (μmol/L)	90(65-119)	93 (70-121)	82 (61-108)	0.019	88 (63-112)	94 (74-124)	76 (55-97)	0.002	0.237
UA (μmol/L)	400±106	395±104	406±110	0.300	402±100	408±101	394±99	0.474	0.729
GLU (mmol/L)	4.72(4.42-5.09)	4.71 (4.41-5.05)	4.76 (4.42-5.17)	0.594	4.86(4.53-5.30)	4.97 (4.55-5.31)	4.83 (4.46-5.20)	0.367	0.047
eGFR	78.24(54.07-103.04)	69.58 (52.74-97.48)	88.49 (65.39-109.08)	0.008	78.20 (56.14-109.07)	68.76 (54.87-86.45)	96.11 (61.64-117.34)	0.003	0.204
hsCRP (mg/L)	0.74(0.35-1.83)	0.76 (0.35-1.70)	0.71 (0.35-2.32)	0.739	1.11 (0.56-2.67)	1.74 (0.57-3.91)	0.85 (0.56-1.91)	0.115	0.004
ASO (IU/mL)	57.3(26.1-113.1)	78.4 (36.0-127.1)	36.8 (19.2-82.2)	<0.001	58.6 (33.8-103.7)	72.1 (37.4-131.3)	46.0 (30.0-82.6)	0.053	0.747
24hU-TP (g/24h)	1.58(0.70-3.76)	1.08 (0.62-2.10)	3.76 (1.21-6.16)	<0.001	1.71 (0.74-3.36)	1.39 (0.57-2.20)	2.28 (1.23-4.70)	0.001	0.001
U-TP/Cr (g/L)	1.12(0.56-2.43)	0.83 (0.48-1.48)	2.15 (0.71-4.74)	<0.001	1.24 (0.5-2.05)	0.80 (0.42-1.32)	1.82 (0.88-3.54)	<0.001	0.003

Abbreviations: IgG, immunoglobulin G; IgM, immunoglobulin M; IgA, immunoglobulin A; C3, complement 3; C4, complement 4; CH50, total complement activity; Ca, serum calcium; Mg, serum magnesium; P, serum phosphorus; ALT, alanine transaminase; TBil, total bilirubin; DBil, direct bilirubin; ALB, serum albumin; TP, serum total protein; TC, total cholesterol; TG, triglyceride; LDL-C, low-density lipoprotein cholesterol; HDL-C, high-density lipoprotein cholesterol; SCr, serum creatinine; BUN, blood urea nitrogen; UA, uric acid; GLU, blood glucose; eGFR, estimated glomerular filtration rate; hsCRP, hypersensitivity C-reactive protein; ASO,

antistreptolysin O; 24hUpro, 24-h proteinuria; uPCR, random urinary protein creatinine ratio; a, IgAN compared with non-IgAN in the training cohorts or independent test cohorts; b, training cohorts compared with independent test cohorts.

Pathological Types

The training set consisted of 243 (56.38%) patients with IgAN and 188 (43.62%) patients with non-IgAN. Of the 188 patients with non-IgAN, 107 (56.91%) were diagnosed with primary membranous nephropathy (MN); 34 (18.09%) were diagnosed with membranoproliferative glomerulonephritis; 15 (7.98%) were diagnosed with focal segmental glomerulosclerosis (FSGS); 15 (7.98%) were diagnosed with minimal change disease; and 17 (9.04%) were diagnosed with other renal diseases.

The independent test set consisted of 61 (53.98%) patients with IgAN and 52 (46.02%) patients with non-IgAN. Of the 52 patients with non-IgAN, 30 (57.69%) were diagnosed with MN; 9 (17.31%) were diagnosed with membranoproliferative glomerulonephritis; 5 (9.62%) were diagnosed with FSGS; 7 (13.46%) were diagnosed with minimal change disease; and 1 (1.92%) were diagnosed with other renal diseases.

Development of clinical diagnosis nomogram model

Of the clinical features, 33 candidates were reduced to 5 potential candidates subjected to stepwise regression analysis on the basis of 431 patients in the training set. Multivariable logistic regression analysis showed that 5 potential candidates through stepwise selection were all statistically significant ($p < 0.05$), thus all 5 candidates were put into the nomogram as the independent variable. Table 2 shows the regression coefficients and risk scores for each variable. These potential features were age, IgA, ALB, P and 24hUpro. A model including the 5 potential features was developed and presented as the nomogram (shown in Fig. 2). It presents each variable in the nomogram which was assigned a specific score, and the total score which was used to obtain probability for predicting IgAN from primary GN (shown in Fig. 2).

Table 2 Regression coefficients and risk scores for each variable			
Features	Prediction Model		
	β	OR (95% CI)	<i>p</i> -value
IgA (g/L)	1.115	3.048 (2.237-4.153)	<0.001
ALB (g/L)	0.069	1.071 (1.020-1.124)	0.006
P (mmol/L)	-3.136	0.043 (0.007-0.283)	0.001
24hUpro (g/24h)	-0.277	0.758 (0.645-0.890)	0.001
age, years	-0.064	0.938 (0.915-0.961)	<0.001

Apparent Performance of the model

The multivariate logistic regression had an AUC of 0.89 (95% CI 0.86-0.92) in the training set and 0.88 (95% CI 0.85-0.92) in the validation set through 5-fold cross-validation, which indicated the model we developed had an excellent prediction power (shown in Fig. 3. a). The model obtained 0.85 sensitivity (recall), 0.81 specificity, 0.83 accuracy and 0.85 precision in the training set. In the validation set, it had a sensitivity (recall) of 0.85, specificity of 0.81, accuracy of 0.83 and precision of 0.85. Furthermore, The independent test cohort obtained a high AUC value of 0.82 (95% CI 0.74-0.90) (shown in Fig. 3. a). Also, the model had 0.74 sensitivity (recall), 0.79 specificity, 0.76 accuracy and 0.80 precision in the independent test. Furthermore, in the training set, validation set and independent test set, the calibration plot indicated a consistency between the observed result and prediction for the probability of IgAN (shown in Fig. 3. b).

Decision curve analysis

The result of DCA for the nomogram is presented in Figure 4. It indicated that the nomogram had an excellent performance in clinical practice. There was a broad spectrum of alternative threshold probability, suggesting that the model was a good

assessment tool.

Discussion

Increasingly, nomogram is used by clinicians to estimate and predict the risk of disease for patients, and include easy-to-use digital interfaces that improve diagnostic and predictive efficiency by incorporating multiple independent predictors. In this study, we developed and validated a non-invasive diagnostic prediction model that integrates 5 independent features (IgA, ALB, P, 24hUpro and age) for the prediction of IgAN in primary GN. This model displayed an excellent level of discrimination and a high AUC of 0.89. Besides, the model was validated by 5-fold cross validation with an AUC of 0.88. The independent test also certified a good level of discrimination with an AUC of 0.82. The calibration plot indicated good consistency between the actual and predicted diagnoses.

Currently, renal biopsy still is the gold standard for clinical diagnosis and evaluation the degrees of IgA nephropathy^[17]. However, renal biopsy is an invasive method and is not suitable to biopsy patients with mild disease and in countries and districts lacking screening programs disease^[5]. Therefore, study of non-invasive clinical features related to the diagnosis of IgAN and development of non-invasive diagnosis model have important clinical significance for the early diagnosis of IgAN. Due to the lack of non-invasive methods to predict IgAN, we developed a model with non-invasive clinical features to predict IgAN and used nomogram to calculate risk for individual patient.

As early as in 1995, the joint committee of the special study group on progressive glomerular diseases, the Ministry of Health and Welfare of Japan, and the Japanese Society of Nephrology reported serum IgA of more than 350 mg/dl in adults as one of the diagnostic criteria for IgA nephropathy^[18]. Since then, several researchers have reported serum IgA as diagnostic and prognostic markers of IgAN^[19, 20]. Recently, studies have found that serum IgA/C3 ratio had better diagnostic and prognostic value for IgAN^[8, 21, 22]. At the beginning of the model building, we also considered including serum IgA/C3 ratio in the nomogram model, but unfortunately it did not improve the diagnostic value of the model. So, in this nomogram, serum IgA was included instead of IgA/C3 ratio.

Galactose deficient IgA1 (Gd-IgA1) is a critical molecule in the pathogenesis of IgAN^[23]. Galactose deficiency of O-linked glycans in the hinge region of IgA1 is the beginning of a sequence of events that may lead to renal injuries. The formation of galactose-deficient IgA1 (Gd-IgA1) is the pivotal step of multi-hit pathogenesis of IgAN^[24]. Meanwhile, Gd-IgA1 was suggested as a potential disease-specific biomarker that predicts disease activity and prognosis^[25, 26]. However, due to the limitation of testing methods, Gd-IgA1 has not been used as a routine test in clinic. So, Gd-IgA1 was not considered to be included in this nomogram model.

In this study, serum IgG levels is significantly increased in IgAN patients compared with non-IgAN patients. Serum IgG concentration at baseline is a predictive marker for the prognosis of IgAN. Serum IgG level is an independent risk factor for poor outcomes in IgAN at the time of renal biopsy. Every 1 g/L decrease in serum IgG level was associated with a 1.74-fold increased risk of the incidence of composite renal outcomes^[27]. Furthermore, IgG deposits in the mesangium and capillary loops predict adverse renal outcome in patients with IgAN^[28]. However, serum IgG level did not distinguish IgAN well from primary GN and not improve the efficiency of the diagnostic model.

The strength of our study suggested by including few features and applying a non-invasive model we can obtain a good prediction power, thus minimize the tests needed for patients and could easily be adopted in clinical practice. In addition, we divided eligible patients into training and validation datasets using two independent set to evaluate the multivariate logistic regression model both internally and externally. In addition, the predictive power of the novel model was very high (AUC of 0.82) in independent test cohorts, suggesting its high diagnostic value.

However, this study had several limitations. Firstly, the model has been developed and validated in monocentric retrospective cohort of primary GN patients. So, multicenter retrospective cohort study is needed to verify the generalizability of the results in

other primary GN patients. Second, the retrospective study design could not exclude the confounding effects. Further prospective cohort study needed to be done to increase the generalizability and deal with confounding effect.

In conclusion, this study developed a nomogram that incorporating IgAN, ALB, P, 24hUpro, and age. Our nomogram displayed excellent performance in both training and validation sets for predicting IgAN from primary GN, and thereby can be used conveniently either in triage or as a replacement for renal biopsy to facilitate the individualized prediction of IgAN diagnosis in patients with primary GN.

Abbreviations

IgAN: immunoglobulin A nephropathy, GN: glomerulonephritis, ALB: serum albumin, P: serum phosphorus, 24hUpro: 24-hour urinary protein, ROC: receiver operating characteristic, AUC: area under the receiver operating characteristic curve, CI: confidence interval, DCA: decision curve analysis, TP: serum total protein, SCr: serum creatinine, BUN: blood urea nitrogen, UA: uric acid, GLU: blood glucose, TC: total cholesterol, TG: triglyceride, LDL-C: low-density lipoprotein cholesterol, HDL-C: high-density lipoprotein cholesterol, IgG: immunoglobulin G, C3: complement 3, CH50: total complement activity, Ca: serum calcium, Mg: serum magnesium, CRP: C-reactive protein, ASO: antistreptolysin O, uPCR: random urinary protein creatinine ratio, eGFR: estimated glomerular filtration rate, AIC: Akaike information criterion, OR: odds ratio, FSGS: focal segmental glomerulosclerosis, Gd-IgA1: galactose deficient IgA1.

Declarations

Ethics approval and consent to participate

The protocol of using human blood samples in vitro and patients data was approved by the Institutional Biosafety Committees at Peking University Shenzhen Hospital. All studies were performed in accordance with the Declaration of Helsinki and complied with the guidelines of the Office for Human Research Protections. All patients provided written informed consent before initiation of study.

Consent for publication

Not applicable.

Availability of data and materials

The datasets generated and analysed during the current study are not publicly available due to privacy of patients but are available from the corresponding author on reasonable request.

Competing interests

We declare that there are no competing interests related to this study.

Funding

Not applicable.

Authors' contributions

H. L., Y.X. and L.J. designed the study. H.L., L.C., and M.L. acquired the clinical data and implemented the study. H.L., L.C., and M.L. acquired the pathological data. H.L., C.H. and W.L. performed the statistical analysis, interpreted the results and drafted the manuscript. C.H. and L.J. provided critical revision of the manuscript for important intellectual content. C.H., Y.X. and L.J. reviewed the paper.

Acknowledgements

We appreciate the study subjects for their assistance.

References

1. Paunas FTI, Finne K, Leh S, Osman TA, Marti HP, Berven F, et al. **Characterization of glomerular extracellular matrix in IgA nephropathy by proteomic analysis of laser-captured microdissected glomeruli.** *BMC Nephrol* 2019, 20(1):410.
2. Wyatt RJ, Julian BA. **IgA nephropathy.** *New Engl J Med* 2013, 368(25):2402-2414.
3. Schena FP, Nistor I. **Epidemiology of IgA Nephropathy: A Global Perspective.** *Semin Nephrol* 2018, 38(5):435-442.
4. Perše M, Večerić-Haler Ž. **The Role of IgA in the Pathogenesis of IgA Nephropathy.** *Int J Mol Sci* 2019, 20(24).
5. Rodrigues JC, Haas M, Reich HN. **IgA Nephropathy.** *Clin J Am Soc Nephrol* 2017, 12(4):677-686.
6. Li LS, Liu ZH. **Epidemiologic data of renal diseases from a single unit in China: analysis based on 13,519 renal biopsies.** *Kidney Int* 2004, 66(3):920-923.
7. Berthouix FC, Mohey H, Afiani A. **Natural history of primary IgA nephropathy.** *Semin Nephrol* 2008, 28(1):4-9.
8. Gong WY, Liu M, Luo D, Liu FN, Yin LH, Li YQ, et al. **High serum IgA/C3 ratio better predicts a diagnosis of IgA nephropathy among primary glomerular nephropathy patients with proteinuria ≤ 1 g/d: an observational cross-sectional study.** *BMC Nephrol* 2019, 20(1):150.
9. Liu L, Wang H, Zhao B, Liu X, Sun Y, Mao Y. **Nomogram to predict the progression of patients with primary membranous nephropathy and nephrotic syndrome.** *Int Urol Nephrol* 2021.
10. Li F, Wei RB, Wang Y, Su TY, Li P, Huang MJ, et al. **Nomogram prediction model for renal anaemia in IgA nephropathy patients.** *Open medicine (Warsaw, Poland)* 2021, 16(1):718-727.
11. Hu Y, Shi R, Mo R, Hu F. **Nomogram for the prediction of diabetic nephropathy risk among patients with type 2 diabetes mellitus based on a questionnaire and biochemical indicators: a retrospective study.** *Aging (Albany NY)* 2020, 12(11):10317-10336.
12. Liu LL, Zhu LB, Zheng JN, Bi TD, Ma JF, Wang LN, et al. **Development and assessment of a predictive nomogram for the progression of IgA nephropathy.** *Sci Rep* 2018, 8(1):7309.
13. Levey AS, Stevens LA, Schmid CH, Zhang YL, Castro AF, 3rd, Feldman HI, et al. **A new equation to estimate glomerular filtration rate.** *Ann Intern Med* 2009, 150(9):604-612.
14. Trimarchi H, Barratt J, Cattran DC, Cook HT, Coppo R, Haas M, et al. **Oxford Classification of IgA nephropathy 2016: an update from the IgA Nephropathy Classification Working Group.** *Kidney Int* 2017, 91(5):1014-1021.
15. Kramer AA, Zimmerman JE. **Assessing the calibration of mortality benchmarks in critical care: The Hosmer-Lemeshow test revisited.** *Crit Care Med* 2007, 35(9):2052-2056.
16. Fitzgerald M, Saville BR, Lewis RJ. **Decision curve analysis.** *Jama* 2015, 313(4):409-410.
17. Maixnerova D, Reily C, Bian Q, Neprasova M, Novak J, Tesar V. **Markers for the progression of IgA nephropathy.** *J Nephrol* 2016, 29(4):535-541.
18. Sakai H, Abe K, Kobayashi Y, Koyama A, Shigematsu H, Harada T, et al. **Clinical guidelines of IgA nephropathy.** *Nihon Jinzo Gakkai shi* 1995, 37(8):417-421.
19. Tomino Y, Suzuki S, Imai H, Saito T, Kawamura T, Yorioka N, et al. **Measurement of serum IgA and C3 may predict the diagnosis of patients with IgA nephropathy prior to renal biopsy.** *J Clin Lab Anal* 2000, 14(5):220-223.
20. Yanagawa H, Suzuki H, Suzuki Y, Kiryluk K, Gharavi AG, Matsuoka K, et al. **A panel of serum biomarkers differentiates IgA nephropathy from other renal diseases.** *PLoS One* 2014, 9(5):e98081.
21. Ishiguro C, Yaguchi Y, Funabiki K, Horikoshi S, Shirato I, Tomino Y. **Serum IgA/C3 ratio may predict diagnosis and prognostic grading in patients with IgA nephropathy.** *Nephron* 2002, 91(4):755-758.
22. Lang Y, Song S, Zhao L, Yang Y, Liu T, Shen Y, et al. **Serum IgA/C3 ratio and glomerular C3 staining predict progression of IgA nephropathy in children.** *Translational pediatrics* 2021, 10(3):666-672.
23. Coppo R, Amore A. **Aberrant glycosylation in IgA nephropathy (IgAN).** *Kidney Int* 2004, 65(5):1544-1547.

24. Suzuki H. **Biomarkers for IgA nephropathy on the basis of multi-hit pathogenesis.** *Clin Exp Nephrol* 2019, 23(1):26-31.
25. Zhang K, Li Q, Zhang Y, Shang W, Wei L, Li H, et al. **Clinical Significance of Galactose-Deficient IgA1 by KM55 in Patients with IgA Nephropathy.** *Kidney Blood Press Res* 2019, 44(5):1196-1206.
26. Moldoveanu Z, Wyatt RJ, Lee JY, Tomana M, Julian BA, Mestecky J, et al. **Patients with IgA nephropathy have increased serum galactose-deficient IgA1 levels.** *Kidney Int* 2007, 71(11):1148-1154.
27. Tang F, Hu H, Xu R, Tao C, Wan Q. **Association Between Serum IgG Concentrations and Prognosis in IgA Nephropathy.** *Iran J Kidney Dis* 2020, 14(6):454-462.
28. Peng S, Lu W, Jiang X, Xu X, Wu Y. **IgG deposits in the mesangium and capillary loops predict poor renal outcome in patients with IgA nephropathy: a single-center retrospective study.** *Ren Fail* 2020, 42(1):902-911.

Figures

Training cohorts

Independent test cohorts

Figure 1

Figure 2

Logistic model predicting IgAN. The model was developed in the training dataset, with age, serum IgA, serum albumin, serum phosphorus, and 24hUpro incorporated. Points of each variable were acquired by drawing a straight line upward from the corresponding value to the “Points” line. Then sum the points received from each variable and located the number on the “Total Points” axis. To conclude the patient’s sort of probability for having IgAN, draw a straight line down to the corresponding “Probability of IgAN” axis.

Figure 3

ROC curve and calibration plot for the diagnostic model of IgAN. (a) ROC curve. It displayed that the AUC of the model for prediction of IgAN was 0.89, 0.88 and 0.82 in the training set, validation set and independent test cohorts, respectively. The x-axis represents false positive rate. The y-axis represents true positive rate. (b) Calibration plot. The x-axis represents the predicted probability of IgAN. The y-axis represents the actual diagnosed IgAN. The dashed diagonal line represents a perfect prediction of an ideal model. The solid line represents the performance of the model, and the closer the fitting with the dashed diagonal line is, the better the prediction effect will be. Train, training set, val, validation set, test, independent test cohorts.

Figure 4

Decision curve analysis (DCA) for the model. The x-axis represents risk threshold of IgAN. The y-axis represents the net benefit. The red, green and blue solid line represents the model for prediction of IgAN in the training set, validation set and independent test cohorts, respectively. The solid grey line represents non-IgAN patients. The solid black line represents IgAN patients.