Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Association of Clinicopathological Features with IgA Nephropathy: A Principal Component Analysis

Mei Li
Second Xiangya Hospital
Liang Peng
Second Xiangya Hospital
Li-Li Wan
Second Xiangya Hospital
Hai-Yang Liu
Second Xiangya Hospital
Ming Xia
Second Xiangya Hospital
Ling-Zhi Wu
Second Xiangya Hospital
Yan Li
second xiangya hospital
Chang Wang
Second Xiangya Hospital
Guo-Chun Chen
second xiangya hospital
Yu Liu
second xiangya hospital
Hong Liu (liuhong618@csu.edu.cn)
Second Xiangya Hospital of central south university https://orcid.org/0000-0001-6040-5323

Research article

Keywords: IgAN, clinical and pathological features, principal component analysis

Posted Date: November 10th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-103691/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Abstract

Background: IgA nephropathy(IgAN) is the leading form of glomerular disease worldwide. Currently, the pathogenesis of IgAN is unclear and IgAN can only be diagnosed by renal biopsy, which lacks non-invasive methods. This study aims to analyze the association between clinicopathological characteristics and IgAN by principal component analysis.

Methods: Based on a combination of z-test and PCA, the fit of this model was evaluated with logistic regression analyses.

Results: Data from 847 patients with biopsy-proven IgAN in 1395 cases from May 2008 to April 2013 were analyzed. the average age is 33.156±12.308 years old and males account for 43%. Z-test selected 27 clinical and pathological indicators related to IgAN, and the principal component prediction model was established based on these 27 indicators. Logistic regression model providing 91.93% IgAN renal recall rate and 71.29% overall accuracy, which shows that the PCA model has high reliability.

Conclusions: As the model result shows, the higher level spheroid hardening rate, serum creatinine, blood uric acid and lower level eGFR might promote the occurrence of IgAN, which also provides more information for non-invasive diagnosis of IgAN patients.

Background

IgAN is a common primary chronic glomerular disease associated with end-stage renal disease[1]. The pathogenic mechanism of IgAN is not clear. It is reported that IgAN eventually results in end-stage renal disease in approximately 30% of patients within 20 years of diagnosis[2]. However, the requirement for diagnostic kidney biopsies has prevented the description of the full consequences of this disease[3]. It is valuable to use non-invasive methods to facilitate early diagnosis and risk prediction in patients with IgAN.

A number of risk factors affecting the prognosis of IgAN have been identified, including hypertension, higher proteinuria, decreased glomerular filtration rate (GFR), hyperuricemia, sex, and severe pathological score, which have been used to build various scoring systems to predict the prognosis of IgAN [2, 4]. However, these scoring systems are limited by the following factors: small sample size, multiple pathological scoring criteria, relatively few variables included, and poor clinical applicability[4, 5, 6]. Machine learning better identify variables related to kidney outcomes, predictive performance and to learn from multiple modules of data compared with conventional statistical methods and establishing accurate prediction models by machine learning has begun to be applied in medicine[6, 7]. Recently, a model have been developed to assist in predicting risk for kidney disease progression in IgAN patients based on the combination of survival analysis and machine learning algorithms[8], Although the abovementioned model show better prediction power than the absolute renal risk and may be a more beneficial tool to enhance individualized treatment and management of IgAN patients, it has its limitations such as significant effects of different therapeutic interventions on prognosis.

To alleviate such difficulties, we proposes a new method for risk factor prediction and diagnostic decision based on the principal component analysis (PCA) model [9] ,which based on the results of multivariate analysis. In most cases, predictive factors of disease occurrence (such as hypertension and proteinuria) are interrelated, and it is not easily to assess their individual contribution to overall risk from a statistical perspective. PCA model can address this problem by reducing dimensions [10], from which we can obtain correlation coefficients for each factor reflecting their respective contributions. Logistic regression analysis is used to test the fitness of the PCA model. Based on the above methods, we try to explore the predictive value of clinical pathological parameters for the risk factors of IgAN, so as to lay the foundation for the establishment of a more suitable prediction and diagnostic model for IgAN patients.

Methods

Study Population.

Data from 1395 patients undergone kidney biopsy from May 2008 to April 2013 were first screened from the Department of Nephrology at the Second Xiangya Hospital. Such cases like secondary causes of mesangial IgA deposits, lack of pathological results or computer errors occurred were excluded. 1042 patients data were finally included in the analysis, eventually 847 cases were confirmed as IgAN, and the remaining 195 data with Other types of kidney disease were in the non-IgAN ggroup. All studies were conducted in accordance with the guiding principles of the Helsinki Declaration and the approval of the ethics committee of Xiangya Medical College of Central South University.

Evaluation of clinicopathological parameters

Clinical data obtained during hospitalization of patients undergoing kidney biopsy. Urine samples were collected for 24 hours to detect urine protein excretion. Reference standards for semi-quantitative scoring of pathological indicators of glomeruli, tubules-interstitial, and renal vessels [11]: mesangial proliferation (0–4 points), interstitial fibrosis (0–3 points), and renal tubular atrophy (0-3points). Renal insufficiency, gross hematuria, and Interstitial inflammatory lesions were assessed by the presence(rated 1 point) or absence (rated 0 point) of such lesions. Grades of tonsil enlargement are I to III,no tonsil enlargement is recorded as 0 points, grade I is recorded as 1 point, unilateral Grade II was recorded as 2 points, bilateral Grade II was recorded as 3 points, and Grade III was recorded as 4 points.

Principal Component Analysis

PCA is a linear dimensionality reduction technique that reduces a set of potentially correlated variables(p assumed) to fewer variables that still contains most of the original information, these fewer variables also called principal component (PC) that are linearly uncorrelated. principles and steps of PCA are reported as previous[9]. In short, PCA looks for linear combinations of variables in order to extract the maximum variance. PCA then removes this variance and seeks a second linear combination that accounts for the largest proportion of the remaining variance. there may be a third, a fourth, ..., n th linear combination. In addition, it involves calculating the eigenvalues and eigenvectors of the covariance

matrix, then sorting these eigenvectors in descending order of eigenvalues, and finally projecting the actual data in the direction of the eigenvectors. Usually the mathematical processing is to linearly combine the original P indicators as a new comprehensive indicator. The most classic approach is to use the variance of PC1 (the first linear combination selected, that is, the first comprehensive indicator) to express it, that is, the larger Va (PC1), the more information PC1 contains.Therefore,PC1 selected in all linear combinations should have the largest variance, so PC1 is called the first principal component. If the first principal component is not sufficient to represent the information of the original P indicators, Then select PC2, which is the second linear combination. In order to effectively reflect the original information, the existing information of PC1 does not need to appear in PC2 again. To express in mathematical language requires Cov (PC1, PC2) = 0, then PC2 is The second principal component, and so on, can be used to construct the third, fourth, ..., p-th principal component.

Statistical Analysis.

Datas were expressed as mean ± standard deviation(SD) analyzed using python version 3.6.3. the z test function was first used to calculate the difference between the biochemical indicators of the IgA group (including all subtypes group) and control group. Logistic regression analysis is used to test the fitness of the PCA model. P value < 0.05 was considered as statistically significant.

Results

Clinical and pathological characteristics of the population

847 patients with IgAN were enrolled in this study(Table 1). We can conclude from the statistical data that the average age is 33 years old (13 to 73 years old), and males account for 43%. The average estimated glomerular filtration rate (eGFR) is 95.59 ml / min. The eGFR value of normal people is about 125 ml / min in the analysis of related data, so the average level of eGFR is in the normal range, but the variance is 95.83, indicating that the eGFR of patients with kidney disease There are large fluctuations. According to statistics, 96 patients were below 50 ml / min, and 33 patients were above 150 ml / min.The average value of urinary protein is 1.58 g / 24 h, the statistical difference of urinary protein is very large, the lowest is 0.0028 g / 24 h, and the highest is 16.331 g / 24 h. We divided the tonsil abnormalities into five grades (0, 1, 2, 3, 4) with an average score of 0.5. The average serum albumin content is 35.47 g / L, and the average blood urea nitrogen content is 9.20 mmol / L (normal value is 40–55 g / L).

Normal values of serum creatinine are different in different hospitals. Generally speaking, the standard value of normal serum creatinine is: 44–133 μmol / L. When the blood creatinine exceeds 133 μmol / L, it means that kidney damage has occurred, and renal insufficiency and renal failure have already occurred. In the statistical data, the average blood creatinine content is 101.11umol / L, the lowest is 3.35umol / L, and the highest is 1333.0umol / L. Therefore, the average renal function of the patients in the statistics is normal, but some patients have symptoms such as uremia. The mean and variance of other related factors are shown in Table 1.

Table 1
The demographic, clinical, laboratory data. Abbreviations: MAP: mean arterial pressure, eGFR: estimated glomerular filtration rate, BUN: blood urea nitrogen.

clinicopathological features	Effective statistics(N)	statistical data
renal insufficiency(%)	847	0.150 ± 0.360
age(years old)	847	33.156 ± 12.308
body weight(kg)	818	58.912 ± 11.154
systolic pressure(mmHg)	847	126.947 ± 20.357
diastolic pressure(mmHg)	847	80.170 ± 12.438
gross hematuria	847	0.145 ± 0.352
tonsil abnormalities	847	0.501 ± 0.801
serum creatinine(umol/L)	798	101.110 ± 101.674
eGFR	773	95.590 ± 95.834
BUN(mmol/L)	796	9.209 ± 31.479
serum uric acid (umol/L)	785	346.797 ± 113.470
blood triglyceride(mmol/L)	721	1.916 ± 1.787
total cholesterol(mmol/L)	721	5.638 ± 2.832
serum IgA(g/L)	665	2.562 ± 1.118
serum IgM(g/L)	664	1.891 ± 11.676
serum C3(g/L)	691	1.021 ± 0.314
serum C4(g/L)	673	0.243 ± 0.116
urine protein(g/24h)	734	1.586 ± 2.367
number of glomeruli under light microscope	847	15.949 ± 7.488
spherical hardening number	847	2.167 ± 2.725
spheroid hardening rate(%)	847	11.723 ± 16.724
segment hardening number	847	1.401 ± 1.668
segment hardening rate(%)	847	8.228 ± 11.874
total crescent formation	847	0.422 ± 1.052

clinicopathological features	Effective statistics(N)	statistical data
cell crescent	847	0.129 ± 0.492
cell fiber crescent	847	0.207 ± 0.796
fiber crescent	847	0.086 ± 0.477
glomerular adhesion	847	0.194 ± 0.656
mesangial cells and mesangial matrix hyperplasia	847	2.375 ± 0.819
basement membrane condition	847	1.407 ± 2.402
capillary cavity opening(1,2,3)	847	1.815 ± 1.136
capillary endothelial hyperplasia	847	0.129 ± 0.339
tubule atrophy(0,1,2,3)	847	1.279 ± 0.835
interstitial fibrosis(0,1,2,3)	847	1.290 ± 0.830
whether there is inflammatory cell infiltration in renal interstitial	847	0.761 ± 0.426

Correlation analysis of various biochemical characteristics and IgAN

In order to study the correlation between biochemical characteristics and IgAN, this study intends to use a z-test to perform a bivariate analysis to compare the IgAN group with the healthy group. According to the statistical values in Table 2 below, the following 27 indicators have significant differences in the data of the sick and healthy groups: serum IgA, serum C3, urinary protein quantification, the number of glomeruli under light microscope, the number of sclerosis, spheric sclerosis rate, number of segment sclerosis, total number of crescent formation, number of cell crescent, number of cell fiber crescent, number of fiber crescent, number of glomerular adhesion, mesangial cell and mesangium stromal hyperplasia, basement membrane condition, capillary cavity opening, tubule atrophy, interstitial fibrosis, renal stromal infiltration of inflammatory cells, serum creatinine, tonsil abnormality, gross hematuria history, systolic blood pressure, age, renal insufficiency, eGFR, blood urea nitrogen and blood uric acid. That is, the above factors have an impact on IgAN. Whereas serum IgM, serum C4, segment sclerosis rate ,capillary endothelial hyperplasia, total cholesterol, body weight, gender, blood triglycerides and diastolic blood pressure were not significantly different between the IgAN group and the healthy group.

Table 2

the correlation between biochemical characteristics and IgAN. P value < 0.05 or z value > 1.96 was considered as significant, p value < 0.01 or z value > 2.58 was considered as very significant. Abbreviations:eGFR: estimated glomerular filtration rate, BUN: blood urea nitrogen.

clinicopathological features	z	p	significance
serum IgA(g/L)	4.624	0.000	very
serum C3(g/L)	4.399	0.000	very
urinary protein(g/24h)	5.763	0.000	very
number of glomeruli under light microscope	5.697	0.000	very
spherical hardening number	5.011	0.000	very
spheroid hardening rate(%)	6.466	0.000	very
segment hardening number	3.971	0.000	very
total crescent formation	7.945	0.000	very
cell crescent	4.707	0.000	very
cell fiber crescent	5.243	0.000	very
fiber crescent	4.713	0.000	very
glomerular adhesion	1.974	0.048	yes
mesangial cells and mesangial matrix hyperplasia	4.836	0.000	very
basement membrane condition	8.916	0.000	very
capillary cavity opening(1,2,3)	8.817	0.000	very
tubule atrophy(0,1,2,3)	6.561	0.000	very
interstitial fibrosis(0,1,2,3)	6.441	0.000	very
whether there is inflammatory cell infiltration in renal interstitial	3.225	0.001	very
serum creatinine(umol/L)	13.33	0.000	very
tonsil abnormalities	8.032	0.000	very
gross hematuria	6.269	0.000	very
systolic pressure(mmHg)	4.578	0.000	very
age(years old)	7.061	0.000	very
renal insufficiency	20.91	0.000	very
eGFR	4.868	0.000	very
BUN(mmol/L)	2.612	0.008	very

clinicopathological features	z	p	significance
blood uric acid(umol/L)	9.276	0.000	very

Performance of the PCA model

We extracted the factors mentioned in Table2 that are obviously related to IgAN, and then eliminated all the computer error rows, so that the original total of 1395 pieces of data became only 1042 pieces of data, which were divided into IgAN and non-IgAN groups According to the results of kidney biopsy. To better diagnose IgAN, we use principal component analysis to extract the most important n principal components for the next step in constructing a judgment model. In order to determine the number of dimensionality reductions, we draw the number of dimensions and the variance chart of all components (Fig. 1). When the number of reduced dimensions is 20, the sum of the variances of all components is 90%, that is, about 10% of the information is lost. We can observe that when the dimension is equal to 3, the variance of all components is close to 100%, that is, about 2% of the information is lost, and this loss is within our ideal range.

The variance contribution rate of each principal component is 59.25%, 20.44%, 18.14%. The cumulative contribution rate is 97.83% (Fig. 2), so these 3 principal components can represent 97.83% of the information for the judgment of IgAN based on biochemical indicators. According to the indicators with the largest absolute value of each principal component coefficient, 3 representative indicators can be selected instead of 27 indicators. From the absolute value of the data in Table 3, the the biochemical parameters (also called variables) that determine the size of the PC1 are serum creatinine, blood uric acid and eGFR. Judging from the sign of the data, higher level of serum creatinine, blood uric acid and lower level of eGFR could promote the occurrence of IgAN. The biochemical parameters that determine the size of the remaining four main components are as follows: PC2: blood uric acid ; PC3: eGFR. The datas indicate that abnormalities in serum creatinine, blood uric acid and eGFR typically reflect an increased risk of IgAN occurrence.

Logistic regression analysis

Based on the sample data of the 27 indicators mentioned, the principal component analysis (PCA) was used to reduce the dimension of the index data. The dimension-reduced data was randomly divided into 80% of the data as the training set and 20% of the data as the test set. A logistic algorithm was used to construct a diagnosis model of kidney disease, and the results were as follows: The number of data was reduced to 3 dimensions. The accuracy, recall, and accuracy were used to evaluate the model. The results are shown in Fig. 2. A total of 209 cases were selected as the test set, including 124 cases from IgAN patients and 85 cases from control group. A total of 209 cases were selected as the test set, including 124 cases of IgA kidney patients and 85 cases from healthy control groups, of which 114 were correctly judged as IgAN with a recall rate of 91.93%. However, only 35 patients were correctly judged as

non-IgA patients, so the overall accuracy of the model is 71.29%.The above data shows that the PCA predictive model has a good fitness.

Table 3
Component matrix. the larger the absolute value of the variable, the greater the contribution to the principal component. Abbreviations:eGFR: estimated glomerular filtration rate, BUN: blood urea nitrogen.

clinicopathological features	PC1	PC2	PC3
serum IgA(g/L)	0.000	0.000	-0.000
serum C3(g/L)	0.000	0.000	0.000
urinary protein quantification(g/24h)	0.004	0.003	-0.001
glomeruli under light microscope number	-0.007	-0.000	0.002
spherical hardening number	0.003	0.002	-0.002
spheroid hardening rate(%)	0.029	0.009	-0.012
segment hardening number	0.000	0.001	-0.001
total crescent formation	0.005	-0.003	0.001
total crescent formation	0.001	-0.000	-0.000
cell fiber crescent	0.002	-0.002	0.001
fiber crescent	0.001	-0.001	0.000
glomerular adhesion	0.000	0.000	0.000
mesangial cells and mesangial matrix hyperplasia	0.000	0.001	-0.000
basement membrane condition	0.004	0.001	-0.001
capillary cavity opening(1,2,3)	0.002	0.001	-0.001
tubule atrophy(0,1,2,3)	0.001	0.001	-0.001
interstitial fibrosis(0,1,2,3)	0.001	0.001	-0.001
whether there is inflammatory cell infiltration in renal interstitial	0.000	0.000	0.000
serum creatinine(umol/L)	**0.881**	-0.365	0.298
tonsil abnormalities	-0.000	0.000	0.000
gross hematuria	-0.000	-0.000	0.000
systolic pressure(mmHg)	0.036	0.014	-0.006
age(years old)	0.016	-0.009	-0.012
renal insufficiency	0.002	0.001	-0.000
eGFR	**-0.221**	0.240	**0.944**
BUN(mmol/L)	0.022	-0.010	0.041

clinicopathological features	PC1	PC2	PC3
blood uric acid(umol/L)	**0.415**	**0.899**	-0.130

Discussion

IgAN is the most common primary chronic glomerular disease and has a worldwide incidence exceeding 1.5 per 100000 persons per year[12]. However validated tools for predicting disease risk remain limited[13]. The risk factors of IgAN may include both clinical and pathological aspects. Studies in recent decades have found some clinical indicators related to renal outcome. In this study, we constructed and analyzed a risk-prediction model of IgAN. We first compared the clinical and pathological indicators of the IgAN group and the control group, and then selected indicators with significant differences to build a PCA model, and then tested the fitness of this model using logistic regression analysis.The new model is superior to the original model used to predict IgAN because it has the advantages of eliminating the correlation between the evaluation indicators and reflecting the contribution rate through objective and reasonable coefficients of each principal component. The test of regression analysis further proves the importance of these factors.

Establishing a good and widely accepted risk prediction and diagnosis model can help inform prevention and diagnosis to patients [14, 15], which can help clinicians make decisions about precise treatment and follow-up. In addition, predictive factors in risk prediction models can increase the importance and awareness of these factors during a medical examination.

Based on 27 pathological and biochemical indicators with obvious factors for IgAN, 3 indicators (serum creatinine, blood uric acid, eGFR, BUN) were selected as the final evaluation factors by principal component analysis. The results show that the higher level spheroid hardening rate, serum creatinine and lower level eGFR might promote the occurrence of IgAN,which is helpful for the non-invasive diagnosis and risk prediction of IgAN at the same time. According to the results of regression analysis, the IgAN renal recall rate was 91.93%, and the overall accuracy was 71.29%. It shows that the PCA prediction model has high reliability. Although proteinuria is one of the most important risk factors recognized by IgAN[2, 16, 17], there are two data in our data that directly detect proteinuria, which are urine protein quantification (g / 24 h) and protein qualitative. However, our protein qualitative data is seriously insufficient and may affect the model results. According to experience, the onset of gross hematuria is directly related to the number of urinary red blood cells, urine microscopy, and urine occult blood. Our data has counted these four detailed data, but the insufficient number of routine urine microscopy and routine occult blood records may also affect the model results. This is also the shortcoming of this study.

In summary, we explored the predictive value of the clinicopathological parameters of IgAN using the PCA model, which provides evidence for the prevention and non-invasive diagnosis of IgAN.

Abbreviations

IgAN：IgA nephropathy； GFR：decreased glomerular filtration rate；PCA： principal component analysis； PC： principal component；SD：standard deviation；eGFR: estimated glomerular filtration rate; BUN: blood urea nitrogen.

Declarations

Acknowledgment

The authors would like to acknowledge all the study participants and research assistants who were involved in the data collection.

Authors' Contribution

Hong Liu, Mei Li, Guo-Chun Chen, Chang wang and Yu Liu designed the study; Mei Li, Liang Peng ,Li-Li Wan, Hai-Yang Liu, Ming Xia,Yan Li and Ling-Zhi Wu performed the research;Mei Li wrote the paper.

Funding

Funding for the study was supported by National Natural Science Foundation of China (No. 81770714)

Availability of data and materials

All data related to this study are included in the manuscript and additional file.

Ethics approval and consent to participate

All studies were conducted in accordance with the guiding principles of the Helsinki Declaration and the approval of the ethics committee of Xiangya Medical College of Central South University. Study participants were informed clearly about their freedom to opt out of the study at any point of time without justifying for doing so.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Lai, K. N.; Tang, S. C.; Schena, F. P.; Novak, J.; Tomino, Y.; Fogo, A. B.; Glassock, R. J. IgAN. *Nat Rev Dis Primers* **2016**, *2*, 16001.
2. Le W; Liang, S.; Hu, Y.; Deng, K.; Bao, H.; Zeng, C.; Liu, Z. Long-term renal survival and related risk factors in patients with IgAN: results from a cohort of 1155 cases in a Chinese adult population.

Nephrol Dial Transplant **2012**, *27*, 1479-1485.

3. Wyatt, R. J.; Julian, B. A. IgAN. *N Engl J Med* **2013**, *368*, 2402-2414.
4. Goto, M.; Wakai, K.; Kawamura, T.; Ando, M.; Endoh, M.; Tomino, Y. A scoring system to predict renal outcome in IgAN: a nationwide 10-year prospective cohort study. *Nephrol Dial Transplant* **2009**, *24*, 3068-3074.
5. Berthoux, F.; Mohey, H.; Laurent, B.; Mariat, C.; Afiani, A.; Thibaudin, L. Predicting the risk for dialysis or death in IgAN. *Am. Soc. Nephrol.* **2011**, *22*, 752-761.
6. Pesce, F.; Diciolla, M.; Binetti, G.; Naso, D.; Ostuni, V. C.; Di Noia, T.; Vagane, A. M.; Bjorneklett, R.; Suzuki, H.; Tomino, Y.et al. Clinical decision support system for end-stage kidney disease risk estimation in IgAN patients. *Nephrol Dial Transplant* **2016**, *31*, 80-86.
7. Beam, A. L.; Kohane, I. S. Big Data and Machine Learning in Health Care. *JAMA* **2018**, *319*, 1317-1318.
8. Chen, T.; Li, X.; Li, Y.; Xia, E.; Qin, Y.; Liang, S.; Xu, F.; Liang, D.; Zeng, C.; Liu, Z. Prediction and Risk Stratification of Kidney Outcomes in IgAN. *J. Kidney Dis.* **2019**, *74*, 300-309.
9. Jolliffe, I. T.; Cadima, J. Principal component analysis: a review and recent developments. *Philos Trans A Math Phys Eng Sci* **2016**, *374*, 20150202.
10. Zhang, Z.; Castello, A. Principal components analysis in clinical studies. *Ann Transl Med* **2017**, *5*, 351.
11. Lim, B. J.; Joo, D. J.; Kim, M. S.; Kim, Y. S.; Kim, S. I.; Kim, Y.; Huh, K. H.; Koh, M. J.; Jeong, H. J. Usefulness of Oxford classification in assessing immunoglobulin A nephropathy after transplantation. *Transplantation* **2013**, *95*, 1491-1497.
12. Donadio, J. V.; Grande, J. P. IgAN. *N Engl J Med* .**2002**, *347*, 738-748.
13. Barbour, S. J.; Coppo, R.; Zhang, H.; Liu, Z. H.; Suzuki, Y.; Matsuzaki, K.; Katafuchi, R.; Er, L.; Espino-Hernandez, G.; Kim, S. J.et al. Evaluating a New International Risk-Prediction Tool in IgAN. *JAMA Intern. Med.* **2019**, *179*, 942-952.
14. Yang, Y.; Luo, M.; Xiao, L.; Zhu, X. J.; Wang, C.; Fu, X.; Yuan, S. G.; Xiao, F.; Liu, H.; Dong, Z.et al. Exploration of pathological prediction of chronic kidney diseases by a novel theory of bi-directional probability. *Sci Rep* **2016**, *6*, 32151.
15. Coppo, R.; Fervenza, F. C. Persistent Microscopic Hematuria as a Risk Factor for Progression of IgAN: New Floodlight on a Nearly Forgotten Biomarker. *Am. Soc. Nephrol.* **2017**, *28*, 2831-2834.
16. Cravedi, P.; Remuzzi, G. Pathophysiology of proteinuria and its value as an outcome measure in chronic kidney disease. *Br J Clin Pharmacol* . **2013**, *76*, 516-523.
17. Reich, H. N.; Troyanov, S.; Scholey, J. W.; Cattran, D. C. Remission of proteinuria improves prognosis in IgAN. *Am. Soc. Nephrol.* **2007**, *18*, 3177-3183.

Figures

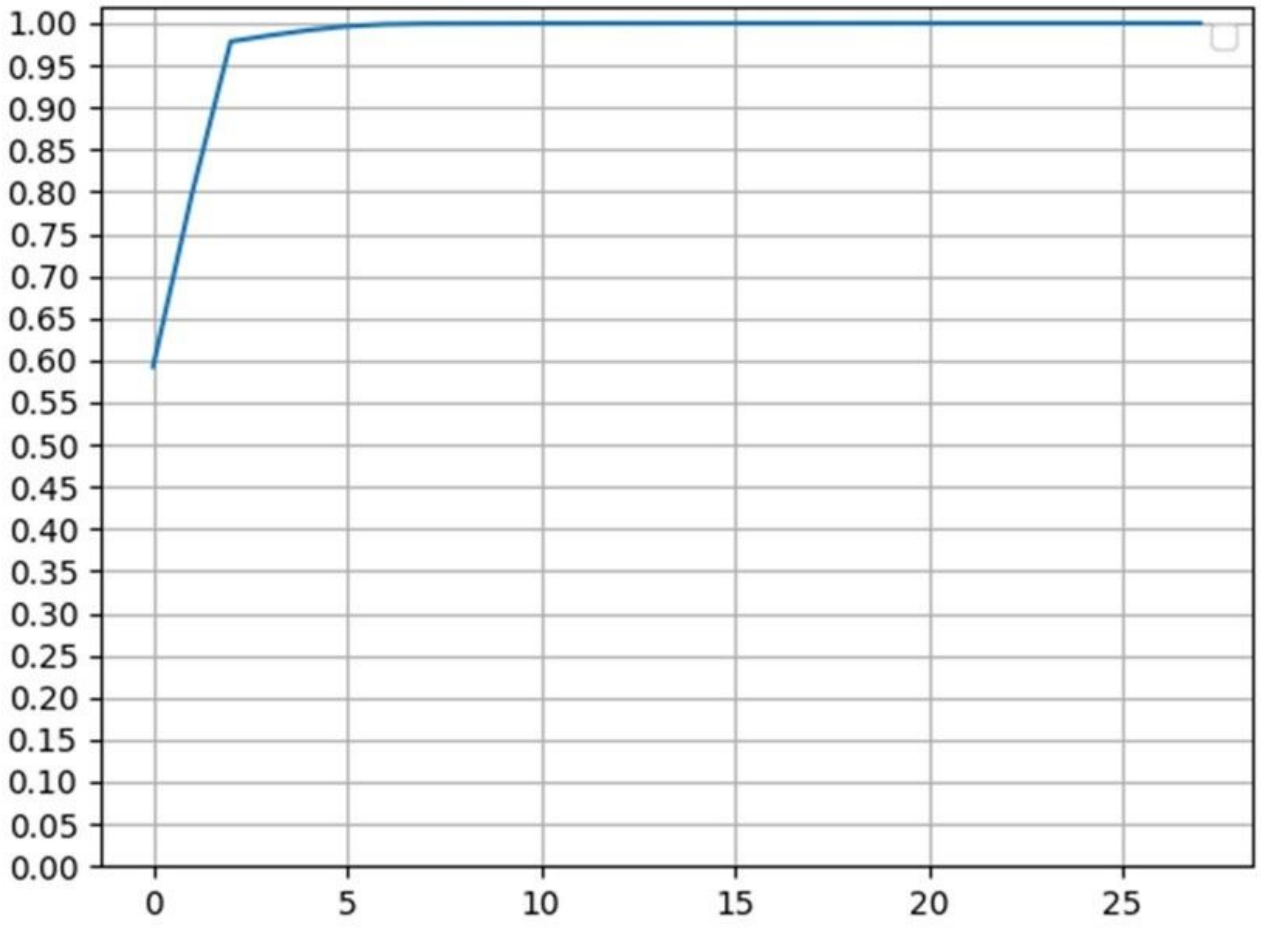

Figure 1

variance chart of all components.

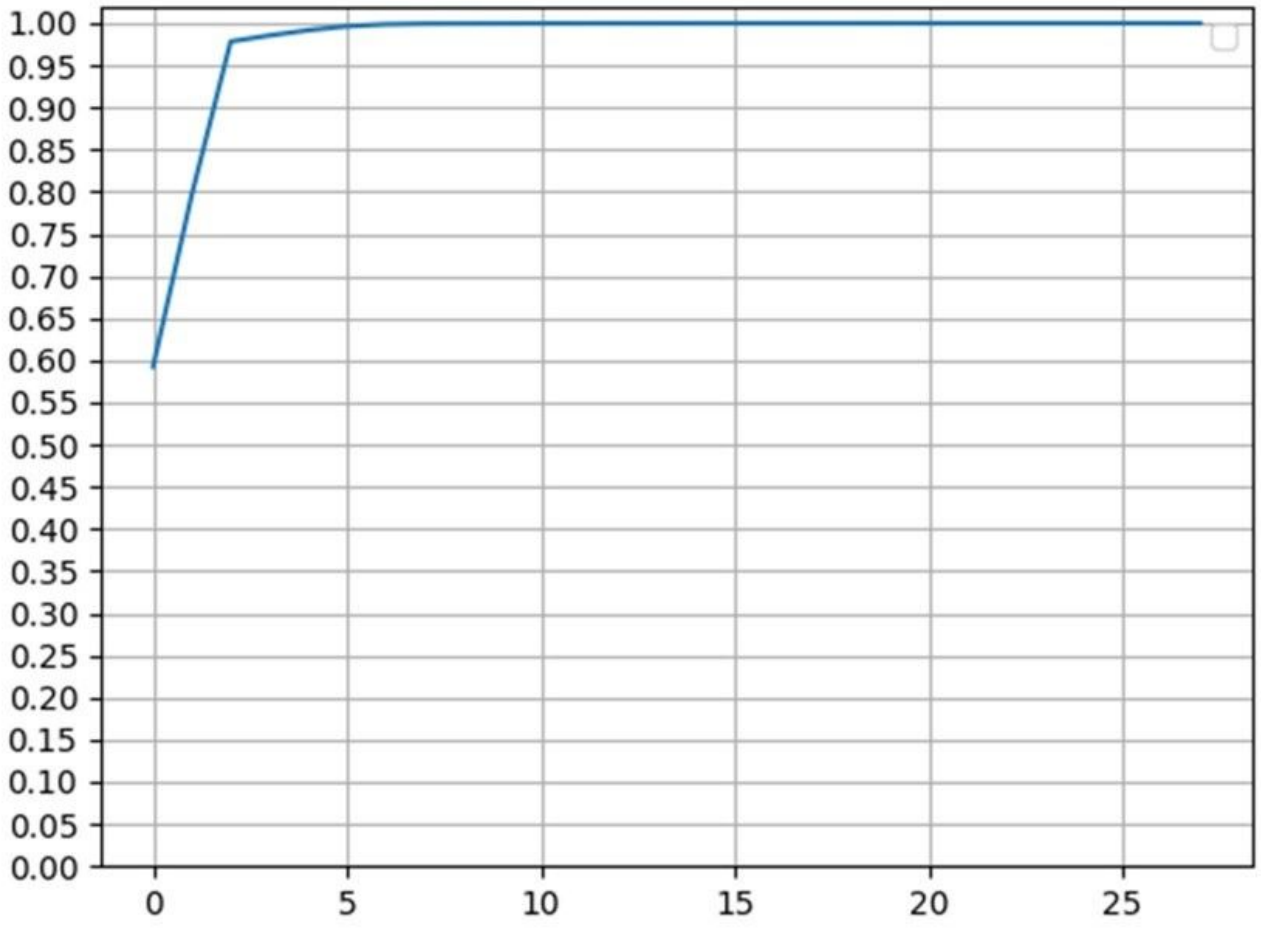

Figure 1

variance chart of all components.

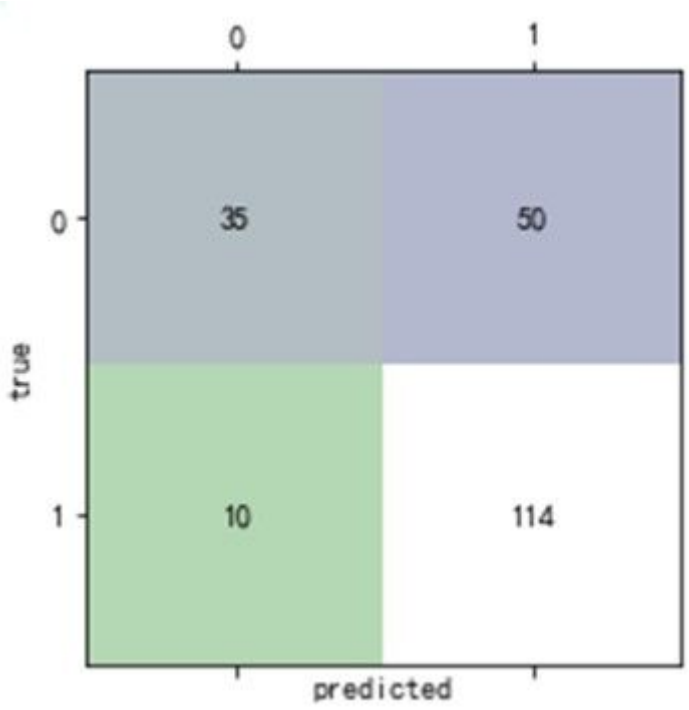

Figure 2

Logistic regression model confusion matrix.

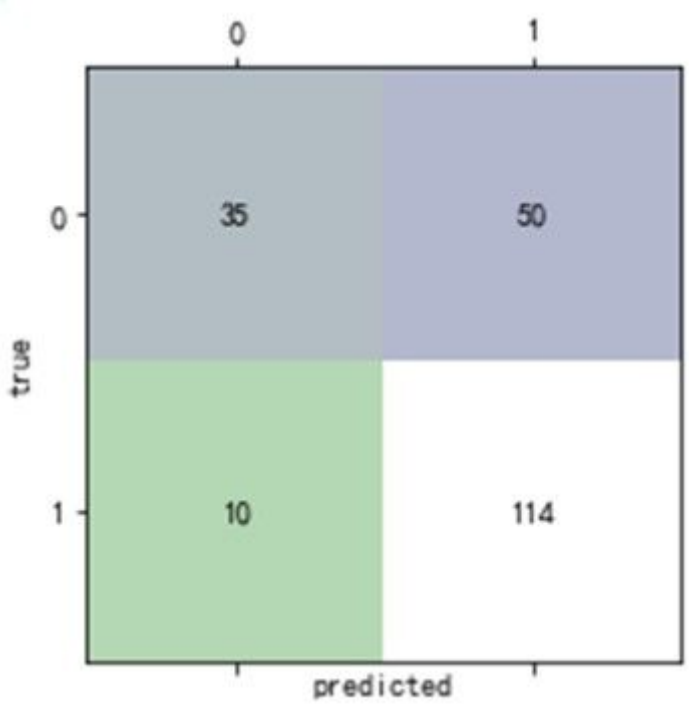

Figure 2

Logistic regression model confusion matrix.

Supplementary Files

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

- PCAdata.xlsx
- PCAdata.xlsx
- modeldata1.xlsx
- modeldata1.xlsx