British Journal of Hospital Medicine

Br. J. Hosp. Med. (Lond) **2026**; 87(11): 55626
https://doi.org/10.31083/BJHM55626

Article

Development and Validation of a Diagnostic Prediction Model for Diabetic Kidney Disease in Overweight and Obese Patients With Type 2 Diabetes Mellitus

Xiling Hu[1,2,3,†], Qunying Lin[3,†], Xinqi Zhuang[1], Yin-Ping Zhang[1,*]

[1]School of Nursing, Xi'an Jiaotong University Health Science Center, 710061 Xi'an, Shaanxi, China
[2]Department of Endocrinology and Metabolism, Guangdong Provincial Key Laboratory of Diabetology & Guangzhou Municipal Key Laboratory of Mechanistic and Translational Obesity Research, The Third Affiliated Hospital of Sun Yat-sen University, 510630 Guangzhou, Guangdong, China
[3]Nursing Department, Zhaoqing Hospital, The Third Affiliated Hospital of Sun Yat-sen University, 526070 Zhaoqing, Guangdong, China
*Correspondence: cathyzh@mail.xjtu.edu.cn (Yin-Ping Zhang)
[†]These authors contributed equally.
Academic Editor: Sudarshan Ramachandran
Submitted: 5 July 2026 Revised: 20 September 2026 Accepted: 8 October 2026 Published: 10 November 2026

Abstract

Aims/Background: Overweight and obesity are common in patients with type 2 diabetes mellitus (T2DM) and increase the risk of diabetic kidney disease (DKD), but evidence on associated factors and prediction models in this high-risk population remains limited. We aimed to develop and evaluate a diagnostic prediction model for prevalent DKD in patients with T2DM and overweight or obesity. **Methods:** This retrospective cross-sectional study included 1903 patients. The development cohort comprised 1438 patients treated at the Third Affiliated Hospital of Sun Yat-sen University between January 2021 and December 2023 and was randomly split into training (n = 1009) and internal validation (n = 429) cohorts. A separate cohort of 465 patients treated at another hospital of the Third Affiliated Hospital of Sun Yat-sen University between January and June 2024 was used for external validation. Predictors were selected using univariable screening, least absolute shrinkage and selection operator (LASSO) regression, and stepwise multivariable logistic regression. The final model was presented as a nomogram. Performance was evaluated using the area under the curve (AUC), calibration curves and statistics, Brier scores, and decision curve analysis. **Results:** Overall, 316 patients (16.6%) had DKD, with 173 (17.1%), 59 (13.8%), and 84 (18.1%) in the training, internal validation, and external validation cohorts, respectively. The final model included five predictors: diabetes duration, systolic blood pressure, serum albumin, total cholesterol, and cystatin C. The AUCs were 0.782 (95% confidence interval [CI], 0.742–0.822), 0.753 (95% CI, 0.668–0.839), and 0.786 (95% CI, 0.728–0.844) in the training, internal validation, and external validation cohorts, respectively. In sex-stratified analyses, AUCs were numerically higher in men than in women across all three cohorts (training: 0.806 vs 0.745; internal validation: 0.760 vs 0.754; external validation: 0.793 vs 0.766), although confidence intervals overlapped. Calibration curves generally followed the ideal line. Decision curve analysis suggested potential clinical benefit over a range of threshold probabilities. **Conclusion:** The five-predictor nomogram demonstrated acceptable discriminative performance for prevalent DKD in patients with T2DM and overweight or obesity, offering a tool to facilitate the identification of patients requiring further diagnostic assessment. Nevertheless, broader external validation and evaluation of its clinical impact are warranted.

Keywords: diabetic kidney disease; type 2 diabetes mellitus; overweight/obesity; diagnostic prediction model

1. Introduction

Type 2 diabetes mellitus (T2DM) is a progressive metabolic disorder characterized by persistent hyperglycemia. An estimated 589 million adults worldwide had diabetes in 2024, and this population has been projected to grow to 852.5 million by 2050 [1]. With approximately 148.0 million adults affected in 2024, China accounted for the largest national population with diabetes. This total is projected to rise to 168.3 million by 2050 [1]. This substantial burden makes the assessment of diabetes-related complications an important clinical priority.

Overweight and obesity are highly prevalent metabolic conditions that are closely linked to the development of T2DM [2]. Changes in lifestyle and dietary patterns have contributed to the increasing prevalence of excess body weight worldwide [3]. By 2030, more than 2.9 billion adults worldwide are projected to be living with overweight, including approximately 1.1 billion with obesity [2]. In China, 41% of adults were found to be living with overweight and 9% had obesity in 2025 [2]. Moreover, the prevalence of diabetes increases progressively with body mass index (BMI), reaching 12.8% among overweight individuals and 18.5% among those with obesity [4]. Overweight and obesity may contribute to the occurrence of diabetes-related complications. As a frequent microvascular complication of diabetes, diabetic kidney disease (DKD) contributes substantially to the burden of chronic kidney disease (CKD) and end-stage renal disease (ESRD). It is estimated that DKD affects up to 36% of individuals with diabetes in China [5,6].

Copyright: © 2026 The Author(s). Published by IMR Press.
This is an open access article under the CC BY 4.0 license.

Publisher's Note: IMR Press stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Evidence suggests that excess body weight may further increase the risk of DKD among patients with T2DM, with each 1 kg/m^2 BMI rise conferring ~3% higher risk [7]. Thus, overweight and obese patients with T2DM may represent a particularly vulnerable population for the development of DKD.

Given the substantial disease burden and adverse clinical consequences of DKD, early identification of patients at increased risk is essential for timely intervention and prevention of disease progression [7]. However, evidence regarding factors associated with DKD among overweight and obese patients with T2DM remains limited, and few prediction models have been developed specifically for this high-risk population. Risk prediction models derived from the general T2DM population may not fully capture the risk profile associated with excess body weight. Therefore, this study aimed to identify factors associated with DKD and to develop and evaluate a diagnostic prediction model for DKD among overweight and obese patients with T2DM.

2. Methods

2.1 Participants and Study Design

Electronic medical records from the Third Affiliated Hospital of Sun Yat-sen University (Guangdong, China) were used for this retrospective cross-sectional study. The model development cohort comprised 1438 eligible patients who attended the Department of Endocrinology and Metabolism between January 2021 and December 2023. Patients were randomly split in a 7:3 ratio, with 1009 assigned to the training cohort and 429 to the internal validation cohort.

The external validation cohort comprised 465 patients treated at the Zhaoqing Hospital between January and June 2024, selected using the same eligibility criteria. Although the Zhaoqing Hospital operates within the same hospital system and follows the same ethical review framework and clinical quality standards, it provided an independent patient cohort from a different campus and a later period. Data from this cohort were used exclusively for external validation. Fig. 1 outlines the participant selection process.

2.2 Eligibility Criteria and Outcome Definition

2.2.1 Inclusion Criteria

Patients were eligible if they were aged ≥18 years, had an established diagnosis of T2DM according to the 2021 American Diabetes Association criteria [8], and had a BMI ≥24 kg/m^2. Overweight was defined as 24 ≤ BMI < 28 kg/m^2, and obesity as BMI ≥28 kg/m^2 [9].

2.2.2 Outcome Definition

Among patients with established diabetes, DKD was diagnosed when kidney abnormalities were attributable to diabetes and at least one of the following two conditions was satisfied: (1) urinary albumin-to-creatinine ratio (UACR) ≥30 mg/g in two or more of three samples obtained over 3–6 months; or (2) estimated glomerular filtration rate (eGFR) <60 mL/min/1.73 m^2 for longer than 3 months [6,8]. The eGFR was calculated using the Chinese-modified version of the Modification of Diet in Renal Disease (MDRD) equation developed by Ma et al. [10]: eGFR = 175 × (Scr)$^{-1.234}$ × (age)$^{-0.179}$ × 0.79 (if female), where Scr is serum creatinine in mg/dL, age is expressed in years, and eGFR is expressed in mL/min/1.73 m^2. Previously documented CKD was an exclusion criterion, whereas DKD status was determined during the study assessment according to the above criteria.

2.2.3 Exclusion Criteria

Any of the following conditions precluded inclusion in this study: (1) a documented diagnosis or history of non-diabetic chronic kidney disease (CKD), or CKD of unknown etiology, prior to the study assessment; (2) kidney disease or renal injury attributable to causes other than diabetes, including urinary tract infection–related renal injury, acute kidney injury, renal malignancy, or other non-diabetic causes; (3) receipt of kidney replacement therapy, including dialysis or kidney transplantation; (4) missing values in over half of the study variables.

2.3 Predictor Variables

In this study, CNKI (https://www.cnki.net/), WanFang Data (https://www.wanfangdata.com.cn/), the VIP database (https://www.cqvip.com/), CBM (SinoMed; https://www.sinomed.ac.cn/), PubMed (https://pubmed.ncbi.nlm.nih.gov/), the Cochrane Library (https://www.cochranelibrary.com/), Web of Science (https://www.webofscience.com/), and Embase (https://www.embase.com/) were searched for relevant publications published between January 2019 and December 2023. Search terms included “diabetic kidney disease” or “diabetic nephropathy”, combined with “predictor”, “risk factor”, “predictive factor”, “associated factor”, or “correlated factor”. No age-specific search term was used; however, the population was restricted to adults aged ≥18 years at the study selection stage. Studies restricted to participants aged <18 years were excluded. The literature search was completed in January 2024.

Following completion of the literature search and review, an expert consultation was conducted in January 2024 to assess the clinical relevance and applicability of the literature-derived candidate predictors. This consultation was conducted as an online structured expert consultation meeting involving brainstorming discussion. Ten healthcare professionals involved in diabetes care were purposively selected, all holding an associate senior professional title or above. The panel comprised three endocrinologists, two nephrologists, three endocrinology nurses, and two nephrology nurses. The experts reviewed the candidate predictors in the context of prevalent DKD in patients with T2DM and overweight or obesity. Each candidate predictor was discussed in terms of its clinical relevance and

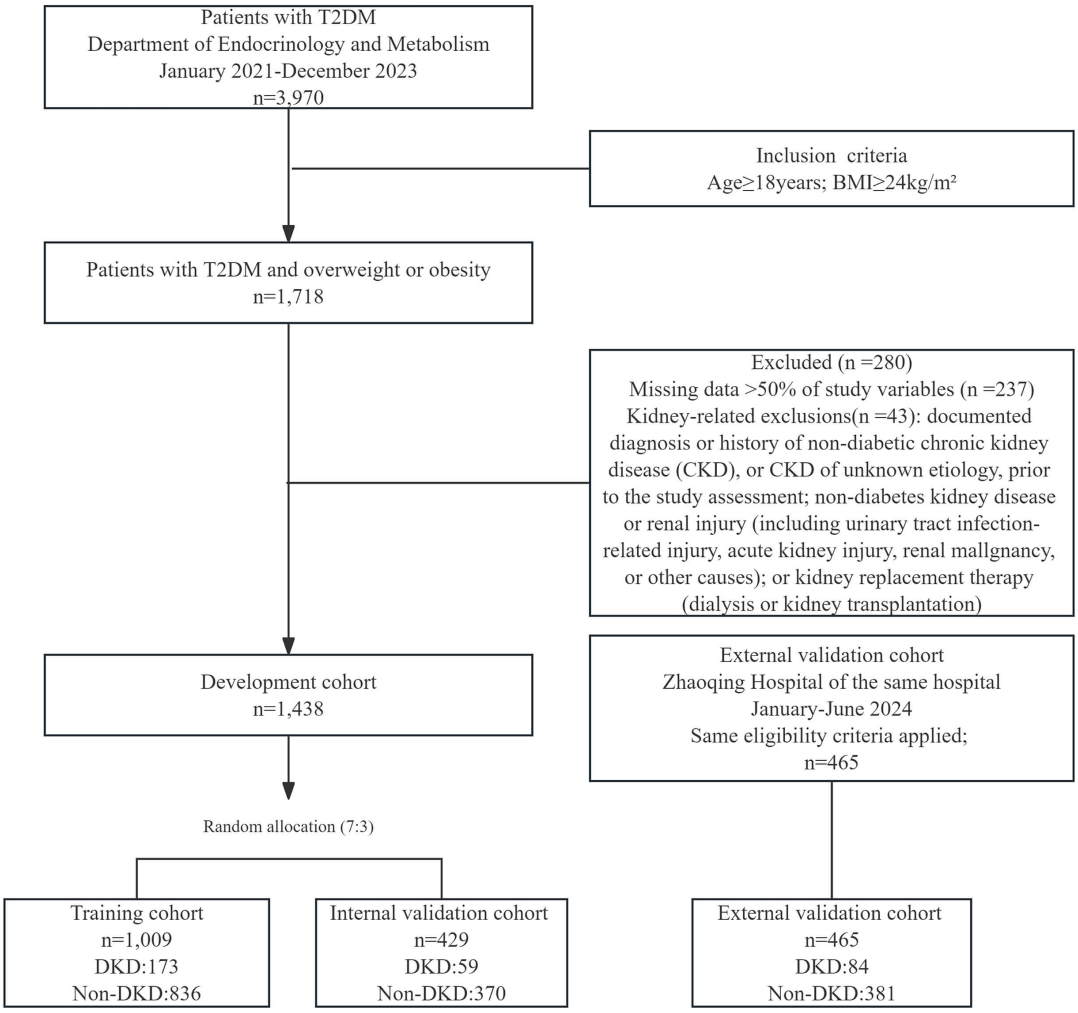


Fig. 1. Participant selection and cohort allocation. **Notes:** A total of 3970 patients with type 2 diabetes mellitus (T2DM) were screened. Of these, 2252 did not meet the inclusion criteria (age ≥18 years and BMI ≥24 kg/m^2), and 280 were excluded (237 for >50% missing data and 43 for kidney-related reasons). The development cohort (n = 1438) was randomly split at a 7:3 ratio into the training cohort (n = 1009; DKD: 173, non-DKD: 836) and the internal validation cohort (n = 429; DKD: 59, non-DKD: 370). The external validation cohort included 465 patients from the Zhaoqing Hospital (DKD: 84, non-DKD: 381). Overall, 1903 participants were included (DKD: 316; non-DKD: 1587). BMI, body mass index; CKD, chronic kidney disease; DKD, diabetic kidney disease.

applicability. Consensus was reached through discussion and agreement among the participating experts; no formal quantitative voting or Delphi procedure was used. Disagreements were resolved through further discussion until agreement was reached on the final list of predictors.

The variables assessed were grouped into demographic and lifestyle characteristics, clinical history, anthropometric and blood pressure measurements, and laboratory measures. Demographic and lifestyle variables included sex, age, smoking status, and alcohol consumption. Clinical history included diabetes duration and recorded comorbidities: hypertension, hyperlipidemia, hyperuricemia, gout, osteoporosis, gallstones, cerebral infarction, coronary heart disease (CHD), hepatitis, tuberculosis, malignancy, nephrolithiasis, chronic obstructive pulmonary disease (COPD), obstructive sleep apnea-hypopnea syndrome (OSAHS), peptic ulcer disease, and other comorbidities. Anthropometric and blood pressure variables included waist circumference, hip circumference, BMI, waist-to-hip ratio (WHR), systolic blood pressure (SBP), and diastolic blood pressure (DBP).

Laboratory variables included alanine aminotransferase (ALT), serum albumin (ALB), gamma-glutamyl transferase (GGT), fasting blood glucose (GLU), blood urea nitrogen (BUN), uric acid (UA), total cholesterol (CHOL), low-density lipoprotein cholesterol (LDL-C), high-density lipoprotein cholesterol (HDL-C), triglycerides (TRIG), creatine kinase (CK), creatine kinase-MB isoenzyme (CK-MB), cystatin C (CysC), glycated hemoglobin A1c (HbA1c), and fasting, 0.5-hour post-load, and 2-hour post-load insulin.

2.4 Data Collection

Study variables were retrieved retrospectively from electronic medical records. BMI was derived from weight divided by height squared, using kilograms and meters, respectively; WHR was calculated from waist and hip circumferences. Glucose and fasting insulin measurements followed an overnight fast of at least 8 hours. The insulin assessment also included samples obtained 0.5 and 2 hours after a 75-gram oral glucose load. ALB was measured using the bromocresol green method.

Comorbidity indicators were based on diagnoses documented in the medical records. Gout and hyperuricemia were coded independently; documentation of one condition did not establish the presence of the other. Study laboratory measurements were not used to retrospectively reclassify these diagnoses.

For laboratory variables evaluated as potential predictors, the first available measurements during admission were used, except for post-load insulin measurements, which corresponded to the specified sampling times. Both campuses followed the quality control standards of the National Center for Clinical Laboratories and used uniform specimen collection and processing procedures. Two researchers independently verified the extracted data. Disagreements were resolved through discussion, and when necessary, a senior researcher would be consulted for resolution.

2.5 Statistical Analysis

Analyses in this study were carried out using R version 4.3.3 (R Foundation for Statistical Computing, Vienna, Austria), with a two-sided significance threshold of 0.05. The Kolmogorov–Smirnov test was performed to assess data distribution. Continuous variables with a normal distribution were expressed as mean ± standard deviation, and differences between two independent groups were assessed using the independent-samples *t*-test. Other continuous variables that did not conform to a normal distribution were expressed as median (Q1, Q3). Comparisons between patients with and without DKD were performed using the Mann–Whitney *U* test, whereas comparisons across the three cohorts were performed using the Kruskal–Wallis *H* test. Categorical data were reported as frequencies and percentages. For 2 × 2 tables, test selection depended on the smallest expected frequency: Pearson's chi-square test for values ≥5, Yates' continuity correction for values from 1 to <5, and Fisher's exact test for values <1. For 2 × 3 tables, Pearson's test required every expected frequency to be ≥1 and at least 80% to be ≥5; tables failing either condition were evaluated using the R × C extension of Fisher's exact test.

After application of the eligibility criteria, missing values accounted for less than 1% of observations for each variable in the development cohort. The development cohort was first randomly split into a training cohort (70%) and an internal validation cohort (30%) using R's sample function. Given the very low proportion of missing values, we used random forest imputation (implemented using the missRanger package) with predictive mean matching to impute missing data. The imputation model was fitted only in the training cohort and subsequently applied to the internal validation cohort. The external validation cohort had no missing values and was used exclusively for evaluating model performance. This approach retained available participant information and maintained a consistent sample across predictor-selection and model-development analyses. No missing values remained after imputation.

Predictor selection was performed exclusively using the training data. Each candidate variable was first examined in a separate logistic regression model. To minimize the risk of omitting potentially relevant predictors, a more liberal threshold of Wald-test p-value < 0.10 was applied at the univariate screening stage. Variables meeting this criterion were subsequently entered into the least absolute shrinkage and selection operator (LASSO) logistic regression model. Ten-fold cross-validation was used to select the optimal penalty parameter. We used lambda.1se, selecting the largest penalty whose cross-validated binomial deviance remained within one standard error of the minimum. Variables retaining nonzero coefficients then underwent bidirectional stepwise selection in a multivariate logistic model. Variance inflation factors (VIFs) were assessed for the retained predictors, with a VIF of 10 or greater indicating substantial multicollinearity. The resulting model was translated into a nomogram.

Model performance was assessed in terms of discrimination, calibration, overall prediction error, and potential clinical utility. Receiver operating characteristic (ROC) analysis provided the area under the curve (AUC) and its 95% confidence interval (CI) for each cohort. Classification thresholds were optimized separately within each cohort using the maximum of sensitivity + specificity – 1. Sensitivity, specificity, and accuracy were then calculated at that cohort's selected threshold. These analyses were exploratory because threshold selection and evaluation used the same participants within each cohort. Calibration was examined graphically and summarized by mean and maximum absolute calibration errors, supplemented by the Hosmer–Lemeshow and Spiegelhalter tests. Brier scores were used to quantify overall prediction error. Decision curve analysis was performed to evaluate potential clinical utility. Sex-stratified analyses were performed within each cohort, with ROC curves, areas under the curves (AUCs), and corresponding 95% CIs reported separately for men and women.

3. Results

3.1 Participant Characteristics

The analysis included 1903 eligible participants: 1009 in the training cohort, 429 in the internal validation cohort,

and 465 in the external validation cohort. Table 1 summarizes their demographic and clinical profiles. Comparisons across cohorts were performed to characterize differences in case mix relevant to the interpretation of external model performance. The three cohorts were generally comparable in demographic and clinical characteristics, with no statistically significant differences for most variables. However, significant between-group differences were observed for hyperuricemia and DBP ($p < 0.05$). Regarding DKD-related outcomes, the proportions of patients with DKD were 17.1% (173/1009), 13.8% (59/429), and 18.1% (84/465) in the training, internal validation, and external validation cohorts, respectively.

Overall, 316 participants (16.6%) had DKD and 1587 (83.4%) did not. Descriptive comparisons according to DKD status in the overall study population are presented in Table 2. Participants with and without DKD differed significantly in several clinical and biochemical characteristics, including age, diabetes duration, hypertension, osteoporosis, CHD, cerebral infarction, nephrolithiasis, peptic ulcer disease, BMI, WHR, waist circumference, SBP, fasting insulin, 0.5-hour post-load insulin, 2-hour post-load insulin, ALT, ALB, BUN, UA, total cholesterol, HDL-C, triglycerides, CK-MB, CysC and other comorbidity (all $p < 0.05$). In the external validation cohort, significant differences were observed between patients with and without DKD for age, diabetes duration, hypertension, cerebral infarction, CHD, nephrolithiasis, SBP, waist, ALB, BUN, UA, total cholesterol, triglycerides, CysC, fasting insulin, and WHR (all $p < 0.05$; **Supplementary Table 1**).

3.2 Selection of Predictors

Univariate logistic regression analyses identified 18 variables (p-value < 0.10) associated with DKD in the training cohort: age, diabetes duration, hypertension, cerebral infarction, CHD, nephrolithiasis, SBP, waist circumference, BMI, ALB, GGT, BUN, UA, total cholesterol, HDL-C, CK-MB, CysC, and fasting insulin (**Supplementary Table 2**). These 18 variables were entered into the LASSO regression. LASSO regression retained six predictors at the penalty parameter selected using the one-standard-error criterion (lambda.1se): diabetes duration, SBP, ALB, BUN, total cholesterol, and CysC (Fig. 2A,B). Subsequent multivariate logistic regression with bidirectional stepwise selection identified diabetes duration (odds ratio [OR] 1.037, 95% confidence interval [CI] 1.009–1.066; $p = 0.009$), SBP (OR 1.020, 95% CI 1.009–1.030; $p < 0.001$), ALB (OR 0.904, 95% CI 0.865–0.944; $p < 0.001$), CHOL (OR 1.286, 95% CI 1.131–1.463; $p < 0.001$), and CysC (OR 2.983, 95% CI 1.402–6.351; $p = 0.005$) as significant predictors of DKD. BUN was not significantly associated with DKD (OR 1.076, 95% CI 0.974–1.188; $p = 0.152$) and was therefore excluded from the final model. The final model comprised diabetes duration, SBP, ALB, CHOL, and CysC. The results of the multivariate logistic regression analysis used for predictor selection are presented in Table 3.

Variance inflation factors (VIFs) for diabetes duration, SBP, ALB, total cholesterol, and CysC were 1.16, 1.05, 1.17, 1.06, and 2.51, respectively, indicating no substantial multicollinearity among the selected predictors (**Supplementary Table 3**).

3.3 Development of the Diagnostic Prediction Model

The final five-predictor logistic regression model was presented as a nomogram to estimate the probability of prevalent DKD in patients with T2DM and overweight or obesity (Fig. 3). Values for diabetes duration, SBP, ALB, total cholesterol, and CysC are converted to points, whose sum determines the estimated probability of prevalent DKD.

3.4 Model Performance in the Training and Internal Validation Cohorts

The model achieved an AUC of 0.782 (95% CI, 0.742–0.822) in the training cohort and 0.753 (95% CI, 0.668–0.839) in the internal validation cohort (Fig. 4A,B). The cohort-specific thresholds determined by maximizing Youden's index were 0.168 for the training cohort and 0.239 for the internal validation cohort. At these respective thresholds, sensitivity, specificity, and accuracy were 66.5%, 77.8%, and 75.8% in the training cohort and 61.0%, 89.7%, and 85.8% in the internal validation cohort.

The calibration curve closely followed the ideal line in the training cohort, whereas some deviation from the ideal line was observed at intermediate predicted probabilities in the internal validation cohort. The apparent and bias-corrected curves closely overlapped in both cohorts (Fig. 5A,B). The mean absolute calibration errors were 0.010 and 0.031, and the maximum absolute calibration errors were 0.054 and 0.145 in the training and internal validation cohorts, respectively. The corresponding Brier scores were 0.109 and 0.081. Spiegelhalter's test did not indicate statistically significant miscalibration in either cohort ($p = 0.710$ and $p = 0.813$, respectively). Similarly, the Hosmer–Lemeshow test showed no statistically significant lack of fit in the training cohort ($\chi^2 = 7.919$, $p = 0.441$) or the internal validation cohort ($\chi^2 = 10.338$, $p = 0.242$).

3.5 Model Performance in the External Validation Cohort

Based on the results of the external validation cohort, an AUC of 0.786 (95% CI, 0.728–0.844; Fig. 6A) was observed. A threshold of 0.200 maximized Youden's index within this cohort, giving a sensitivity of 64.3%, specificity of 83.5%, and accuracy of 80.0%. The calibration curve remained broadly close to the reference diagonal (Fig. 6B). Mean and maximum absolute calibration errors were 0.018 and 0.056, and the Brier score was 0.110. Calibration tests were not statistically significant: Spiegelhalter $Z = 0.028$, $p = 0.978$; Hosmer–Lemeshow $\chi^2 = 6.118$, $p = 0.634$.

Table 1. Demographic and clinical characteristics of participants.

Variables	Training cohort (n = 1009)	Internal validation cohort (n = 429)	External validation cohort (n = 465)	Test statistic	*p*-value
Group (%)				$\chi^2 = 3.448$	0.178
Non-DKD	836 (82.9)	370 (86.2)	381 (81.9)		
DKD	173 (17.1)	59 (13.8)	84 (18.1)		
Sex (%)				$\chi^2 = 0.128$	0.938
Male	622 (61.6)	267 (62.2)	284 (61.1)		
Female	387 (38.4)	162 (37.8)	181 (38.9)		
Age (years)	55.00 (46.00, 65.00)	54.00 (46.00, 64.00)	56.00 (46.00, 65.00)	$H = 1.061$	0.588
Smoker (%)				$\chi^2 = 0.941$	0.625
No	702 (69.6)	297 (69.2)	312 (67.1)		
Yes	307 (30.4)	132 (30.8)	153 (32.9)		
Alcohol (%)				$\chi^2 = 0.467$	0.792
No	809 (80.2)	340 (79.3)	366 (78.7)		
Yes	200 (19.8)	89 (20.7)	99 (21.3)		
Diabetes duration (years)	6.00 (2.00, 11.00)	6.00 (2.00, 11.00)	6.00 (2.00, 11.00)	$H = 0.403$	0.818
Hypertension (%)				$\chi^2 = 1.110$	0.574
No	606 (60.1)	249 (58.0)	267 (57.4)		
Yes	403 (39.9)	180 (42.0)	198 (42.6)		
Hyperlipidemia (%)				$\chi^2 = 0.798$	0.671
No	938 (93.0)	397 (92.5)	437 (94.0)		
Yes	71 (7.0)	32 (7.5)	28 (6.0)		
Hyperuricemia (%)				$\chi^2 = 6.115$	0.047
No	992 (98.3)	416 (97.0)	461 (99.1)		
Yes	17 (1.7)	13 (3.0)	4 (0.9)		
Gout (%)				$\chi^2 = 0.892$	0.640
No	965 (95.6)	414 (96.5)	443 (95.3)		
Yes	44 (4.4)	15 (3.5)	22 (4.7)		
Osteoporosis (%)				$\chi^2 = 0.587$	0.746
No	997 (98.8)	423 (98.6)	461 (99.1)		
Yes	12 (1.2)	6 (1.4)	4 (0.9)		
Gallstone (%)				$\chi^2 = 0.188$	0.910
No	984 (97.5)	420 (97.9)	454 (97.6)		
Yes	25 (2.5)	9 (2.1)	11 (2.4)		
Cerebral infarction (%)				$\chi^2 = 4.284$	0.117
No	963 (95.4)	398 (92.8)	438 (94.2)		
Yes	46 (4.6)	31 (7.2)	27 (5.8)		
CHD (%)				$\chi^2 = 0.963$	0.618
No	916 (90.8)	385 (89.7)	415 (89.2)		
Yes	93 (9.2)	44 (10.3)	50 (10.8)		
Hepatitis (%)				$\chi^2 = 0.239$	0.887
No	904 (89.6)	388 (90.4)	418 (89.9)		
Yes	105 (10.4)	41 (9.6)	47 (10.1)		
Tuberculosis (%)				$\chi^2 = 0.936$	0.626
No	991 (98.2)	424 (98.8)	456 (98.1)		
Yes	18 (1.8)	5 (1.2)	9 (1.9)		
Malignancy (%)				$\chi^2 = 0.484$	0.785
No	989 (98.0)	418 (97.4)	455 (97.8)		
Yes	20 (2.0)	11 (2.6)	10 (2.2)		
Nephrolithiasis (%)				$\chi^2 = 3.866$	0.145
No	957 (94.8)	416 (97.0)	439 (94.4)		
Yes	52 (5.2)	13 (3.0)	26 (5.6)		

Table 1. Continued.

Variables	Training cohort (n = 1009)	Internal validation cohort (n = 429)	External validation cohort (n = 465)	Test statistic	*p*-value
COPD (%)				Fisher	0.605
No	1008 (99.9)	428 (99.8)	464 (99.8)		
Yes	1 (0.1)	1 (0.2)	1 (0.2)		
OSAHS (%)				Fisher	0.584
No	1001 (99.2)	424 (98.8)	459 (98.7)		
Yes	8 (0.8)	5 (1.2)	6 (1.3)		
Peptic ulcer disease (%)				$\chi^2 = 0.290$	0.865
No	996 (98.7)	422 (98.4)	459 (98.7)		
Yes	13 (1.3)	7 (1.6)	6 (1.3)		
Other comorbidity (%)				$\chi^2 = 1.972$	0.373
No	631 (62.5)	252 (58.7)	290 (62.4)		
Yes	378 (37.5)	177 (41.3)	175 (37.6)		
SBP (mmHg)	132.00 (121.00, 145.00)	134.00 (123.00, 148.00)	133.00 (122.00, 145.00)	$H = 2.426$	0.297
DBP (mmHg)	81.00 (74.00, 89.00)	83.00 (75.00, 90.00)	81.00 (74.00, 89.00)	$H = 8.216$	0.016
Waist (cm)	94.00 (90.00, 100.00)	94.00 (89.50, 101.00)	94.50 (90.00, 100.00)	$H = 1.305$	0.521
Hip (cm)	99.00 (95.00, 104.00)	99.00 (95.00, 104.00)	99.00 (96.00, 104.00)	$H = 0.552$	0.759
BMI (kg/m^2)	26.50 (25.06, 28.40)	26.56 (25.18, 28.77)	26.60 (25.02, 28.88)	$H = 0.944$	0.624
ALT (U/L)	23.00 (16.00, 37.00)	23.00 (16.00, 35.00)	23.00 (16.00, 37.00)	$H = 0.043$	0.979
ALB (g/L)	42.00 (39.10, 44.70)	42.30 (39.40, 44.70)	41.70 (39.00, 44.70)	$H = 2.823$	0.244
GGT (U/L)	30.00 (21.00, 46.00)	29.00 (20.00, 45.00)	30.00 (21.00, 52.00)	$H = 1.008$	0.604
GLU (mmol/L)	7.85 (6.21, 10.80)	8.37 (6.45, 11.99)	7.99 (6.31, 11.61)	$H = 4.132$	0.127
BUN (mmol/L)	5.35 (4.38, 6.73)	5.30 (4.30, 6.56)	5.35 (4.33, 6.80)	$H = 0.704$	0.703
UA (μmol/L)	389.00 (324.00, 468.20)	374.00 (317.00, 453.00)	381.00 (316.00, 458.00)	$H = 2.861$	0.239
CHOL (mmol/L)	4.65 (3.96, 5.51)	4.66 (3.90, 5.57)	4.57 (3.87, 5.48)	$H = 0.738$	0.692
LDL-C (mmol/L)	2.85 (2.24, 3.62)	2.90 (2.14, 3.58)	2.82 (2.16, 3.57)	$H = 0.961$	0.618
HDL-C (mmol/L)	1.01 (0.87, 1.18)	1.02 (0.87, 1.20)	0.99 (0.86, 1.17)	$H = 2.830$	0.243
TRIG (mmol/L)	1.54 (1.11, 2.27)	1.54 (1.11, 2.27)	1.56 (1.11, 2.32)	$H = 0.307$	0.858
CK (U/L)	98.00 (70.00, 147.00)	97.00 (69.00, 150.00)	95.00 (70.00, 145.00)	$H = 0.291$	0.865
CK-MB (ng/mL)	7.00 (6.00, 9.00)	7.00 (6.00, 9.00)	7.00 (6.00, 9.00)	$H = 1.098$	0.578
CysC (mg/L)	0.84 (0.72, 1.03)	0.83 (0.71, 1.01)	0.85 (0.72, 1.05)	$H = 2.139$	0.343
HbA1c (%)	8.50 (7.10, 10.50)	8.70 (7.10, 10.70)	8.70 (7.20, 10.60)	$H = 0.693$	0.707
Fasting insulin (μIU/mL)	11.89 (7.63, 17.33)	10.88 (7.29, 16.78)	12.89 (7.95, 17.58)	$H = 5.348$	0.069
Insulin 0.5 h (μIU/mL)	27.06 (16.87, 40.89)	24.69 (15.22, 40.03)	27.44 (17.41, 44.95)	$H = 4.647$	0.098
Insulin 2 h (μIU/mL)	39.99 (24.08, 60.90)	37.94 (21.07, 57.28)	39.33 (25.01, 64.92)	$H = 2.888$	0.236
WHR	0.95 (0.91, 0.99)	0.95 (0.92, 0.98)	0.95 (0.92, 0.99)	$H = 0.559$	0.756

Note: Data are presented as *n* (%) for categorical variables and median (Q1, Q3) for continuous variables.

DKD, diabetic kidney disease; Smoker, smoking status; Alcohol, alcohol consumption status; CHD, coronary heart disease; Nephrolithiasis, kidney stones; COPD, chronic obstructive pulmonary disease; OSAHS, obstructive sleep apnea-hypopnea syndrome; SBP, systolic blood pressure; DBP, diastolic blood pressure; Waist, waist circumference; Hip, hip circumference; BMI, body mass index; ALT, alanine aminotransferase; ALB, serum albumin; GGT, gamma-glutamyl transferase; GLU, fasting blood glucose; BUN, blood urea nitrogen; UA, uric acid; CHOL, total cholesterol; LDL-C, low-density lipoprotein cholesterol; HDL-C, high-density lipoprotein cholesterol; TRIG, triglycerides; CK, creatine kinase; CK-MB, creatine kinase-MB isoenzyme; CysC, cystatin C; HbA1c, glycated hemoglobin A1c; Insulin 0.5 h, 0.5-hour post-load insulin; Insulin 2 h, 2-hour post-load insulin; WHR, waist-to-hip ratio (dimensionless).

3.6 Clinical Utility Assessed by Decision Curve Analysis

Decision curves showed positive net benefit over a range of threshold probabilities in the training, internal validation, and external validation cohorts (Fig. 7). These results suggest potential clinical utility, with interpretation dependent on the threshold range and comparison with the treat-all and treat-none strategies.

Table 2. Demographic and clinical characteristics of the overall study population stratified based on DKD status.

Variables	Total (n = 1903)	Non-DKD (n = 1587)	DKD (n = 316)	Test statistic	*p*-value
Sex (%)				$\chi^2 = 2.227$	0.136
Male	1173 (61.6)	990 (62.4)	183 (57.9)		
Female	730 (38.4)	597 (37.6)	133 (42.1)		
Age (years)	55.00 (46.00, 65.00)	54.00 (45.00, 64.00)	61.00 (51.00, 68.00)	$z = -6.611$	<0.001
Smoker (%)				$\chi^2 = 2.422$	0.120
No	1311 (68.9)	1105 (69.6)	206 (65.2)		
Yes	592 (31.1)	482 (30.4)	110 (34.8)		
Alcohol (%)				$\chi^2 = 0.489$	0.485
No	1515 (79.6)	1268 (79.9)	247 (78.2)		
Yes	388 (20.4)	319 (20.1)	69 (21.8)		
Diabetes duration (years)	6.00 (2.00, 11.00)	5.00 (1.00, 10.00)	10.00 (5.00, 18.00)	$z = -9.020$	<0.001
Hypertension (%)				$\chi^2 = 42.915$	<0.001
No	1122 (59.0)	988 (62.3)	134 (42.4)		
Yes	781 (41.0)	599 (37.7)	182 (57.6)		
Hyperlipidemia (%)				$\chi^2 = 3.109$	0.078
No	1772 (93.1)	1485 (93.6)	287 (90.8)		
Yes	131 (6.9)	102 (6.4)	29 (9.2)		
Hyperuricemia (%)				$\chi^2 = 0.090$	0.764
No	1869 (98.2)	1558 (98.2)	311 (98.4)		
Yes	34 (1.8)	29 (1.8)	5 (1.6)		
Gout (%)				$\chi^2 = 0.019$	0.891
No	1822 (95.7)	1519 (95.7)	303 (95.9)		
Yes	81 (4.3)	68 (4.3)	13 (4.1)		
Osteoporosis (%)				$\chi^2 = 4.914$	0.027
No	1881 (98.8)	1573 (99.1)	308 (97.5)		
Yes	22 (1.2)	14 (0.9)	8 (2.5)		
Gallstone (%)				$\chi^2 = 1.005$	0.316
No	1858 (97.6)	1547 (97.5)	311 (98.4)		
Yes	45 (2.4)	40 (2.5)	5 (1.6)		
Cerebral infarction (%)				$\chi^2 = 18.175$	<0.001
No	1799 (94.5)	1516 (95.5)	283 (89.6)		
Yes	104 (5.5)	71 (4.5)	33 (10.4)		
CHD (%)				$\chi^2 = 8.331$	0.004
No	1716 (90.2)	1445 (91.1)	271 (85.8)		
Yes	187 (9.8)	142 (8.9)	45 (14.2)		
Hepatitis (%)				$\chi^2 = 0.682$	0.409
No	1710 (89.9)	1422 (89.6)	288 (91.1)		
Yes	193 (10.1)	165 (10.4)	28 (8.9)		
Tuberculosis (%)				$\chi^2 = 0.108$	0.742
No	1871 (98.3)	1561 (98.4)	310 (98.1)		
Yes	32 (1.7)	26 (1.6)	6 (1.9)		
Malignancy (%)				$\chi^2 = 0.865$	0.352
No	1862 (97.8)	1555 (98.0)	307 (97.2)		
Yes	41 (2.2)	32 (2.0)	9 (2.8)		
Nephrolithiasis (%)				$\chi^2 = 6.585$	0.010
No	1812 (95.2)	1520 (95.8)	292 (92.4)		
Yes	91 (4.8)	67 (4.2)	24 (7.6)		
COPD (%)				Fisher	1.000
No	1900 (99.8)	1584 (99.8)	316 (100)		
Yes	3 (0.2)	3 (0.2)	0 (0)		
OSAHS (%)				$\chi^2 = 0.165$	0.685
No	1884 (99.0)	1570 (98.9)	314 (99.4)		

Table 2. Continued.

Variables	Total (n = 1903)	Non-DKD (n = 1587)	DKD (n = 316)	Test statistic	*p*-value
Yes	19 (1.0)	17 (1.1)	2 (0.6)		
Peptic ulcer disease (%)				$\chi^2 = 7.563$	0.006
No	1877 (98.6)	1571 (99.0)	306 (96.8)		
Yes	26 (1.4)	16 (1.0)	10 (3.2)		
Other comorbidity (%)				$\chi^2 = 6.930$	0.008
No	1173 (61.6)	999 (62.9)	174 (55.1)		
Yes	730 (38.4)	588 (37.1)	142 (44.9)		
SBP (mmHg)	133.00 (122.00, 145.50)	132.00 (121.00, 144.00)	141.00 (127.00, 156.00)	$z = -7.248$	<0.001
DBP (mmHg)	82.00 (74.00, 89.00)	82.00 (74.00, 89.00)	81.00 (74.00, 89.25)	$z = 0.083$	0.934
Waist (cm)	94.00 (90.00, 100.00)	94.00 (90.00, 100.00)	96.00 (90.00, 102.00)	$z = -3.814$	<0.001
Hip (cm)	99.00 (95.00, 104.00)	99.00 (95.00, 104.00)	100.00 (95.00, 104.50)	$z = -1.812$	0.070
BMI (kg/m^2)	26.56 (25.08, 28.63)	26.49 (25.05, 28.40)	27.03 (25.25, 29.50)	$z = -3.144$	0.002
ALT (U/L)	23.00 (16.00, 36.50)	24.00 (16.00, 37.00)	19.00 (15.00, 33.00)	$z = 4.018$	<0.001
ALB (g/L)	42.00 (39.10, 44.70)	42.40 (39.60, 44.90)	39.20 (35.68, 42.82)	$z = 10.474$	<0.001
GGT (U/L)	30.00 (21.00, 46.00)	29.00 (21.00, 46.00)	30.00 (21.00, 47.00)	$z = -1.229$	0.219
GLU (mmol/L)	7.99 (6.28, 11.18)	7.90 (6.29, 11.04)	8.05 (6.26, 11.97)	$z = -0.920$	0.357
BUN (mmol/L)	5.33 (4.33, 6.70)	5.22 (4.25, 6.34)	6.90 (4.95, 8.87)	$z = -10.489$	<0.001
UA (μmol/L)	382.00 (319.00, 462.00)	375.00 (317.00, 453.00)	414.00 (352.50, 494.00)	$z = -5.223$	<0.001
CHOL (mmol/L)	4.62 (3.92, 5.51)	4.59 (3.92, 5.46)	4.74 (3.96, 5.91)	$z = -2.810$	0.005
LDL-C (mmol/L)	2.85 (2.20, 3.60)	2.85 (2.21, 3.57)	2.87 (2.17, 3.67)	$z = -1.063$	0.288
HDL-C (mmol/L)	1.01 (0.87, 1.18)	1.01 (0.87, 1.18)	0.98 (0.84, 1.18)	$z = 2.277$	0.023
TRIG (mmol/L)	1.54 (1.11, 2.29)	1.51 (1.09, 2.24)	1.70 (1.25, 2.61)	$z = -4.022$	<0.001
CK (U/L)	97.00 (70.00, 147.00)	97.00 (70.00, 145.00)	102.00 (72.75, 159.25)	$z = -1.553$	0.120
CK-MB (ng/mL)	7.00 (6.00, 9.00)	7.00 (6.00, 9.00)	7.50 (6.00, 10.00)	$z = -2.222$	0.026
CysC (mg/L)	0.84 (0.72, 1.03)	0.82 (0.71, 0.97)	1.10 (0.84, 1.55)	$z = -12.933$	<0.001
HbA1c (%)	8.60 (7.10, 10.60)	8.50 (7.10, 10.65)	9.00 (7.30, 10.40)	$z = -0.859$	0.390
Fasting insulin (μIU/mL)	11.79 (7.63, 17.28)	11.09 (7.29, 16.41)	15.23 (9.06, 22.50)	$z = -7.133$	<0.001
Insulin 0.5 h (μIU/mL)	26.42 (16.64, 41.55)	25.37 (15.91, 40.74)	30.33 (18.09, 47.59)	$z = -3.581$	<0.001
Insulin 2 h (μIU/mL)	39.26 (23.86, 61.32)	38.88 (23.18, 60.90)	40.15 (27.94, 64.09)	$z = -1.987$	0.047
WHR	0.95 (0.92, 0.99)	0.95 (0.91, 0.98)	0.96 (0.93, 1.00)	$z = -3.709$	<0.001

Note: Data are presented as *n* (%) for categorical variables and median (Q1, Q3) for continuous variables.

DKD, diabetic kidney disease; Smoker, smoking status; Alcohol, alcohol consumption status; CHD, coronary heart disease; Nephrolithiasis, kidney stones; COPD, chronic obstructive pulmonary disease; OSAHS, obstructive sleep apnea-hypopnea syndrome; SBP, systolic blood pressure; DBP, diastolic blood pressure; Waist, waist circumference; Hip, hip circumference; BMI, body mass index; ALT, alanine aminotransferase; ALB, serum albumin; GGT, gamma-glutamyl transferase; GLU, fasting blood glucose; BUN, blood urea nitrogen; UA, uric acid; CHOL, total cholesterol; LDL-C, low-density lipoprotein cholesterol; HDL-C, high-density lipoprotein cholesterol; TRIG, triglycerides; CK, creatine kinase; CK-MB, creatine kinase-MB isoenzyme; CysC, cystatin C; HbA1c, glycated hemoglobin A1c; Insulin 0.5 h, 0.5-hour post-load insulin; Insulin 2 h, 2-hour post-load insulin; WHR, waist-to-hip ratio (dimensionless).

Table 3. Multivariate logistic regression analysis for predictor selection.

Variable	Diagnostic prediction model			*p*
	β	SE	OR (95% CI)	
Diabetes duration (years)	0.036	0.014	1.037 (1.009, 1.066)	0.009
SBP (mmHg)	0.019	0.005	1.020 (1.009, 1.030)	<0.001
ALB (g/L)	–0.101	0.022	0.904 (0.865, 0.944)	<0.001
BUN (mmol/L)	0.073	0.051	1.076 (0.974, 1.188)	0.152
CHOL (mmol/L)	0.252	0.066	1.286 (1.131, 1.463)	<0.001
CysC (mg/L)	1.093	0.385	2.983 (1.402, 6.351)	0.005

SBP, systolic blood pressure; ALB, serum albumin; BUN, blood urea nitrogen; CHOL, total cholesterol; CysC, cystatin C; SE, standard error; CI, confidence interval; OR, odds ratio.

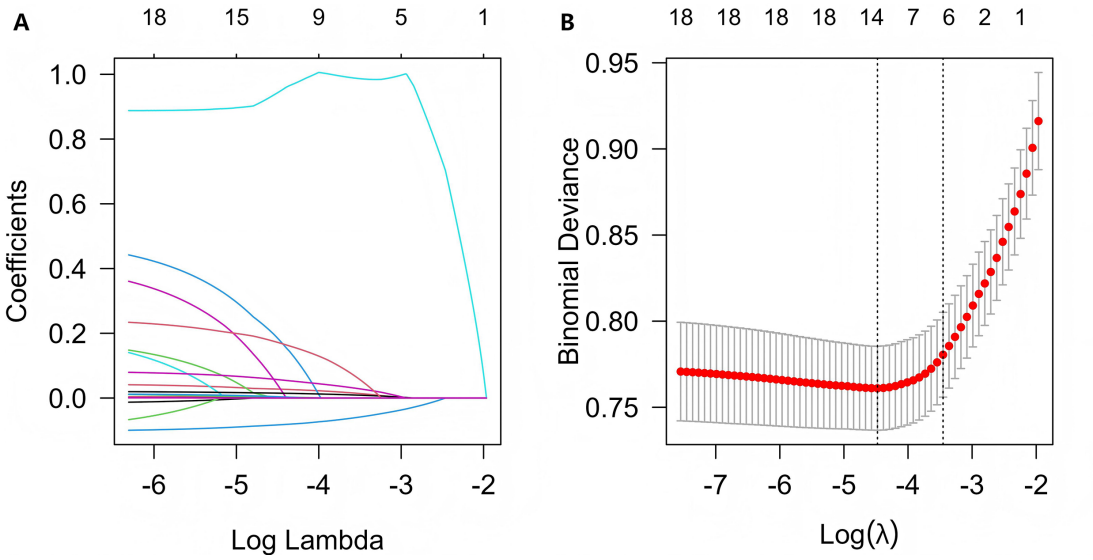


Fig. 2. LASSO-based selection of candidate predictors. (A) Changes in the coefficients of the 18 candidate variables across penalty values. (B) Cross-validated binomial deviance plotted against log(λ), based on 10 folds. Vertical lines identify lambda.min and lambda.1se. At lambda.1se, six variables retained nonzero coefficients. LASSO, least absolute shrinkage and selection operator.

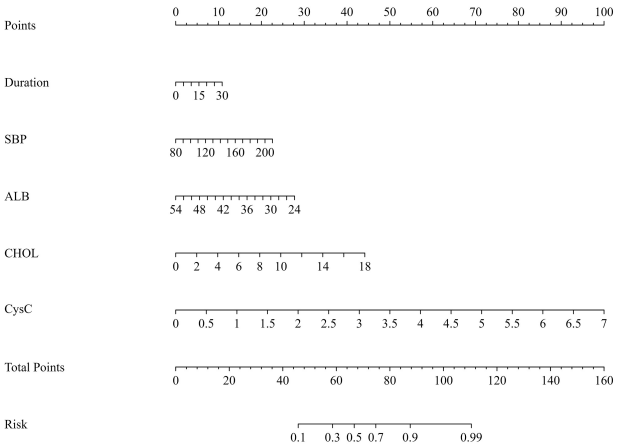


Fig. 3. Nomogram for estimating the probability of prevalent diabetic kidney disease (DKD) in patients with T2DM and overweight or obesity. Locate each patient's value on its predictor scale and determine the corresponding points. Sum the points of all five predictors, and then use the total-points scale to estimate the probability of DKD. SBP, systolic blood pressure; ALB, serum albumin; CHOL, total cholesterol; CysC, cystatin C; T2DM, type 2 diabetes mellitus.

3.7 Model Discrimination by Sex

In the sex-stratified analysis, the model achieved AUCs of 0.806 (95% CI, 0.758–0.855) in men and 0.745 (95% CI, 0.676–0.814) in women in the training cohort. The corresponding AUCs were 0.760 (95% CI, 0.654–0.866) and 0.754 (95% CI, 0.610–0.899) in the internal validation cohort, and 0.793 (95% CI, 0.714–0.872) and 0.766 (95% CI, 0.676–0.857) in the external validation cohort, respectively. AUC point estimates were numerically higher in men across all three cohorts, although the confidence intervals overlapped. Sex-specific ROC curves and detailed results are presented in **Supplementary Fig. 1** and **Supplementary Table 4**.

4. Discussion

This study developed a five-predictor nomogram for identifying prevalent DKD in adults with T2DM and overweight or obesity. The final model included diabetes duration, SBP, serum ALB, total cholesterol, and CysC. The AUCs were 0.782, 0.753, and 0.786 in the training, internal validation, and external validation cohorts, respectively, indicating acceptable model performance in terms of discrimination across the three cohorts. Calibration curves generally followed the ideal line. Although the model was developed in the training cohort, the Brier score was higher in the training cohort (0.109) than in the internal validation cohort (0.081). This is likely due to the higher DKD event rate in the training cohort (17.1% vs 13.8%), as the Brier score is sensitive to outcome prevalence. Given that the AUCs were comparable between the two cohorts (0.782 vs 0.753), this difference in Brier score should not be interpreted as poorer model fit in the training cohort. Decision curve analysis suggested potential clinical benefit over a range of threshold probabilities, while sex-stratified analyses provided additional information on discrimination in men and women. These findings support further evaluation of the model as a tool for identifying patients who may require additional assessment for DKD.

The focus on patients with excess body weight is clinically relevant because obesity may increase renal metabolic and hemodynamic demands. The resulting hyperfiltra-

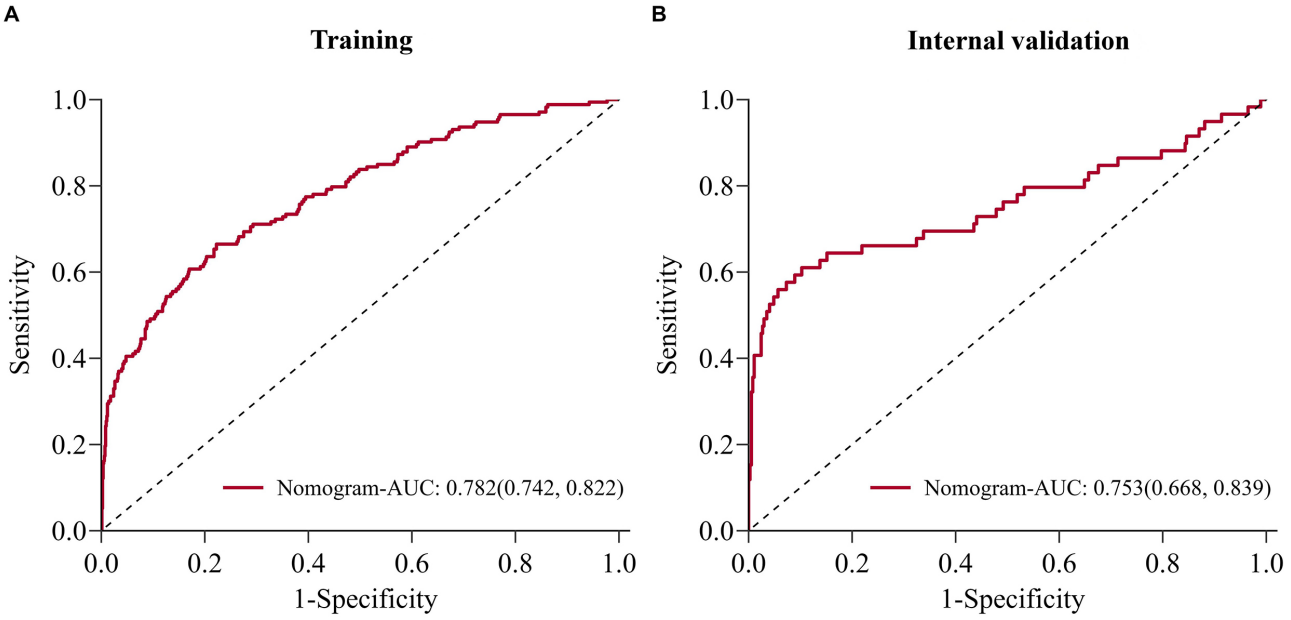


Fig. 4. ROC curves for the nomogram in the training (A) and internal validation sets (B). AUC, area under the curve; ROC, receiver operating characteristic.

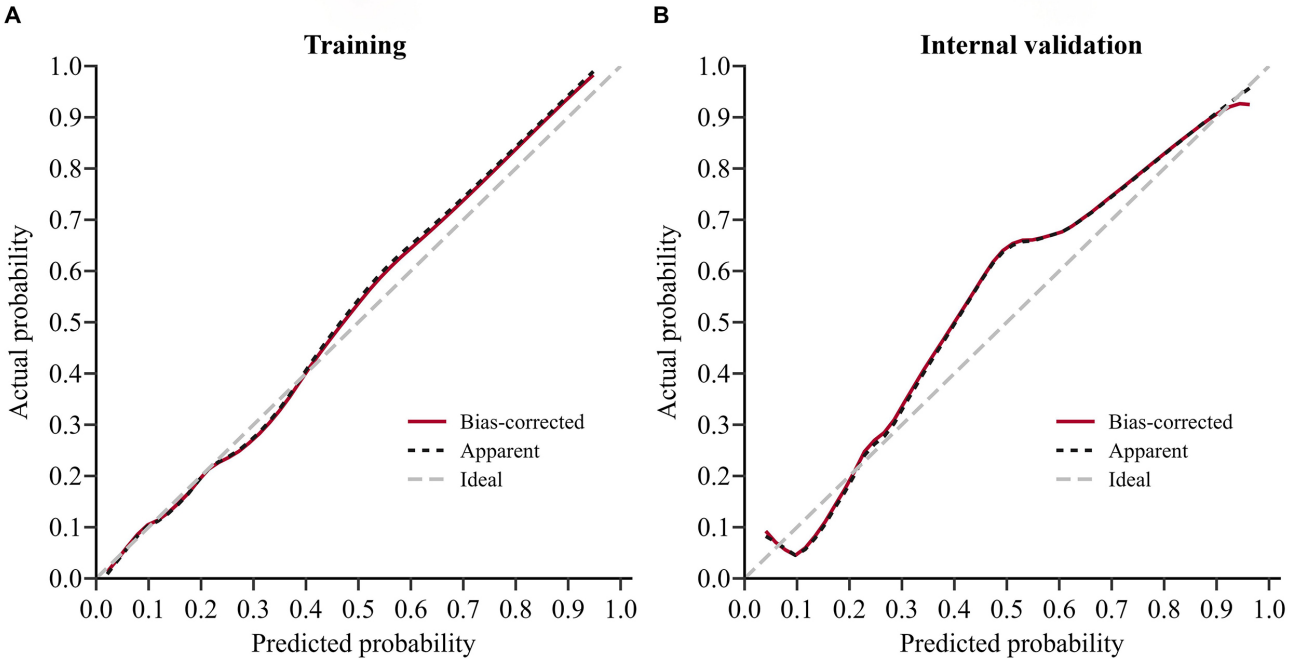


Fig. 5. Calibration curves of the nomogram in the training and internal validation sets. (A) Training set. (B) Internal validation set.

tion and elevated glomerular pressure can place mechanical stress on the filtration barrier and contribute to podocyte injury [11]. These mechanisms provide a rationale for examining DKD in overweight or obese patients with T2DM. In the overall study population, 16.6% of participants met the diagnostic criteria for DKD. A meta-analysis of 30 observational studies involving 79,364 Chinese patients with T2DM reported a pooled prevalence of diabetic nephropathy of 21.8% (95% confidence interval [CI], 18.5–25.4%) [12]. The lower proportion of DKD cases observed in our study may partly reflect the exclusion of patients with previously documented CKD and those with other specified kidney diseases. Differences in diagnostic criteria, study settings, and participant characteristics may also have contributed to said lower proportion, particularly given the heterogeneous definitions of diabetic nephropathy across the studies included in the meta-analysis. Large-scale, population-based epidemiological studies are there-

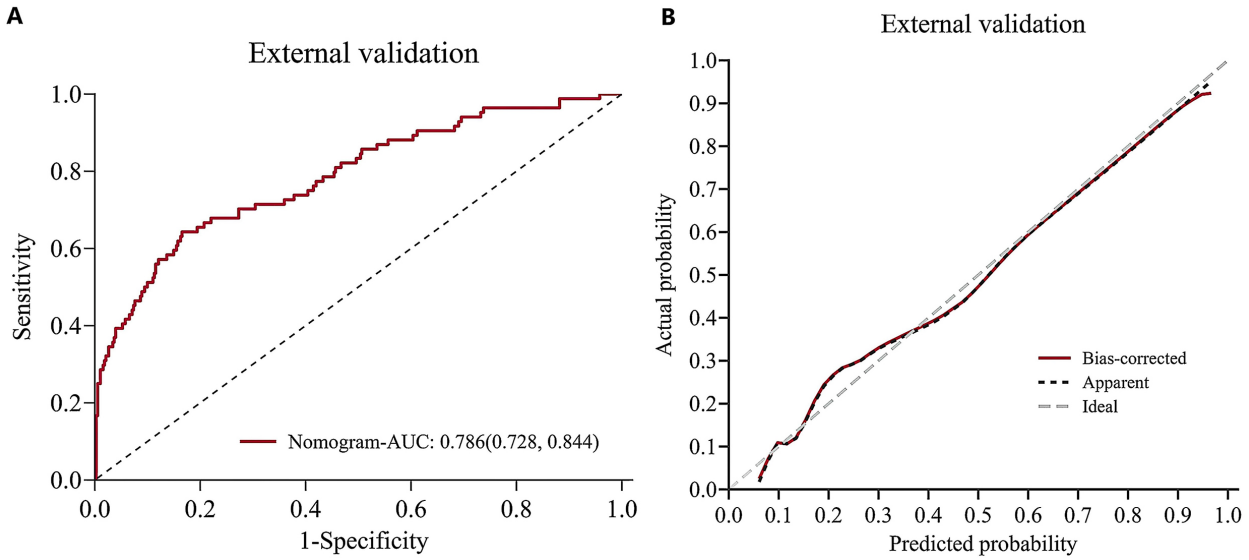


Fig. 6. Model performance based on the external validation cohort. (A) ROC curve for the nomogram predicting DKD in the external validation set. (B) Calibration curve of the nomogram for the external validation set.

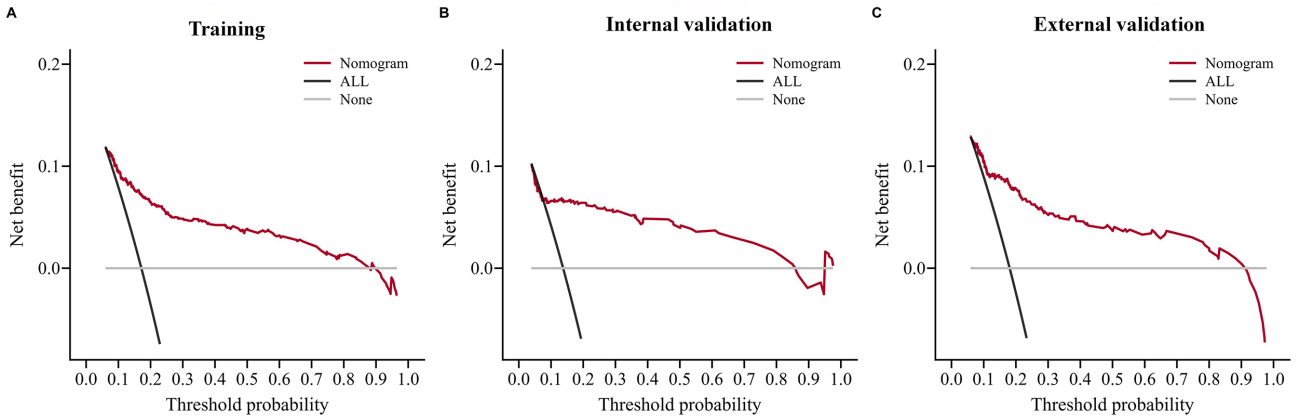


Fig. 7. Decision curve analyses of the nomogram for the training, internal validation, and external validation sets. (A) Training set. (B) Internal validation set. (C) External validation set.

fore needed to better characterize the prevalence of DKD among patients with T2DM and overweight or obesity.

In our study, longer diabetes duration was associated with higher odds of prevalent DKD among overweight or obese patients with T2DM. This finding is consistent with previous reports linking diabetes duration to renal impairment and DKD occurrence or progression [13,14]. Greater proteinuria and higher DKD risk have also been reported among patients with a T2DM duration exceeding 10 years [13]. Longer duration may reflect greater cumulative exposure to metabolic and vascular disturbances, providing a plausible explanation for its contribution to the model.

Higher SBP was associated with greater odds of prevalent DKD and was retained in the final diagnostic model. An Asian meta-analysis reported higher DKD odds in patients with hypertension (OR = 2.58) [14]. Furthermore, a dose-response analysis estimated a relative risk (RR) of 1.06 (95% CI, 1.03–1.09) per 5-mmHg increment in SBP [7]. This association is biologically plausible, as sustained blood pressure elevation may increase intraglomerular pressure and contribute to renal vascular injury and microvascular endothelial damage [14,15,16]. These findings support considering SBP when assessing prevalent DKD and underscore the importance of blood pressure management in patients with T2DM and overweight or obesity.

Higher ALB levels were associated with lower odds of prevalent DKD among overweight or obese patients with T2DM in our study. ALB is the most abundant plasma protein and plays a central role in maintaining colloid osmotic pressure and fluid balance [17]. Our finding is consistent with previous evidence linking lower serum ALB levels to adverse renal outcomes. Cheng et al. [18] found that lower ALB was associated with worse kidney function and less favorable renal outcomes. Similarly, a study linked

hypoalbuminemia to a faster decline in glomerular filtration rate [19]. These findings support the clinical relevance of serum ALB in DKD; however, evidence regarding disease severity and prognosis should be distinguished from its association with prevalent DKD in our study. Lower serum ALB may reflect existing kidney damage rather than a causal contribution to DKD. Thus, serum ALB may contribute to the assessment of prevalent DKD in overweight or obese patients with T2DM, whereas its ability to predict disease progression in this population requires longitudinal validation.

Total cholesterol was retained in the final model, with higher levels associated with a greater estimated probability of prevalent DKD. This finding may reflect the involvement of altered lipid metabolism in diabetic kidney injury. Experimental evidence has linked impaired ATP-binding cassette transporter A1 (ABCA1)-mediated cholesterol efflux to cholesterol accumulation and podocyte injury in DKD, providing a possible biological context for this association [20]. However, serum total cholesterol does not directly measure renal cholesterol accumulation, and these mechanistic findings do not establish a causal role for circulating total cholesterol in our population. Given the cross-sectional design, the temporal relationship between cholesterol levels and DKD could not be determined. Furthermore, data on lipid-lowering medication use were not collected, leaving potential treatment-related confounding unaccounted for. Total cholesterol should therefore be interpreted as a predictor within the combined diagnostic model, rather than as evidence that lowering total cholesterol would prevent DKD.

In our study, higher CysC levels were associated with greater odds of prevalent DKD among overweight or obese patients with T2DM. As an endogenous marker of kidney function, CysC may provide useful information for identifying renal impairment, supporting its potential role in the assessment of DKD in this population. However, this association may reflect existing impairment of kidney function, and its value for early detection or screening requires further validation. In addition, CysC levels can be influenced by nonrenal factors, including glucocorticoid use and hyperthyroidism [21]. Therefore, CysC should be interpreted alongside other clinical findings, with patients' medical history and medication use taken into consideration.

Comparison with previous DKD prediction models helps contextualize the potential applicability of our nomogram. For instance, Deng et al. [22] developed a diagnostic prediction model for prevalent DKD based on a community health check-up cohort of 1049 T2DM patients, incorporating seven predictors, and achieved an AUC of 0.725 (95% CI, 0.680–0.769) in the training set and 0.698 (95% CI, 0.627–0.770) in the validation set. While providing a practical screening tool, its discrimination leaves room for improvement, and it was not specifically tailored to high-risk subgroups. Furthermore, Li et al. [23] developed another model for identifying prevalent DKD using age, sex, triglyceride-glucose (TyG) index, BUN, and CysC. Although it showed strong discrimination in the China Health and Retirement Longitudinal Study (CHARLS) cohort (AUC = 0.868), its AUC decreased to 0.781 (95% CI, 0.648–0.888) in a hospital-based clinical cohort, highlighting the challenge of model transportability across different clinical settings. In the present study, focusing specifically on the high-risk overweight or obese T2DM subpopulation, our final model achieved an AUC of 0.782 (95% CI, 0.742–0.822) in the training cohort, 0.753 (95% CI, 0.668–0.839) in the internal validation cohort, and notably, 0.786 (95% CI, 0.728–0.844) in the external validation cohort. While Li et al.'s model [23] demonstrated a noticeable performance decline in a clinical setting, our model maintained robust and consistent discrimination in the external validation cohort, demonstrating better generalizability. Developed specifically for overweight or obese patients with T2DM in Guangdong, our nomogram showed moderate discrimination and, by combining routine metabolic and renal indicators, suggests potential value in identifying prevalent DKD in this high-risk population.

Metabolic markers such as HbA1c and measures of visceral adiposity were not retained in the final model, as our study population consisted exclusively of overweight and obese patients with T2DM. Within this specific cohort, the variability of these metabolic indicators is inherently limited, since these patients already share a similar metabolic phenotype characterized by insulin resistance and impaired glycemic control. Indeed, emerging evidence suggests that obesity and T2DM are closely intertwined through a common pathophysiological axis of insulin resistance and adipose tissue dysfunction [24]. Other possible explanations include correlations with retained predictors, measurement variability, and model selection procedures. Therefore, the contribution of these markers to distinguishing prevalent DKD in this population warrants further investigation, ideally through direct assessment of their distributions and incremental diagnostic value.

Limitations

Several limitations should be acknowledged. First, this retrospective cross-sectional study assessed prevalent DKD and therefore cannot establish causality or predict future DKD onset. Comorbidities were identified from electronic medical records rather than systematic screening, potentially leading to underascertainment. Data on medication use were not extracted, leaving the potential effects of renoprotective and lipid-lowering therapies unaccounted for. Second, sequential univariate screening, LASSO, and stepwise regression may have increased selection instability and overfitting. The final conventional logistic regression model did not retain LASSO coefficient shrinkage. Third, cohort-specific thresholds were selected and evaluated within the same cohorts, which may have resulted

in overly optimistic classification estimates. No bootstrap-based optimism correction was performed; therefore, the apparent performance in the training cohort may be optimistic. Finally, participants were drawn from two campuses of the same hospital in Guangdong, limiting generalizability to other settings. The single random split and limited subgroup event counts also introduce uncertainty into the validation and sex-stratified estimates. Further validation in independent populations and prospective evaluation of clinical impact are warranted.

5. Conclusion

The five-variable nomogram estimated the probability of prevalent DKD in overweight or obese adults with T2DM and demonstrated acceptable discrimination across the evaluated cohorts. It warrants further assessment as an adjunct to established diagnostic evaluation. Validation outside the participating hospital and prospective studies of its clinical use are needed before implementation.

Key Points

• This study developed a diagnostic prediction model for prevalent diabetic kidney disease (DKD) in adults with type 2 diabetes mellitus (T2DM) and overweight or obesity in Guangdong, China.

• The retrospective cross-sectional study included 1903 patients: 1009 in a training cohort, 429 in an internal validation cohort, and 465 in an external validation cohort.

• The final nomogram incorporated five clinical predictors: diabetes duration, systolic blood pressure, serum albumin, total cholesterol, and cystatin C.

• The model achieved AUCs of 0.782, 0.753, and 0.786 in the training, internal validation, and external validation cohorts, respectively, and decision curve analysis suggested potential clinical benefit, although broader external validation and prospective evaluation of clinical utility are needed.

• The nomogram, based on routinely available clinical variables, can serve as a diagnostic prediction tool to identify prevalent DKD in overweight/obese T2DM patients and support clinical decision-making.

Abbreviations

T2DM, type 2 diabetes mellitus; DKD, diabetic kidney disease; non-DKD, without diabetic kidney disease; CKD, chronic kidney disease; ESRD, end-stage renal disease; eGFR, estimated glomerular filtration rate; UACR, urinary albumin-to-creatinine ratio; SBP, systolic blood pressure; DBP, diastolic blood pressure; BMI, body mass index; WHR, waist-to-hip ratio; HbA1c, glycated hemoglobin A1c; GLU, fasting blood glucose; ALT, alanine aminotransferase; ALB, serum albumin; GGT, gamma-glutamyl transferase; Scr, serum creatinine; BUN, blood urea nitrogen; UA, uric acid; CHOL, total cholesterol; LDL-C, low-density lipoprotein cholesterol; HDL-C, high-density lipoprotein cholesterol; TRIG, triglycerides; CK, creatine kinase; CK-MB, creatine kinase-MB isoenzyme; CysC, cystatin C; CHD, coronary heart disease; COPD, chronic obstructive pulmonary disease; OSAHS, obstructive sleep apnea-hypopnea syndrome; LASSO, least absolute shrinkage and selection operator; VIF, variance inflation factor; ROC, receiver operating characteristic; AUC, area under the curve; CI, confidence interval; OR, odds ratio; SE, standard error; SD, standard deviation.

Availability of Data and Materials

Due to ethical restrictions and privacy protection requirements, the raw data from this study are not publicly available. Access may be granted upon reasonable request to the corresponding author, subject to institutional ethics approval.

Author Contributions

XLH: Conceptualization, study design, data collection, and drafting of the original manuscript. QYL: Data collection, data curation, and formal analysis. XQZ: Data collection, data interpretation, and investigation. YPZ: Conceptualization, study design, project administration, supervision. All authors contributed to revising the manuscript critically for important intellectual content. All authors approved the final version, and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Ethics Approval and Consent to Participate

The study was approved by the Ethics Committee of the Third Affiliated Hospital of Sun Yat-sen University (approval no. II2024-099-01). The study was conducted in compliance with the Declaration of Helsinki, and all participants provided written informed consent before taking part.

Acknowledgment

We thank all patients and the Third Affiliated Hospital of Sun Yat-sen University for their support for this research.

Funding

This work was partly supported by the National Natural Science Foundation of China (72204277); the Guangdong Basic and Applied Basic Research Foundation (2025A1515012706); and the 3rd Affiliated Hospital of Sun Yat-sen University, Clinical Research Program (YHJH202404).

Conflicts of Interest

The authors declare no conflicts of interest.

Supplementary Material

Supplementary material associated with this article can be found, in the online version, at https://doi.org/10.31083/BJHM55626.

References

[1] International Diabetes Federation (IDF). IDF Diabetes Atlas. 11th edn. 2025. Available at: https://diabetesatlas.org/resources/idf-diabetes-atlas-2025/ (Accessed: 11 April 2026).

[2] World Obesity Federation. World Obesity Atlas 2025. 2025. Available at: https://www.worldobesity.org/resources/resource-library/world-obesity-atlas-2025 (Accessed: 5 September 2026).

[3] Wang Y, Zhao L, Gao L, Pan A, Xue H. Health policy and public health implications of obesity in China. The Lancet. Diabetes & Endocrinology. 2021; 9: 446–461. https://doi.org/10.1016/S2213-8587(21)00118-2

[4] Bureau of Disease Prevention and Control, National Health Commission of the People's Republic of China. Report on Chinese Residents' Nutrition and Chronic Disease Status (2020). People's Medical Publishing House: Beijing. 2021. (In Chinese)

[5] Ye X, Yu R, Jiang F, Hou X, Wei L, Bao Y, et al. Osteocalcin and Risks of Incident Diabetes and Diabetic Kidney Disease: A 4.6-Year Prospective Cohort Study. Diabetes Care. 2022; 45: 830–836. https://doi.org/10.2337/dc21-2113

[6] Chinese Diabetes Society. Guideline for the prevention and treatment of diabetes mellitus in China (2024 edition). Chinese Journal of Diabetes Mellitus. 2025; 17: 16–139. https://doi.org/10.3760/cma.j.cn115791-20241203-00705 (In Chinese)

[7] Guo J, Liu C, Wang Y, Shao B, Fong TL, Lau NC, et al. Dose-response association of diabetic kidney disease with routine clinical parameters in patients with type 2 diabetes mellitus: a systematic review and meta-analysis. EClinicalMedicine. 2024; 69: 102482. https://doi.org/10.1016/j.eclinm.2024.102482

[8] American Diabetes Association. 2. Classification and Diagnosis of Diabetes: *Standards of Medical Care in Diabetes-2021*. Diabetes Care. 2021; 44: S15–S33. https://doi.org/10.2337/dc21-S002

[9] Pan XF, Wang L, Pan A. Epidemiology and determinants of obesity in China. The Lancet. Diabetes & Endocrinology. 2021; 9: 373–392. https://doi.org/10.1016/S2213-8587(21)00045-0

[10] Ma YC, Zuo L, Chen JH, Luo Q, Yu XQ, Li Y, et al. Modified glomerular filtration rate estimating equation for Chinese patients with chronic kidney disease. Journal of the American Society of Nephrology. 2006; 17: 2937–2944. https://doi.org/10.1681/ASN.2006040368

[11] Zbrzeźniak-Suszczewicz J, Winiarska A, Perkowska-Ptasińska A, Stompór T. Obesity-Related Glomerulosclerosis-How Adiposity Damages the Kidneys. International Journal of Molecular Sciences. 2025; 26: 6247. https://doi.org/10.3390/ijms26136247

[12] Zhang XX, Kong J, Yun K. Prevalence of Diabetic Nephropathy among Patients with Type 2 Diabetes Mellitus in China: A Meta-Analysis of Observational Studies. Journal of Diabetes Research. 2020; 2020: 2315607. https://doi.org/10.1155/2020/2315607

[13] Kim CS, Suh SH, Choi HS, Bae EH, Ma SK, Kim B, et al. Impact of diabetes duration and hyperglycemia on the progression of diabetic kidney disease: Insights from the KNHANES 2019-2021. World Journal of Diabetes. 2025; 16: 102094. https://doi.org/10.4239/wjd.v16.i5.102094

[14] Zheng Y, Ma S, Chen Y, Li C. Risk factors for diabetic kidney disease in type 2 diabetes mellitus in Asia: a meta-analysis. Frontiers in Endocrinology. 2026; 17: 1703533. https://doi.org/10.3389/fendo.2026.1703533

[15] Adler AI, Stratton IM, Neil HA, Yudkin JS, Matthews DR, Cull CA, et al. Association of systolic blood pressure with macrovascular and microvascular complications of type 2 diabetes (UKPDS 36): prospective observational study. BMJ (Clinical Research Ed.). 2000; 321: 412–419. https://doi.org/10.1136/bmj.321.7258.412

[16] Durante A, Mazzapicchi A, Baiardo Redaelli M. Systemic and Cardiac Microvascular Dysfunction in Hypertension. International Journal of Molecular Sciences. 2024; 25: 13294. https://doi.org/10.3390/ijms252413294

[17] di Masi A. Human Serum Albumin: From Molecular Aspects to Biotechnological Applications. International Journal of Molecular Sciences. 2023; 24: 4081. https://doi.org/10.3390/ijms24044081

[18] Cheng T, Wang X, Han Y, Hao J, Hu H, Hao L. The level of serum albumin is associated with renal prognosis and renal function decline in patients with chronic kidney disease. BMC Nephrology. 2023; 24: 57. https://doi.org/10.1186/s12882-023-03110-8

[19] Lang J, Katz R, Ix JH, Gutierrez OM, Peralta CA, Parikh CR, et al. Association of serum albumin levels with kidney function decline and incident chronic kidney disease in elders. Nephrology, Dialysis, Transplantation. 2018; 33: 986–992. https://doi.org/10.1093/ndt/gfx229

[20] Merscher-Gomez S, Guzman J, Pedigo CE, Lehto M, Aguillon-Prada R, Mendez A, et al. Cyclodextrin protects podocytes in diabetic kidney disease. Diabetes. 2013; 62: 3817–3827. https://doi.org/10.2337/db13-0399

[21] Patel S, Dhalani M, Reddy M, Ruhela N, Jain R, Kaur G. Cystatin C in chronic kidney disease: enhancing diagnostic accuracy and patient outcomes. Postgraduate Medicine. 2026; 138: 136–145. https://doi.org/10.1080/00325481.2026.2634426

[22] Deng Z, Yang J, Zhou H. Development and validation of a model that predicts the risk of diabetic kidney disease in type 2 diabetes mellitus patients: a retrospective study. Frontiers in Endocrinology. 2026; 16: 1708419. https://doi.org/10.3389/fendo.2025.1708419

[23] Li NN, Liu L, Cao GH, Tang Y, Peng W, Zhou HY. Development and External Validation of a Clinically Applicable Model for Identifying Diabetic Kidney Disease in Population-Based and Clinical Cohorts. International Journal of General Medicine. 2026; 19: 637578. https://doi.org/10.2147/IJGM.S637578

[24] Hu F, Zhang T. Study on Risk Factors of Diabetic Nephropathy in Obese Patients with Type 2 Diabetes Mellitus. International Journal of General Medicine. 2020; 13: 351–360. https://doi.org/10.2147/IJGM.S255858