

# Prediction of Diabetic Kidney Disease in Newly Diagnosed Type 2 Diabetes Mellitus

Xiaodie Mu , Aihua Wu , Huiyue Hu, Hua Zhou \*, Min Yang\*

Department of Nephrology, The Third Affiliated Hospital of Soochow University, Changzhou, People's Republic of China

\*These authors contributed equally to this work

Correspondence: Hua Zhou; Min Yang, Department of Nephrology, The Third Affiliated Hospital of Soochow University, Changzhou, 213003, People's Republic of China, Tel +86 0519 68872082, Email zhouhua2323@suda.edu.cn; yangmin1516@suda.edu.cn

**Background:** Diabetic kidney disease (DKD), a common microvascular complication of diabetes mellitus (DM), is always asymptomatic until it develops to the advanced stage. Thus, we aim to develop a nomogram prediction model for progression to DKD in newly diagnosed type 2 diabetes mellitus (T2DM).

**Methods:** This was a single-center analysis of prospective data collected from 521 newly diagnosed patients with T2DM. All related clinical records were incorporated, including the triglyceride-glucose index (TyG index). The least absolute shrinkage and selection operator (LASSO) was used to build a prediction model. In addition, discrimination, calibration, and clinical practicality of the nomogram were evaluated.

**Results:** In this study, 156 participants were incorporated as the validation set, while the remaining 365 were incorporated into the training set. The predictive factors included in the individualized nomogram prediction model included 5 variables. The area under the curve (AUC) for the prediction model was 0.826 (95% CI 0.775 to 0.876), indicating excellent discrimination performance. The model performed exceptionally well in terms of predictive accuracy and clinical applicability, according to calibration curves and decision curve analysis.

**Conclusion:** The predictive nomogram for the risk of DKD in newly diagnosed T2DM patients had outstanding discrimination and calibration, which could help in clinical practice.

**Keywords:** diabetic kidney disease, type 2 diabetes mellitus, nomogram, prediction model, risk assessment

## Introduction

The latest global diabetes mellitus (DM) map demonstrated by the International Diabetes Federation (IDF) showed that the incidence of DM patients (aged 20–79) would rise to 12.2% (783.2 million people) in 2045.<sup>1</sup> China has the most DM currently, and the number is expected to exceed 174 million by 2045.<sup>1</sup> Diabetic kidney disease (DKD), characterized by persistent microalbuminuria or decreased estimated glomerular filtration rate (eGFR), is one of the most significant microvascular complications of diabetes. DKD is now more common than primary glomerulonephritis and is the major cause of chronic kidney disease (CKD) in China, where it affects 40% of those with type 2 diabetes mellitus (T2DM) who are hospitalized.<sup>2–4</sup>

DKD is always asymptomatic until it develops to the advanced stage. Currently, the main methods for early diagnosis of DKD are urinary albumin creatinine ratio (UACR) and eGFR. However, on the one hand, these diagnostic methods have their limitations. UACR does not increase in the early stage (stages I and II) of DKD,<sup>5</sup> meanwhile eGFR increases in the early stage.<sup>6</sup> On the other hand, the increasing incidence of DKD is placing a burden on the healthcare system, making it essential to find an approach to screen DKD in the medical environment. Hence, early detection and treatment of DKD are critical.

Previous studies have demonstrated multiple clinically related factors of DKD, such as age, UACR, haemoglobin A1c (HbA1c) and high-density lipoprotein (HDL).<sup>7,8</sup> Furthermore, related studies have shown that the development of DKD is strongly associated with insulin resistance (IR), which may accelerate the progression of DKD.<sup>9–11</sup> Hyperinsulinemic

euglycemic glucose clamp (HEGC), a significant method for evaluating IR, is expensive and difficult to operate.<sup>12</sup> Additionally, the homeostasis model for insulin resistance (HOMA-IR) index is also a common tool to assess IR in clinical practice.<sup>13</sup> However, serum insulin or C-peptide tests are not appropriate for all inpatients or outpatients owing to the impact of insulin treatment. Triglyceride-glucose index (TyG index), as well as triglyceride and fasting glucose products, demonstrates greater potential in evaluating IR.<sup>14,15</sup> Previous clinical studies have demonstrated the performance of the TyG index in evaluating IR, and it was superior to the HOMA-IR index.<sup>10</sup> In conclusion, these easily available clinical risk factors may be used to build an accurate prediction model of DKD development in individuals with newly diagnosed T2DM, facilitating DKD screening. In addition, we intended to use the least absolute shrinkage and selection operator (LASSO) regression to screen variables and construct a predictive model. LASSO regression, a compressive estimation method for selecting relevant features for various classes of data, helps deal with biased estimates with complex collinearity data.<sup>16</sup> It can perform variable screening and complexity adjustment while fitting generalized linear models, and is widely applied in variable screening of various prediction models, including commonly used clinical variables and related omics data.<sup>17,18</sup> LASSO regression has contributed to constructing some prediction models, such as tumors, metabolic diseases, kidney diseases, etc.<sup>19,20</sup>

In this study, we designed a long-term cohort trial in patients with newly diagnosed T2DM for more than 10 years, aiming to explore the risk factors related to DKD, building and validating the prediction model.

## Materials and Methods

### Study Population

This was a retrospective analysis on prospectively collected T2DM patients' data from the Third Affiliated Hospital of Soochow University. We consecutively incorporated 863 newly diagnosed T2DM patients in our hospital from December 2010 to January 2014 for long-term follow-up.

Inclusion criteria: (1) meet the American Diabetes Association (ADA) classification criteria for T2DM and were newly diagnosed patients;<sup>14</sup> (2) age > 18 years old.

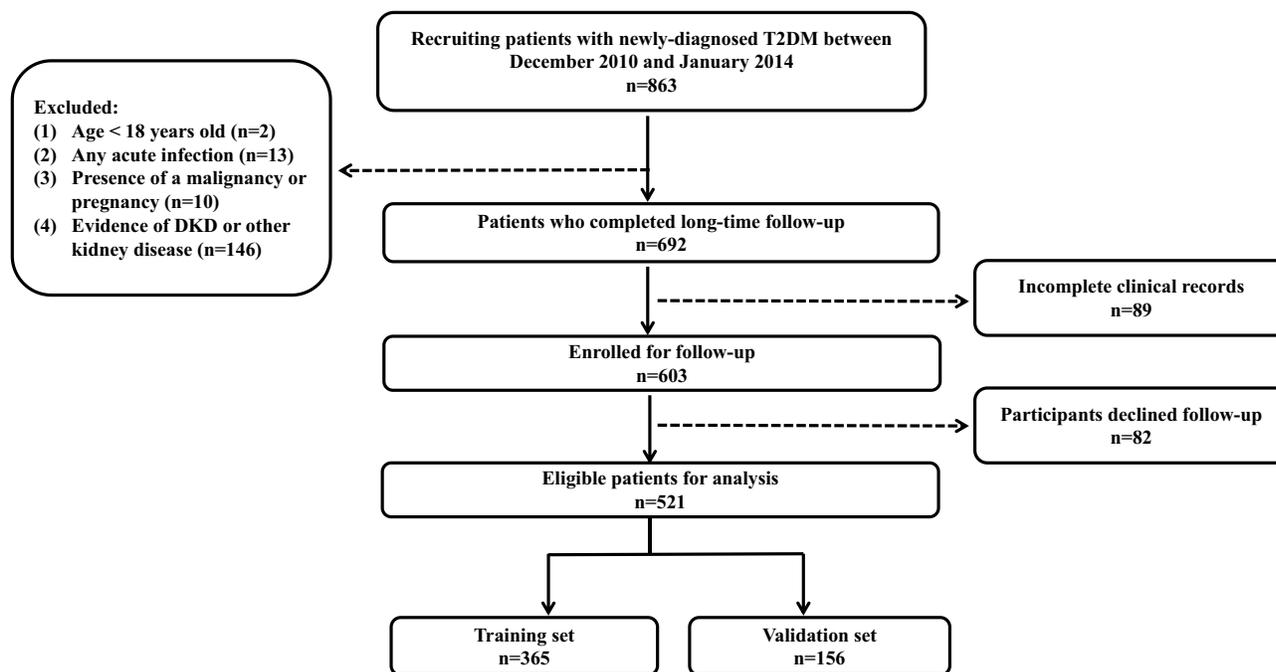
Exclusion criteria: (1) a history of DKD or other kidney diseases; (2) UACR  $\geq$  30mg/g, eGFR < 60 mL/min/1.73m<sup>2</sup> (CKD-EPI),<sup>21</sup> or urine dipstick test  $\geq$  1+; (3) presence of an acute infection; (4) presence of a malignancy or pregnancy; (5) incomplete clinical records.

The endpoint event was the occurrence of DKD,<sup>22</sup> defined as (1) persistent albuminuria (UACR  $\geq$  30mg/g) over 3 months (2) eGFR < 60 mL/min/1.73 m<sup>2</sup>; (3) excluded other kidney diseases.

Participants were followed up until DKD was diagnosed for the first time. Otherwise, patients without DKD events were followed up until June 2022. [Figure 1](#) shows the procession of subject inclusion.

### Data Collection

Baseline demographic data (sex and age) and related clinical data were from the electronic medical record system. The risk prediction model was trained using variables selected from medical record system and published literature,<sup>23,24</sup> including age, sex, body mass index (BMI), history of smoking and hypertension, systolic blood pressure (SBP), fasting blood glucose (FBG), postprandial blood glucose (PBG), haemoglobin A1c (Hb1Ac), fasting C-peptide (FCP), white blood cells (WBC), hemoglobin (Hb), neutrophil to lymphocyte ratio (NLR), fibrinogen (FIB), D-dimer, alanine aminotransferase (ALT), aspartate aminotransferase (AST), serum albumin (ALB), bile acid (BA), bicarbonate, blood urea nitrogen (BUN), uric acid (UA), eGFR, total cholesterol (TC), triglyceride (TG), HDL, low density lipoprotein (LDL), apolipoprotein a1 (Apoa1), apolipoprotein b (ApoB), UACR, TyG index, HOMA-IR, thyroid stimulating hormone (TSH), free triiodothyronine (FT3), free tetraiodothyronine (FT4), diagnosis of diabetic retinopathy (DR) and diabetic peripheral neuropathy (DPN), as well as antidiabetic and antihypertensive treatment. The superior indicator of the HOMA-IR and TyG index for predicting DKD was also determined by comparing the area under the curve (AUC).



**Figure 1** Flow chart of participants.

**Note:** The flow chart shows the entire research process.

**Abbreviations:** T2DM, type 2 diabetes mellitus; DKD, diabetic kidney disease.

## Statistical Analysis

Statistical analyses were performed using SPSS software version 25.0, R software version 4.2.2 and GraphPad Prism software version 9.0.0. The median (interquartile range) [M (P25, P75)] was used to represent data with a non-normal distribution, whereas the mean  $\pm$  standard deviation (SD) was used to express data with a quantitative normal distribution. The qualitative variable was selected as percentages (%). A random number table generated using R software was assigned to the training and validation set in a ratio of 7: 3. There was no statistical difference between the training and validation set for all variables, so the validation set can fully validate the prediction model based on the training set. We screened the variables and built the model based on the training set, and presented the model in a nomogram visualization. In addition, we utilized the validation set to internally validate the performance and clinical utility of the model constructed based on the training set. Spearman correlation was employed for correlation analysis. For further analysis, a nomogram was developed based on multivariate analysis. The logistic regression results were presented as odds ratio (OR), 95% confidence intervals (95% CI), and P values. To obtain the perfect prediction model, we used “glmnet” package for LASSO regression screening to further filter the predictor variables derived from the traditional univariate logistic regression. The nomogram was drawn using the R package “regplot”, and receiver operating characteristic (ROC) analysis was performed to assess how well the nomogram predicted the risk of developing DKD. Use 1000 bootstrap replications for validation. Calibration curves were conducted using the “calibrate” R package. Draw a decision curve using the “rmda” R package.  $P < 0.05$  was considered statistically significant.

## Results

### Baseline Characteristics

The final queue size was 521 persons after follow-up (Figure 1). The average follow-up time of the T2DM participants was  $7.8 \pm 1.8$  years. There were 127 DKD events (24.4%) among the 521 individuals identified in our long-time follow-up cohort.

Table 1 presents the baseline demographic and medication of 521 T2DM patients, meanwhile, Table 2 presents the laboratory measurement indicators. The R package randomly divided the whole individuals into the training set and

**Table 1** Demographic and Medication of Newly Diagnosed T2DM Patients

Variable	Training Set (n=365)			Validation Set (n=156)		
	Non-DKD (n=275)	DKD (n=90)	P value	Non-DKD (n=119)	DKD (n=37)	P value
Age (years)	47.0 (38.0, 58.0)	53.0 (45.0, 62.0)	<0.001	49.0 (41.0, 57.0)	59.0 (53.0, 64.5)	<0.001
Gender (male, %)	175 (63.6)	60.0 (66.7)	0.602	73.0 (61.3)	25.0 (67.6)	0.494
BMI (kg/m <sup>2</sup> )	24.6 (22.9, 26.5)	25.8 (23.2, 28.1)	0.026	24.2 (22.3, 26.2)	25.3 (23.2, 27.0)	0.069
Smoking, n (%)	98.0 (35.6)	45.0 (50.0)	0.015	43.0 (36.1)	17.0 (45.9)	0.284
Hypertension, n (%)	103 (37.5)	51.0 (56.7)	0.001	39.0 (32.8)	23.0 (62.2)	0.001
SBP (mmHg)	134.0 (123.0, 144.0)	136.5 (126.3, 149.8)	0.090	135.0 (125.0, 144.0)	139.0 (127.0, 149.5)	0.258
DR, n (%)	50.0 (18.2)	24.0 (26.7)	0.082	23.0 (19.3)	10.0 (27.0)	0.317
DPN, n (%)	42.0 (15.3)	24.0 (26.7)	0.015	21.0 (17.6)	10.0 (27.0)	0.212
Antidiabetic treatment						
Biguanides, n (%)	209.0 (76.0)	72.0 (80.0)	0.434	89.0 (74.8)	30.0 (81.1)	0.432
Sulfonylureas, n (%)	26.0 (9.5)	11.0 (12.2)	0.450	10.0 (8.4)	7.0 (18.9)	0.073
Thiazolidinediones, n (%)	81.0 (29.5)	28.0 (31.1)	0.766	27.0 (22.7)	12.0 (32.4)	0.232
α-glucosidase inhibitors, n (%)	145.0 (52.7)	47.0 (52.2)	0.934	59.0 (49.6)	20.0 (54.1)	0.634
Nateglinides, n (%)	26.0 (9.5)	15.0 (16.7)	0.060	10.0 (8.4)	2.0 (5.4)	0.550
Intensive insulin therapy, n (%)	169.0 (61.5)	53.0 (58.9)	0.665	67.0 (56.3)	26.0 (70.3)	0.130
Pancreatic kininogenase enteric-coated tablets, n (%)	50.0 (18.2)	26.0 (28.9)	0.030	33.0 (27.7)	8.0 (21.6)	0.461
Antihypertensive treatment						
CCBs, n (%)	51.0 (18.5)	25.0 (27.8)	0.061	24.0 (20.2)	10.0 (27.0)	0.377
ACEIs/ARBs, n (%)	77.0 (28.0)	39.0 (43.3)	0.007	26.0 (21.8)	16.0 (43.2)	0.010
Other antihypertensive drugs, n (%)	9.0 (3.3)	7.0 (7.8)	0.070	3.0 (2.5)	2.0 (5.4)	0.384

**Abbreviations:** T2DM, type 2 diabetes mellitus; BMI, body mass index; SBP, systolic blood pressure; DR, diabetic retinopathy; DPN, diabetic peripheral neuropathy; CCB, calcium channel blocker; ACEI, angiotensin-converting enzyme inhibitor; ARB, angiotensin receptor blocker.

**Table 2** Laboratory Measurement Indicators of Newly Diagnosed T2DM Patients

Variable	Training Set (n=365)			Validation Set (n=156)		
	Non-DKD (n=275)	DKD (n=90)	P value	Non-DKD (n=119)	DKD (n=37)	P value
FBG (mg/dl)	176.4 (140.4, 216.0)	168.3 (142.7, 225.9)	0.359	165.6 (144.0, 207.0)	171.0 (141.3, 208.8)	0.945
PBG (mg/dl)	265.5 ± 72.4	267.1 ± 77.2	0.588	259.5 ± 68.3	256.0 ± 65.3	0.783
HbA1c (%)	10.9 ± 2.5	10.6 ± 2.5	0.332	10.7 ± 2.4	10.7 ± 2.1	0.885
FCP (ng/mL)	1.6 (1.1, 2.3)	2.2 (1.6, 2.8)	0.001	1.7 (1.1, 2.3)	1.6 (1.0, 2.3)	0.621
WBC (10 <sup>9</sup> /L)	6.4 (5.4, 7.6)	6.4 (5.5, 7.6)	0.639	6.5 (5.4, 7.6)	6.5 (5.7, 7.3)	0.924
Hb (g/L)	143.0 (132.0, 153.0)	141.7 (133.0, 150.0)	0.898	144.8 (136.0, 152.3)	140.0 (134.0, 147.0)	0.072
NLR	1.6 (1.2, 2.0)	1.6 (1.2, 2.3)	0.431	1.5 (1.2, 2.0)	1.4 (1.0, 1.8)	0.339
FIB (g/L)	2.5 (2.2, 2.9)	2.4 (2.1, 3.0)	0.857	2.4 (2.2, 2.8)	2.5 (2.2, 3.0)	0.456
D-dimer (ug/L)	140.0 (100.0, 258.8)	146.5 (100.0, 310.0)	0.453	140.0 (100.0, 270.0)	170.0 (100.0, 256.5)	0.924
ALT (U/L)	28.0 (18.0, 43.8)	27.5 (19.5, 52.8)	0.586	25.0 (19.0, 43.0)	22.0 (18.0, 36.5)	0.627
AST (U/L)	16.0 (12.0, 24.0)	15.0 (13.0, 25.0)	0.864	15.0 (12.0, 21.0)	15.0 (11.0, 26.3)	0.754
ALB (g/L)	37.0 (34.4, 39.3)	36.6 (34.5, 40.1)	0.761	37.5 (35.4, 40.0)	35.8 (34.9, 38.0)	0.023
Bicarbonate (mmol/L)	23.7 (22.3, 25.7)	23.9 (22.4, 25.8)	0.542	23.6 (21.9, 25.7)	23.6 (21.3, 25.0)	0.288
BA (mmol/L)	3.6 (2.6, 5.2)	4.6 (2.9, 6.4)	0.012	4.4 (2.9, 6.5)	4.3 (2.9, 5.8)	0.829
BUN (mmol/L)	4.8 (4.0, 5.8)	4.7 (3.9, 6.1)	0.223	4.7 (3.8, 5.7)	5.1 (3.8, 6.6)	0.153
UA (μmol/L)	268.7 (216.6, 317.6)	275.3 (212.3, 357.4)	0.145	255.0 (210.8, 320.5)	288.7 (226.0, 374.5)	0.110
eGFR (mL/min/1.73 m <sup>2</sup> )	113.2 (94.4, 135.5)	86.9 (77.5, 104.6)	<0.001	110.4 (92.5, 136.9)	87.8 (75.5, 105.7)	<0.001
TC (mg/dl)	181.5 (156.6, 211.0)	194.1 (161.4, 224.1)	0.056	191.0 (165.5, 216.5)	191.0 (165.3, 211.7)	0.905
TG (mg/dl)	212.2 (145.7, 292.0)	340.7 (196.0, 542.9)	<0.001	204.7 (151.5, 308.3)	213.5 (155.1, 419.1)	0.412
HDL (mg/dl)	38.1 (32.1, 44.1)	35.4 (30.9, 40.6)	0.001	39.4 (34.0, 45.2)	39.4 (33.6, 43.7)	0.557

(Continued)

**Table 2** (Continued).

Variable	Training Set (n=365)			Validation Set (n=156)		
	Non-DKD (n=275)	DKD (n=90)	P value	Non-DKD (n=119)	DKD (n=37)	P value
LDL (mg/dl)	87.0 (74.3, 107.8)	86.2 (71.2, 108.3)	0.738	92.8 (76.5, 109.8)	94.3 (77.1, 106.9)	0.850
Apoa1 (g/L)	1.2 (1.0, 1.2)	1.2 (1.1, 1.3)	0.306	1.2 (1.1, 1.3)	1.2 (1.1, 1.2)	0.588
Apob (g/L)	1.0 (0.9, 1.2)	1.1 (0.9, 1.4)	<0.001	1.0 (0.9, 1.2)	1.0 (0.9, 1.2)	0.957
UACR (mg/g)	10.0 (6.9, 16.8)	10.7 (7.6, 21.1)	0.023	9.1 (6.4, 14.6)	13.4 (8.6, 23.2)	0.002
TyG index	9.8 (9.4, 10.3)	10.3 (9.9, 10.8)	<0.001	9.8 (9.4, 10.2)	10.0 (9.4, 10.6)	0.032
HOMA-IR	4.3 (3.1, 6.3)	5.9 (4.0, 7.5)	<0.001	4.5 (3.1, 6.2)	4.0 (2.7, 6.5)	0.353
TSH (uIU/mL)	1.9 (1.3, 2.6)	2.1 (1.6, 3.2)	0.077	2.0 (1.4, 2.9)	2.1 (1.3, 2.4)	0.630
FT3 (pmol/L)	4.2 (3.8, 4.7)	4.2 (3.8, 4.6)	0.306	4.3 (3.9, 4.8)	4.1 (3.6, 4.4)	0.036
FT4 (pmol/L)	16.8 (15.0, 18.2)	15.9 (14.5, 18.2)	0.039	16.5 (15.0, 18.6)	16.8 (15.5, 18.2)	0.769

**Abbreviations:** T2DM, type 2 diabetes mellitus; FBG, fasting blood glucose; PBG, postprandial blood glucose; HbA1c, haemoglobin A1c; FCP, fasting c-peptide; WBC, white blood cell; Hb, haemoglobin; NLR, neutrophil to lymphocyte ratio; FIB, fibrinogen; ALT, alanine aminotransferase; AST, aspartate aminotransferase; ALB, albumin; BA, bile acid; BUN, blood urea nitrogen; UA, uric acid; eGFR, estimated glomerular filtration rate; TC, total cholesterol; TG, triglyceride; HDL, high-density lipoprotein; LDL, low-density lipoprotein; Apoa1, apolipoprotein a1; Apob, apolipoprotein b; UACR, urinary albumin creatinine ratio; TyG index, triglyceride-glucose index; TSH, thyroid stimulating hormone; FT3, free triiodothyronine; FT4, free thyroxine.

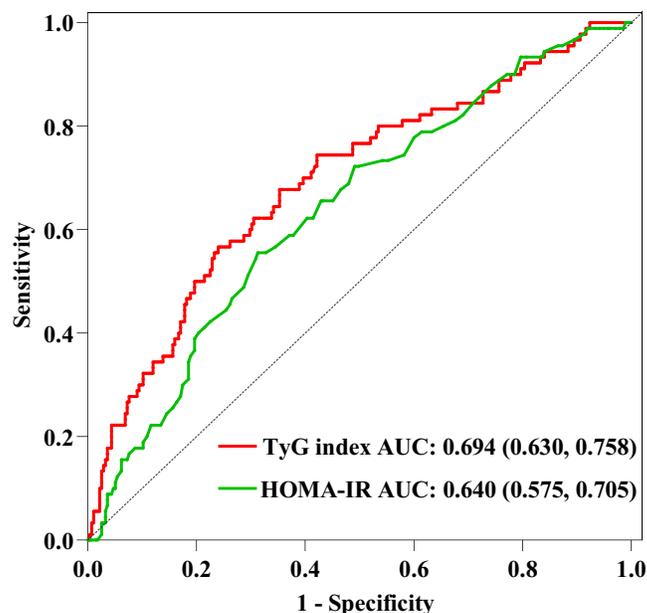
validation set in a ratio of 7: 3. 365 participants (70%) were randomly assigned to the training set and 156 participants (30%) to the validation set. In the training set, 90 patients (24.7%) progressed to DKD, while 37 patients (23.7%) progressed to DKD in the validation set. There was no significant difference in demographic and clinical characteristics between the training and the validation set ( $P > 0.05$ ). However, variables of significance include DKD participants being on average older in age, having higher FCP, higher BMI, higher BA, higher TG, higher Apob, higher UACR, higher HOMA-IR, and TyG index, and having greater proportion with a history of smoking and hypertension, a higher incidence of DPN, and a higher utilization rate of some antidiabetic and antihypertensive drugs while having lower eGFR and HDL (Table 1 and Table 2). In addition, we also develop logistic and LASSO regression, aiming to identify the related factors with DKD.

## Comparison of TyG Index and HOMA-IR

To evaluate the performance of the HOMA-IR and TyG index in predicting DKD, we plotted AUC. HOMA-IR and TyG index are both tools to assess IR. We determined the AUCs for the TyG index and HOMA-IR in the training set, were 0.694 (95% CI 0.630 to 0.758) and 0.640 (95% CI 0.575 to 0.705), respectively (Figure 2). The findings were following those reported in other studies, which found that TyG index was more accurate than HOMA-IR at predicting the risk of IR and DM-related complications.<sup>25,26</sup> In addition, multivariate analysis revealed that the HOMA-IR was no statistical significance after adjusting confounding variables (Table 3).

## LASSO Regression Analysis

To identify the independent significant factors, we used logistic regression and LASSO regression for analysis. We conducted a correlation analysis and univariate logistic regression (Figure 3 and Table 3). Correlation analysis found a strong collinearity correlation among some variables. LASSO regression, a greedy algorithm, can be used to select the most significant variables to avoid overfitting, then establish a more accurate linear regression model.<sup>27</sup> According to previous reports and univariate analysis, we selected 37 variables related to DKD (age, gender, BMI, smoking, hypertension, SBP, HbA1c, FBG, PBG, ALB, BUN, UA, TC, TG, HDL, LDL, Apoa1, Apob, UACR, eGFR, FCP, FT4, TyG index, HOMA-IR, DR, DPN, antihypertensive and antidiabetic treatment) for LASSO analysis. Figure 4A shows the cross-validation curve result, and the coefficient distribution plot is shown in Figure 4B. Through 10-fold cross-validation, we chose the minimum value within 1 standard error and obtained 5 independent variables that were finally screened out (Table 4).



**Figure 2** ROC curve of TyG index and HOMA-IR.

**Notes:** ROC curve of TyG index and HOMA-IR drawn according to the training set. We calculated the AUC of the TyG index and HOMA-IR, which were 0.694 (95% CI 0.630 to 0.758) and 0.640 (95% CI 0.575 to 0.705), respectively.

**Abbreviations:** ROC, receiver operating characteristic; AUC, area under the curve; TyG index, triglyceride-glucose index; HOMA-IR, homeostasis model assessment for insulin resistance.

## Development of an Individualized Prediction Model

To display the prediction model vividly, we applied an individualized dynamic nomogram to visualize it (Figure 5). Then, we applied ROC to analyze the performance of the prediction model. Figure 6 shows the AUC of the prediction model. For the training set, the predicted AUC of the nomogram was 0.826 (95% CI 0.775 to 0.876), meanwhile, the AUC of the validation set was 0.803 (95% CI 0.719 to 0.887).

**Table 3** Logistic Regression Analysis of DKD-Related Risk Factor

Variables	Univariate Logistics Regression		Multivariate Logistics Regression	
	OR (95% CI)	P value	OR (95% CI)	P value
Age	1.037 (1.017, 1.057)	<0.001	1.007 (0.973, 1.042)	0.687
Gender	1.143 (0.691, 1.889)	0.602		
BMI	1.099 (1.021, 1.183)	0.012	1.072 (0.960, 1.198)	0.217
Smoking	1.806 (1.116, 2.923)	0.016	1.819 (0.909, 3.643)	0.091
Hypertension	2.184 (1.347, 3.540)	0.002	0.847 (0.359, 1.998)	0.704
SBP	1.013 (0.999, 1.028)	0.062		
FBG	1.004 (0.999, 1.008)	0.117		
PBG	1.001 (0.998, 1.004)	0.587		
HbA1c	0.953 (0.864, 1.051)	0.331		
FCP	1.538 (1.180, 2.004)	0.001	2.130 (0.592, 7.658)	0.247
WBC	1.055 (0.877, 1.152)	0.943		
Hb	0.994 (0.979, 1.010)	0.487		
NLR	1.076 (0.844, 1.371)	0.557		
FIB	0.988 (0.727, 1.342)	0.937		
D-dimer	1.000 (1.000, 1.001)	0.342		

(Continued)

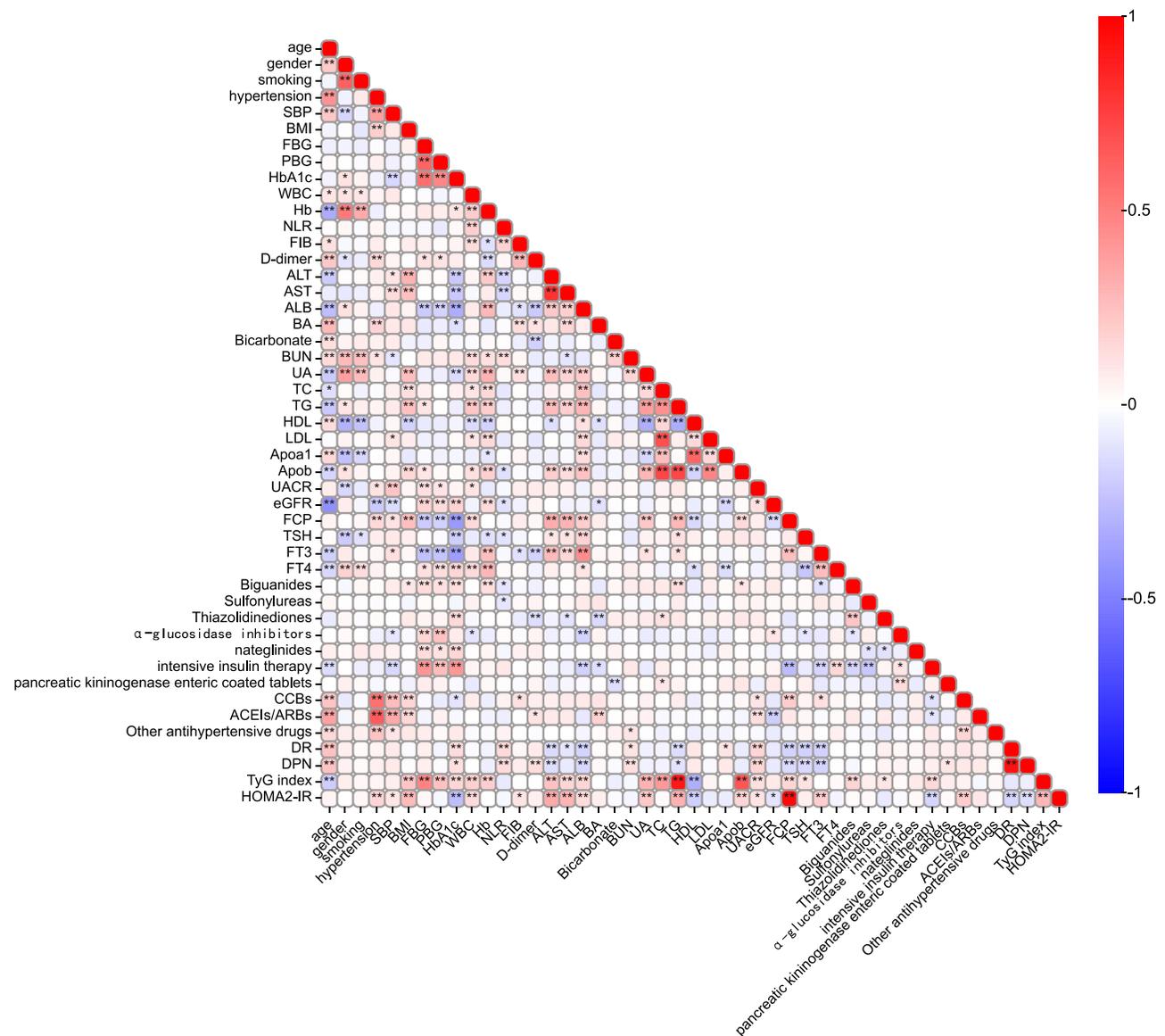
Table 3 (Continued).

Variables	Univariate Logistics Regression		Multivariate Logistics Regression	
	OR (95% CI)	P value	OR (95% CI)	P value
ALT	0.999 (0.991, 1.007)	0.790		
AST	0.991 (0.974, 1.008)	0.302		
ALB	0.991 (0.934, 1.052)	0.773		
Bicarbonate	1.031 (0.949, 1.120)	0.467		
BA	1.108 (1.032, 1.189)	0.005	1.037 (0.942, 1.142)	0.453
BUN	1.081 (0.943, 1.239)	0.263		
UA	1.003 (1.000, 1.006)	0.073		
eGFR	0.970 (0.960, 0.981)	<0.001	0.962 (0.946, 1.017)	<0.001
TC	1.006 (1.001, 1.011)	0.021	0.999 (0.996, 1.002)	0.549
TG	1.002 (1.001, 1.003)	<0.001	1.001 (0.990, 1.012)	0.835
HDL	0.953 (0.924, 0.982)	0.002	0.981 (0.945, 1.017)	0.300
LDL	1.000 (0.992, 1.009)	0.930		
Apoa1	2.387 (0.493, 11.550)	0.626		
Apob	3.512 (1.688, 7.306)	0.001		
UACR	1.049 (1.016, 1.084)	0.004	1.036 (0.998, 1.086)	0.145
TyG index	2.682 (1.875, 3.836)	<0.001	4.021 (1.158, 13.969)	0.029
HOMA-IR	1.194 (1.084, 1.316)	<0.001	0.820 (0.512, 1.315)	0.411
TSH	1.154 (1.003, 1.327)	0.046		
FT3	0.903 (0.675, 1.207)	0.490		
FT4	0.888 (0.806, 0.979)	0.017	0.882 (0.757, 1.028)	0.107
DR	1.636 (0.936, 2.861)	0.084		
DPN	2.017 (1.140, 3.571)	0.016	2.890 (1.198, 6.972)	0.018
Antidiabetic treatment				
Biguanides	1.263 (0.703, 2.270)	0.485		
Sulfonylureas	1.333 (0.631, 2.820)	0.451		
Thiazolidinediones	1.082 (0.646, 1.812)	0.766		
$\alpha$ -glucosidase inhibitors	0.980 (0.608, 1.578)	0.934		
Nateglinides	1.915 (0.965, 3.803)	0.063		
Intensive insulin therapy	0.898 (0.553, 1.459)	0.665		
Pancreatic kininogenase enteric-coated tablets	1.828 (1.056, 3.166)	0.031	1.579 (0.678, 3.676)	0.290
Antihypertensive treatment				
CCBs	1.689 (0.972, 2.935)	0.063		
ACEIs/ARBs	1.966 (1.201, 3.220)	0.007	1.602 (0.942, 1.142)	0.292
Other antihypertensive drugs	2.493 (0.901, 6.899)	0.079		

**Abbreviations:** DKD, diabetic kidney disease; BMI, body mass index; SBP, systolic blood pressure; FBG, fasting blood glucose; PBG, postprandial blood glucose; HbA1c, haemoglobin A1c; FCP, fasting c-peptide; WBC, white blood cell; Hb, haemoglobin; NLR, neutrophil to lymphocyte ratio; FIB, fibrinogen; ALT, alanine aminotransferase; AST, aspartate aminotransferase; ALB, albumin; BA, bile acid; BUN, blood urea nitrogen; UA, uric acid; eGFR, estimated glomerular filtration rate; TC, total cholesterol; TG, triglyceride; HDL, high-density lipoprotein; LDL, low-density lipoprotein; Apo a1, apolipoprotein a1; Apob, apolipoprotein b; UACR, urinary albumin creatinine ratio; TyG index, triglyceride-glucose index; TSH, thyroid stimulating hormone; FT3, free triiodothyronine; FT4, free thyroxine; DR, diabetic retinopathy; DPN, diabetic peripheral neuropathy; CCB, calcium channel blocker; ACEI, angiotensin-converting enzyme inhibitor; ARB, angiotensin receptor blocker.

## Calibration of the Model

To validate the nomogram, we used 1000 bootstrap analysis and draw a calibration curve. The calibration curves of the model revealed a strong correlation between the predicted probability and the actual probability, indicating that our model had been properly calibrated (Figure 7).



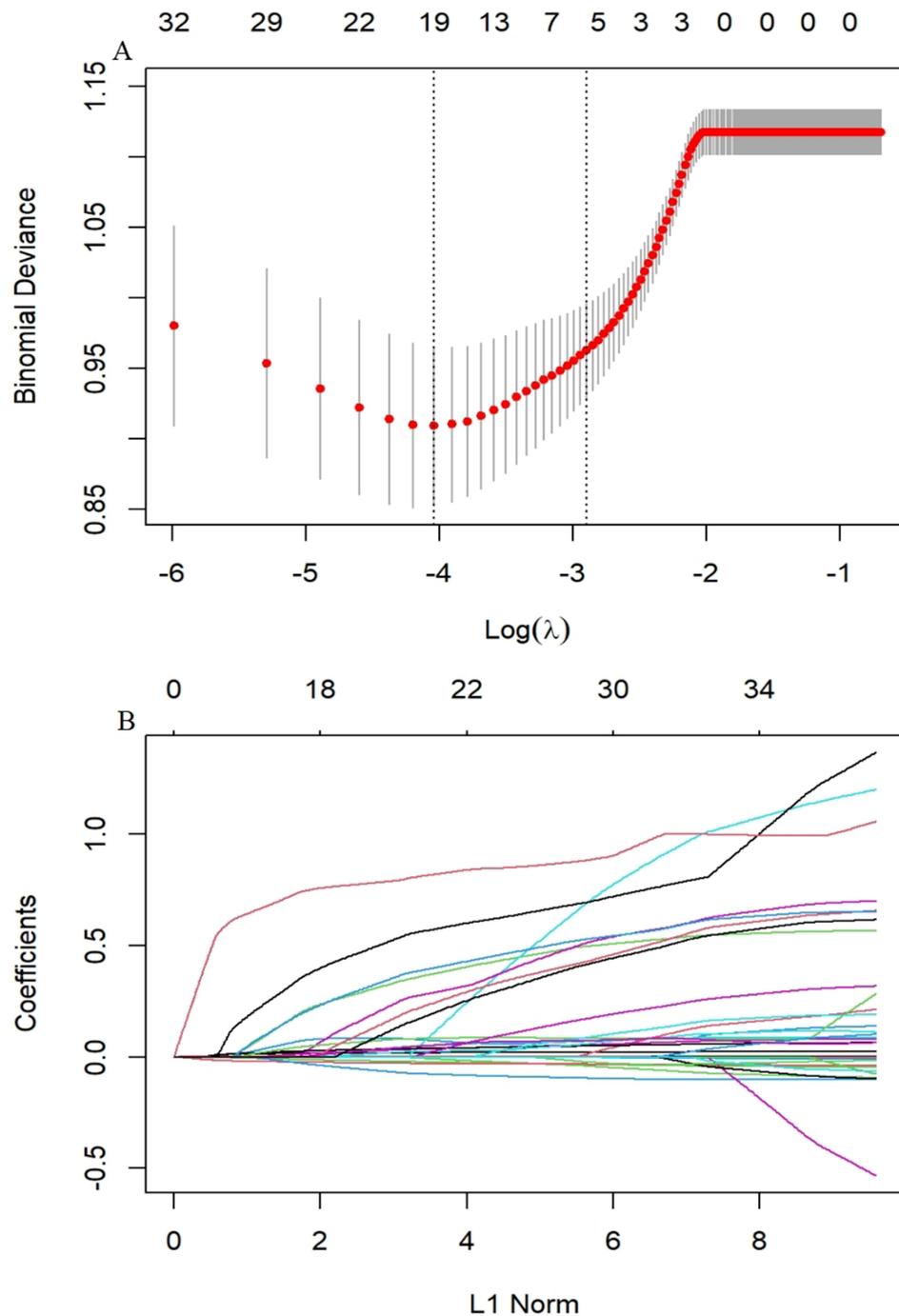
**Figure 3** Clinical indicators correlation analysis.

**Notes:** \*P value < 0.05, \*\*P value < 0.001, significant correlation between variables.

**Abbreviations:** BMI, body mass index; SBP, systolic blood pressure; FBG, fasting blood glucose; PBG, postprandial blood glucose; HbA1c, haemoglobin A1c; FCP, fasting c-peptide; WBC, white blood cell; Hb, haemoglobin; NLR, neutrophil to lymphocyte ratio; FIB, fibrinogen; ALT, alanine aminotransferase; AST, aspartate aminotransferase; ALB, albumin; BA, bile acid; BUN, blood urea nitrogen; UA, uric acid; eGFR, estimated glomerular filtration rate; TC, total cholesterol; TG, triglyceride; HDL, high-density lipoprotein; LDL, low-density lipoprotein; Apoa1, apolipoprotein a1; Apob, apolipoprotein b; UACR, urinary albumin creatinine ratio; TyG index, triglyceride-glucose index; TSH, thyroid stimulating hormone; FT3, free triiodothyronine; FT4, free thyroxine; DR, diabetic retinopathy; DPN, diabetic peripheral neuropathy; CCB, calcium channel blocker; ACEI, angiotensin-converting enzyme inhibitor; ARB, angiotensin receptor blocker.

## Clinical Use of the Model

To prove the clinical practicability of the nomogram, we carried out a decision curve analysis (DCA). Figure 8 presents the DCA for the DKD nomogram. Between 20% and 40% of T2DM patients progressed to DKD,<sup>28</sup> and within this cutoff, both decision curves were above the None and All lines. This demonstrated that predicting DKD in newly diagnosed T2DM using the nomogram model provides more clinical benefits than either treatment or no treatment for all patients.



**Figure 4** LASSO regression analysis.

**Notes:** (A) Cross-validation curve. (B) Coefficient distribution plot of 37 variables. The cross-validated error within one standard error of the minimum is shown on the right vertical line, while the minimum error is shown on the left vertical line.

**Abbreviation:** LASSO, less absolute shrinkage and selection operator.

## Discussion

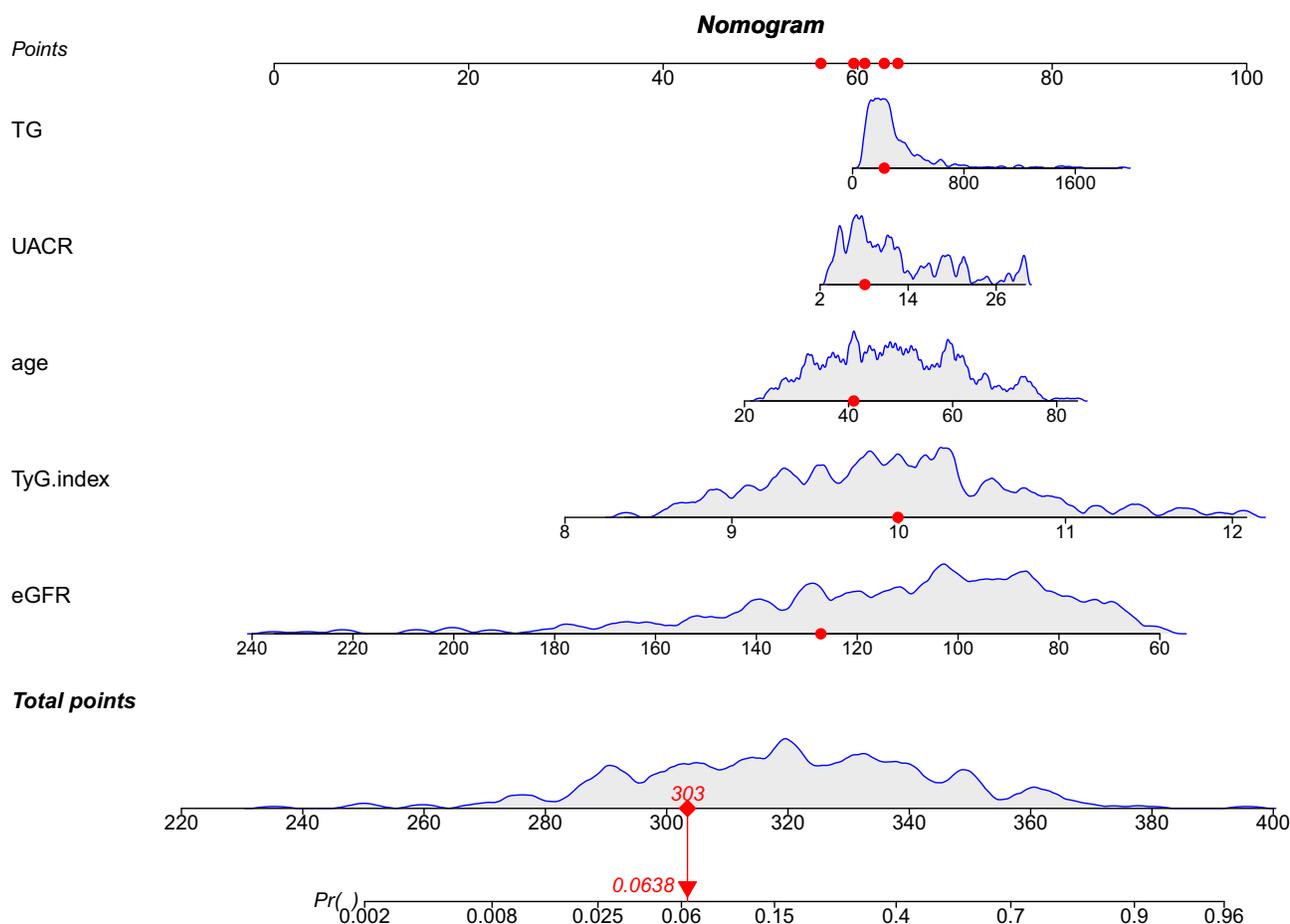
Our study was a single-center analysis of prospectively collected data from 521 newly diagnosed T2DM patients, 127 of whom had a DKD event. We screened variables through LASSO regression and incorporated 5 variables to construct a prediction model. Furthermore, we carried out internal validation, discrimination ability, calibration ability, and clinical practicability of the model. These results revealed that the prediction model performed well and had clinical applicability.

**Table 4** LASSO Regression Analysis of DKD-Related Risk Factor

Variables	Coefficient
Age	0.0242
TG	-0.0002
UACR	0.0634
eGFR	0.0377
TyG index	1.0575

**Abbreviations:** LASSO, less absolute shrinkage and selection operator; TG, triglyceride; UACR, urinary albumin creatinine ratio; eGFR, estimated glomerular filtration rate; TyG index, triglyceride-glucose index.

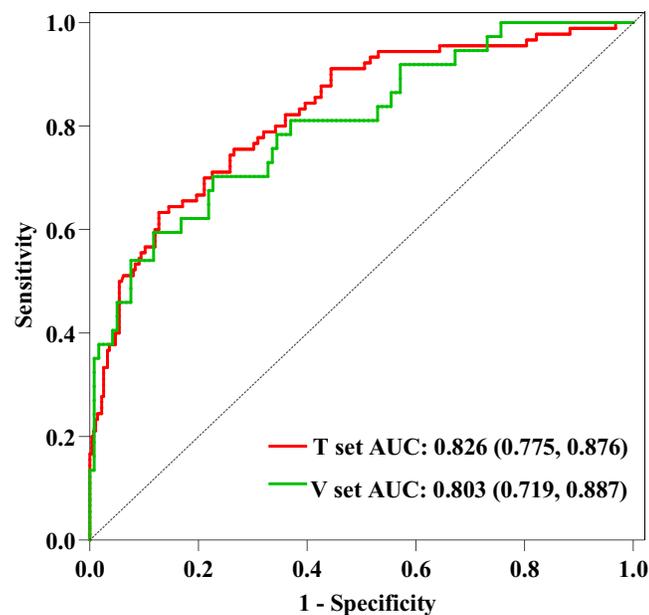
Independent prediction indicators were screened by logistic and LASSO regression, and finally, a DKD prediction model for predicting newly diagnosed T2DM was developed. Nomogram revealed that age, renal excretion, lipid metabolism, and IR may be potential good predictors of DKD. Age is recognized as an independent risk factor for T2DM and DKD.<sup>29</sup> Persistent microalbuminuria (UACR > 30mg/g) can appear in the early stage of the DKD hyperfiltration state, and metabolic substances such as serum creatinine will accumulate with the decrease of renal filtration level in the later period.<sup>30</sup> At present, some



**Figure 5** Dynamic nomogram developed for predicting DKD in patients with newly diagnosed T2DM.

**Notes:** Dynamic nomogram can reflect the score of each variable of patients in actual situation, so as to obtain the corresponding DKD risk probability. The risk probability of DKD calculated for the first patient in the cohort by each variable score is shown in the figure above.

**Abbreviations:** TG, triglyceride; eGFR, estimated glomerular filtration rate; UACR, urinary albumin creatinine ratio; TyG index, triglyceride-glucose index; DKD, diabetic kidney disease; T2DM, type 2 diabetes mellitus.



**Figure 6** ROC curves of the predictive model.

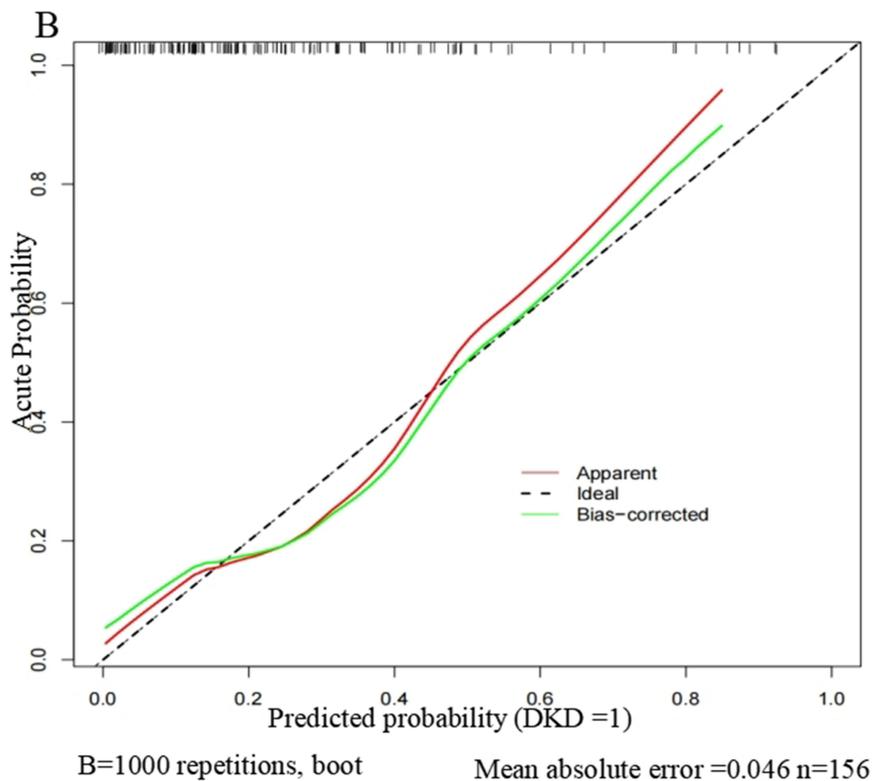
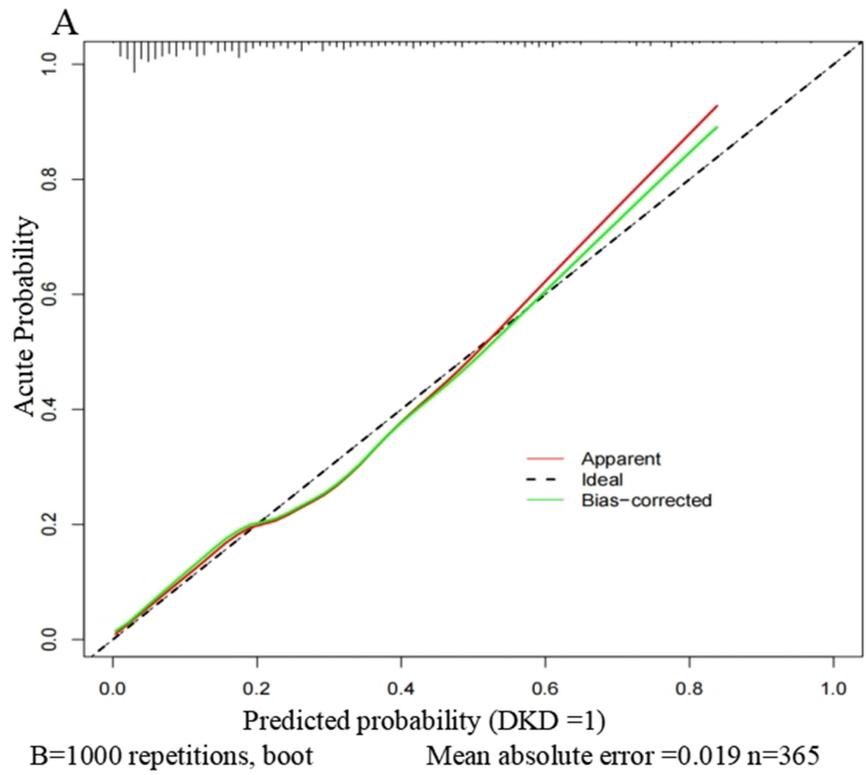
**Notes:** For the training set, the predicted AUC of the nomogram was 0.826 (95% CI 0.775 to 0.876), meanwhile, the AUC of the validation set was 0.803 (95% CI 0.719 to 0.887).

**Abbreviations:** ROC, Receiver operating characteristic; AUC, area under the curve; V set, validation set; T set, training set.

scholars considered that mitochondrial dysfunction played an indispensable role in DN, which will cause fatty acid oxidation disorder, resulting in abnormal lipid deposition in the kidney.<sup>31</sup> Furthermore, many lipid metabolites have been confirmed to be related to the complications of DM. As an index of IR, the TyG index performs excellently in predicting complications of DM.<sup>32,33</sup> TyG index can accurately evaluate whether IR exists only by combining two simple indexes: serum lipid and glucose. Moreover, an Australian study investigated the relationship between the TyG index and end-stage renal disease (ESRD). The results revealed a positive correlation between the TyG index and ESRD risk.<sup>9</sup> Therefore, researchers should not ignore the relationship between IR and DKD in the future.

The onset of DKD is hidden, and it may be difficult for patients newly diagnosed with T2DM to realize the risk of developing DKD. However, if DKD occurs, we can only take symptomatic treatment (control albuminuria and serum glucose), then wait for ESRD to evolve slowly. Therefore, incorporating the clinical variables described above into the model and developing a nomogram as a screening tool can well identify newly diagnosed T2DM patients with DKD risk. For the identification of our model, the AUC of the training set and the validation set were 0.826 and 0.803, respectively. These results revealed that the model had an excellent predictive ability in identifying DKD and non-DKD in T2DM patients. The model's calibration curve revealed a strong correlation between the predicted probability and the actual probability, indicating that our model had been properly calibrated. DCA results showed that using this model to predict DKD had clinical practicability and benefits patients.

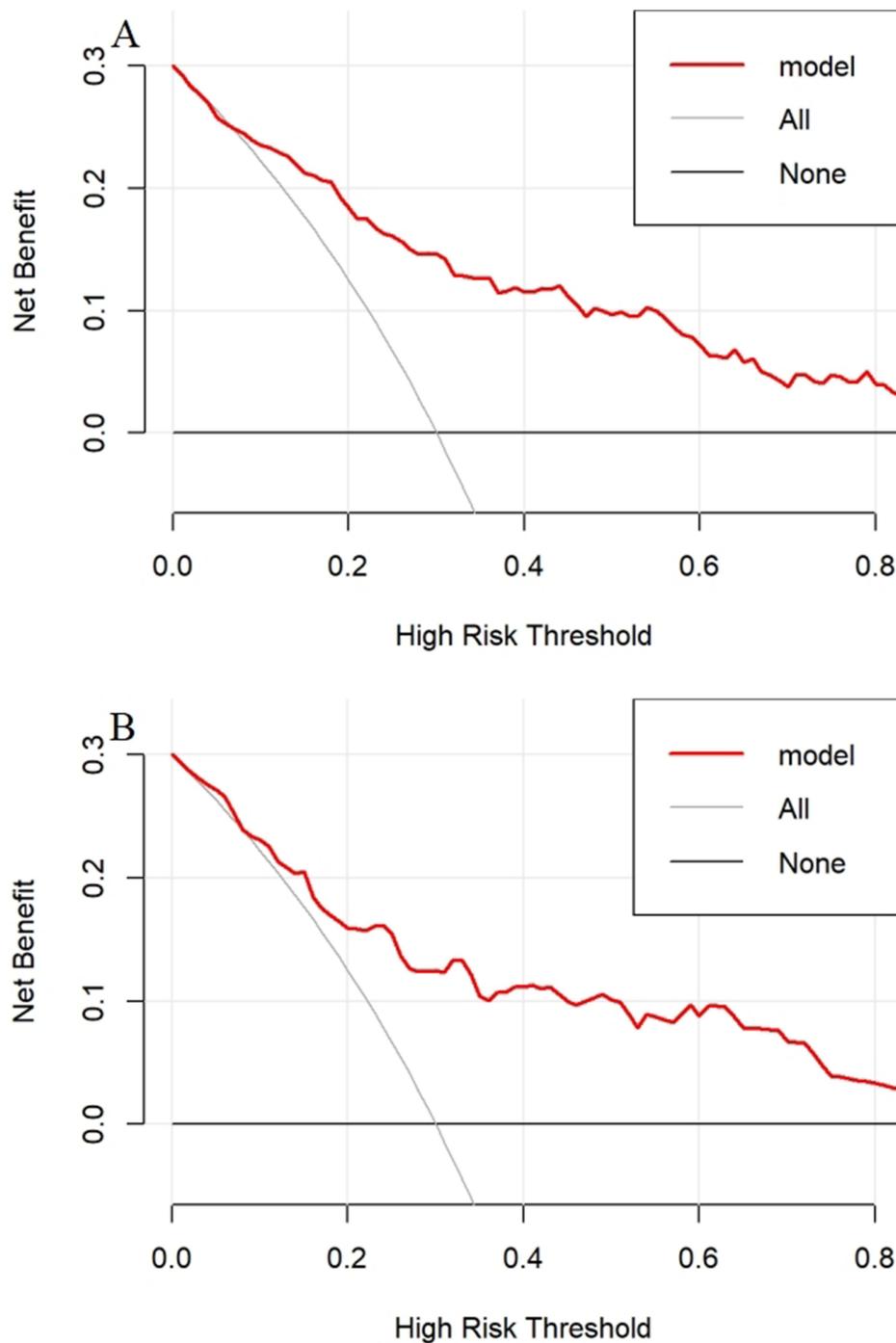
In comparison to previous studies on developing DKD prediction models, our study had some strengths. First, our study was a retrospective cohort study with a long-time follow-up. Some patients with DKD have a DM duration of 10 years or more, so it may be difficult to collect enough participants with positive outcomes in short-term follow-up and underestimate the prevalence rate, which ultimately affects the predictive performance. Secondly, the participants in our study were newly diagnosed T2DM patients compared with other research focus individuals, who frequently overlooked the danger of developing DKD. Newly diagnosed T2DM patients are a concern in this study, and the prediction model of DKD is constructed to find all DKD events in DM patients earlier. Finally, to incorporate all related variables, we initially investigated the pathological process and risk factors for T2DM and its complications. IR is a major pathogenic factor and characteristic of T2DM, causing concomitant injury to various target organs, and DKD is no exception.<sup>34,35</sup> It is crucial to find an index for assessing IR since it is the chief culprit of glucose metabolism disorders. The performance of the TyG index and the HOMA-IR as IR evaluation indexes varies throughout other studies. Hence, we used AUC to distinguish the performance of the TyG



**Figure 7** Calibration curves of the nomogram.

**Notes:** (A) Calibration curve of the model in the training group. (B) Calibration curve of the model in the validation group. The calibration curve reflects the consistency between the predicted probability and the actual probability. Y-axis = actual probability of DKD. X-axis = predicted probability of DKD. The shadow line represents a perfect prediction by an ideal model. The red line represents the performance of our model, which coincides well with the shadow line.

**Abbreviation:** DKD, diabetic kidney disease.



**Figure 8** DCA for the nomogram.

**Notes:** (A) Decision curve of the model in the training group. (B) Decision curve of the model in the validation group. The y-axis measures the net benefit. The solid red line represents the nomogram. The solid gray line represents the assumption that all patients have DKD. The solid black line represents the assumption that no patients have DKD.

**Abbreviations:** DCA, decision curve analysis; DKD, diabetic kidney disease.

index from the HOMA-IR and finally revealed that the TyG index performs better. Therefore, we incorporated significant correlation factors including the TyG index, constructing a more accurate prediction model.

There were still some limitations in this trial. First, the lack of external validation was one of the important limitations of our study. To duplicate and externally verify the results of this study, more investigation is required. Secondly, this was

a cohort study with limited sample size, so the endpoint event was defined as the occurrence of DKD, and no subgroup analysis was carried out. If DKD was divided into subgroups with different renal function levels according to eGFR stages, the research results would be more abundant. In addition, the family history of DM was missing from baseline data, and it is undeniable that it is also a significant factor in the development of DKD. Finally, since this was a retrospective cohort study, it may suffer from recollection bias and loss of follow-up. In this cohort study, there were 82 participants with loss to follow-up, which was controlled within 10% and had no significant impact on the outcome.

## Conclusion

In conclusion, the new DKD progression model for T2DM patients constructed in this study had excellent identification and calibration ability, which was helpful for clinical practice. In the future, we need to integrate large-scale T2DM cohort studies to further verify this nomogram prediction model.

## Data Sharing Statement

The data underlying this article will be shared on reasonable request to the corresponding author.

## Ethics Statement

This research study was complied with the Declaration of Helsinki, and approved by the Ethics Committee of the Third Affiliated Hospital of Soochow University (2013#27). All individuals provided written informed consent for this research study.

## Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

## Funding

This work was supported by grants from the National Natural Science Foundation of China (82000684) and Changzhou Sci & Tech Program (CE20215024) and Top Talent of Changzhou “The 14th Five-Year Plan” High-Level Health Talents Training Project (2022260).

## Disclosure

The authors report no conflicts of interest in this work.

## References

1. Sun H, Saeedi P, Karuranga S, et al. IDF diabetes atlas: global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045. *Diabetes Res Clin Pract.* 2022;183:109119. doi:10.1016/j.diabres.2021.109119
2. Afkarian M, Zelnick LR, Hall YN, et al. Clinical manifestations of kidney disease among US adults with diabetes, 1988–2014. *JAMA.* 2016;316(6):602–610. doi:10.1001/jama.2016.10924
3. Alicic RZ, Rooney MT, Tuttle KR. Diabetic kidney disease: challenges, progress, and possibilities. *Clin J Am Soc Nephrol.* 2017;12(12):2032–2045. doi:10.2215/CJN.11491116
4. Colhoun HM, Marcovecchio ML. Biomarkers of diabetic kidney disease. *Diabetologia.* 2018;61(5):996–1011. doi:10.1007/s00125-018-4567-5
5. Di Bonito P, Mozzillo E, Rosanio FM, et al. Albuminuric and non-albuminuric reduced eGFR phenotypes in youth with type 1 diabetes: factors associated with cardiometabolic risk. *NMCD.* 2021;31(7):2033–2041. doi:10.1016/j.numecd.2021.03.019
6. Tonnejck L, Muskiet MHA, Smits MM, et al. Glomerular hyperfiltration in diabetes: mechanisms, clinical significance, and treatment. *JASN.* 2017;28(4):1023–1039. doi:10.1681/ASN.2016060666
7. Chan DT, Watts GF, Irish AB, Dogra GK. Insulin resistance and vascular dysfunction in chronic kidney disease: mechanisms and therapeutic interventions. *Nephrol Dial Transplant.* 2017;32(8):1274–1281. doi:10.1093/ndt/gfv326
8. Nakashima A, Kato K, Ohkido I, Yokoo T. Role and treatment of insulin resistance in patients with chronic kidney disease: a review. *Nutrients.* 2021;13(12):4349. doi:10.3390/nu13124349
9. Chamroonkiadtikun P, Ananchaisarp T, Wanichanon W. The triglyceride-glucose index, a predictor of type 2 diabetes development: a retrospective cohort study. *Prim Care Diabetes.* 2020;14(2):161–167. doi:10.1016/j.pcd.2019.08.004

10. Liu L, Xia R, Song X, et al. Association between the triglyceride-glucose index and diabetic nephropathy in patients with type 2 diabetes: a cross-sectional study. *J Diabetes Investig.* 2021;12(4):557–565. doi:10.1111/jdi.13371
11. Luciano RL, Moeckel GW. Update on the native kidney biopsy: core curriculum 2019. *Am J Kidney Dis.* 2019;73(3):404–415. doi:10.1053/j.ajkd.2018.10.011
12. Elahi D. In praise of the hyperglycemic clamp. A method for assessment of beta-cell sensitivity and insulin resistance. *Diabetes Care.* 1996;19(3):278–286. doi:10.2337/diacare.19.3.278
13. Matthews DR, Hosker JP, Rudenski AS, Naylor BA, Treacher DF, Turner RC. Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. *Diabetologia.* 1985;28(7):412–419. doi:10.1007/BF00280883
14. Association AD. 11. Microvascular complications and foot care. *Diabetes Care.* 2019;42(Suppl 1):S124–S138. doi:10.2337/dc19-S011
15. Tao L-C, J-n X, Wang -T-T, Hua F, J-j L. Triglyceride-glucose index as a marker in cardiovascular diseases: landscape and limitations. *Cardiovasc Diabetol.* 2022;21(1):68. doi:10.1186/s12933-022-01511-x
16. Tang G, Qi L, Sun Z, et al. Evaluation and analysis of incidence and risk factors of lower extremity venous thrombosis after urologic surgeries: a prospective two-center cohort study using LASSO-logistic regression. *Int J Surg.* 2021;89:105948. doi:10.1016/j.ijvsu.2021.105948
17. Shao Y, Jia H, Huang L, et al. An original ferroptosis-related gene signature effectively predicts the prognosis and clinical status for colorectal cancer patients. *Front Oncol.* 2021;11:711776. doi:10.3389/fonc.2021.711776
18. He J, Chen Z, Xue Q, et al. Identification of molecular subtypes and a novel prognostic model of diffuse large B-cell lymphoma based on a metabolism-associated gene signature. *J Transl Med.* 2022;20(1):186. doi:10.1186/s12967-022-03393-9
19. McEligot AJ, Poynor V, Sharma R, Panangadan A. Logistic LASSO regression for dietary intakes and breast cancer. *Nutrients.* 2020;12(9):2652. doi:10.3390/nu12092652
20. Wu L, Zhou B, Liu D, et al. LASSO regression-based diagnosis of acute ST-Segment Elevation Myocardial Infarction (STEMI) on Electrocardiogram (ECG). *J Clin Med.* 2022;11(18):5408. doi:10.3390/jcm11185408
21. Levey AS, Stevens LA, Schmid CH, et al. A new equation to estimate glomerular filtration rate. *Ann Intern Med.* 2009;150(9):604–612. doi:10.7326/0003-4819-150-9-200905050-00006
22. Lameire NH, Levin A, Kellum JA, et al. Harmonizing acute and chronic kidney disease definition and classification: report of a Kidney Disease: improving Global Outcomes (KDIGO) consensus conference. *Kidney Int.* 2021;100(3):516–526. doi:10.1016/j.kint.2021.06.028
23. Dong Z, Wang Q, Ke Y, et al. Prediction of 3-year risk of diabetic kidney disease using machine learning based on electronic medical records. *J Transl Med.* 2022;20(1):143. doi:10.1186/s12967-022-03339-1
24. Yamanouchi M, Furuichi K, Shimizu M, et al. Serum hemoglobin concentration and risk of renal function decline in early stages of diabetic kidney disease: a nationwide, biopsy-based cohort study. *Nephrol Dial Transplant.* 2022;37(3):489–497. doi:10.1093/ndt/gfab185
25. Gao W, Wang J, Chen Y, et al. Discordance between the triglyceride glucose index and HOMA-IR in incident albuminuria: a cohort study from China. *Lipids Health Dis.* 2021;20(1):176. doi:10.1186/s12944-021-01602-w
26. Wang S, Shi J, Peng Y, et al. Stronger association of triglyceride glucose index than the HOMA-IR with arterial stiffness in patients with type 2 diabetes: a real-world single-centre study. *Cardiovasc Diabetol.* 2021;20(1):82. doi:10.1186/s12933-021-01274-x
27. J-y H, Wang Y, Tong X-M, Yang T. When to consider logistic LASSO regression in multivariate analysis? *Eur J Surg Oncol.* 2021;47(8):2206. doi:10.1016/j.ejso.2021.04.011
28. Persson F, Rossing P. Diagnosis of diabetic kidney disease: state of the art and future perspective. *Kidney Int Suppl.* 2018;8(1):2–7. doi:10.1016/j.kisu.2017.10.003
29. Jiang W, Wang J, Shen X, et al. 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(4):925–933. doi:10.2337/dc19-1897
30. Ricciardi CA, Gnudi L. Kidney disease in diabetes: from mechanisms to clinical presentation and treatment strategies. *Metabolism.* 2021;124:154890. doi:10.1016/j.metabol.2021.154890
31. Forbes JM, Thorburn DR. Mitochondrial dysfunction in diabetic kidney disease. *Nat Rev Nephrol.* 2018;14(5):291–312. doi:10.1038/nrneph.2018.9
32. Park K, Ahn CW, Lee SB, et al. Elevated TyG index predicts progression of coronary artery calcification. *Diabetes Care.* 2019;42(8):1569–1573. doi:10.2337/dc18-1920
33. Wu Z, Zhou D, Liu Y, et al. Association of TyG index and TG/HDL-C ratio with arterial stiffness progression in a non-normotensive population. *Cardiovasc Diabetol.* 2021;20(1):134. doi:10.1186/s12933-021-01330-6
34. Penno G, Solini A, Orsi E, et al. Insulin resistance, diabetic kidney disease, and all-cause mortality in individuals with type 2 diabetes: a prospective cohort study. *BMC Med.* 2021;19(1):66. doi:10.1186/s12916-021-01936-3
35. Wang Z, Xie J, Wang J, Feng W, Liu N, Liu Y. Association between a novel metabolic score for insulin resistance and mortality in people with diabetes. *Front Cardiovasc Med.* 2022;9:895609. doi:10.3389/fcvm.2022.895609

## Diabetes, Metabolic Syndrome and Obesity

Dovepress

### Publish your work in this journal

Diabetes, Metabolic Syndrome and Obesity is an international, peer-reviewed open-access journal committed to the rapid publication of the latest laboratory and clinical findings in the fields of diabetes, metabolic syndrome and obesity research. Original research, review, case reports, hypothesis formation, expert opinion and commentaries are all considered for publication. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/diabetes-metabolic-syndrome-and-obesity-journal>