

World Journal of *Diabetes*

World J Diabetes 2024 January 15; 15(1): 1-128



EDITORIAL

- 1 Effects of Tai Chi in diabetes patients: Insights from recent research
Hamasaki H
- 11 Individualized intensive insulin therapy of diabetes: Not only the goal, but also the time
Hu Y, Chen HJ, Ma JH

MINIREVIEWS

- 15 Management of monogenic diabetes in pregnancy: A narrative review
Jeeyavudeen MS, Murray SR, Strachan MWJ

ORIGINAL ARTICLE**Retrospective Cohort Study**

- 24 Prediabetes: An overlooked risk factor for major adverse cardiac and cerebrovascular events in atrial fibrillation patients
Desai R, Katukuri N, Goguri SR, Kothawala A, Alle NR, Bellamkonda MK, Dey D, Ganesan S, Biswas M, Sarkar K, Prattipati P, Chauhan S

Retrospective Study

- 34 Predictive value of bilirubin and serum γ -glutamyltranspeptidase levels in type-2 diabetes mellitus patients with acute coronary syndrome
Chen J, Zhang WC, Tang XQ, Yin RH, Wang T, Wei XY, Pan CJ
- 43 Clinical study of different prediction models in predicting diabetic nephropathy in patients with type 2 diabetes mellitus
Cai SS, Zheng TY, Wang KY, Zhu HP
- 53 Heterogeneously elevated branched-chain/aromatic amino acids among new-onset type-2 diabetes mellitus patients are potentially skewed diabetes predictors
Wang M, Ou Y, Yuan XL, Zhu XF, Niu B, Kang Z, Zhang B, Ahmed A, Xing GQ, Su H
- 72 Investigating the relationship between intracranial atherosclerotic plaque remodelling and diabetes using high-resolution vessel wall imaging
Mo YQ, Luo HY, Zhang HW, Liu YF, Deng K, Liu XL, Huang B, Lin F

Observational Study

- 81 Body composition and metabolic syndrome in patients with type 1 diabetes
Zeng Q, Chen XJ, He YT, Ma ZM, Wu YX, Lin K

Basic Study

- 92 Urinary exosomal microRNA-145-5p and microRNA-27a-3p act as noninvasive diagnostic biomarkers for diabetic kidney disease
Han LL, Wang SH, Yao MY, Zhou H
- 105 Myricetin induces M2 macrophage polarization to alleviate renal tubulointerstitial fibrosis in diabetic nephropathy *via* PI3K/ Akt pathway
Xu WL, Zhou PP, Yu X, Tian T, Bao JJ, Ni CR, Zha M, Wu X, Yu JY

LETTER TO THE EDITOR

- 126 Nutrition interventions and clinical outcomes of pregnant women with gestational diabetes mellitus: More than meets the eye
Sinha S, Nishant P, Sinha RK, Morya AK, Prasad R

ABOUT COVER

Editorial Board Member of *World Journal of Diabetes*, Yan-Kai Xia, BMed, PhD, Professor, Vice President, Nanjing Medical University, Nanjing 211166, Jiangsu Province, China. yankaixia@nimu.edu.cn

AIMS AND SCOPE

The primary aim of *World Journal of Diabetes* (*WJD*, *World J Diabetes*) is to provide scholars and readers from various fields of diabetes with a platform to publish high-quality basic and clinical research articles and communicate their research findings online.

WJD mainly publishes articles reporting research results and findings obtained in the field of diabetes and covering a wide range of topics including risk factors for diabetes, diabetes complications, experimental diabetes mellitus, type 1 diabetes mellitus, type 2 diabetes mellitus, gestational diabetes, diabetic angiopathies, diabetic cardiomyopathies, diabetic coma, diabetic ketoacidosis, diabetic nephropathies, diabetic neuropathies, Donohue syndrome, fetal macrosomia, and prediabetic state.

INDEXING/ABSTRACTING

The *WJD* is now abstracted and indexed in Science Citation Index Expanded (SCIE, also known as SciSearch®), Current Contents/Clinical Medicine, Journal Citation Reports/Science Edition, PubMed, PubMed Central, Reference Citation Analysis, China Science and Technology Journal Database, and Superstar Journals Database. The 2023 Edition of Journal Citation Reports® cites the 2022 impact factor (IF) for *WJD* as 4.2; IF without journal self cites: 4.1; 5-year IF: 4.5; Journal Citation Indicator: 0.69; Ranking: 51 among 145 journals in endocrinology and metabolism; and Quartile category: Q2.

RESPONSIBLE EDITORS FOR THIS ISSUE

Production Editor: *Yin-Xi Chen*; Production Department Director: *Xu Guo*; Editorial Office Director: *Jia-Ru Fan*.

NAME OF JOURNAL

World Journal of Diabetes

ISSN

ISSN 1948-9358 (online)

LAUNCH DATE

June 15, 2010

FREQUENCY

Monthly

EDITORS-IN-CHIEF

Lu Cai, Md. Shahidul Islam, Michael Horowitz

EDITORIAL BOARD MEMBERS

<https://www.wjgnet.com/1948-9358/editorialboard.htm>

PUBLICATION DATE

January 15, 2024

COPYRIGHT

© 2024 Baishideng Publishing Group Inc

INSTRUCTIONS TO AUTHORS

<https://www.wjgnet.com/bpg/gerinfo/204>

GUIDELINES FOR ETHICS DOCUMENTS

<https://www.wjgnet.com/bpg/GerInfo/287>

GUIDELINES FOR NON-NATIVE SPEAKERS OF ENGLISH

<https://www.wjgnet.com/bpg/gerinfo/240>

PUBLICATION ETHICS

<https://www.wjgnet.com/bpg/GerInfo/288>

PUBLICATION MISCONDUCT

<https://www.wjgnet.com/bpg/gerinfo/208>

ARTICLE PROCESSING CHARGE

<https://www.wjgnet.com/bpg/gerinfo/242>

STEPS FOR SUBMITTING MANUSCRIPTS

<https://www.wjgnet.com/bpg/GerInfo/239>

ONLINE SUBMISSION

<https://www.f6publishing.com>

Retrospective Study

Clinical study of different prediction models in predicting diabetic nephropathy in patients with type 2 diabetes mellitus

Sha-Sha Cai, Teng-Ye Zheng, Kang-Yao Wang, Hui-Ping Zhu

Specialty type: Endocrinology and metabolism**Provenance and peer review:** Unsolicited article; Externally peer reviewed.**Peer-review model:** Single blind**Peer-review report's scientific quality classification**Grade A (Excellent): 0
Grade B (Very good): B, B
Grade C (Good): C
Grade D (Fair): D
Grade E (Poor): 0**P-Reviewer:** Dąbrowski M, Poland; Di Ciaula A, Italy; Mostafavinia A, Iran**Received:** August 24, 2023**Peer-review started:** August 24, 2023**First decision:** November 9, 2023**Revised:** November 25, 2023**Accepted:** December 25, 2023**Article in press:** December 25, 2023**Published online:** January 15, 2024**Sha-Sha Cai, Teng-Ye Zheng, Kang-Yao Wang, Hui-Ping Zhu,** Department of Nephrology, The First People's Hospital of Wenling, Wenling 317500, Zhejiang Province, China**Corresponding author:** Hui-Ping Zhu, MM, Associate Chief Physician, Reader in Health Technology Assessment, Department of Nephrology, The First People's Hospital of Wenling, No. 333 Chuan'an South Road, Chengxi Street, Wenling 317500, Zhejiang Province, China. zhuhiping2261@163.com**Abstract****BACKGROUND**

Among older adults, type 2 diabetes mellitus (T2DM) is widely recognized as one of the most prevalent diseases. Diabetic nephropathy (DN) is a frequent complication of DM, mainly characterized by renal microvascular damage. Early detection, aggressive prevention, and cure of DN are key to improving prognosis. Establishing a diagnostic and predictive model for DN is crucial in auxiliary diagnosis.

AIM

To investigate the factors that impact T2DM complicated with DN and utilize this information to develop a predictive model.

METHODS

The clinical data of 210 patients diagnosed with T2DM and admitted to the First People's Hospital of Wenling between August 2019 and August 2022 were retrospectively analyzed. According to whether the patients had DN, they were divided into the DN group (complicated with DN) and the non-DN group (without DN). Multivariate logistic regression analysis was used to explore factors affecting DN in patients with T2DM. The data were randomly split into a training set ($n = 147$) and a test set ($n = 63$) in a 7:3 ratio using a random function. The training set was used to construct the nomogram, decision tree, and random forest models, and the test set was used to evaluate the prediction performance of the model by comparing the sensitivity, specificity, accuracy, recall, precision, and area under the receiver operating characteristic curve.

RESULTS

Among the 210 patients with T2DM, 74 (35.34%) had DN. The validation dataset showed that the accuracies of the nomogram, decision tree, and random forest models in predicting DN in patients with T2DM were 0.746, 0.714, and 0.730,

respectively. The sensitivities were 0.710, 0.710, and 0.806, respectively; the specificities were 0.844, 0.875, and 0.844, respectively; the area under the receiver operating characteristic curve (AUC) of the patients were 0.811, 0.735, and 0.850, respectively. The Delong test results revealed that the AUC values of the decision tree model were lower than those of the random forest and nomogram models ($P < 0.05$), whereas the difference in AUC values of the random forest and column-line graph models was not statistically significant ($P > 0.05$).

CONCLUSION

Among the three prediction models, random forest performs best and can help identify patients with T2DM at high risk of DN.

Key Words: Type 2 diabetes mellitus; Diabetic nephropathy; Random forest; Decision-making tree; Nomogram; Forecast

©The Author(s) 2024. Published by Baishideng Publishing Group Inc. All rights reserved.

Core Tip: Machine learning is widely used in medical prediction models. Logistic regression (nomogram), decision tree, and random forest models are three important machine learning techniques. However, few studies have compared the predictive efficacies of these three models in patients with type 2 diabetes mellitus and diabetic nephropathy. Here, we established three risk prediction models-nomogram, decision tree, and random forest-for comparison and found that random forest has the strongest combined predictive power.

Citation: Cai SS, Zheng TY, Wang KY, Zhu HP. Clinical study of different prediction models in predicting diabetic nephropathy in patients with type 2 diabetes mellitus. *World J Diabetes* 2024; 15(1): 43-52

URL: <https://www.wjgnet.com/1948-9358/full/v15/i1/43.htm>

DOI: <https://dx.doi.org/10.4239/wjd.v15.i1.43>

INTRODUCTION

Type 2 diabetes mellitus (T2DM) is one of the most common diseases affecting the older population. However, its incidence among children, adolescents, and young adults is increasing because of obesity, physical inactivity, and poor dietary habits. According to the International Diabetes Federation, approximately 537 million people (20-79 years old) worldwide currently have diabetes, with more than 90% of cases being T2DM[1]. Approximately 4.2 million people died of diabetes and its complications in 2019. Diabetic nephropathy (DN) is a chronic kidney disease induced by DM and a frequent microvascular complication of DM[2]. Relevant research data show that approximately 1/100 patients with diabetes develop end-stage renal disease yearly, and approximately 3/50 patients with massive albuminuria eventually develop end-stage renal disease yearly[3]. Patients with DN have a higher risk of death than those with diabetes alone or without comorbid DN. However, recent studies have shown that the prevalence of DN in patients with T2DM ranges between 20% and 40% [4,5]. To reduce the death rate of patients, early identification, prevention, and slowing down the development of DN are important. However, random urine measurement of the urinary albumin/creatinine ratio and 24 h urinary albumin quantification have shortcomings in the diagnosis of DN. Renal biopsy is the gold standard for the diagnosis of DN, but the acceptance of the examination is often low, and the economic cost is high[6]. Therefore, constructing a diagnostic and predictive model of DN plays a significant role in auxiliary diagnosis. At present, machine learning has been widely used in medical prediction models. Logistic regression (nomogram), decision tree, and random forest models are three important techniques in machine learning, all of which can quickly mine effective information from data; However, their application effects differ for different data types[7]. Little research has compared the predictive efficacies of the three models for DN in patients with T2DM. Therefore, this study established a prediction model for DN in patients with T2DM based on a nomogram, decision tree, and random forest and compared the prediction efficacy of the three models, providing a basis for the clinical identification of high-risk populations.

MATERIALS AND METHODS

Data sources

First, this was a retrospective study. A total of 210 patients admitted to the First People's Hospital of Wenling with a clear diagnosis of T2DM between August 2019 and August 2022 were selected for this study. According to the diagnostic information, 74 of the 210 patients with T2DM complicated by DN were defined as the DN group, and the remaining 136 patients with T2DM without concurrent DN were defined as the non-DN group. The inclusion criteria were as follows: (1) Age 18 to 75 years; (2) T2DM diagnosed according to the diagnostic criteria; and (3) Complete clinical data, including demographic data and laboratory test results. The exclusion criteria were as follows: (1) Definite diagnosis of primary kidney disease or secondary kidney disease of the immune system, blood system, or drug; (2) Complication with severe

primary diseases of the digestive, respiratory, cardiovascular, hematological, and nervous systems, accompanied by more than one malignant tumor; and (3) Rapidly progressive hypertension or diseases other than cerebrovascular diseases within the last 3 mo.

Diagnostic criteria

Standards criteria for DM: Based on “Standards for the Diagnosis and Treatment of Diabetes (2023 Edition)” of ADA[8] and “Guidelines for the Diagnosis, Prevention and Treatment of Type 2 Diabetes”[9]. Diagnostic standards were developed as follows: (1) Common symptoms of diabetes such as polydipsia, polyuria, polyphagia, unexplained weight loss, and random blood sugar ≥ 11.1 mmol/L; (2) Fasting blood sugar ≥ 7.0 mmol/L; or (3) Blood glucose ≥ 11.1 mmol/L 2 h after dextrose load. Patients with no typical symptoms of diabetes need to be reexamined on another day to confirm. The fasting state refers to not eating any calories for > 8 h; random blood glucose levels were measured at any time of the day, regardless of the time of the last meal.

Diagnostic criteria for DN: The diagnostic criteria were established based on the ADA “Standards for the Diagnosis and Treatment of Diabetes (2023 Edition)”, Guidelines for the Diagnosis, Prevention and Treatment of Type 2 Diabetes”, and “Guidelines for primary management of diabetic kidney disease in China”[10]. That is, patients with diabetes with renal impairment and urinary microalbumin/creatinine ratio (ACR) ≥ 30 mg/g (or ≥ 3 mg/mmol) or glomerular filtration rate < 60 mL/min/1.73 m² for a total duration of > 3 mo can be diagnosed with DN.

Indicators of observation

General information: Age, sex, body mass index (kg/m²), diabetes duration, and history of high blood pressure, diabetic retinopathy (DR), and coronary heart disease. Laboratory indicators (collected within 24 h after admission): Fasting blood glucose (FBG, mmol/L), serum creatinine (Scr, μ mol/L), glycosylated hemoglobin (HbA_{1c}, %, 7% = 53 mmol/mol), blood urea nitrogen (BUN, mmol/L), total cholesterol (mmol/L), triglyceride (mmol/L), high-density lipoprotein cholesterol (mmol/L), and low-density lipoprotein cholesterol (mmol/L).

Statistical analysis

Data were analyzed and processed using the SPSS software (version 23.0). Quantitative data in accordance with the normal distribution were expressed as mean \pm SD, and the *t*-test was used for comparison among groups. Count data were expressed as (%), and the χ^2 test was used for comparison among groups. The DN and non-DN groups were subjected to univariate analysis, and variables with statistically significant variances were incorporated into a multifactor logistic regression analysis model to screen for highly relevant predictive variables. The prediction model was constructed using R language, and the data were randomly split into training and validation sets at a ratio of 7:3 for model construction and validation, respectively. A nomogram model was constructed using the “rms” package; in this model, multiple predictors were integrated, and line segments with a certain scale were drawn on the same plane to express the relationship between variables. A decision tree was constructed using the “rpart” package. In the construction of the decision tree, the Gini coefficient minimization criterion was adopted to select features, and a binary tree including root, internal, and leaf nodes was generated. The estimation error identification complexity parameter of cross-validation was automatically calculated, and the dataset was classified through multiple conditional discrimination processes to finally obtain the required results. A random forest was constructed using the “Random Forest” package, which is used to construct multiple decision trees. When a certain sample needs to be predicted, the predicted results of each tree in the forest are counted, and then the final result is selected from these predicted results by voting method. The sensitivity, specificity, accuracy, recall, precision, and area under the receiver operating characteristic curve (AUC) were used to compare the effectiveness of the models, and the best prediction effect of the model was selected. The DeLong test was used for AUC comparisons. Differences were considered statistically significant at $P < 0.05$.

RESULTS

Concurrent DN condition

A total of 210 patients with T2DM were included in this study, of whom 87 were men and 123 were women. Seventy-four patients had T2DM complicated with DN, and the incidence of DN was 35.24%. There were 74 patients in the DN group, with a mean age of 56.01 ± 9.41 years. There were 136 patients in the non-DN group, with a mean age of 57.42 ± 8.15 years.

Univariate analysis of DN in patients with T2DM

The clinical data of the patients in the DN and non-DN groups were compared. The duration of diabetes, FBG, Scr, HbA_{1c}, and BUN levels were higher in the DN group than in the non-DN group, and the proportion of patients with DR was also higher in the DN group than in the non-DN group ($P < 0.05$). More details are shown in Table 1.

Multivariate analysis of DN in patients with T2DM

Regression analysis was performed by taking the patients with T2DM complicated with DN to be the dependent variables and taking the duration of diabetes, FBG, Scr, HbA_{1c}, BUN, and DR as the independent variables (there was no collinearity problem between the diagnosed variables), as shown in Tables 2 and 3. Multivariate analysis showed that the duration of diabetes, FBG, Scr, HbA_{1c}, and DR were factors influencing DN in patients with T2DM ($P < 0.05$).

Table 1 Results of univariate analysis of diabetic nephropathy in patients with type 2 diabetes [n (%)/mean ± SD]

Indicators	DN group (n = 74)	Non-DN group (n = 136)	t/ χ^2	P value
Age	56.01 ± 9.41	57.42 ± 8.15	-1.129	0.260
Sex			1.150	0.284
Male	27 (36.49)	60 (44.12)		
Female	47 (63.51)	76 (55.88)		
BMI (kg/m ²)	24.60 ± 3.92	23.73 ± 2.94	1.673	0.097
Duration of diabetes (yr)	5.28 ± 1.34	4.86 ± 0.76	2.507	0.014
History of hypertension			0.465	0.495
Yes	24 (32.43)	38 (27.94)		
No	50 (67.57)	98 (72.06)		
DR			8.761	0.003
Yes	34 (45.95)	24 (17.65)		
No	40 (54.05)	112 (82.35)		
Coronary heart disease			0.350	0.554
Yes	19 (25.68)	30 (22.06)		
No	55 (74.32)	106 (77.94)		
FBG (mmol/L)	8.18 ± 1.67	7.71 ± 1.14	2.160	0.033
Scr (μmol/L)	91.25 ± 14.72	84.61 ± 9.80	3.485	0.001
HbA1c (%)	7.04 ± 1.33	6.29 ± 1.05	4.239	< 0.001
BUN (mmol/L)	7.10 ± 0.96	6.78 ± 1.13	2.111	0.036
TC (mmol/L)	4.81 ± 0.89	4.83 ± 0.97	-0.189	0.851
TG (mmol/L)	1.78 ± 0.4	1.72 ± 0.38	1.090	0.277
HDL-C (mmol/L)	1.41 ± 0.31	1.45 ± 0.32	-0.972	0.332
LDL-C (mmol/L)	3.31 ± 0.57	3.33 ± 0.86	-0.185	0.853

BMI: Body mass index; FBG: Fasting blood glucose; Scr: Serum creatinine; HbA1c: Glycosylated hemoglobin; BUN: Blood urea nitrogen; DR: Diabetic retinopathy; TC: Total cholesterol; TG: Triglyceride; HDL-C: High-density lipoprotein cholesterol; LDL-C: Low-density lipoprotein cholesterol.

Nomogram model

Based on the results of the multifactor logistic regression analysis, the obtained independent predictors (FBG, Scr, HbA1c, DR, and duration of diabetes) were used to construct a nomogram model for predicting DN in patients with T2DM (Figure 1).

Decision tree model

A decision-tree prediction model of DN in patients with T2DM was constructed, and four explanatory variables were screened: FBG, HbA1c, Scr, and duration of diabetes. The results showed that the duration of diabetes was the first factor influencing DN in patients with T2DM. The model identified six judgment rules, of which three judged concurrent DN and three judged no concurrent DN. The incidence of DN in patients with T2DM with duration of diabetes ≥ 7 years was 6%. The incidence of DN in patients with T2DM with duration of diabetes < 7 years, HbA1c ≥ 7.6%, and Scr ≥ 89 μmol/L was 5%. The incidence of DN in patients with T2DM with duration of diabetes < 7 years, HbA1c < 7.6%, FBG ≥ 7.5 mmol/L, and Scr ≥ 98 μmol/L was 5% (Figure 2).

Random forest model

Based on the overall change in the prediction precision of the constructed random forest model, the variables affecting DN in patients with T2DM were HbA1c, Scr, FBG, duration of diabetes, and DR (Figure 3).

Evaluation of prediction effect of three models

In the validation set, the overall evaluation of the efficacy of the nomogram in predicting concurrent DN in patients with T2DM was not significantly different from that of the random forest model. In contrast, the overall efficacy of the decision tree model was significantly lower than that of the nomogram and random forest model, with significant differences (all

Table 2 Variable assignment

Factors	Variables of interest	Assignment of value
Concurrent DN or not	Y	Yes = 1, No = 0
Duration of diabetes (yr)	× 1	Numerical value
FBG (mmol/L)	× 2	Numerical value
Scr (μmol/L)	× 3	Numerical value
HbA1c (%)	× 4	Numerical value
BUN (mmol/L)	× 5	Numerical value
DR	× 6	Yes = 1, No = 0

FBG: Fasting blood glucose; Scr: Serum creatinine; HbA1c: Glycosylated hemoglobin; BUN: Blood urea nitrogen; DR: Diabetic retinopathy.

Table 3 Results of multivariate analysis of diabetic nephropathy in patients with type 2 diabetes

Factors	B	SE	Wald	P value	OR	95%CI
Duration of diabetes (yr)	0.352	0.164	4.619	0.032	1.421	1.031-1.959
FBG (mmol/L)	0.272	0.128	4.518	0.034	1.312	1.021-1.685
Scr (μmol/L)	0.063	0.015	16.745	< 0.001	1.065	1.033-1.097
HbA1c (%)	0.707	0.161	19.31	< 0.001	2.029	1.480-2.781
BUN (mmol/L)	0.250	0.158	2.507	0.113	1.283	0.942-1.748
DR	0.883	0.360	6.035	0.014	2.419	1.196-4.894

FBG: Fasting blood glucose; Scr: Serum creatinine; HbA1c: Glycosylated hemoglobin; BUN: Blood urea nitrogen; DR: Diabetic retinopathy; OR: Odds ratio; CI: Confidence interval.

$P > 0.05$) (Table 4 and Figure 4).

DISCUSSION

According to the National Kidney Foundation, persistent proteinuria is the primary indicator of kidney injury. ACR > 30 μg/mg in random urine samples was defined as renal injury, of which 30-300 μg/mg was the microalbuminuria stage, suggesting that the kidney had increased capillary permeability, which was the earliest sign of renal injury in patients with diabetes. In patients with T2DM, abnormal blood glucose levels frequently exist prior to diagnosis, and patients with T2DM may already have microalbuminuria (or clinical albuminuria) at the time of initial diagnosis[11,12]. Nephropathy is one of the most frequent complications of diabetes, and its insidious characteristics provide an intervention node for studying DN. Interventions for early detection, early detection diagnosis, and treatment can prevent the occurrence of DN and reduce or delay the occurrence of end-stage renal disease[13]. Therefore, identifying and intervening in clinically modifiable factors for the occurrence and progression of DN remains the primary strategy for its prevention and cure of DN.

The incidence of DN complications in patients with T2DM in this study was 35.24%. Wagnew *et al*[14] revealed that the incidence of DN was 35.3%, which is consistent with the results of this study. However, Zhang *et al*[15] revealed that the incidence of DN was only 21.8%, which was significantly lower than that reported in the present study. The incidence of DN varies greatly among studies and may be related to factors such as population characteristics and regions. The findings of this study revealed that FBG, Scr, HbA1c, DR, and duration of diabetes were factors affecting DN in patients with T2DM. The levels of FBG, Scr, HbA1c, and duration of diabetes in the DN group were higher than those in the non-DN group, and the proportion of DR was higher than that in the non-DN group, which was similar to the results of previous studies[16-18]. This suggests that patients with DN have worse glycemic control and poorer renal function[19, 20]. Moreover, compared with patients with T2DM without DR, patients with DR are more likely to develop DN.

Nagel *et al*[21] followed high-risk populations for > 20 years and found that fasting hyperglycemia was a predictor of high albumin leakage rate. Hyperglycemia causes renal damage through the activation of multiple pathways, including the formation of glycosylated complexes by the interaction of glucose with proteins outside the cell, metabolism to sorbitol through the polyol pathway, and metabolism to glucosamine through the hexosamine biosynthesis process, thereby mediating hyperglucose-induced renal damage. HbA1c can cause microvascular damage. In a tissue environment

Table 4 Efficacy of three model validation sets in predicting concurrent diabetic nephropathy in patients with type 2 diabetes

Model	Accuracy	Sensitivity	Specificity	Rate of recall	Rate of precision	AUC (95%CI)
Nomogram	0.746	0.710	0.844	0.906	0.690	0.811 (0.700-0.923)
Decision tree	0.714	0.710	0.875	0.906	0.659	0.735 (0.602-0.869)
Random forest	0.730	0.806	0.844	0.906	0.674	0.850 (0.750-0.950)

AUC: Area under the receiver operating characteristic curve; 95%CI: 95% confidence interval.

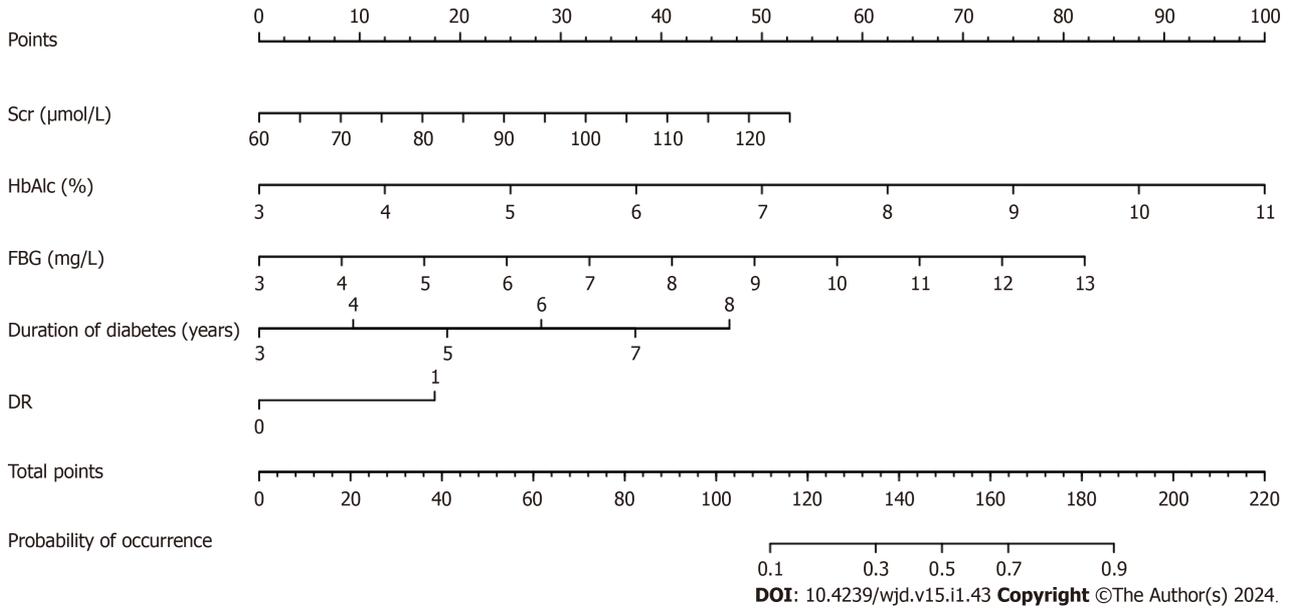


Figure 1 Nomogram prediction model for diabetic nephropathy in patients with type 2 diabetes mellitus. Scr: Serum creatinine; HbA1c: Glycosylated hemoglobin; DR: Diabetic retinopathy; FBG: Fasting blood glucose.

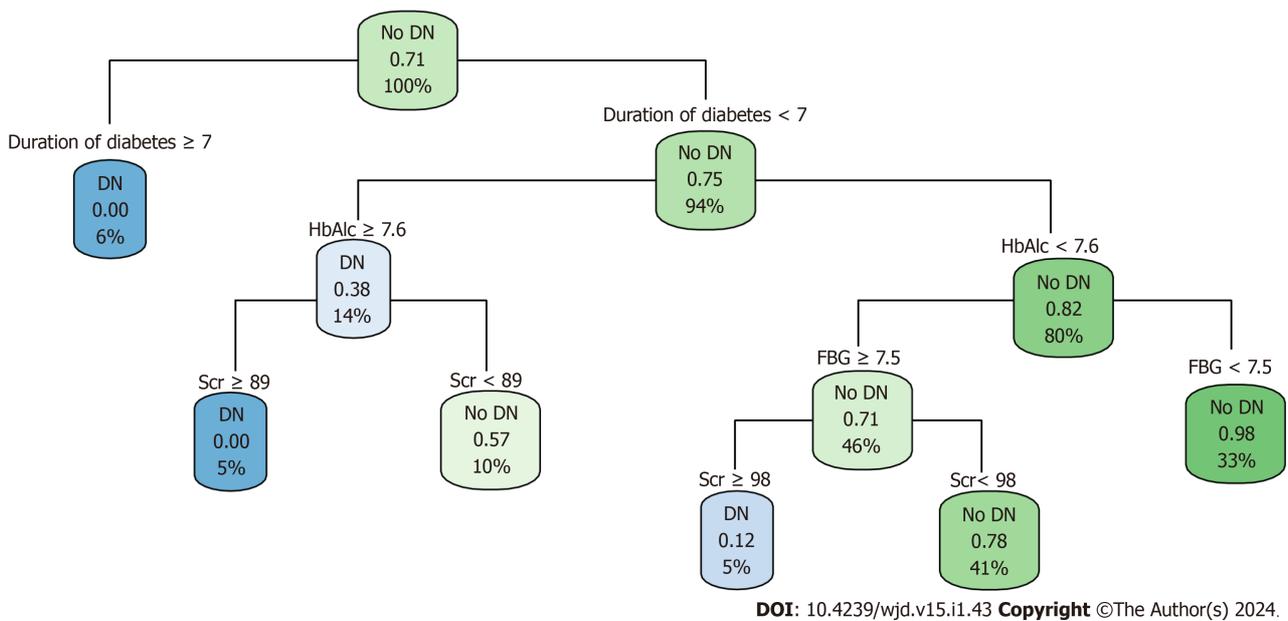


Figure 2 Decision tree model of diabetic nephropathy in patients type 2 diabetes. Scr: Serum creatinine; HbA1c: Glycosylated hemoglobin; FBG: Fasting blood glucose.

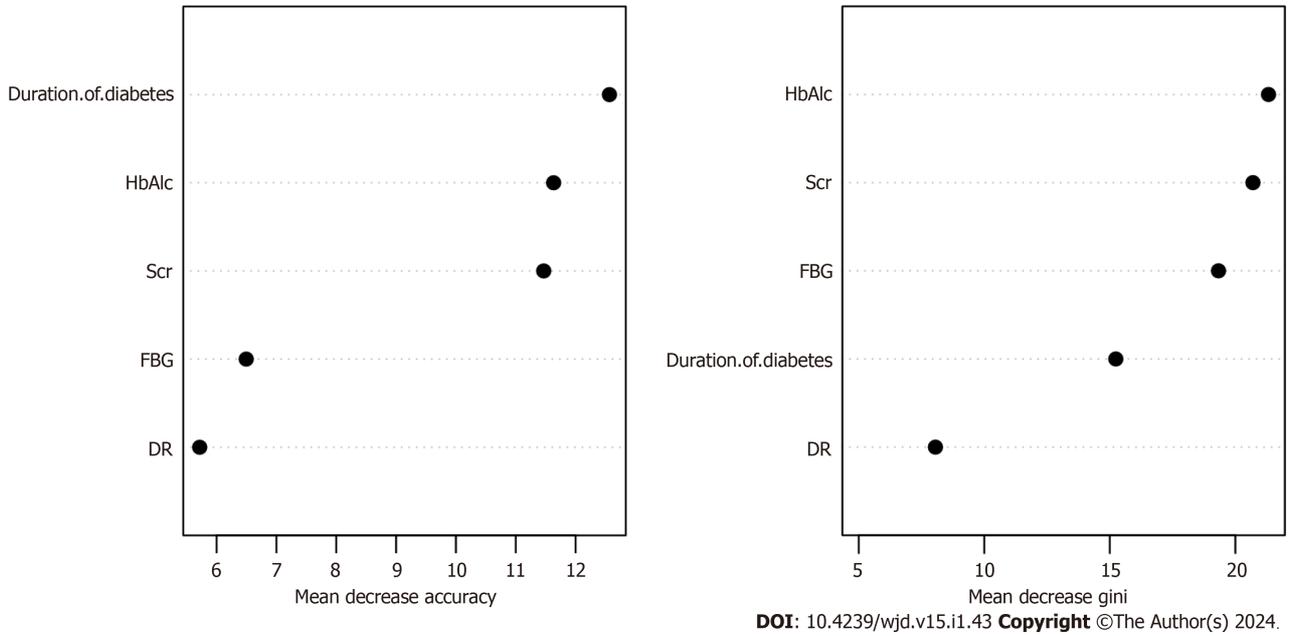


Figure 3 Random forest model of diabetic nephropathy in patients with type 2 diabetes. Scr: Serum creatinine; HbA1c: Glycosylated hemoglobin; DR: Diabetic retinopathy; FBG: Fasting blood glucose.

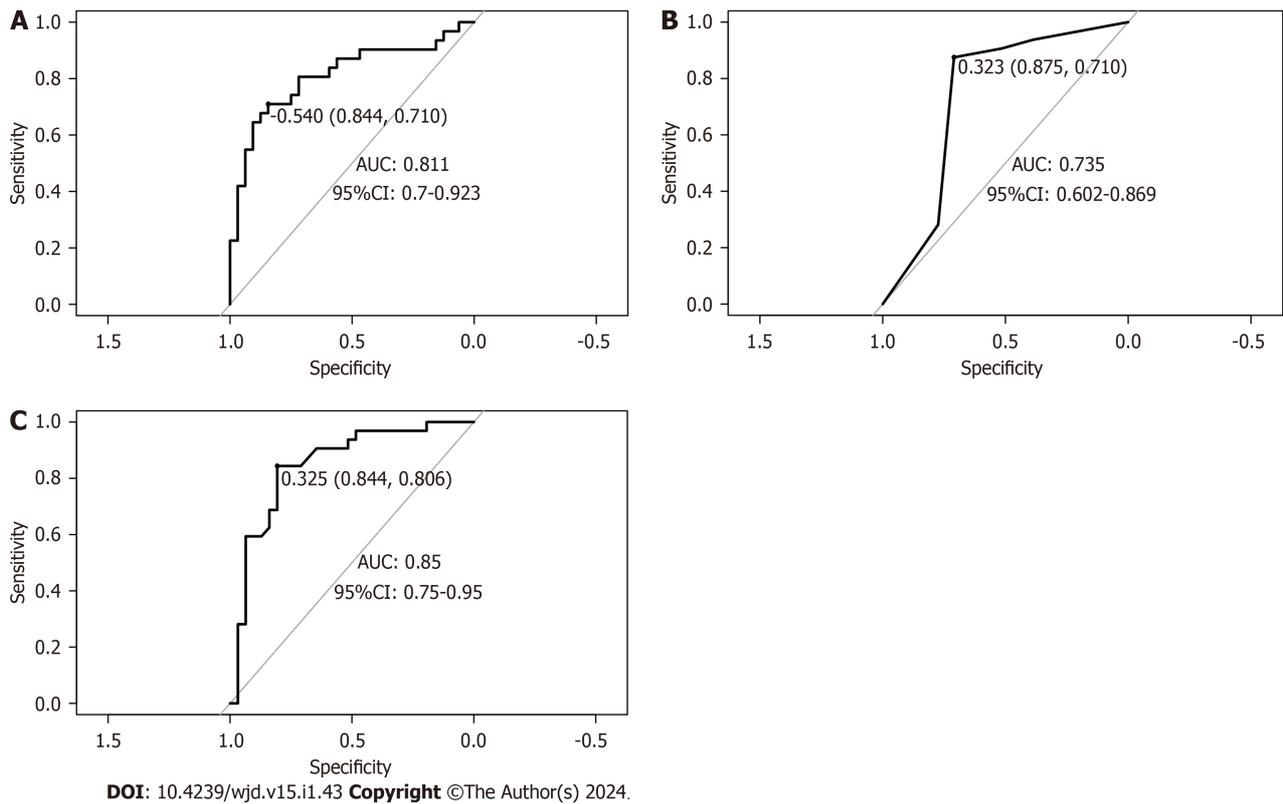


Figure 4 Receiver operating characteristic curves for the validation sets of the three models. A: Nomogram; B: Decision-making tree; C: Random forest. AUC: Area under the receiver operating characteristic curve; CI: Confidence interval.

with high glucose levels, the non-enzymatic catalytic process of the glycation reaction is accelerated, manifesting as a continuous increase in HbA1c. After glycation, abnormal hemoglobin levels cause chronic damage to the microcirculation vessels, damaging the basement membrane charge barrier and urine proteins. This indicates that microvascular damage to diabetic kidney tissues is a risk factor for microvascular or macrovascular lesions in patients with diabetes[22]. This study also confirmed that high blood glucose and HbA1c levels are associated with DN. Since Scr is mainly excreted in urine *via* glomerular filtration from the blood and is almost not reabsorbed by renal tubules, the output of creatinine is constant when renal function is normal or slightly damaged, and an increase in Scr indicates that renal function is

damaged[23]. Diabetes duration is the main risk factor for various complications in patients with T2DM, especially the primary risk factor for renal complications. If the duration of diabetes is > 5 years, microalbuminuria can occur. Without active intervention, it can progress to DN[24]. Studies have shown that DN and DR have the same risk factors, such as diabetes course and FBG[25]. Diabetic retinal abnormalities are associated with glomerular injury. Studies have shown that changes in retinal arteriole and venule diameter are associated with renal histological changes, such as basement membrane thickness and mesangial matrix volume increase[26]. In patients with diabetes with macroalbuminuria, retinopathy is rare; Therefore, DR combined with macroalbuminuria is more likely to be diagnosed as DN. For patients with DR, early intervention of risk factors and renal pathology examination should be performed to control the risk of DN.

Nomograms, decision trees, and random forest prediction models were established based on the above indicators. The results showed that the random forest model performed better than the nomogram and decision tree models in terms of AUC, sensitivity, and other evaluation indicators. In contrast, the overall evaluation indicators of the nomogram model were better than those of the decision tree model. The statistical difference between the AUC of the random forest and nomogram model was not significant; However, the statistical difference between the AUC of the decision tree and nomogram and the random forest model was significant. In general, the comprehensive predictive abilities of the three prediction models for concurrent DN in patients with T2DM constructed in this study were as follows: Decision tree < nomogram < random forest. This may be because the decision tree model prefers variables with higher values; however, these variables are not necessarily the best predictive variables, which reduces the predictive ability of the decision tree model. In addition, eliminating some candidate variables in the pruning process of the decision-tree model reduces its predictive ability to a certain extent[27]. A nomogram is a model that integrates multiple related factors to predict the probability of an event; it is intuitive, visual, and has good accuracy[28]. However, owing to its easy form (very much like a linear model), it is difficult to capture complex relationships and handle the problem of data imbalance and is sensitive to multicollinear data. Therefore, the prediction efficiency was slightly lower than that of the random forest model. However, random forest integrates the output of an individual decision tree to produce the final prediction result, which has the characteristics of robust operation. It is not easily affected by collinearity between variables and has a significant effect on reducing the variance of the model[29]. Tseng *et al*[30] found that the nomogram and random forest models outperformed the decision tree model in the diagnosis of postoperative cardiac patients with acute kidney injury. Hu *et al* [31] showed that the random forest model has better predictive power than the nomogram model and can effectively predict the risk of cognitive impairment in older adults. These results further demonstrate that the random forest model has a strong comprehensive prediction ability. As this was a retrospective study, all data were obtained from the same hospital, and the sample representation was limited.

CONCLUSION

This study established nomogram, random forest, and decision tree prediction models for T2DM complicated by DN using a machine learning algorithm. The result showed that the random forest model had good prediction and stability, thus providing a reference for the clinical identification of T2DM complicated by DN. Future studies need to further validate the model through prospective multi-center data and include more variables and samples to further improve the predictive ability of the model to better guide clinical practice.

ARTICLE HIGHLIGHTS

Research background

Hyperglycemia is the main pathophysiological feature of diabetes, and its complications are the key factors of death and disability in patients with diabetes. Diabetic nephropathy (DN) is a microvascular complication and is one of the main complications of diabetes. The initial prediction of DN is beneficial for taking measures to prevent and delay the occurrence and progression of corresponding complications. Machine learning has been widely used to construct predictive models for diabetic complications.

Research motivation

Patients with type 2 diabetes mellitus (T2DM) complicated by DN are at high risk of mortality. We explored the factors affecting the complications of DN to establish three prediction models commonly used in medicine, compared the prediction effects, and selected the optimal model to provide a basis for clinical identification of patients with T2DM complicated with DN.

Research objectives

This study aimed to explore the factors influencing T2DM complicated with DN and use these factors to construct a prediction model for DN. The prediction effect of random forest is the best among the three models of nomogram, decision tree, and random forest and may become a useful tool for the early recognition of the risk of DN.

Research methods

We retrospectively analyzed the clinical data of 210 patients with T2DM treated at our hospital between August 2019 and August 2022. Factors influencing DN were analyzed, and nomograms, decision trees, and random forest prediction models were established to compare their prediction efficiency. These three prediction methods are widely used in the medical field and have advantages and limitations. At the same time, through research, we can select a more suitable model to predict the complication risk of DN.

Research results

Fasting blood glucose, serum creatinine, glycosylated hemoglobin, diabetic retinopathy, and the duration of diabetes were independent factors influencing DN. Among the established nomograms, decision trees, and random forest prediction models, random forest has the best predictive ability and can be applied to the prevention and early screening of DN. Future studies should validate the model using prospective and multi-center data and include more samples and variables to further improve the prediction ability of the model. In addition, existing algorithms should be further improved, and a combination of multiple algorithms should be considered to improve the prediction accuracy.

Research conclusions

In this study, the predictive performances of three models were compared. The random forest model performed best in predicting the risk of DN in patients with T2DM and may be a useful alternative tool for diagnosing T2DM.

Research perspectives

Future studies should include larger and more comprehensive samples, conduct multi-center studies, further improve existing algorithms, and consider the combination of multiple algorithms to construct a more complete and accurate prediction model.

FOOTNOTES

Author contributions: Cai SS contributed to the conception and design of this study; Zheng TY and Wang KY participated in the administrative support; Zhu HP took part in the provision of study materials or patients; and all authors approved the final manuscript.

Institutional review board statement: The study was reviewed and approved by the First People's Hospital of Wenling (Approval No. KY-2023-2034-01).

Informed consent statement: Approved exemption for informed consent.

Conflict-of-interest statement: All the authors report no relevant conflicts of interest for this article.

Data sharing statement: The clinical data for research can be obtained from the corresponding author.

Open-Access: This article is an open-access article that was selected by an in-house editor and fully peer-reviewed by external reviewers. It is distributed in accordance with the Creative Commons Attribution NonCommercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <https://creativecommons.org/licenses/by-nc/4.0/>

Country/Territory of origin: China

ORCID number: Sha-Sha Cai 0000-0002-3297-4702; Teng-Ye Zheng 0009-0008-8444-9054; Kang-Yao Wang 0009-0003-3898-2388; Hui-Ping Zhu 0009-0000-9631-1951.

S-Editor: Wang JJ

L-Editor: A

P-Editor: Xu ZH

REFERENCES

- 1 **Magliano DJ**, Boyko EJ; IDF Diabetes Atlas 10th edition scientific committee. IDF DIABETES ATLAS [Internet]. Brussels: International Diabetes Federation; 2021– [PMID: 35914061]
- 2 **American Diabetes Association**. 2. Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes-2020. *Diabetes Care* 2020; **43**: S14-S31 [PMID: 31862745 DOI: 10.2337/dc20-S002]
- 3 **Rosolowsky ET**, Skupien J, Smiles AM, Niewczas M, Roshan B, Stanton R, Eckfeldt JH, Warram JH, Krolewski AS. Risk for ESRD in type 1 diabetes remains high despite renoprotection. *J Am Soc Nephrol* 2011; **22**: 545-553 [PMID: 21355053 DOI: 10.1681/ASN.2010040354]
- 4 **Luk AOY**, Hui EMT, Sin MC, Yeung CY, Chow WS, Ho AYY, Hung HF, Kan E, Ng CM, So WY, Yeung CK, Chan KS, Chan KW, Chan PF, Siu SC, Tiu SC, Yeung VTF, Chan JCN, Chan FWK, Cheung C, Cheung NT, Ho ST, Lam KSL, Yu LWL, Chao D, Lau IT. Declining Trends of Cardiovascular-Renal Complications and Mortality in Type 2 Diabetes: The Hong Kong Diabetes Database. *Diabetes Care* 2017; **40**:

- 928-935 [PMID: [28490423](#) DOI: [10.2337/dc16-2354](#)]
- 5 **de Boer IH**, Rue TC, Hall YN, Heagerty PJ, Weiss NS, Himmelfarb J. Temporal trends in the prevalence of diabetic kidney disease in the United States. *JAMA* 2011; **305**: 2532-2539 [PMID: [21693741](#) DOI: [10.1001/jama.2011.861](#)]
- 6 **Afkarian M**, Zelnick LR, Hall YN, Heagerty PJ, Tuttle K, Weiss NS, de Boer IH. Clinical Manifestations of Kidney Disease Among US Adults With Diabetes, 1988-2014. *JAMA* 2016; **316**: 602-610 [PMID: [27532915](#) DOI: [10.1001/jama.2016.10924](#)]
- 7 **Handelman GS**, Kok HK, Chandra RV, Razavi AH, Lee MJ, Asadi H. eDoctor: machine learning and the future of medicine. *J Intern Med* 2018; **284**: 603-619 [PMID: [30102808](#) DOI: [10.1111/joim.12822](#)]
- 8 **ElSayed NA**, Aleppo G, Aroda VR, Bannuru RR, Brown FM, Bruemmer D, Collins BS, Hilliard ME, Isaacs D, Johnson EL, Kahan S, Khunti K, Leon J, Lyons SK, Perry ML, Prahalad P, Pratley RE, Seley JJ, Stanton RC, Gabbay RA; on behalf of the American Diabetes Association. 2. Classification and Diagnosis of Diabetes: Standards of Care in Diabetes-2023. *Diabetes Care* 2023; **46**: S19-S40 [PMID: [36507649](#) DOI: [10.2337/dc23-S002](#)]
- 9 **Tsur A**, Harman-Bohem I, Buchs AE, Raz I, Wainstein J. [The guidelines for the diagnosis prevention and treatment of type 2 diabetes mellitus-2005]. *Harefuah* 2006; **145**: 583-586, 630 [PMID: [16983842](#)]
- 10 **Li L**, Yang Y, Zhu X, Xiong X, Zeng L, Xiong S, Jiang N, Li C, Yuan S, Xu H, Liu F, Sun L. Design and validation of a scoring model for differential diagnosis of diabetic nephropathy and nondiabetic renal diseases in type 2 diabetic patients. *J Diabetes* 2020; **12**: 237-246 [PMID: [31602779](#) DOI: [10.1111/1753-0407.12994](#)]
- 11 **Qi C**, Mao X, Zhang Z, Wu H. Classification and Differential Diagnosis of Diabetic Nephropathy. *J Diabetes Res* 2017; **2017**: 8637138 [PMID: [28316995](#) DOI: [10.1155/2017/8637138](#)]
- 12 **Samsu N**. Diabetic Nephropathy: Challenges in Pathogenesis, Diagnosis, and Treatment. *Biomed Res Int* 2021; **2021**: 1497449 [PMID: [34307650](#) DOI: [10.1155/2021/1497449](#)]
- 13 **Kanwar YS**, Sun L, Xie P, Liu FY, Chen S. A glimpse of various pathogenetic mechanisms of diabetic nephropathy. *Annu Rev Pathol* 2011; **6**: 395-423 [PMID: [21261520](#) DOI: [10.1146/annurev.pathol.4.110807.092150](#)]
- 14 **Wagnew F**, Eshetie S, Kibret GD, Zegeye A, Dessie G, Mulugeta H, Alemu A. Diabetic nephropathy and hypertension in diabetes patients of sub-Saharan countries: a systematic review and meta-analysis. *BMC Res Notes* 2018; **11**: 565 [PMID: [30081966](#) DOI: [10.1186/s13104-018-3670-5](#)]
- 15 **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. *J Diabetes Res* 2020; **2020**: 2315607 [PMID: [32090116](#) DOI: [10.1155/2020/2315607](#)]
- 16 **Tziomalos K**, Athyros VG. Diabetic Nephropathy: New Risk Factors and Improvements in Diagnosis. *Rev Diabet Stud* 2015; **12**: 110-118 [PMID: [26676664](#) DOI: [10.1900/RDS.2015.12.110](#)]
- 17 **Hu Y**, Shi R, Mo R, Hu F. Nomogram for the prediction of diabetic nephropathy risk among patients with type 2 diabetes mellitus based on a questionnaire and biochemical indicators: a retrospective study. *Aging (Albany NY)* 2020; **12**: 10317-10336 [PMID: [32484786](#) DOI: [10.18632/aging.103259](#)]
- 18 **Shi R**, Niu Z, Wu B, Zhang T, Cai D, Sun H, Hu Y, Mo R, Hu F. Nomogram for the Risk of Diabetic Nephropathy or Diabetic Retinopathy Among Patients with Type 2 Diabetes Mellitus Based on Questionnaire and Biochemical Indicators: A Cross-Sectional Study. *Diabetes Metab Syndr Obes* 2020; **13**: 1215-1229 [PMID: [32368114](#) DOI: [10.2147/DMSO.S244061](#)]
- 19 **Zhou DM**, Wei J, Zhang TT, Shen FJ, Yang JK. Establishment and Validation of a Nomogram Model for Prediction of Diabetic Nephropathy in Type 2 Diabetic Patients with Proteinuria. *Diabetes Metab Syndr Obes* 2022; **15**: 1101-1110 [PMID: [35431563](#) DOI: [10.2147/DMSO.S357357](#)]
- 20 **Xi C**, Wang C, Rong G, Deng J. A Nomogram Model that Predicts the Risk of Diabetic Nephropathy in Type 2 Diabetes Mellitus Patients: A Retrospective Study. *Int J Endocrinol* 2021; **2021**: 6672444 [PMID: [33897777](#) DOI: [10.1155/2021/6672444](#)]
- 21 **Nagel G**, Zitt E, Peter R, Pompella A, Concin H, Lhotta K. Body mass index and metabolic factors predict glomerular filtration rate and albuminuria over 20 years in a high-risk population. *BMC Nephrol* 2013; **14**: 177 [PMID: [23962027](#) DOI: [10.1186/1471-2369-14-177](#)]
- 22 **Zhu Y**, Wang X, Wang W, Wang H, Zhang F. Expression and influence of pentraxin-3, HbA1c and ApoA1/ApoB in serum of patients with acute myocardial infarction combined with diabetes mellitus type 2. *Exp Ther Med* 2018; **15**: 4395-4399 [PMID: [29731826](#) DOI: [10.3892/etm.2018.5930](#)]
- 23 **Hu H**, Nakagawa T, Honda T, Yamamoto S, Okazaki H, Yamamoto M, Miyamoto T, Eguchi M, Kochi T, Shimizu M, Murakami T, Tomita K, Ogasawara T, Sasaki N, Uehara A, Kuwahara K, Kabe I, Mizoue T, Sone T, Dohi S; Japan Epidemiology Collaboration on Occupational Health Study Group. Low serum creatinine and risk of diabetes: The Japan Epidemiology Collaboration on Occupational Health Study. *J Diabetes Investig* 2019; **10**: 1209-1214 [PMID: [30756513](#) DOI: [10.1111/jdi.13024](#)]
- 24 **Ninković V**, Ninković S, Zivojinović D. [Cardiovascular autonomous dysfunction in diabetics: the influence of disease duration, glycoregulation degree and diabetes type]. *Srp Arh Celok Lek* 2008; **136**: 488-493 [PMID: [19069339](#) DOI: [10.2298/sarh0810488n](#)]
- 25 **Saini DC**, Kochar A, Poonia R. Clinical correlation of diabetic retinopathy with nephropathy and neuropathy. *Indian J Ophthalmol* 2021; **69**: 3364-3368 [PMID: [34708806](#) DOI: [10.4103/ijo.IJO_1237_21](#)]
- 26 **Li Y**, Su X, Ye Q, Guo X, Xu B, Guan T, Chen A. The predictive value of diabetic retinopathy on subsequent diabetic nephropathy in patients with type 2 diabetes: a systematic review and meta-analysis of prospective studies. *Ren Fail* 2021; **43**: 231-240 [PMID: [33478336](#) DOI: [10.1080/0886022X.2020.1866010](#)]
- 27 **Bamber JH**, Evans SA. The value of decision tree analysis in planning anaesthetic care in obstetrics. *Int J Obstet Anesth* 2016; **27**: 55-61 [PMID: [27026589](#) DOI: [10.1016/j.ijoa.2016.02.007](#)]
- 28 **Lv J**, Liu YY, Jia YT, He JL, Dai GY, Guo P, Zhao ZL, Zhang YN, Li ZX. A nomogram model for predicting prognosis of obstructive colorectal cancer. *World J Surg Oncol* 2021; **19**: 337 [PMID: [34857001](#) DOI: [10.1186/s12957-021-02445-6](#)]
- 29 **Uddin S**, Khan A, Hossain ME, Moni MA. Comparing different supervised machine learning algorithms for disease prediction. *BMC Med Inform Decis Mak* 2019; **19**: 281 [PMID: [31864346](#) DOI: [10.1186/s12911-019-1004-8](#)]
- 30 **Tseng PY**, Chen YT, Wang CH, Chiu KM, Peng YS, Hsu SP, Chen KL, Yang CY, Lee OK. Prediction of the development of acute kidney injury following cardiac surgery by machine learning. *Crit Care* 2020; **24**: 478 [PMID: [32736589](#) DOI: [10.1186/s13054-020-03179-9](#)]
- 31 **Hu M**, Shu X, Yu G, Wu X, Välimäki M, Feng H. A Risk Prediction Model Based on Machine Learning for Cognitive Impairment Among Chinese Community-Dwelling Elderly People With Normal Cognition: Development and Validation Study. *J Med Internet Res* 2021; **23**: e20298 [PMID: [33625369](#) DOI: [10.2196/20298](#)]



Published by **Baishideng Publishing Group Inc**
7041 Koll Center Parkway, Suite 160, Pleasanton, CA 94566, USA
Telephone: +1-925-3991568
E-mail: office@baishideng.com
Help Desk: <https://www.f6publishing.com/helpdesk>
<https://www.wjgnet.com>

