Am J Clin Exp Urol 2026;14(1):1-16
www.ajceu.us /ISSN:2330-1910/AJCEU0170330

Original Article

A non-invasive predictive model for identifying non-diabetic kidney disease in type 2 diabetes mellitus: development and multicenter validation

Yuyan Yang[1*], Yijiang Song[2*], Pinning Feng[3], Xianlian Deng[4], Ya Li[1], Peijia Liu[1], Bin Peng[2], Yuanrui Liu[2], Youlin Liu[4], Jin Li[4], Peng Zhang[2], Feng Hu[1]

[1]Department of Nephrology, Shenzhen Third People's Hospital, The Second Affiliated Hospital of Southern University of Science and Technology, Shenzhen 518112, Guangdong, China; [2]Department of Laboratory Medicine, Nanfang Hospital, Southern Medical University, Guangzhou 510515, Guangdong, China; [3]Department of Laboratory Medicine, The First Affiliated Hospital of Sun Yat-sen University, No. 58 Zhongshan Second Road, Yuexiu District, Guangzhou 510080, Guangdong, China; [4]Shenzhen Mindray Biomedical Electronics Co., LTD., 9 Keji South 12 Road, Yuehai Street, Nanshan District, Shenzhen 518055, Guangdong, China. [*]Equal contributors.

Received November 26, 2025; Accepted January 18, 2026; Epub February 15, 2026; Published February 28, 2026

Abstract: Background: This study aimed to develop a non-invasive, simple, and rapid predictive model for identifying non-diabetic kidney disease (NDKD) in patients with type 2 diabetes mellitus (T2DM). Methods: We performed a retrospective analysis of clinical data from 117 T2DM patients who underwent renal biopsy at a single medical institution between 2017 and 2022; candidate variables were first prioritized based on clinical relevance, followed by the construction of a predictive framework using logistic regression. Dubbed the RICH model, the final framework integrated four key parameters: red blood cell (RBC) count, immunoglobulin A (IgA) level, cystatin C-derived estimated glomerular filtration rate (eGFR_2), and glycated hemoglobin A1c (HbA1c). Results: External validation was conducted across three independent centers involving 299 T2DM patients (2018-2024), achieving area under the receiver operating characteristic curve (AUC-ROC) values of 0.755, 0.764, and 0.755, which complemented the internal validation AUC-ROC of 0.847; at an optimal threshold probability of 0.559, approximately 20% of patients obtained clinical net benefit from the model, and notably, applying the RICH model for early NDKD screening has the potential to reduce the renal biopsy rate by 42.05%. Conclusions: The RICH model exhibits robust performance in predicting NDKD among T2DM patients with renal impairment, providing a practical tool for clinical decision-making.

Keywords: Type-2 diabetes mellitus, diabetic nephropathy, non-diabetic kidney disease, predictive model, diagnosis

Introduction

According to 2021 statistics, roughly 537 million adults (aged 20-79) globally were living with diabetes mellitus (DM). Within this global population, around 141 million patients were located in China - making up 26.2% of the world's total diabetic cases. Critically, more than 90% of diabetes cases in China are type 2 diabetes mellitus (T2DM) [1-3]. An analysis found that 21.8% of Chinese T2DM patients develop diabetic kidney disease (DKD), which stands as diabetes' most severe microvascular complication. In China, DKD is responsible for 30%-50% of all end-stage renal disease (ESRD) cases, solidifying its status as the top cause of ESRD [4, 5]. Alarmingly, the burden of DKD in China continues to rise each year: its incidence hit 101.6 per 100,000 people in 2015, marking a 2.7-fold surge from 2000 levels [6]. As a result, DKD has become a prominent public health issue that hinders both social progress and economic growth.

Diabetic nephropathy is a common disease in clinical practice. However, whether diabetes is the etiology of renal injury is a critical prerequisite affecting clinical diagnoses and treatment decisions [7, 8]. NDKD is completely different from DKD in pathogenesis, treatment method, and prognosis assessment. Furthermore, some studies have shown that most NDKD patients

https://doi.org/10.62347/UYEP2269

exhibit better prognoses and longer survival time after treatment [9, 10]. Therefore, accurate identification of whether diabetic patients with renal injury are DKD or NDKD patients is of critical importance for devising therapeutic plans and making prognoses for patients. As of today, renal biopsy remains a "golden standard" in the clinical identification of DKD and NDKD. However, renal biopsy provides no sensitive early diagnostic indicators, with the following limitations in its practical implementation: (1) Renal biopsy is invasive and presents some risks, with low cost-effectiveness and poor acceptance among some patients; (2) Renal biopsy can be only used in distinct indications, making it not a suitable method for all diabetic patients; (3) Pathological diagnosis involved in renal biopsy puts a high demand for professional skills and experiences of operators, with significant differences in diagnostic accuracy among hospitals at different levels; (4) Renal biopsy involves complex operational procedures, which generally limit the application of this method in basic-level hospitals. Therefore, under most circumstances, renal biopsy is only applied in those clinical cases with a strong suspicion of NDKD. In addition, there don't have official practical guidelines published regarding the timing of renal biopsy applications among DM patients [11]. This indicates that it's necessary to investigate a non-invasive, simple, and rapid method for early identification and diagnosis of diabetic patients with renal injury, thus providing guidance for timely and accurate diagnoses in clinics.

The application of machine learning in big data analysis has emerged as a key approach in the revolution of precision medicine [12-14], which is also involved in the field of DKD and NDKD identification and diagnosis [14-16]. Wang et al. conducted a study involving 132 diabetic patients who underwent renal biopsy at a single center [17]. However, their research focused on tubulointerstitial markers (e.g., Neutrophil Gelatinase-Associated Lipocalin and β2-microglobulin), with insufficient attention to routine clinical tests [17]. Zhao et al. developed a differential diagnosis risk scoring model, identifying diabetic retinopathy (DR), diabetes duration, eGFR, 24-hour urinary protein, and hematuria as independent risk factors for DKD - but the model was not visualized for clinical application [18]. What's more, previous studies are primarily single-center retrospective studies involving many model indicators, with some projects affected by such subjective factors as diabetes duration, the sudden appearance of massive proteinuria, and family history, which have limited the promotion and application of these models [19, 20]. Therefore, the development of a diagnostic model for NDKD that balances broad applicability and clinical utility is essential for advancing its translation into clinical practice.

To this end, this study was designed with a focus on generalizability, leveraging multicenter cohort data for analysis. To further enhance clinical applicability, we intend to incorporate objective laboratory indicators and convert the model into a visual nomogram, thereby maximizing its accessibility for clinicians, particularly those practicing in primary care settings. On this basis, the present study aims to develop a non-invasive predictive model for the differential diagnosis of renal injury in T2DM patients, with concurrent model visualization to be carried out. This user-friendly clinical decision support tool is anticipated to reduce unnecessary renal biopsies and improve the diagnostic and therapeutic efficacy for patients who cannot undergo renal biopsy due to various contraindications.

Material and methods

Clinical subjects

The three cohorts included in this study were respectively from Shenzhen Third People's Hospital, the First Affiliated Hospital of Sun Yat-sen University, and Nanfang Hospital of Southern Medical University. All of these are representative tertiary first-class hospitals in China, covering medical centers at different levels in the Pearl River Delta region. The details are as follows: (1) The cohort of Shenzhen Third People's Hospital involving 170 cases (including 67 DKD cases, 26 mixed-type kidney disease (MIX) cases and 77 NDKD cases); (2) The cohort of the First Affiliated Hospital of Sun Yat-sen University involving 78 cases (including 28 DKD cases, 9 MIX cases and 41 NDKD cases); (3) The cohort of Nanfang Hospital of Southern Medical University involving 168 cases (including 35 DKD cases, 13 MIX cases and 120 NDKD cases). This study was approved by the ethics committees of the participating hospi-

tals, and all patients who underwent renal biopsy provided written informed consent.

Modeling cohort: Adult T2DM patients who underwent renal biopsy due to renal injury complications at Shenzhen Third People's Hospital from January 2017 to May 2022 were collected and included in the modeling cohort of this study. The inclusion criteria are as follows: Patients clinically diagnosed with T2DM, who were ≥ 18 years old at the time of renal biopsy. The exclusion criteria are as follows: Patients with incomplete medical records or unclear medical materials; patients with complications such as renal calculus, severe infections, and malignant tumors; patients with systemic diseases such as systemic lupus erythematosus and vasculitis; patients who had undergone renal replacement therapy before renal biopsy. Routine optical microscopy was performed on all biopsy specimens. The measurement results obtained were interpreted by at least 2 renal pathologists, and at least 2 nephrologists were involved in making comprehensive clinical diagnoses based on the patients' medical history, symptoms, physical signs, and pathological test results. Any inconsistent results were resolved through discussions. According to the comprehensive clinical diagnosis results, 117 patients were finally included in the modeling cohort and divided into DKD, NDKD, and MIX groups. However, patients in the MIX group were excluded from the subsequent modeling analysis, with the remaining DKD and NDKD patients forming the final modeling cohort.

Validation cohort: Fifty-three patients who met the eligibility criteria at Shenzhen Third People's Hospital from June 2022 to September 2022 were collected and included in Validation Cohort 1. Among these patients, there were 24 DKD patients, 4 MIX patients, and 25 NDKD patients. Seventy-eight patients who met the eligibility criteria at the First Affiliated Hospital of Sun Yat-sen University from July 2023 to February 2024 were collected and included in Validation Cohort 2, including 28 DKD patients, 9 MIX patients, and 41 NDKD patients. A total of 168 patients who met the eligibility criteria at Nanfang Hospital of Southern Medical University from May 2018 to December 2020 were enrolled in Validation Cohort 3, including 35 DKD patients, 13 MIX patients, and 120 NDKD patients.

Study methods

Data collection: Risk variables incorporated in the model construction include gender, age, height, weight, body mass index (BMI), diabetes duration (Time), pulse pressure (PP), systolic blood pressure (SBP), diastolic blood pressure (DBP), uric acid (UA), urea, creatinine (Crea), Cystatin C (CysC), eGFR_1 (calculated using Crea), eGFR_2 (calculated using CysC), and eGFR_3 (calculated based on the combined formula of Crea and CysC) [21-23], IgA, high-density lipoprotein-cholesterol (HDL-C), low-density lipoprotein-cholesterol (LDL-C), albumin (ALB), total cholesterol (TC), triglyceride (TG), glucose (Glu), HbA1c, urine kappa (U_κ), urine lambda (U_λ), 24-hour urinary total protein (24 h UTP), RBC, Hemoglobin (HGB), and hematocrit (HCT). All variables were measured upon hospitalization of patients undergoing renal biopsy. Variables with a missing value rate of over 30% were excluded from the analysis, while those with a missing rate of < 30% were handled using the mean imputation method.

Indicator definition: The diagnosis of T2DM should comply with the standards formulated by the American Diabetes Association [24]. The diagnostic criterion for chronic kidney disease (CKD) is the presence of persistent renal structural damage and functional disorders caused by multiple factors for at least three months, or a consistent eGFR level below 60 mL/min/1.73 m^2 for three months without obvious causes [22]. Renal biopsy indicators listed in domestic and foreign studies and guidelines on diabetic kidney disease were adopted in this study [25]. The diagnosis of DKD is based on the pathological classification standard for DKD from the American Society of Nephrology [25-27], and the diagnosis of NDKD is based on the Mayo Clinic Classification Standards [28].

Model construction and validation

Variable screening and model construction: Data analyses were performed using SPSS version 27.0 and R software version 4.3.1. First, the Shapiro-Wilk test was conducted in SPSS 27.0 to assess the normality of continuous variables. For continuous variables with a normal distribution ($P > 0.05$), independent samples t-test was used for intergroup comparisons, and the results were expressed as mean ± standard deviation ($\bar{x} \pm SD$). For continuous vari-

ables without a normal distribution, the Mann-Whitney U nonparametric test was applied for intergroup analysis, with the results presented as median (interquartile range) [M (Q_1, Q_3)]. For categorical variables, the Pearson's chi-square test was adopted for intergroup comparisons when the theoretical frequency $T < 5$ and the total sample size $n \geq 40$, and the results were expressed as frequency (percentage). Variables with statistically significant differences ($P < 0.05$) in the univariate analysis were screened out for further investigation. Subsequently, the variables with significant results in the univariate analysis were imported into R software for further analysis. Combined with multivariate Logistic regression analysis (forward/backward stepwise method), the core variables to be finally included in the model were identified. Based on the screened variables, four parameters were ultimately incorporated to construct the predictive model, with the aim of balancing model performance, simplicity and objectivity without compromising predictive accuracy. The predictive model was constructed as a multivariate Logistic regression model using the glm function in R software. The occurrence of NDKD was defined as the dependent variable (Yes =1; No =0), while the screened core variables were set as independent variables. The regression coefficients, odds ratios (ORs) and 95% confidence intervals (95% CIs) of each variable were calculated.

Model performance validation: Model validation and evaluation were implemented in R software through the following steps: 1) Receiver Operating Characteristic (ROC) Curve Analysis: The ROC curve was plotted, and the area under the curve (AUC) was calculated to evaluate the model's discriminative ability for the outcome event (an AUC > 0.7 indicates good discriminative value, while an AUC > 0.8 denotes excellent discriminative ability). 2) Hosmer-Lemeshow Goodness-of-Fit Test: This test was performed to assess the model's fitting effect, where a $P > 0.05$ suggests a good model fit. 3) Calibration Curve Verification: The calibration curve was plotted and verified using the Bootstrap method (with 1000 resampling iterations) to evaluate the consistency between the predicted probabilities and the actual incidence rates of the model. 4) Brier Score Calculation: The Brier score (ranging from 0 to 1, with lower values indicating higher predictive accuracy) was computed to quantify the overall predictive error of the model. 5) Net Benefit Analysis and Nomogram Construction: Based on the final Logistic regression model, the ggplot2 package in R software was used to plot the net benefit analysis curve. Taking the risk probability of NDKD as the threshold variable, the net reduction rate of intervention under different thresholds was analyzed. Additionally, the rms package was applied to construct a nomogram for visualizing the model. The concordance index (C-index) was calculated to further verify the predictive efficacy of the nomogram (a C-index > 0.7 indicates favorable predictive performance), and the decision curve analysis (DCA) was plotted to evaluate the clinical net benefit of the model.

Results

Pathological and clinical baseline characteristics

Patient pathological characteristics: A total of 416 adult T2DM patients with renal injury were included in this study. Pathological analysis of renal biopsy specimens revealed that among the 117 patients in the modeling cohort, 43 were diagnosed with DKD (36.8%), 52 with NDKD (44.4%), and 22 with MIX (18.8%). Among all NDKD patients, minimal change disease (MCD) was the most common pathological type (25%), followed by membranous nephropathy (MN; 19.2%), IgA nephropathy (13.5%), and hypertensive nephropathy (13.5%). Among all MIX patients, diabetic kidney disease complicated with hypertensive kidney injury ranked as the most prevalent pathological type (59.1%), followed by MN (9.1%) and IgA nephropathy (9.1%). Detailed data are summarized in **Table 1**.

Clinical baseline characteristics of patients: In order to avoid the introduction of confounding factors, the MIX group was excluded from the model study. A total of 95 patients were ultimately included in the modeling cohort. Among them, 77 were male patients (81.1%), aged 46-56 years, with no statistically significant differences in age (P=0.500) and gender (P=0.938) between the NDKD and DKD groups.

The median diabetes duration across the entire cohort was 48 months, with the NDKD group demonstrating a significantly shorter median duration compared to the DKD group

Table 1. Pathological type distribution among NDKD and MIX patients in research cohorts [n (%)]

Pathological type	Modeling cohort			Validation cohort 1			Validation cohort 2			Validation cohort 3		
	MIX (n=22)	NDKD (n=52)	ALL (n=117)	MIX (n=4)	NDKD (n=25)	ALL (n=53)	MIX (n=9)	NDKD (n=41)	ALL (n=78)	MIX (n=13)	NDKD (n=120)	ALL (n=168)
Diabetic kidney disease			21 (17.9)			6 (11.3)			19 (24.4)			34 (20.2)
Diabetic nephropathy with ischemic or renal tubulointerstitial injury			22 (18.8)			18 (34.0)			9 (11.5)			1 (0.6)
Membranous nephropathy	2 (9.1)	10 (19.2)	12 (10.3)	2 (50.0)	2 (8.0)	4 (7.5)	4 (44.4)	18 (43.9)	22 (28.2)	4 (30.8)	62 (51.7)	66 (39.3)
IgA nephropathy	2 (9.1)	7 (13.5)	9 (7.7)	1 (25.0)	8 (32.0)	9 (17.0)	1 (11.1)	1 (2.4)	2 (2.6)	1 (7.7)	18 (15.0)	19 (11.3)
Hypertensive nephropathy	13 (59.1)	7 (13.5)	20 (17.1)	1 (25.0)		1 (1.9)	2 (22.2)	7 (17.1)	9 (11.5)	5 (38.5)		5 (3.0)
Minimal change disease	1 (4.5)	13 (25.0)	14 (12.0)		5 (20.0)	5 (9.4)		1 (2.4)	1 (1.3)		10 (8.3)	10 (6.0)
Focal segmental glomerulosclerosis		1 (1.9)	1 (0.9)				1 (11.1)	3 (7.3)	4 (5.1)		7 (5.8)	7 (4.2)
Proliferative glomerulonephritis	2 (9.1)		2 (1.7)				1 (11.1)	4 (9.8)	5 (6.4)		3 (2.5)	3 (1.8)
Mesangial proliferative glomerulonephritis	1 (4.5)	1 (1.9)	2 (1.7)		1 (4.0)	1 (1.9)		2 (4.9)	2 (2.6)			
Ischemic kidney injury		6 (11.5)	6 (5.1)		4 (16.0)	4 (7.5)						
Renal tubulointerstitial nephropathy		2 (3.8)	2 (1.7)		2 (8.0)	2 (3.8)		2 (4.9)	2 (2.6)	2 (15.4)	4 (3.3)	6 (3.6)
Hyperuricemia nephropathy with ischemic kidney injury		3 (5.8)	3 (2.6)		2 (8.0)	2 (3.8)						
Lupus nephritis											7 (5.8)	7 (4.2)
ANCA-associated vasculitis kidney injury										1 (7.7)		1 (0.6)
Henoch-Schönlein purpura nephritis		2 (3.8)	2 (1.7)									
Obesity-related glomerular hypertrophy											1 (0.8)	1 (0.6)
Amyloidosis nephropathy								1 (2.4)	1 (1.3)		4 (3.3)	4 (2.4)
Benign small artery sclerosis											3 (2.5)	3 (1.8)
Malignant small artery sclerosis											1 (0.8)	1 (0.6)
Lipoprotein glomerulopathy					1 (4.0)	1 (1.9)						
Glomerular podocyte lesions								2 (4.9)	2 (2.6)			

Abbreviations: NDKD: Nondiabetic kidney disease; MIX: Diabetic kidney disease combined with nondiabetic kidney disease.

(30 months vs. 84 months; $P < 0.001$). The overall median pulse pressure was 49 mmHg (interquartile range [IQR], 39-63 mmHg), and the NDKD group exhibited a marginally lower median pulse pressure than the DKD group (44 mmHg vs. 54 mmHg; $P=0.004$).

When compared to the NDKD cohort, the DKD group displayed markedly worse conventional renal function markers (urea: 7.70 mmol/L vs. 5.45 mmol/L; $P=0.004$; CREA: 125 μmol/L vs. 102 μmol/L; $P=0.019$; CysC: 1.663 mg/L vs. 1.130 mg/L; $P < 0.001$), more severe glucose metabolic dysregulation (Glu: 7.85 mmol/L vs. 6.19 mmol/L; $P=0.003$; HbA1c: 7.8% vs. 6.6%; $P=0.004$), and elevated proteinuria levels (U_κ: 66.08 mg/L vs. 50.91 mg/L; $P=0.008$; U_λ: 34.67 mg/L vs. 26.45 mg/L; $P=0.002$; 24 h UTP: 1,404.0 mg/L vs. 215.0 mg/L; $P=0.003$).

In contrast, the DKD group had significantly reduced eGFR (mL/min/1.73 m^2): eGFR_1 (52.01 vs. 69.81; $P=0.012$), eGFR_2 (38.08 vs. 62.74; $P < 0.001$), and eGFR_3 (37.63 vs. 60.25; $P=0.003$). The DKD group also showed lower IgA levels (2.28 g/L vs. 2.58 g/L; $P=0.038$) and notable anemia-associated abnormalities (RBC count: 3.44 × 10^{12}/L vs. 3.89 × 10^{12}/L; $P=0.001$; HGB: 120.21 g/L vs. 135.58 g/L; $P=0.002$; HCT: 35.79% vs. 40.24%; $P=0.001$) relative to the NDKD group.

No statistically meaningful differences were observed between the two groups in terms of height, weight, BMI, SBP, ALB, DBP, TC, TG, HDL-C, or LDL-C. Complete clinical baseline characteristics are compiled in **Table 2**.

Development of a predictive model for differential diagnosis

Univariate Logistic regression screening yielded 12 candidate variables for further assessment: Time, PP, UREA, CysC, eGFR_1, eGFR_2, IgA, Glu, HbA1c, RBC count, HGB, and HCT. Subsequent multivariate Logistic regression narrowed these down to four statistically significant predictors: RBC, IgA, eGFR_2 and HbA1c.

Additional screening and model fitting were performed on clinically relevant empirical indicators (BMI, eGFR_1, eGFR_2) using forward and backward stepwise selection methods, with detailed results summarized in **Table 3**. To balance model performance, simplicity, and objectivity - without compromising predictive accuracy - the final predictive model incorporated four parameters: RBC (OR=12.577; 95% confidence interval [CI]: 1.144-169.707; $P=0.043$), IgA (OR=2.232; 95% CI: 1.096-5.241; $P=0.040$), eGFR_2 (OR=1.056; 95% CI: 0.978-1.177; $P=0.211$), and HbA1c (OR=0.484; 95% CI: 0.275-0.749; $P=0.003$). The formula is: $P = (\exp(A))/(1 + \exp(A))$ ($A = 1.11921 \times RBC + 0.61602 \times IgA + 0.03180 \times eGFR\_2 - 0.63858*HbA1c - 2.30750$). This model is abbreviated as the "RICH" model.

For the modeling cohort, the RICH model yielded an AUC of 0.847 (95% CI: 0.766-0.929), with a cutoff value of 0.559. Its diagnostic performance included a sensitivity (Se) of 76.9%, specificity (Sp) of 83.7%, positive predictive value (PPV) of 85.1%, and negative predictive value (NPV) of 75.0% (**Table 4**). The Hosmer-Lemeshow test verified strong model calibration ($P=0.433$), indicating no statistically significant discrepancy between predicted and actual probabilities.

DCA - a key method for evaluating clinical utility - showed that the model delivered meaningful net clinical benefit for patients when the threshold probability fell between 0.10 and 0.80. At the optimal threshold of 0.559, roughly 20% of patients gained net benefit (**Figure 1A**). Notably, using the RICH model for early NDKD detection could potentially reduce the renal biopsy rate by 42.05% (**Figure 1B**).

We validated the RICH model using external validation cohorts, which showed favorable discrimination and calibration (results in **Figure 2**). As detailed in **Table 3**: Validation Cohort 1 had an AUC of 0.755 (95% CI: 0.611-0.899), with 76% Se, 75% Sp, 76% PPV, and 75% NPV; Validation Cohort 2 exhibited an AUC of 0.764 (95% CI: 0.654-0.874), with 66.0% Se, 75.0% Sp, 79.4% PPV, and 60.0% NPV; Validation Cohort 3 had an AUC of 0.755 (95% CI: 0.665-0.845), with 60.8% Se, 82.9% Sp, 92.4% PPV, and 38.2% NPV.

Hosmer-Lemeshow test results confirmed no statistically significant discrepancies between predicted and observed probabilities across the three validation cohorts (P-values: 0.244, 0.281, 0.838, respectively). Additionally, all cohorts had a Brier score below 0.25, indicating good predictive accuracy.

Table 2. Baseline characteristics of patients in study cohorts

Clinical indicators	Modeling cohort					Validation Cohort 1					Validation Cohort 2					Validation Cohort 3				
	Total (n=95)	DKD group (n=43)	NDKD group (n=52)	Statistics (χ^2/F/Z)	*P* value	Total (n=49)	DKD group (n=24)	NDKD group (n=25)	Statistics (χ^2/F/Z)	*P* value	Total (n=69)	DKD group (n=28)	NDKD group (n=41)	Statistics (χ^2/F/Z)	*P* value	Total (n=155)	DKD group (n=35)	NDKD group (n=120)	Statistics (χ^2/F/Z)	*P* value
Male [case (%)]	77 (81.1)	35 (81.4)	42 (80.8)	0.006[a]	0.938	46 (93.9)	23 (95.8)	23 (92.0)	0.313[a]	0.576	50 (72.5)	25 (89.3)	25 (61.0)	6.683[a]	0.010	104 (67.1)	20 (57.1)	84 (70.0)	2.029[a]	0.154
Age (year)	51 (46, 56)	51 (47, 54)	49 (45, 56)	-0.674	0.500	51 (43, 57)	49 (42, 56)	52 (43, 58)	-0.961	0.336	55 (43, 62)	54 (41, 64)	56 (45, 62)	-0.673	0.501	53 (46, 62)	51 (43, 57)	55 (47, 63)	-2.038	0.042
Height (m)	1.66± 0.07	1.65± 0.08	1.67± 0.06	-0.995	0.322	1.68± 0.06	1.68± 0.07	1.67± 0.05	0.590	0.558	1.66± 0.08	1.67± 0.07	1.65± 0.08	0.808	0.422	1.63± 0.08	1.63± 0.11	1.63± 0.08	-0.383	0.702
Weight (kg)	70.0 (61.0, 78.0)	68.0 (60.5, 78.0)	70.8 (63.1, 78.0)	-1.022	0.307	69.0 (65.0, 80.0)	70.0 (65.0, 79.0)	69.0 (64.4, 81.0)	-0.500	0.617	66 (58, 76)	69 (61, 74)	66 (56, 78)	-0.275	0.783	68.6 (62.2, 75.7)	66.2 (53.8, 74.5)	70.1 (63.5, 76.1)	-1.834	0.067
Body Mass Index (kg/m^2)	25.71 (22.72, 27.34)	25.39 (22.60, 27.34)	25.80 (23.67, 27.32)	-0.482	0.630	24.89 (22.87, 27.37)	24.79 (22.63, 27.20)	25.06 (23.46, 27.63)	-1.070	0.285	24.24 (22.39, 26.76)	24.46 (22.72, 25.60)	24.24 (22.31, 27.67)	-0.336	0.737	26.07 (23.45, 27.54)	24.91 (21.28, 27.24)	26.09 (24.07, 27.88)	-2.198	0.028
Duration of diabetes (month)	48 (12, 120)	84 (36, 168)	30 (7, 60)	-3.894	0.000	36 (4, 93)	66 (27, 129)	12 (0, 42)	-3.085	0.002	36 (3, 120)	120 (39, 177)	12 (1, 39)	-4.155	0.000	12 (1, 60)	60 (4, 100)	6 (1, 39)	-3.861	0.000
Systolic Blood Pressure (mmHg)	139±25	143±25	135±24	1.491	0.139	132±24	139±28	126±17	1.877	0.068	146±23	146±24	147±22	-0.124	0.902	142±24	146±25	141±23	1.137	0.257
Diastolic Blood Pressure (mmHg)	87±12	85±13	87±11	-0.768	0.445	86±13	86±13	86±12	-0.021	0.983	87±13	81±12	91±12	-3.293	0.002	82±15	80±12	83±15	-0.815	0.417
Pulse Pressure (mmHg)	49 (39, 63)	54 (42, 67)	44 (35, 54)	-2.842	0.004	44 (35, 55)	53 (40, 59)	42 (33, 49)	-2.671	0.008	54 (48, 73)	63 (50, 78)	52 (47, 65)	-2.049	0.040	57 (47, 74)	68 (49, 80)	56 (46, 69)	-2.203	0.028
Blood Uric Acid (μmol/L)	406± 121	406± 105	406± 134	-0.001	0.999	409± 107	395±97	116±23	-0.880	0.383	436± 112	423± 95	445± 123	-0.801	0.426	441± 116	458± 130	437± 111	0.960	0.339
Blood Urea Nitrogen (mmol/L)	6.49 (4.71, 9.10)	7.70 (5.69, 10.82)	5.45 (4.36, 8.35)	-2.890	0.004	7.04 (5.91, 9.32)	7.64 (6.50, 11.80)	6.71 (5.24, 8.60)	-1.920	0.055	11.10 (7.75, 15.97)	14.09 (11.18, 23.64)	8.80 (5.65, 12.81)	-3.513	0.000	7.60 (5.20, 11.00)	10.20 (7.60, 14.20)	6.85 (5.10, 9.85)	-3.336	0.001
Blood Creatinine (μmol/L)	108 (79, 167)	125 (94, 193)	102 (75, 135)	-2.352	0.019	122 (96, 145)	125 (97, 146)	114 (90, 145)	-0.910	0.363	162 (96, 329)	223 (131, 441)	134 (79, 201)	-2.976	0.003	91 (73, 156)	129 (89, 186)	87 (72, 141)	-2.350	0.019
Blood Cystatin C (mg/L)	1.380 (1.032, 1.978)	1.663 (1.304, 2.264)	1.130 (0.892, 1.767)	-3.682	0.000	1.31 (1.13, 1.60)	1.38 (1.18, 1.78)	1.29 (1.04, 1.55)	-0.860	0.390	1.89 (1.39, 3.03)	2.48 (1.70, 3.67)	1.65 (1.16, 2.54)	-2.603	0.009	1.52 (1.06, 2.06)	1.82 (1.52, 2.70)	1.35 (1.03, 1.99)	-3.067	0.002
Glomerular Filtration Rate 1 (mL/min/1.73 m^2)	64.63 (39.78, 92.92)	52.01 (33.15, 74.09)	69.81 (48.09, 103.99)	-2.501	0.012	60.03 (48.82, 81.47)	57.07 (46.37, 72.90)	63.36 (48.82, 85.83)	-0.870	0.384	35.37 (16.02, 68.36)	22.14 (9.81, 48.92)	46.38 (25.46, 84.25)	-3.043	0.002	60.99 (32.21, 82.73)	41.81 (26.63, 74.81)	67.19 (35.82, 82.73)	-1.932	0.053
Glomerular Filtration Rate 2 (mL/min/1.73 m^2)	47.60 (30.82, 72.65)	38.08 (25.28, 54.85)	62.74 (35.11, 90.10)	-3.619	0.000	53.99 (39.74, 63.10)	50.36 (34.98, 61.35)	56.19 (43.68, 72.17)	-0.900	0.368	34.29 (17.06, 50.55)	23.51 (13.57, 38.24)	37.51 (21.63, 60.54)	-2.640	0.008	43.88 (28.27, 68.98)	33.62 (21.11, 44.49)	48.52 (30.15, 71.95)	-2.932	0.003

Glomerular Filtration Rate 3 (mL/min/1.73 m^2)	50.18 (30.37, 73.26)	37.63 (24.56, 55.53)	60.25 (36.53, 83.30)	-2.991	0.003	47.12 (37.04, 63.85)	46.88 (32.55, 58.58)	52.82 (38.66, 67.51)	-1.120	0.263	35.35 (15.36, 54.79)	21.53 (11.37, 40.95)	39.07 (22.50, 68.84)	-2.908	0.004	48.75 (28.00, 73.58)	37.61 (22.53, 61.73)	54.54 (28.61, 74.65)	-2.519	0.012
Immunoglobulin A (g/L)	2.39 (1.80, 3.13)	2.28 (1.63, 2.72)	2.58 (2.02, 3.43)	-2.075	0.038	2.75 (2.15, 3.31)	2.57 (2.07, 3.13)	2.88 (2.27, 3.50)	-1.540	0.124	2.48 (1.89, 3.22)	2.94 (2.10, 3.55)	2.00 (1.66, 2.87)	-2.261	0.024	2.38 (1.71, 3.26)	2.56 (1.81, 3.15)	2.37 (1.69, 3.29)	-0.058	0.954
Serum Albumin (g/L)	39.3± 7.3	37.8± 6.1	40.6± 8.0	-1.880	0.063	40.7± 8.1	37.5± 9.2	43.8± 5.5	-2.959	0.005	30.4± 8.3	33.9± 6.5	28.0± 8.5	3.078	0.003	28.77± 9.25	30.25± 7.02	28.34± 9.79	1.076	0.283
Total Cholesterol (mmol/L)	5.02 (3.96, 5.82)	5.28 (3.96, 5.82)	4.65 (3.90, 5.74)	-0.923	0.356	4.95 (4.11, 5.65)	4.88 (4.08, 5.31)	5.11 (4.18, 5.92)	-1.120	0.263	5.20 (4.05, 6.40)	4.15 (3.13, 6.18)	5.40 (4.95, 6.80)	-2.483	0.013	5.83 (4.50, 8.40)	5.26 (3.97, 6.78)	6.16 (4.59, 9.14)	-2.011	0.044
Triglyceride (mmol/L)	2.04 (1.25, 3.15)	2.21 (1.25, 3.14)	1.79 (1.21, 3.20)	-0.976	0.329	1.95 (1.28, 2.34)	1.65 (1.10, 2.20)	2.17 (1.85, 2.66)	-2.140	0.032	1.71 (1.11, 2.37)	1.28 (0.93, 2.11)	1.91 (1.36, 2.54)	-2.145	0.032	2.14 (1.43, 3.13)	1.61 (1.21, 2.35)	2.21 (1.61, 3.42)	-2.737	0.006
High-density lipoprotein (mmol/L)	1.04 (0.86, 1.20)	1.05 (0.93, 1.20)	1.04 (0.84, 1.23)	-0.075	0.940	1.02 (0.86, 1.24)	1.19 (0.93, 1.34)	0.90 (0.84, 1.14)	-2.381	0.017	1.28 (1.04, 1.59)	1.05 (0.85, 1.28)	1.45 (1.19, 1.85)	-4.504	0.000	1.13 (0.87, 1.47)	1.10 (0.86, 1.32)	1.18 (0.87, 1.63)	-1.787	0.074
Low-density lipoprotein (mmol/L)	2.81 (2.15, 3.54)	3.01 (2.12, 3.70)	2.77 (2.16, 3.24)	-1.088	0.277	2.97 (2.31, 3.89)	2.56 (2.15, 3.81)	3.29 (2.50, 4.14)	-1.480	0.139	3.19 (2.53, 4.17)	2.61 (1.86, 4.03)	3.38 (2.95, 4.61)	-2.616	0.009	3.57 (2.67, 5.35)	3.39 (2.56, 4.39)	3.63 (2.77, 5.63)	-1.648	0.099
Fasting Blood Glucose (mmol/L)	6.66 (5.64, 8.49)	7.85 (6.12, 9.11)	6.19 (5.53, 7.40)	-2.938	0.003	6.50 (5.52, 9.79)	7.99 (5.38, 10.72)	6.39 (5.78, 8.15)	-0.640	0.522	6.10 (4.90, 8.10)	6.80 (5.48, 10.25)	5.60 (4.90, 7.20)	-1.919	0.055	5.88 (4.98, 7.09)	6.59 (5.28, 7.84)	5.72 (4.63, 6.84)	-2.178	0.029
Glycated hemoglobin (%)	6.9 (6.2, 8.7)	7.8 (6.5, 8.8)	6.6 (6.1, 7.4)	-2.907	0.004	7.5 (6.6, 8.8)	8.43 (7.30, 9.40)	7.10 (6.55, 8.40)	-2.002	0.045	6.47 (5.95, 7.06)	6.93 (5.90, 7.39)	6.43 (6.00, 6.81)	-1.015	0.310	6.60 (6.00, 7.30)	7.00 (6.10, 8.30)	6.50 (6.00, 6.90)	-2.634	0.008
Urinary κ Chain (mg/L)	50.91 (18.10, 66.08)	66.08 (34.68, 74.20)	50.91 (11.48, 60.94)	-2.653	0.008	58.40 (20.55, 104.84)	98.29 (19.22, 129.00)	43.00 (20.65, 62.31)	-1.790	0.073	65.20 (42.49, 124.99)	64.41 (52.41, 89.56)	65.20 (38.00, 134.15)	-0.361	0.718	76.00 (29.20, 132.70)	97.10 (39.60, 150.00)	60.50 (22.53, 122.50)	-1.896	0.058
Urinary λ Chain (mg/L)	26.45 (6.17, 34.67)	34.67 (16.95, 48.80)	26.45 (4.06, 27.61)	-3.170	0.002	27.90 (6.82, 54.55)	46.25 (6.32, 72.35)	16.50 (6.82, 32.21)	-1.470	0.141	37.86 (19.47, 78.12)	33.49 (19.28, 49.29)	48.70 (20.18, 84.10)	-1.222	0.222	48.20 (18.30, 81.60)	61.10 (23.70, 90.90)	35.40 (12.30, 72.98)	-2.049	0.041
Urinary Total Protein (mg/L)	514.5 (87.0, 1894.0)	1404.0 (222.0, 2322.0)	215.0 (72.5, 717.3)	-2.991	0.003	382.0 (90.0, 1776.5)	1274.5 (188.3, 2858.8)	302.0 (70.0, 936.0)	-2.090	0.037	339.8 (131.9, 687.4)	311.7 (119.4, 469.5)	390.1 (131.9, 719.7)	-1.192	0.233	4030.0 (1820.0, 7820.0)	4030.0 (2190.0, 7200.0)	3990.0 (1560.0, 7835.0)	-0.218	0.827
Red Blood Cell Count (10^{12}/L)	3.69± 0.69	3.44± 0.59	3.89± 0.71	-3.297	0.001	3.80± 0.73	3.55± 0.72	4.04± 0.68	-2.412	0.020	3.97± 0.88	3.43± 0.65	4.34± 0.84	-4.820	0.000	4.25± 0.87	3.75± 0.92	4.40± 0.80	-4.121	0.000
Hemoglobin Concentration (g/L)	128.62± 24.71	120.21± 22.70	135.58± 24.33	-3.158	0.002	134.18± 27.08	123.88± 27.52	144.08± 23.08	-2.789	0.008	112.65± 25.70	98.43± 21.32	122.37± 24.03	-4.249	0.000	121.99± 22.47	102.03± 14.30	127.81± 21.06	-6.791	0.000
Hematocrit (%)	38.22± 6.68	35.79± 6.12	40.24± 6.51	-3.407	0.001	39.79± 7.60	37.03± 7.97	42.45± 6.29	-2.651	0.011	34.12± 7.60	29.57± 5.92	37.23± 7.09	-4.703	0.000	36.91± 6.79	30.76± 4.30	38.70± 6.33	-6.967	0.000

a: One cell had an expected count less than 5.

Table 3. Univariate and multivariate Logistic regression analysis results of modeling cohort

Influencing factor	Univariate analysis		Multivariate analysis	
	OR (95% CI)	*P* value	OR (95% CI)	*P* value
Gender (male/female)	1.042 (0.371, 3.000)	0.938		
Age (year)	0.982 (0.936, 1.027)	0.430		
Height (m)	17.985 (0.064, 6102.140)	0.319		
Weight (kg)	1.022 (0.988, 1.061)	0.220		
Body mass index (kg/m^2)	1.041 (0.930, 1.171)	0.490		
Diabetes duration (month)	0.987 (0.979, 0.994)	0.001	0.988 (0.976, 0.999)	0.046
Systolic blood pressure (mmHg)	0.987 (0.970, 1.004)	0.141		
Diastolic blood pressure (mmHg)	1.013 (0.980, 1.048)	0.440		
Pulse pressure (mmHg)	0.972 (0.948, 0.994)	0.019		
Blood uric acid (μmol/L)	1.000 (0.997, 1.003)	0.999		
Blood urea nitrogen (mmol/L)	0.887 (0.781, 0.991)	0.046		
Blood creatinine (μmol/L)	0.997 (0.992, 1.002)	0.249		
Blood cystatin C (mg/L)	0.384 (0.183, 0.713)	0.006	0.014 (0.000, 1.000)	0.076
Glomerular filtration rate 1 (mL/min/1.73 m^2)	1.017 (1.004, 1.032)	0.013		
Glomerular filtration rate 2 (mL/min/1.73 m^2)	1.030 (1.014, 1.049)	0.001	1.056 (0.978, 1.177)	0.211
Glomerular filtration rate 3 (mL/min/1.73 m^2)	1.014 (1.001, 1.029)	0.051	0.946 (0.842, 0.989)	0.071
Immunoglobulin A (g/L)	1.688 (1.104, 2.716)	0.021	2.232 (1.096, 5.241)	0.040
Serum albumin (g/L)	1.056 (0.998, 1.121)	0.067		
Total cholesterol (mmol/L)	0.915 (0.673, 1.235)	0.559		
Triglyceride (mmol/L)	0.970 (0.836, 1.124)	0.682		
High-density lipoprotein (mmol/L)	1.525 (0.403, 6.197)	0.539		
Low-density lipoprotein (mmol/L)	0.883 (0.598, 1.287)	0.516		
Fasting blood glucose (mmol/L)	0.796 (0.639, 0.955)	0.029		
Glycated hemoglobin (%)	0.738 (0.560, 0.927)	0.018	0.484 (0.275, 0.749)	0.003
Urinary κ chain (mg/L)	0.993 (0.984, 1.002)	0.131		
Urinary λ chain (mg/L)	0.989 (0.973, 1.004)	0.157		
Urinary microalbumin (mg/L)	1.000 (1.000, 1.000)	0.080		
Red blood cell count (10^{12}/L)	2.893 (1.497, 6.071)	0.003	12.577 (1.144, 169.707)	0.043
Hemoglobin concentration (g/L)	1.028 (1.010, 1.049)	0.004		
Hematocrit (%)	1.118 (1.045, 1.207)	0.002		

Table 4. Performance of the RICH model in research cohorts

Indicator	Modeling cohort	Validation cohort 1	Validation cohort 2	Validation cohort 3
Case	95	49	69	155
Discrimination degree				
C statistics (AUC 95% CI)	0.847 (0.766-0.929)	0.755 (0.611-0.899)	0.764 (0.654-0.874)	0.755 (0.665-0.845)
Sensitivity (%)	76.9	76.0	66.0	60.8
Specificity (%)	83.7	75.0	75.0	82.9
PPV	85.1	76.0	79.4	92.4
TP	40	19.0	27	73
NPV	75	75.0	60.0	38.2
TN	36	18.0	21	29
Calibration degree				
Hosmer-Lemeshow test *P* value	0.433	0.244	0.281	0.838
Composite indicator				
Brier score	0.154	0.208	0.202	0.175

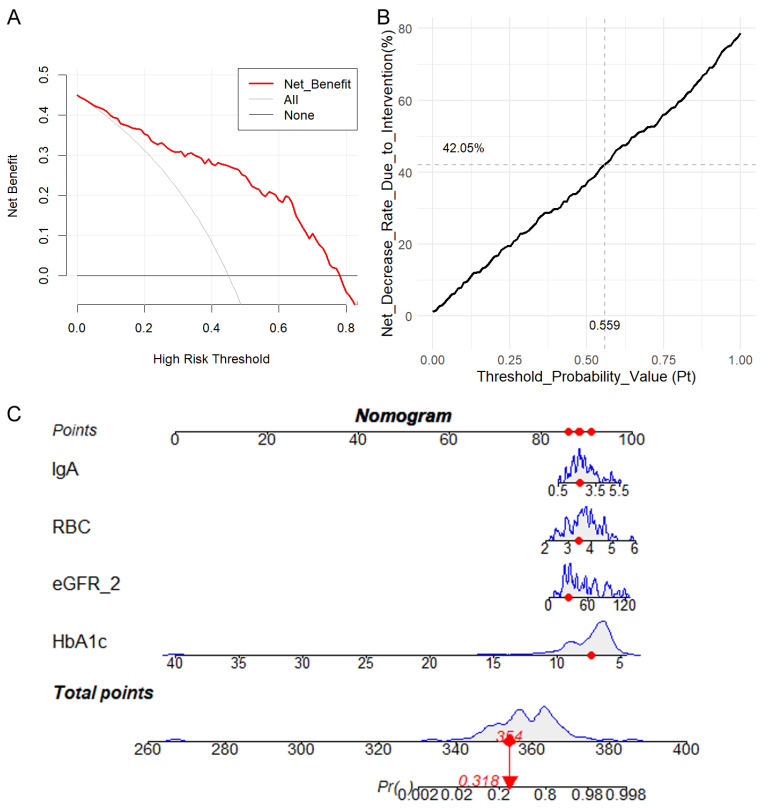

Figure 1. Clinical decision curve analysis for early diagnosis of NDKD and prediction probability nomogram of T2DM patients developing NDKD through the RICH model. A. Net benefit analysis of the Nomogram model. When the threshold probability ranges between 0.10-0.80, patients exhibit consistently high clinical net benefit. At a threshold probability of 0.559, approximately 20% of patients achieve net benefit. All: Net benefit rate when all patients received NDKD treatment; None: Net benefit rate when all patients received no treatment; Net_Benefit: Net benefit rate when treatment was guided by the diagnostic model. B. Net reduction analysis of the intervention. When the threshold probability is set at 0.559, using this model for the early diagnosis of NDKD may potentially reduce the renal biopsy rate by 42.05%. C. Nomogram for the RICH model.

Application of nomogram

To facilitate the clinical application of the RICH model for predicting the probability of NDKD in T2DM patients, this study used R software to assign weights to each risk factor in the model and visualize these values in a nomogram (**Figure 1C**). This tool can also be implemented in application programs (e.g., mobile apps, mini-programs, and web-based calculators).

Discussion

With the continuous increase in the global diabetic population each year, DKD has become a leading cause of ESRD [4, 29]. The pathogenesis, treatment strategies, and prognosis of DKD patients are different from those of NDKD patients [30, 31]. Therefore, it is necessary to use renal biopsy to diagnose DKD and NDKD patients [9]. However, renal biopsy has some limitations. For instance, it is an invasive procedure with potential risks; the indications for renal biopsy in diabetic patients remain unclear [32]; and patients generally have a low acceptance of this procedure [33].

A total of 416 T2DM patients were enrolled in this study. Among these patients, 31.3% were diagnosed with DKD, 57.2% with NDKD, and 11.5% with MIX. Among all NDKD and MIX patients, MN, IgA nephropathy, hypertensive nephropathy, MCD, and focal segmental glomerulosclerosis were the most common pathological types, which is consistent with previous studies' results [14, 15, 34-38]. This indicates that the cohort used in this study has good representativeness.

Based on objective laboratory test indicators collected from three medical centers, we constructed and validated a RICH predictive model to identify adult T2DM patients with renal injury. The results indicate that high levels of eGFR_2, IgA, and RBC, along with a low level of HbA1c, are important factors for NDKD identification. Relevant studies have been reported previously. For instance, Zhou et al. [20, 39], established a diagnostic model involving five indicators, namely disease duration, HbA1c, hematuria, and DR. Another study updated this model by introducing HGB into it [40]. Additionally, several studies have shown that younger age, longer disease duration, presence of DR, absence of hematuria, diagnosis of anemia, lower levels of HGB, TG/CysC, UA, ALB and eGFR, as well as higher levels of Glu, HbA1c, CysC and DBP, are important independent predictive factors for DKD diagnosis and progression, which can also be used as key indicators

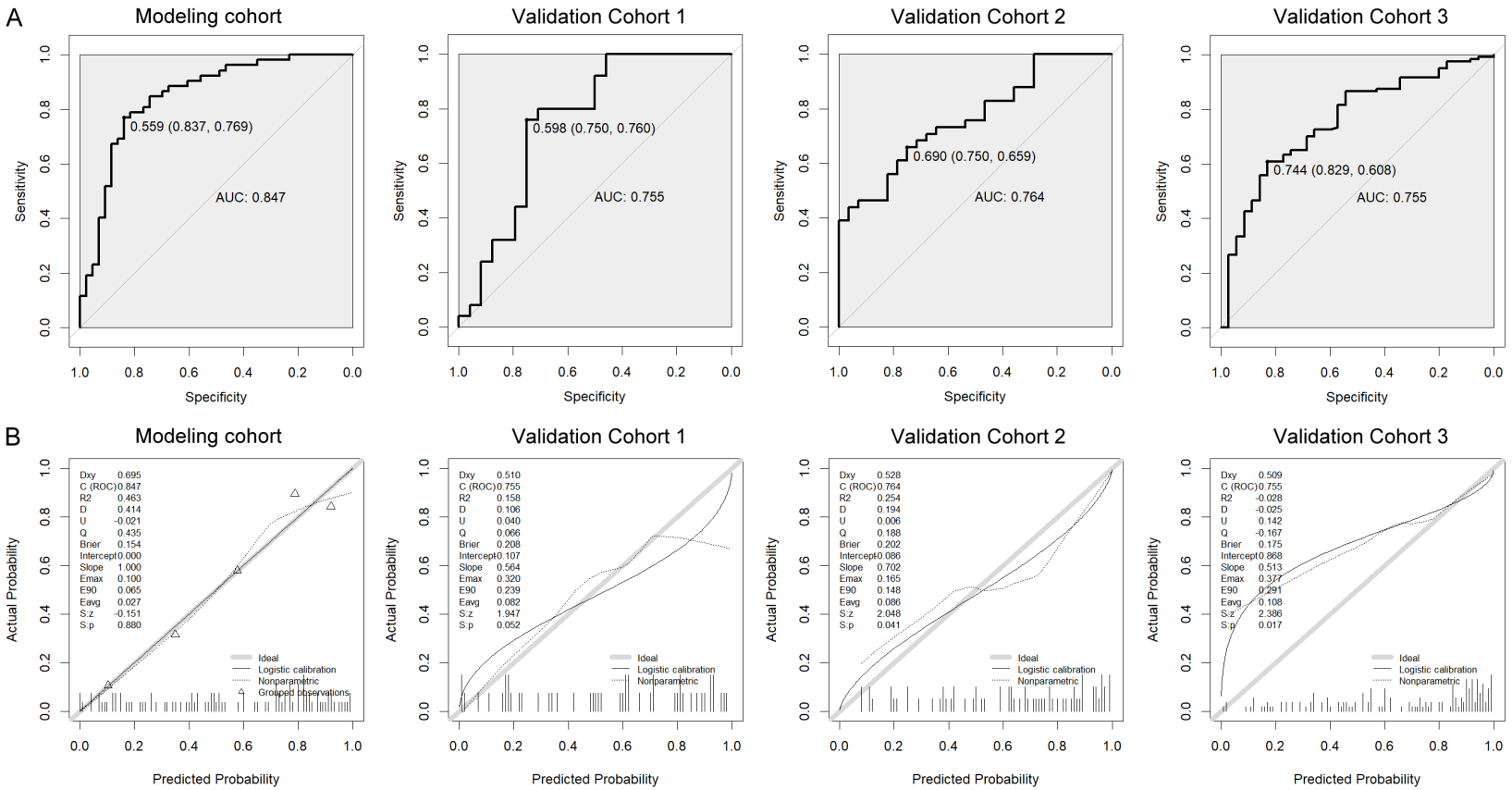

Figure 2. ROC and calibration curves of the RICH model. A. ROC curves of the modeling cohort and Validation Cohorts 1, 2, 3. B. Calibration curves of the modeling cohort and Validation Cohorts 1, 2, 3.

to distinguish DKD from NDKD [13-15, 28, 30, 31].

The RICH predictive model incorporates four biomarkers: RBC, IgA, eGRF_2 (CysC-based estimated glomerular filtration rate), and HbA1c. As the most abundant blood cells, RBCs primarily function to transport oxygen and carbon dioxide, while also participating in maintaining acid-base balance in the human body, thus serving as a critical cellular component for sustaining the normal metabolism of tissues and organs. In the context of renal diseases, RBCs can act as an indicator of anemia [41]; additionally, the morphology of urinary RBCs helps identify the lesion site in renal diseases [42]. Studies have revealed that in DKD, excessive secretion of amylin - a type of amyloid hormone co-synthesized with insulin - contributes to elevated erythropoietin levels and subsequent RBC production [43]. Consistent with this mechanism, our study observed that RBC levels were significantly higher in the DKD group than in the NDKD group. This differential characteristic enables RBCs to serve as a potential and effective predictor for distinguishing DKD from NDKD, providing biological mechanistic support for model construction. As a key member of the immunoglobulin family, serum IgA levels are not only closely associated with susceptibility to renal diseases [44], but also positively correlated with the risk of T2DM in the adult population [45]. This dual association provides a crucial theoretical basis for enhancing the discriminative efficacy of the DKD differential diagnosis model. CysC exhibits extremely high value in the clinical diagnosis of DKD. Compared with traditional indicators such as eGFR and urinary albumin-to-creatinine ratio (UACR), CysC can reflect the risk of DKD onset at an earlier stage [46]. Clinical data have confirmed that higher baseline CysC levels are associated with an increased probability of developing DKD. This unique property offers a key basis for DKD differential diagnosis and justifies CysC as an important component of the predictive model established in this study [46, 47]. Glycated HbA1c is a well-recognized key indicator for the early screening of diabetes, with its diagnostic value widely validated in clinical practice [48]. Furthermore, extensive research data have demonstrated that in T2DM patients, higher HbA1c levels combined with greater fluctuations are correlated with an elevated risk of adverse outcomes, including stroke and renal disease [49]. This independent risk association identifies HbA1c as an indispensable core indicator in our DKD differential diagnosis model, laying a solid foundation for its clinical applicability and predictive accuracy.

Notably, a key innovation of this study, in contrast to prior research, lies in the pioneering inclusion of IgA and eGFR_2 as two pivotal predictive factors in the model. This design was driven by compelling biological and clinical evidence: the majority of NDKD patients in this cohort were diagnosed with MN, IgA nephropathy, MCD, and LN - diseases etiologically linked to intrinsic immune system dysfunction. Supporting this link, renal pathological analyses demonstrated IgA deposition in glomerular mesangial areas or capillary loops [22], and given that IgA is a core component of the immune system, its expression levels are inherently susceptible to immune system perturbations [50]. From a clinical perspective, eGFR serves as an irreplaceable index for renal function assessment, with superior screening sensitivity compared to Serum Creatinine (Scr) and CysC [51]. However, Scr measurements are confounded by variables such as diet, muscle mass, and exercise, which limit their utility for accurate renal function evaluation. Previous investigations have validated CysC as a valuable marker for distinguishing DKD from NDKD; as an optimal endogenous indicator of renal impairment, CysC surpasses Scr in diagnostic precision, and eGFR_2 further enhances this diagnostic capacity [21, 52]. Reinforcing this choice, recent clinical guidelines advocate for the supplementary measurement of CysC and the calculation of eGFR following the latest KDIGO recommendations [53]. For these reasons, eGFR_2 was integrated into the predictive model constructed in this study.

Compared with previous models for the identification and diagnosis of DKD and NDKD, the RICH model proposed in this study exhibits the following distinctive characteristics: Firstly, there are few existing studies focusing on predictive models for DKD and NDKD differentiation. Moreover, most of these studies are single-center retrospective investigations, whose results lack clear representativeness and reliability [15]. In contrast, the model constructed in this study was validated using external cohort

data from two hospitals for the first time, thus verifying the favorable scalability and applicability of the RICH model. Secondly, the RICH model incorporates only objective serum laboratory test indicators, excluding subjective or controversial variables such as diabetes duration and DR. This design effectively avoids the confounding influences of heterogeneous patient awareness levels and variable hospital diagnostic and treatment standards on model performance. Lastly, a relatively large sample size of patients was enrolled in this study, which ensures good stability of model validation. In addition, the RICH model involves only a small number of indicators, all of which can be measured via a single blood test. Furthermore, the application of nomograms in this study provides a more intuitive visualization approach, facilitating convenient clinical implementation [15, 20, 40]. Although DKD can be diagnosed relatively easily in accordance with clinical guidelines, how to avoid misdiagnosis of NDKD and missed diagnosis of DKD - without unnecessarily relying on invasive renal biopsy - is a research topic that warrants further exploration. The RICH model can assist clinicians in effectively identifying disease subtypes among patients at an early stage. Combined with clinical guidelines, this model provides a more robust evidence base for physicians to make clinical decisions. In particular, it serves as a convenient and efficient predictive tool for patients with contraindications to renal biopsy or medical institutions lacking renal biopsy facilities. By leveraging this model, more individualized initial treatment regimens can be formulated for diabetic patients with different disease types, thereby improving their renal prognosis.

There are some limitations in this research. Firstly, most patients who underwent renal biopsy and were enrolled in this study were those with a strong clinical suspicion of NDKD. Therefore, selection bias inherent in this patient cohort could influence the study results. Secondly, like all risk prediction models, the RICH model is constrained by the limitations of clinical sample size. Thus, its reliability needs to be further validated by more external validation cohorts and independent investigations. Thirdly, due to the small sample size and complex pathological mechanisms of patients in the MIX group, this subgroup was excluded from the model construction to avoid introducing confounding biases. In future research, we will expand the sample size to explore the feasibility of incorporating this subgroup into the model. Fourthly, owing to the limitations of heterogeneous diagnostic protocols across hospitals and inconsistent laboratory testing panels in this retrospective study, several potentially important parameters - such as UACR, neutrophil gelatinase-associated lipocalin, and β_2-microglobulin - were not included in the current analysis. Therefore, prior to the clinical translation and practical application of the RICH model, it is imperative to conduct larger-scale, even international multicenter real-world prospective cohort studies to further optimize and validate the model for clinical implementation.

Conclution

Our findings demonstrate that NDKD accounts for a notable proportion (57.2%) of kidney disease cases among T2DM patients who underwent renal biopsy, highlighting the imperative of recognizing and addressing NDKD in this patient population. The RICH model, which is built on four parameters - RBC, IgA, eGFR_2, and HbA1c - exhibits robust performance in predicting NDKD occurrence in T2DM patients with renal impairment. Undoubtedly, larger-scale cohort studies are required to verify these findings and further validate the RICH model. The development of the RICH model not only furnishes clinicians with a potent auxiliary diagnostic tool but also offers a novel perspective for deepening the exploration of the pathogenesis and management strategies of diabetes-related renal complications.

Disclosure of conflict of interest

None.

Address correspondence to: Feng Hu, Department of Nephrology, Shenzhen Third People's Hospital, The Second Affiliated Hospital of Southern University of Science and Technology, Shenzhen 518112, Guangdong, China. ORCID: 0009-0006-1502-4779; E-mail: jfjhf1504@126.com; Peng Zhang, Department of Laboratory Medicine, Nanfang Hospital, Southern Medical University, Guangzhou 510515, Guangdong, China. ORCID: 0000-0002-7779-8474; E-mail: nfyyzp@126.com

References

[1] GBD 2021 Diabetes Collaborators. Global, regional, and national burden of diabetes from 1990 to 2021, with projections of prevalence to 2050: a systematic analysis for the Global Burden of Disease Study 2021. Lancet 2023; 402: 203-234.

[2] Sun H, Saeedi P, Karuranga S, Pinkepank M, Ogurtsova K, Duncan BB, Stein C, Basit A, Chan JCN, Mbanya JC, Pavkov ME, Ramachandaran A, Wild SH, James S, Herman WH, Zhang P, Bommer C, Kuo S, Boyko EJ and Magliano DJ. IDF Diabetes Atlas: Global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045. Diabetes Res Clin Pract 2022; 183: 109119.

[3] Li Y, Teng D, Shi X, Qin G, Qin Y, Quan H, Shi B, Sun H, Ba J, Chen B, Du J, He L, Lai X, Li Y, Chi H, Liao E, Liu C, Liu L, Tang X, Tong N, Wang G, Zhang JA, Wang Y, Xue Y, Yan L, Yang J, Yang L, Yao Y, Ye Z, Zhang Q, Zhang L, Zhu J, Zhu M, Ning G, Mu Y, Zhao J, Teng W and Shan Z. Prevalence of diabetes recorded in mainland China using 2018 diagnostic criteria from the American Diabetes Association: national cross sectional study. BMJ 2020; 369: m997.

[4] Hou JH, Zhu HX, Zhou ML, Le WB, Zeng CH, Liang SS, Xu F, Liang DD, Shao SJ, Liu Y and Liu ZH. Changes in the spectrum of kidney diseases: an analysis of 40,759 biopsy-proven cases from 2003 to 2014 in China. Kidney Dis (Basel) 2018; 4: 10-19.

[5] Yang C, Wang H, Zhao X, Matsushita K, Coresh J, Zhang L and Zhao MH. CKD in China: evolving spectrum and public health implications. Am J Kidney Dis 2020; 76: 258-264.

[6] Cheng HT, Xu X, Lim PS and Hung KY. Worldwide epidemiology of diabetes-related end-stage renal disease, 2000-2015. Diabetes Care 2021; 44: 89-97.

[7] Anders HJ, Huber TB, Isermann B and Schiffer M. CKD in diabetes: diabetic kidney disease versus nondiabetic kidney disease. Nat Rev Nephrol 2018; 14: 361-377.

[8] Fiorentino M, Bolignano D, Tesar V, Pisano A, Biesen WV, Tripepi G, D'Arrigo G and Gesualdo L; ERA-EDTA Immunonephrology Working Group. Renal biopsy in patients with diabetes: a pooled meta-analysis of 48 studies. Nephrol Dial Transplant 2017; 32: 97-110.

[9] Li L, Zhang X, Li Z, Zhang R, Guo R, Yin Q, Yang L, Yue R, Su B, Huang S, Xu H, He C and Liu F. Renal pathological implications in type 2 diabetes mellitus patients with renal involvement. J Diabetes Complications 2017; 31: 114-121.

[10] Heerspink HJL, Jongs N, Neuen BL, Schloemer P, Vaduganathan M, Inker LA, Fletcher RA, Wheeler DC, Bakris G, Greene T, Chertow GM and Perkovic V. Effects of newer kidney protective agents on kidney endpoints provide implications for future clinical trials. Kidney Int 2023; 104: 181-188.

[11] Umanath K and Lewis JB. Update on diabetic nephropathy: core curriculum 2018. Am J Kidney Dis 2018; 71: 884-895.

[12] El-Guindi MA, Sira MM, Sira AM, Salem TA, El-Abd OL, Konsowa HA, El-Azab DS and Allam AA. Design and validation of a diagnostic score for biliary atresia. J Hepatol 2014; 61: 116-123.

[13] Wu X, Huang H, Hou H, Shen G, Yu J, Zhou Y, Bosco MJ, Mao L, Wang F and Sun Z. Diagnostic performance of a 5-marker predictive model for differential diagnosis between intestinal tuberculosis and Crohn's disease. Inflamm Bowel Dis 2018; 24: 2452-2460.

[14] Zhang W, Liu X, Dong Z, Wang Q, Pei Z, Chen Y, Zheng Y, Wang Y, Chen P, Feng Z, Sun X, Cai G and Chen X. New diagnostic model for the differentiation of diabetic nephropathy from non-diabetic nephropathy in Chinese patients. Front Endocrinol (Lausanne) 2022; 13: 913021.

[15] Li L, Yang Y, Zhu X, Xiong X, Zeng L, Xiong S, Jiang N, Li C, Yuan S, Xu H, Liu F and Sun L. Design and validation of a scoring model for differential diagnosis of diabetic nephropathy and nondiabetic renal diseases in type 2 diabetic patients. J Diabetes 2020; 12: 237-246.

[16] Zhang H, Hu J, Zhu J, Li Q and Fang L. Machine learning-based metabolism-related genes signature and immune infiltration landscape in diabetic nephropathy. Front Endocrinol (Lausanne) 2022; 13: 1026938.

[17] Wang X, Ren L, Huang Y, Feng Z, Zhang G and Dai H. The role of tubulointerstitial markers in differential diagnosis and prognosis in patients with type 2 diabetes and biopsy proven diabetic kidney disease. Clin Chim Acta 2023; 547: 117448.

[18] Zhao Y, Liu L, Zuo L, Zhou X, Wang S, Gao H, Yu F, Zhang X, Wang M, Chen L, Zhang R, Zhang F, Bi S, Bai Q, Ding J, Yang Q, Xin S, Chai S, Chen M and Zhang J. A novel risk score model for the differential diagnosis of type 2 diabetic nephropathy: a multicenter study. J Diabetes Res 2023; 2023: 5514767.

[19] Jiang W, Wang J, Shen X, Lu W, Wang Y, Li W, Gao Z, Xu J, Li X, Liu R, Zheng M, Chang B, Li J, Yang J and Chang B. Establishment and validation of a risk prediction model for early diabetic kidney disease based on a systematic review and meta-analysis of 20 cohorts. Diabetes Care 2020; 43: 925-933.

[20] Zhou J, Chen X, Xie Y, Li J, Yamanaka N and Tong X. A differential diagnostic model of diabetic nephropathy and non-diabetic renal diseases. Nephrol Dial Transplant 2008; 23: 1940-1945.

[21] Ferguson TW, Komenda P and Tangri N. Cystatin C as a biomarker for estimating glomerular filtration rate. Curr Opin Nephrol Hypertens 2015; 24: 295-300.
[22] Kidney Disease: Improving Global Outcomes (KDIGO) Glomerular Diseases Work Group. KDIGO 2021 clinical practice guideline for the management of glomerular diseases. Kidney Int 2021; 100: S1-S276.
[23] Levey AS, Stevens LA, Schmid CH, Zhang YL, Castro AF 3rd, Feldman HI, Kusek JW, Eggers P, Van Lente F, Greene T and Coresh J; CKD-EPI (Chronic Kidney Disease Epidemiology Collaboration). A new equation to estimate glomerular filtration rate. Ann Intern Med 2009; 150: 604-612.
[24] American Diabetes Association Professional Practice Committee. 2. Classification and diagnosis of diabetes: standards of medical care in diabetes-2022. Diabetes Care 2022; 45 Suppl 1: S17-S38.
[25] National Kidney Foundation. KDOQI clinical practice guideline for diabetes and CKD: 2012 update. Am J Kidney Dis 2012; 60: 850-886.
[26] Shimizu M, Furuichi K, Toyama T, Kitajima S, Hara A, Kitagawa K, Iwata Y, Sakai N, Takamura T, Yoshimura M, Yokoyama H, Kaneko S and Wada T; Kanazawa Study Group for Renal Diseases and Hypertension. Long-term outcomes of Japanese type 2 diabetic patients with biopsy-proven diabetic nephropathy. Diabetes Care 2013; 36: 3655-3662.
[27] Tervaert TW, Mooyaart AL, Amann K, Cohen AH, Cook HT, Drachenberg CB, Ferrario F, Fogo AB, Haas M, de Heer E, Joh K, Noël LH, Radhakrishnan J, Seshan SV, Bajema IM and Bruijn JA; Renal Pathology Society. Pathologic classification of diabetic nephropathy. J Am Soc Nephrol 2010; 21: 556-563.
[28] Sethi S, Haas M, Markowitz GS, D'Agati VD, Rennke HG, Jennette JC, Bajema IM, Alpers CE, Chang A, Cornell LD, Cosio FG, Fogo AB, Glassock RJ, Hariharan S, Kambham N, Lager DJ, Leung N, Mengel M, Nath KA, Roberts IS, Rovin BH, Seshan SV, Smith RJ, Walker PD, Winearls CG, Appel GB, Alexander MP, Cattran DC, Casado CA, Cook HT, De Vriese AS, Radhakrishnan J, Racusen LC, Ronco P and Fervenza FC. Mayo clinic/renal pathology society consensus report on pathologic classification, diagnosis, and reporting of GN. J Am Soc Nephrol 2016; 27: 1278-1287.
[29] Ruiz-Ortega M, Rodrigues-Diez RR, Lavoz C and Rayego-Mateos S. Special issue "Diabetic nephropathy: diagnosis, prevention and treatment". J Clin Med 2020; 9: 813.
[30] Bermejo S, González E, López-Revuelta K, Ibernon M, López D, Martín-Gómez A, Garcia-Osuna R, Linares T, Díaz M, Martín N, Barros X, Marco H, Navarro MI, Esparza N, Elias S, Coloma A, Robles NR, Agraz I, Poch E, Rodas L, Lozano V, Fernández B, Hernández E, Martínez MI, Stanescu RI, Moirón JP, García N, Goicoechea M, Calero F, Bonet J, Galceran JM, Liaño F, Pascual J, Praga M, Fulladosa X and Soler MJ. Risk factors for non-diabetic renal disease in diabetic patients. Clin Kidney J 2020; 13: 380-388.
[31] Tonneijck L, Muskiet MH, Smits MM, van Bommel EJ, Heerspink HJ, van Raalte DH and Joles JA. Glomerular hyperfiltration in diabetes: mechanisms, clinical significance, and treatment. J Am Soc Nephrol 2017; 28: 1023-1039.
[32] Bermejo S, Pascual J and Soler MJ. The current role of renal biopsy in diabetic patients. Minerva Med 2018; 109: 116-125.
[33] Chou YH, Lin WC and Chen YM. Clinical outcomes in patients with biopsy-proved diabetic nephropathy compared to isolated lupus or crescentic glomerulonephritis. Diabetes Res Clin Pract 2019; 148: 144-151.
[34] Liu XM, Dong ZY, Zhang WG, Liu MY, Zhou JH, Wang Q, Sun XF, Wang Y, Cao XY and Chen XM. Validation of the 2007 kidney disease outcomes quality initiative clinical practice guideline for the diagnosis of diabetic nephropathy and nondiabetic renal disease in Chinese patients. Diabetes Res Clin Pract 2019; 147: 81-86.
[35] Dong Z, Wang Y, Qiu Q, Zhang X, Zhang L, Wu J, Wei R, Zhu H, Cai G, Sun X and Chen X. Clinical predictors differentiating non-diabetic renal diseases from diabetic nephropathy in a large population of type 2 diabetes patients. Diabetes Res Clin Pract 2016; 121: 112-118.
[36] Wang J, Han Q, Zhao L, Zhang J, Wang Y, Wu Y, Wang T, Zhang R, Grung P, Xu H and Liu F. Identification of clinical predictors of diabetic nephropathy and non-diabetic renal disease in Chinese patients with type 2 diabetes, with reference to disease course and outcome. Acta Diabetol 2019; 56: 939-946.
[37] Akimoto T, Ito C, Saito O, Takahashi H, Takeda S, Ando Y, Muto S and Kusano E. Microscopic hematuria and diabetic glomerulosclerosis--clinicopathological analysis of type 2 diabetic patients associated with overt proteinuria. Nephron Clin Pract 2008; 109: c119-26.
[38] Liu S, Guo Q, Han H, Cui P, Liu X, Miao L, Zou H and Sun G. Clinicopathological characteristics of non-diabetic renal disease in patients with type 2 diabetes mellitus in a northeastern Chinese medical center: a retrospective analysis of 273 cases. Int Urol Nephrol 2016; 48: 1691-1698.
[39] Yang Z, Feng L, Huang Y and Xia N. A differential diagnosis model for diabetic nephropathy

and non-diabetic renal disease in patients with type 2 diabetes complicated with chronic kidney disease. Diabetes Metab Syndr Obes 2019; 12: 1963-1972.

[40] Liu M, Chen X, Sun X, Zhou J, Zhang X, Zhu H, Chen Y, Liu S, Wei R, Tang L, Cai G, Zhang L and Bai X. Validation of a differential diagnostic model of diabetic nephropathy and non-diabetic renal diseases and the establishment of a new diagnostic model. J Diabetes 2014; 6: 519-526.

[41] Almashjary MN, Sahl AM, Hazzazi MS, Alhashmi MH, Bawazir WM, Basabrain AA, Borai A, Khalil H, Almohammadi M, Altayar MA, Alqarni AM, Jamalallail SA, Bahashwan SM, Qanash H and Yasin EB. Diagnostic utility of reticulocyte hemoglobin for iron-restricted anemia in patients with end-stage kidney disease on hemodialysis. Front Med (Lausanne) 2025; 12: 1689529.

[42] Lee AJ, Bang HI, Lee SM, Won D, Kang MS, Choi HJ, Lee IH and Jeon CH. Comparison of urinary red blood cell distribution (URD) and dysmorphic red blood cells for detecting glomerular hematuria: a multicenter study. J Clin Lab Anal 2025; 39: e25159.

[43] Verma N, Liu M, Ly H, Loria A, Campbell KS, Bush H, Kern PA, Jose PA, Taegtmeyer H, Bers DM, Despa S, Goldstein LB, Murray AJ and Despa F. Diabetic microcirculatory disturbances and pathologic erythropoiesis are provoked by deposition of amyloid-forming amylin in red blood cells and capillaries. Kidney Int 2020; 97: 143-155.

[44] Liu L, Khan A, Sanchez-Rodriguez E, Zanoni F, Li Y, Steers N, Balderes O, Zhang J, Krithivasan P, LeDesma RA, Fischman C, Hebbring SJ, Harley JB, Moncrieffe H, Kottyan LC, Namjou-Khales B, Walunas TL, Knevel R, Raychaudhuri S, Karlson EW, Denny JC, Stanaway IB, Crosslin D, Rauen T, Floege J, Eitner F, Moldoveanu Z, Reily C, Knoppova B, Hall S, Sheff JT, Julian BA, Wyatt RJ, Suzuki H, Xie J, Chen N, Zhou X, Zhang H, Hammarström L, Viktorin A, Magnusson PKE, Shang N, Hripcsak G, Weng C, Rundek T, Elkind MSV, Oelsner EC, Barr RG, Ionita-Laza I, Novak J, Gharavi AG and Kiryluk K. Genetic regulation of serum IgA levels and susceptibility to common immune, infectious, kidney, and cardio-metabolic traits. Nat Commun 2022; 13: 6859.

[45] Zhang L, Li Y, Wang H, Guo Y, Wang X, Wu H, Zhang Q, Liu L, Meng G, Zhang S, Sun S, Zhou M, Jia Q, Song K, Stubbendorff A, Gu Y and Niu K. Serum immunoglobulin concentrations and risk of type 2 diabetes mellitus in adults: a prospective cohort study from the TCLSIH study. BMC Immunol 2024; 25: 52.

[46] Wang N, Lu Z, Zhang W, Bai Y, Pei D and Li L. Serum cystatin C trajectory is a marker associated with diabetic kidney disease. Front Endocrinol (Lausanne) 2022; 13: 824279.

[47] Kang Y, Jin Q, Zhou M, Zheng H, Li X, Li A, Zhou JW, Lv J and Wang Y. Diagnostic value of serum TGF-β1 and CysC in type 2 diabetic kidney disease: a cross-sectional study. Front Med (Lausanne) 2025; 12: 1529648.

[48] Unnikrishnan R, Shaw JE, Chan JCN, Wild SH, Peters AL, Orrange S, Roden M and Mohan V. Prediabetes. Nat Rev Dis Primers 2025; 11: 49.

[49] Handley D, Gillett AC, Bala R, Tyrrell J and Lewis CM. Latent class growth mixture modeling of HbA1C trajectories identifies individuals at high risk of developing complications of type 2 diabetes mellitus in the UK Biobank. BMJ Open Diabetes Res Care 2025; 13: e004826.

[50] Pabst O. New concepts in the generation and functions of IgA. Nat Rev Immunol 2012; 12: 821-832.

[51] Inker LA, Schmid CH, Tighiouart H, Eckfeldt JH, Feldman HI, Greene T, Kusek JW, Manzi J, Van Lente F, Zhang YL, Coresh J and Levey AS; CKD-EPI Investigators. Estimating glomerular filtration rate from serum creatinine and cystatin C. N Engl J Med 2012; 367: 20-29.

[52] Pottel H, Björk J, Rule AD, Ebert N, Eriksen BO, Dubourg L, Vidal-Petiot E, Grubb A, Hansson M, Lamb EJ, Littmann K, Mariat C, Melsom T, Schaeffner E, Sundin PO, Åkesson A, Larsson A, Cavalier E, Bukabau JB, Sumaili EK, Yayo E, Monnet D, Flamant M, Nyman U and Delanaye P. Cystatin C-based equation to estimate GFR without the inclusion of race and sex. N Engl J Med 2023; 388: 333-343.

[53] Stevens PE, Ahmed SB, Carrero JJ, Foster B, Francis A, Hall RK, Herrington WG, Hill G, Inker LA, Kazancıoğlu R, Lamb E, Lin P, Madero M, McIntyre N, Morrow K, Roberts G, Sabanayagam D, Schaeffner E, Shlipak M, Shroff R, Tangri N, Thanachayanont T, Ulasi I, Wong G, Yang CW, Zhang L and Levin A. KDIGO 2024 clinical practice guideline for the evaluation and management of chronic kidney disease. Kidney Int 2024; 105: S117-S314.