



## OPEN ACCESS

## EDITED BY

Ramkumar Kunka Mohanram,  
SRM Institute of Science  
and Technology, India

## REVIEWED BY

Godfrey Mutashambara Rwegerera,  
University of Botswana, Botswana  
Milena Raffi,  
University of Bologna, Italy

## \*CORRESPONDENCE

Yongjiang Cai

✉ caiyj2000@sina.cn

Weihua Yang

✉ benben0606@139.com

Jiantao Wang

✉ wangjiantao65@126.com

<sup>†</sup>These authors share first authorship

RECEIVED 19 May 2023

ACCEPTED 29 September 2023

PUBLISHED 26 October 2023

## CITATION

Gong D, Chen X, Yang L, Zhang Y,  
Zhong Q, Liu J, Yan C, Cai Y, Yang W and  
Wang J (2023) From normal population to  
prediabetes and diabetes: study of  
influencing factors and prediction models.  
*Front. Endocrinol.* 14:1225696.  
doi: 10.3389/fendo.2023.1225696

## COPYRIGHT

© 2023 Gong, Chen, Yang, Zhang, Zhong,  
Liu, Yan, Cai, Yang and Wang. This is an  
open-access article distributed under the  
terms of the [Creative Commons Attribution  
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or  
reproduction in other forums is permitted,  
provided the original author(s) and the  
copyright owner(s) are credited and that  
the original publication in this journal is  
cited, in accordance with accepted  
academic practice. No use, distribution or  
reproduction is permitted which does not  
comply with these terms.

# From normal population to prediabetes and diabetes: study of influencing factors and prediction models

Di Gong<sup>1,2†</sup>, Xiaohong Chen<sup>3†</sup>, Lin Yang<sup>3†</sup>, Yongjian Zhang<sup>3</sup>,  
Qianqian Zhong<sup>3</sup>, Jing Liu<sup>3</sup>, Chen Yan<sup>1</sup>, Yongjiang Cai<sup>3\*</sup>,  
Weihua Yang<sup>4\*</sup> and Jiantao Wang<sup>4\*</sup>

<sup>1</sup>Shenzhen Eye Hospital, Jinan University, Shenzhen, Guangdong, China, <sup>2</sup>The First Affiliated Hospital of Jinan University, Jinan University, Guangzhou, Guangdong, China, <sup>3</sup>Center of Health Management, Peking University Shenzhen Hospital, Shenzhen, Guangdong, China, <sup>4</sup>Shenzhen Eye Hospital, Jinan University, Shenzhen Eye Institute, Shenzhen, Guangdong, China

**Objective:** The purpose of this study is to investigate the independent influencing factors of the transition from normal population to prediabetes, and from prediabetes to diabetes, and to further construct clinical prediction models to provide a basis for the prevention and management of prediabetes and diabetes.

**Materials and methods:** The data for this study were based on clinical information of participants from the Health Management Center of Peking University Shenzhen Hospital. Participants were classified into normal group, prediabetes group, and diabetes group according to their functional status of glucose metabolism. Spearman's correlation coefficients were calculated for the variables, and a matrix diagram was plotted. Further, univariate and multivariate logistic regression analysis were conducted to explore the independent influencing factors. The independent influencing factors were used as predictors to construct the full-variable prediction model (Full.model) and simplified prediction model (Simplified.model).

**Results:** This study included a total of 5310 subjects and 22 variables, among which there were 1593(30%) in the normal group, 3150(59.3%) in the prediabetes group, and 567(10.7%) in the diabetes group. The results of the multivariable logistic regression analysis showed that there were significant differences in 9 variables between the normal group and the prediabetes group, including age (Age), body mass index(BMI), systolic blood pressure(SBP), urinary glucose (U.GLU), urinary protein(PRO), total protein(TP), globulin(GLB), alanine aminotransferase(ALT), and high-density lipoprotein cholesterol(HDL-C). There were significant differences in 7 variables between the prediabetes group and the diabetes group, including Age, BMI, SBP, U.GLU, PRO, triglycerides(TG), and HDL.C. The Full.model and Simplified.model constructed based on the above influencing factors had moderate discriminative power in both the training set and the test set.

**Conclusion:** Age, BMI, SBP, U.GLU, PRO, TP, and ALT are independent risk factors, while GLB and HDL.C are independent protective factors for the development of prediabetes in the normal population. Age, BMI, SBP, U.GLU, PRO, and TG are independent risk factors, while HDL.C is an independent protective factor for the progression from prediabetes to diabetes. The Full.model and Simplified.model developed based on these influencing factors have moderate discriminative power.

#### KEYWORDS

prediabetes, diabetes, influencing factors, prediction model, odds ratio (OR)

## 1 Introduction

Prediabetes, also known as impaired glucose regulation (IGR), is a pathological state where blood glucose levels are higher than normal but have not yet reached the diagnostic criteria for diabetes. It includes two types: impaired fasting glucose (IFG) and impaired glucose tolerance (IGT). IFG is fasting blood glucose between the normal range and diabetes criteria (6.1–6.9 mmol/L), while IGT is elevated blood glucose during a 2-hour oral glucose tolerance test (OGTT) but not meeting diabetes criteria (7.8–11.0 mmol/L) (1). According to a global epidemiological survey, nearly 400 million adults worldwide have prediabetes, with a prevalence of approximately 6.4% for IFG, 7.5% for IGT, and 2.4% for both IFG and IGT (2). In China, a cross-sectional survey found a high prediabetes prevalence of 35.7% among adults (3), much higher than in other regions around the world (4–6). Prediabetes is a high-risk state for developing diabetes, with an annual conversion rate of 5%–10%. However, some studies have indicated that a proportion of patients can return to normal glucose metabolism (1, 7–9). Therefore, prediabetic patients should take early control measures to prevent further development towards diabetes.

Diabetes is a chronic metabolic disease characterized by persistently high blood sugar levels, leading to damage to various organs and tissues in the body. According to the standards of the World Health Organization (WTO) and the American Diabetes Association (ADA), the diagnostic criteria for diabetes are a fasting blood sugar level  $\geq 7.0$  mmol/L or a random blood sugar level  $\geq 11.1$  mmol/L, or a blood sugar level  $\geq 11.1$  mmol/L after an oral glucose tolerance test (OGTT), or an HbA1c level of  $\geq 6.5\%$  measured by standardized DCCT analysis (10). Over the past 30 years, the number of diabetes patients worldwide has doubled, and there is a concerning trend of its occurrence among younger individuals (11, 12). The global prevalence of diabetes was estimated to be 9.3% (463 million people) in 2019 and is projected to increase to 10.2% (578 million people) by 2030 and 10.9% (700 million people) by 2045 (2). In populous countries, the estimated prevalence of diabetes among adults is 10.9% in China (3), 12–14% in the United States (13), and approximately 7.3% in all 15 states of India (5). Diabetes has undoubtedly become a significant challenge for global public health in the 21st century, particularly in developing countries like China and India (11).

Prediabetes, as a precursor to diabetes, can cause a range of health issues, even though IFG and IGT themselves should not be considered clinical entities. Research has shown that prediabetes is significantly associated with an increased risk of obstructive sleep apnea, composite cardiovascular disease, coronary heart disease, stroke, and all-cause mortality (14, 15). Moreover, if timely and effective measures are not taken to control blood glucose levels, prediabetes may progress into type 2 diabetes (1, 9). As a chronic metabolic disease, the long-term hyperglycemic state of diabetes patients can affect the structure and function of blood vessels through multiple pathways (16). Additionally, the hyperglycemic state can induce oxidative stress, activate the inflammatory response, and affect the coagulation system, leading to a series of pathological and physiological changes that ultimately result in the occurrence of various complications (17). According to a WHO report, diabetes has become the seventh leading cause of death in humans, with cardiovascular disease being the main cause of death and morbidity in diabetes patients (12, 18). Therefore, prevention and treatment of diabetes should be highly emphasized.

Current research indicates that prediabetes and the development of diabetes may be related to many risk factors, including age, family history, race, genetic mutations, lack of physical activity, unhealthy dietary habits, obesity, hypertension, lipoprotein, high cholesterol, and hypertriglyceridemia (1, 3, 19–22). By analyzing the disease risk factors of prediabetes and diabetes, building a clinical prediction model can help identify high-risk patients, but currently, there is no widely used prediction model in clinical practice. Wu et al. found that waist circumference, family history of diabetes, HbA1c, and fasting blood glucose levels were independently associated with the risk of prediabetes. A prediabetes prediction model was constructed by incorporating these four indicators, with an Area Under the Curve (AUC) of 0.70236, indicating a moderately low level of discrimination (23). Yokota et al. conducted a retrospective longitudinal study and found that family history of diabetes, male gender, elevated systolic blood pressure, blood glucose levels, HbA1c, and alanine aminotransferase were important independent predictors for the conversion of prediabetes to diabetes. The prediction model constructed using these variables had a Receiver Operating Characteristic (ROC) Curve of 0.8037, indicating a moderate level of discrimination (24).

Diabetes has emerged as a global public health concern, with its incidence and mortality rates steadily increasing. Prediabetes serves as a warning sign for diabetes, and early detection with effective interventions can prevent its progression, thus reducing the incidence and mortality rates of diabetes. Therefore, the purpose of this study is to conduct a statistical analysis of cross-sectional data from a population undergoing medical examinations, to explore the independent influencing factors associated with the transition from normal individuals to prediabetes and from prediabetes to diabetes. Additionally, the study aims to develop a clinical prediction model for these diseases. In this study, we aim to use blood glucose and HbA1c as the diagnostic gold standards, while considering other risk indicators as predictive factors. The goal is to identify and provide early warning of individuals at risk of prediabetes and diabetes among the population undergoing health examinations. The findings of this research have the potential to offer valuable insights into the influencing factors of prediabetes and diabetes, which could be of significance in enhancing our understanding of these conditions. Clinical practitioners may find the information helpful in making more informed decisions while diagnosing and treating diabetes patients. By identifying specific risk factors, tailored interventions may be developed to improve patient outcomes and enhance their quality of life. Furthermore, this study will serve as a crucial reference for public health workers in devising effective strategies for diabetes prevention and control, empowering them to better manage and prevent the occurrence of diabetes.

## 2 Materials and methods

### 2.1 Data source and collection

The original data for this study were collected from individuals who underwent health examinations at the Health Management Center of Peking University Shenzhen Hospital between January 2020 and March 2023. All participants underwent fasting blood glucose, random blood glucose, OGTT, and HbA1c testing according to WHO standards. The diagnostic criteria for diabetes are a fasting blood sugar level  $\geq 7.0$  mmol/L or a random blood sugar level  $\geq 11.1$  mmol/L, or a blood sugar level  $\geq 11.1$  mmol/L after oral glucose tolerance test (OGTT), or an HbA1c level of  $\geq 6.5\%$  measured by standardized DCCT analysis. The diagnostic criteria for prediabetes are a fasting blood glucose level in the range of 6.1–6.9 mmol/L, or a blood glucose level in the range of 7.8–11.0 mmol/L after OGTT. Participants were categorized into normal, prediabetes, and diabetes groups based on their glucose metabolism status.

### 2.2 Variable selection

Relevant literature was searched using keywords such as “prediabetes” and “diabetes” on databases including PUBMED, EMBASE, and Web of Science to determine the variables to be included in the study. The variables extracted from participants

were glucose metabolism status (Status), age (Age), gender (Gender), body mass index (BMI), systolic blood pressure (SBP), urinary glucose (U.GLU), urinary protein (PRO), total protein (TP), albumin (ALB), globulin (GLB), total bilirubin (T.BIL), direct bilirubin (DB), indirect bilirubin (IB), alanine aminotransferase (ALT), aspartate aminotransferase (AST), blood urea nitrogen (BUN), serum creatinine (SCr), uric acid (UA), total cholesterol (TC), triglycerides (TG), high-density lipoprotein cholesterol (HDL-C), and low-density lipoprotein cholesterol (LDL-C) (a total of 22 variables). The extracted data were then compiled and merged into a single file based on the participants' ID numbers.

### 2.3 Variable assignment

All categorical variables including Status, Gender, U.GLU, and PRO were assigned values. The remaining continuous variables were not assigned values. [Table 1](#) shows the assigned values for each variable.

### 2.4 Data processing and statistical analysis

This study used R 4.2.3 software for data processing and statistical analysis. Differences were considered statistically significant at  $P < 0.05$ . Firstly, the `complete.cases()` function was used to clean missing data. Then, the `summary()` function was used to perform descriptive statistical analysis on the variables in the dataset. The `cor()` function was used to calculate the Spearman correlation coefficient between variables and the matrix plot was generated using the `ggplot2` library. The `glm()` function was used to perform univariate regression analysis on all independent variables. Variables with statistically significant differences in univariate regression analysis were included in the multivariate regression analysis to identify independent influencing factors in the development from normal to prediabetes, and from prediabetes to diabetes. The dataset was randomly divided into training and testing sets at an 8:2 ratio. The `glm()` function was used to build a full variable prediction model (Full.model) and a simplified prediction model (Simplified.model) using the training set. The `roc()` function and `function()` function were used to calculate the discrimination, accuracy, precision, and recall of the Full.model and Simplified.model.

The specific explanations of the R language functions used above are as follows: `complete.cases()`: It is a function used for data processing to check if each row in a data frame or matrix contains complete data (without missing values). `summary()`: It is a function used to summarize statistical data, returning descriptive statistics for each variable, such as mean, median, minimum, maximum, and quantiles. `cor()`: It is a function used to calculate correlation coefficients, computing the correlation between columns of a data frame or matrix. `glm()`: It is a function used to fit Generalized Linear Models, allowing fitting various models, such as linear regression, logistic regression, and Poisson regression. `roc()`: It is a function used to compute ROC curves, which are graphical methods to

TABLE 1 Variable assignment explanation.

Variable name	Meaning of variables	Type of variables	Assignment description
Status	Glucose metabolic status	Categorical variable	1=Normal
			2=Prediabetes
			3=Diabetes
Age	Age	Numerical variable	
Gender	Gender	Categorical variable	0=Female
			1=Male
BMI	BMI Index	Numerical variable	
SBP	Systolic blood pressure	Numerical variable	
U.GLU	Urine glucose	Categorical variable	0=Negative
			1=Positive
PRO	Urine protein	Categorical variable	0=Negative
			1=Positive
TP	Total protein	Numerical variable	
ALB	Albumin	Numerical variable	
GLB	Globulin	Numerical variable	
T.BIL	Total bilirubin	Numerical variable	
DB	Direct bilirubin	Numerical variable	
IB	Indirect bilirubin	Numerical variable	
ALT	Glutathione aminotransferase	Numerical variable	
AST	Glutathione transaminase	Numerical variable	
BUN	Blood urea nitrogen	Numerical variable	
SCr	Blood creatinine	Numerical variable	
UA	Uric acid	Numerical variable	
TC	Total Cholesterol	Numerical variable	
TG	Triglycerides	Numerical variable	
HDL.C	High-density lipoprotein cholesterol	Numerical variable	
LDL.C	Low-density lipoprotein cholesterol	Numerical variable	

evaluate the performance of binary classification models. `function()`: It is a keyword used to create custom functions, enabling operations based on user-defined logic and returning calculated results.

## 3 Results

### 3.1 Detection rate of prediabetes, diabetes

The health examination data of the subjects were summarized and organized, and individuals with missing variables were excluded. Ultimately, 5310 participants were included in the study and divided into three groups based on their glucose metabolism status: normal group (1593 cases, 30%), prediabetes group (3150 cases, 59.3%), and diabetes group (567 cases, 10.7%).

### 3.2 Correlation analysis between variables

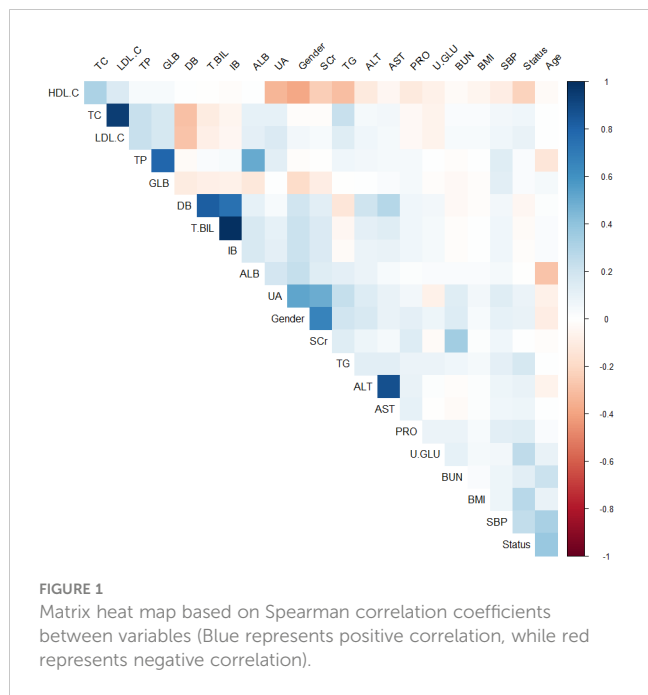
Calculate the Spearman correlation coefficient between each variable and use the “ggplot2” library to create a matrix heatmap. The color of each cell represents the degree of correlation between the corresponding variables. Blue represents positive correlation, while red represents negative correlation. The color depth varies according to the different correlation coefficients. The deeper the color, the stronger the correlation, while the lighter the color, the weaker the correlation. Variables with an absolute value of correlation coefficient  $> 0.8$  are considered to have a strong correlation. The results of the correlation analysis in this study indicate that there is a strong correlation between TC and LDL.C, DB and T.BIL, IB and T.BIL, and AST and ALT, as shown in [Figure 1](#).

### 3.3 Independent factors analysis of normal group and prediabetes group

Univariate regression analysis was performed on the independent variables of the normal group and prediabetes group. Variables with  $P < 0.05$  in the Univariate regression analysis were included in the multivariate regression analysis to analyze the influencing factors of the normal population developing into prediabetes. The results of multivariate regression analysis showed that Age, BMI, SBP, U.GLU, PRO, TP, GLB, ALT, and HDL.C were the independent influencing factors for the development of prediabetes in the normal population. Among them, Age, BMI, SBP, U.GLU, PRO, TP, and ALT were independent risk factors, while GLB and HDL.C were independent protective factors, as shown in [Table 2](#).

### 3.4 Analysis of independent influencing factors in the prediabetes and diabetes groups

We conducted a single-factor logistic regression analysis of the independent variables in the prediabetes and diabetes groups.



Variables with  $P < 0.05$  in the Univariate regression analysis were included in the multivariate regression analysis to analyze the factors influencing the development of diabetes in the prediabetic population. The results of the multivariate regression analysis showed that Age, BMI, SBP, U.GLU, PRO, TG, and HDL.C were independent factors influencing the development of diabetes in the prediabetic population, with Age, BMI, SBP, U.GLU, PRO, and TG being independent risk factors, and HDL.C being an independent protective factor, as shown in [Table 3](#).

### 3.5 Construction and evaluation of a predictive model for the development of pre-diabetes in normal population

According to the regression analysis results in [Table 2](#), a full-variable prediction model (Full.model) was constructed for the independent risk factors that contribute to the development of prediabetes in the normal population, including Age, BMI, SBP, U.GLU, PRO, TP, GLB, ALT, and HDL.C. Based on the regression analysis results in [Tables 2, 3](#), a simplified prediction model

**TABLE 2** Univariate and multivariate regression analysis for the normal and pre-diabetic groups.

Variable name	Univariate regression analysis			Multivariate regression analysis		
	OR	95%CI	P	OR	95%CI	P
Age	1.09	1.08-1.10	<0.001***	1.10	1.09-1.11	<0.001***
Gender(Male)	1.40	1.24-1.58	<0.001***	0.85	0.69-1.04	0.1131
BMI	1.22	1.2-1.25	<0.001***	1.22	1.19-1.25	<0.001***
SBP	1.03	1.03-1.04	<0.001***	1.01	1.00-1.01	<0.001***
U.GLU(Positive)	6.59	3.05-14.24	<0.001***	4.23	1.82-9.83	<0.001***
PRO(Positive)	2.56	1.71-3.84	<0.001***	2.14	1.31-3.50	0.0023**
TP	1.02	1.00-1.03	0.03*	1.12	1.08-1.16	<0.001***
ALB	1.00	0.98-1.03	0.73			
GLB	1.02	1.00-1.04	0.02*	0.90	0.86-0.93	<0.001***
T.BIL	0.98	0.97-0.99	<0.001***	1.25	0.27-5.93	0.7749
DB	0.82	0.77-0.88	<0.001***	0.65	0.13-3.18	0.5954
IB	0.99	0.97-1.00	0.02*	0.79	0.17-3.71	0.7646
ALT	1.01	1.01-1.02	<0.001***	1.02	1.01-1.02	<0.001***
AST	1.02	1.01-1.02	<0.001***	0.99	0.98-1.00	0.0873
BUN	1.20	1.14-1.26	<0.001***	0.98	0.92-1.04	0.5493
SCr	1.00	1.00-1.01	0.07			
UA	1.00	1.00-1.00	<0.001***	1.00	1.00-1.00	0.1226
TC	1.24	1.17-1.31	<0.001***	1.42	0.95-2.12	0.0859
TG	1.51	1.41-1.62	<0.001***	0.99	0.91-1.09	0.9029
HDL.C	0.28	0.23-0.34	<0.001***	0.21	0.13-0.32	<0.001***
LDL.C	1.42	1.32-1.54	<0.001***	0.87	0.54-1.39	0.5508

Statistically significant ORs and 95% CIs are shown in italics; \* $P < 0.01$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ .

TABLE 3 Univariate and multivariate regression analysis for the pre-diabetes group and the diabetes group.

Variable name	Univariate regression analysis			Multivariate regression analysis		
	OR	95%CI	P	OR	95%CI	P
Age	<i>1.03</i>	<i>1.02-1.04</i>	<0.001***	<i>1.03</i>	<i>1.02-1.04</i>	<0.001***
Gender(Male)	<i>1.41</i>	<i>1.17-1.71</i>	<0.001***	1.05	0.83-1.33	0.6705
BMI	<i>1.06</i>	<i>1.03-1.08</i>	<0.001***	<i>1.07</i>	<i>1.04-1.1</i>	<0.001***
SBP	<i>1.02</i>	<i>1.01-1.02</i>	<0.001***	<i>1.01</i>	<i>1.00-1.01</i>	0.0013**
U.GLU(Positive)	<i>10.33</i>	<i>7.75-13.77</i>	<0.001***	<i>9.32</i>	<i>6.88-12.64</i>	<0.001***
PRO(Positive)	<i>3.21</i>	<i>2.39-4.31</i>	<0.001***	<i>2.30</i>	<i>1.65-3.22</i>	<0.001***
TP	0.99	0.97-1.02	0.57			
ALB	0.99	0.95-1.03	0.53			
GLB	1.00	0.97-1.02	0.84			
T.BIL	1.01	1.00-1.03	0.08			
DB	<i>1.12</i>	<i>1.03-1.21</i>	0.01*	1.01	0.92-1.12	0.7827
IB	1.01	1.00-1.03	0.15			
ALT	<i>1.00</i>	<i>1.00-1.00</i>	0.04*	1.00	1.00-1.01	0.1646
AST	1.00	1.00-1.01	0.08			
BUN	<i>1.15</i>	<i>1.08-1.23</i>	<0.001***	1.02	0.95-1.1	0.5739
SCr	1.00	0.99-1.00	0.14			
UA	1.00	1.00-1.00	0.19			
TC	<i>0.90</i>	<i>0.83-0.97</i>	0.01*	0.86	0.58-1.26	0.4305
TG	<i>1.12</i>	<i>1.07-1.18</i>	<0.001***	<i>1.09</i>	<i>1.02-1.16</i>	0.0154*
HDL.C	<i>0.32</i>	<i>0.23-0.45</i>	<0.001***	<i>0.43</i>	<i>0.26-0.73</i>	0.0016**
LDL.C	<i>0.89</i>	<i>0.80-0.99</i>	0.03*	1.34	0.85-2.11	0.2141

Statistically significant ORs and 95% CIs are shown in italics; \*P <0.01, \*\*P <0.01, \*\*\*P <0.001.

(Simplified.model) was constructed using six variables, including Age, BMI, SBP, U.GLU, PRO, and HDL.C. The ROC curve and AUC was used to evaluate the discrimination of the predictive model, with an AUC range of 0-1, where 1 indicates complete consistency and 0.5 indicates poor consistency. The evaluation results of the models show that Full.model has a training set AUC of 0.81 (Figure 2A), an accuracy of 0.78, precision of 0.80, and recall of 0.89; and a training set AUC of 0.82 (Figure 2A), an accuracy of 0.79, precision of 0.80, and recall of 0.91. Simplified.model has a training set AUC of 0.80 (Figure 2B), an accuracy of 0.77, precision of 0.79, and recall of 0.89; and a training set AUC of 0.81 (Figure 2B), an accuracy of 0.77, precision of 0.78, and recall of 0.91.

### 3.6 Construction and evaluation of a predictive model for the progression of pre-diabetes to diabetes

Based on the regression analysis results in Table 3, a full-variable prediction model (Full.model) was constructed for independent factors predicting the development of prediabetes to

diabetes in prediabetic individuals, which included Age, BMI, SBP, U.GLU, PRO, TG, and HDL.C. A simplified prediction model (Simplified.model) was also constructed using Age, BMI, SBP, U.GLU, PRO, and HDL.C as independent factors. The model evaluation results showed that Full.model had moderate discrimination for identifying high-risk individuals for developing diabetes in the prediabetic population. In the training set, the AUC of Full.model was 0.73 (Figure 3A), with an accuracy of 0.86, precision of 0.59, and recall of 0.20, and the AUC was 0.71 (Figure 3A), with an accuracy of 0.86, precision of 0.71, and recall of 0.22. Simplified.model also had moderate discrimination, with an AUC of 0.73 (Figure 3B), accuracy of 0.86, precision of 0.59, and recall of 0.20 in the training set, and an AUC of 0.70 (Figure 3B), accuracy of 0.86, precision of 0.71, and recall of 0.21.

## 4 Discussion

Diabetes has become a global public health problem, with the incidence and mortality rates increasing year by year. Prediabetes is a precursor of diabetes. Detecting prediabetes and taking effective interventions can prevent the further development of diabetes,

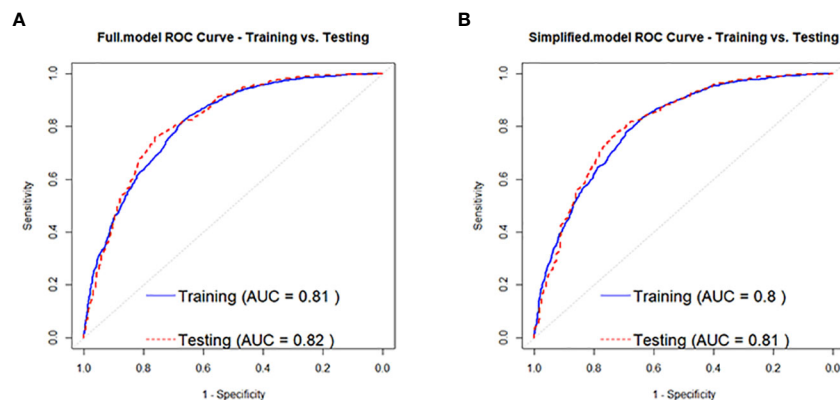


FIGURE 2

ROC curves for the normal population developing to pre-diabetes Full.model (A) and Simplified.model (B).

thereby reducing the incidence and mortality rates of diabetes. In this study, we investigated the risk factors for progression from normal individuals to prediabetes and from prediabetes to diabetes, and analyzed the independent factors using multivariable logistic regression. We found that Age, BMI, SBP, U.GLU, PRO, TP, GLB, ALT, and HDL.C were independent risk factors for progression from normal individuals to prediabetes, while Age, BMI, SBP, U.GLU, PRO, TG, and HDL.C were independent risk factors for progression from prediabetes to diabetes. Among them, Age, BMI, SBP, U.GLU, and PRO were common independent risk factors for both progressions, while HDL.C was a common independent protective factor. We constructed full variable models (Full.model) and simplified models (Simplified.model) for both progressions using the above factors. The evaluation of the models indicated moderate discriminative ability and could assist in the clinical identification of individuals at high risk of developing prediabetes and diabetes.

In previous studies, many researchers have investigated the related risk factors for prediabetes and diabetes, such as Age, BMI, SBP, U.GLU, PRO, etc. Among them, Age and BMI are considered the two strongest risk factors for prediabetes (25, 26). According to estimates from the National Health and Nutrition Examination

Survey (NHANES) in the United States, the overall prevalence of diabetes was 5.0% in adults under 45 years old, 17.5% in adults aged 45-64, and 33.0% in adults aged 65 and older (13). This shows that the risk of diabetes increases significantly with age. In the NHANES study, more than 80% of self-reported prediabetic patients had a BMI  $\geq 25.0$ , indicating that the prevalence of prediabetes is much higher in obese populations (27). A study conducted in China during a median follow-up period of 4.5 years among non-diabetic hypertensive individuals found that compared with individuals with an SBP in the range of 120-130 mmHg, those with an SBP in the range of 130-140 mmHg had a 24% increased risk of developing diabetes and a 29% reduced rate of fasting blood glucose recovery (28). Currently, it is believed that the biological mechanism between blood pressure control and the development of diabetes may be due to hypertension leading to endothelial dysfunction, which limits insulin delivery to metabolically active insulin-sensitive muscle tissue, and optimal blood pressure control can improve endothelial function and enhance microvascular perfusion, thus leading to a reduced risk of diabetes (28, 29). This study found that U.GLU positive and PRO positive in urine tests were independent risk factors for prediabetes and diabetes. These two indicators reflect the damage to the renal function of the subjects, and even

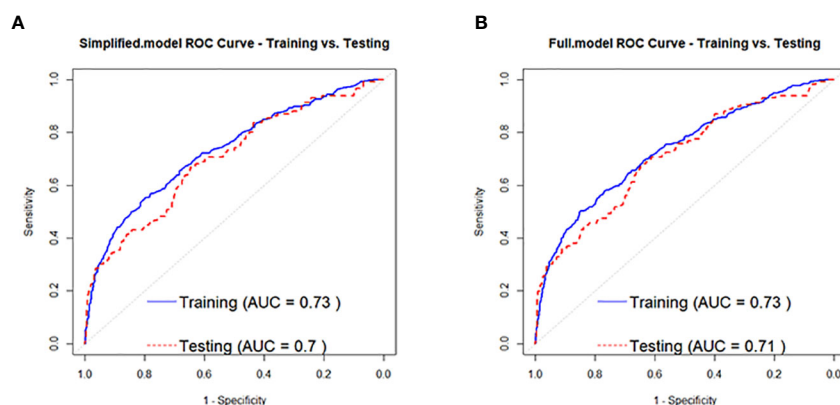


FIGURE 3

ROC curves for people with pre-diabetes developing to diabetes Full.model (A) and Simplified.model (B).

if the patient's blood glucose level returns to normal, the renal function damage caused by diabetes will continue to develop (30, 31). Studies have shown that the sensitivity of prediabetes and diabetes screening through U.GLU testing is 83.5%, and the combined use of U.GLU and FPG can significantly improve the effectiveness of diabetes screening, indicating a high correlation between U.GLU positivity and the development of diabetes (32). In any eGFR category of the general population, the incidence of diabetes and metabolic syndrome increases with increasing levels of urine protein (PRO) (33). Furthermore, studies have pointed out that observing changes in the urinary albumin-to-creatinine ratio (UACR) can predict changes in clinical outcomes and mortality risks for type 2 diabetes patients (34). Therefore, abnormal urine test results not only serve as efficient indicators for prediabetes and diabetes screening, but also serve as important indicators reflecting the level of renal function damage in diabetic patients.

However, our findings are not entirely consistent with previous studies, as we identified new indicators, including TP and ALT, as independent risk factors for the development of prediabetes in the general population. Current research has identified multiple proteins in serum that are related to the occurrence and development of prediabetes, including C-reactive protein (35), lipopolysaccharide-binding protein (36), among others. These results suggest that proteins, as the main carriers of life activities, play a complex role in the occurrence and development of prediabetes. Our study found that ALT is an independent risk factor for the development of prediabetes in the general population. Previous studies have shown that elevated ALT is associated with type 2 diabetes, indicating that ALT may be involved in insulin resistance and the development of diabetes (37, 38). Additionally, some studies have found a negative correlation between early AST/ALT levels in pregnant women and the risk of gestational diabetes mellitus (GDM), suggesting that these levels can serve as predictive factors for GDM (39). Although these findings do not directly support our results, they provide a new perspective for better understanding the occurrence and development of prediabetes.

In addition, we found that TG and LDL-C are independent risk factors for the development of diabetes in individuals with prediabetes. Our results are consistent with previous studies which have shown a positive correlation between TG and LDL-C levels and the progression of diabetes, and the predictive value of these markers for prediabetes and diabetes (40–42). Studies have also found that the TG/HDL ratio is positively correlated with  $\beta$ -cell dysfunction, prediabetes, and diabetes, and is an important risk assessment factor for cardiovascular disease in diabetic patients (43, 44). Campos Muniz C proposed the concept of the triglyceride glucose (TyG) index and found that it is a good predictor of DM2 (45). These findings strongly support our results and indicate that triglycerides are an important risk factor for diabetes. Abnormal blood lipids are also recognized as controllable risk factors in patients with type 2 diabetes, and their management is an important part of preventing cardiovascular disease. Studies have shown that statin therapy can significantly reduce cardiovascular events (46). Our results provide a more comprehensive and in-depth understanding of the mechanisms underlying the

development of diabetes from prediabetes and provide some guidance for the diagnosis and treatment of diabetes.

This study also suggests that GLB and HDL.C are independent protective factors for the development of prediabetes in normal population, and HDL.C is an independent protective factor for the progression from prediabetes to diabetes. However, no significant evidence has been found in previous relevant studies to support the protective effect of elevated levels of GLB on the occurrence of prediabetes. A study on elderly prediabetic and elderly male populations in China found that lower levels of sex hormone-binding globulin (SHBG) were independently associated with metabolic syndrome (47). However, other studies have found that levels of alpha-fetoprotein (AFP) in gestational diabetes mellitus patients were significantly higher than in normal pregnant women, suggesting that AFP may play a role in insulin resistance and metabolic changes in gestational diabetes mellitus (48). Therefore, further research is needed on the specific association between globulin and prediabetes. HDL.C, as an independent protective factor for prediabetes and diabetes, is consistent with previous research findings (40). Studies have shown that HDL.C can not only play an anti-atherosclerotic role against endothelial cells and foam cells, but also have an anti-diabetic effect on the  $\beta$ -cells of the endocrine pancreas, especially by effectively inhibiting stress-induced cell death and enhancing insulin secretion stimulated by glucose (49). The increase in HDL-C levels is not only related to the reduction of cardiovascular disease risk, but also a potential strategy for preventing the occurrence and development of diabetes in the future (50). Therefore, increasing the levels of these protective factors may help prevent the occurrence of prediabetes and diabetes.

This study further constructed full variable prediction models (Full.model) and simplified prediction models (Simplified.model) for the development from normal to prediabetes and from prediabetes to diabetes, respectively, based on the independent influencing factors identified above. The model evaluations showed moderate discrimination, with the AUC of the prediabetes prediction model reaching above 0.8 and the diabetes prediction model reaching 0.7. Compared with the risk prediction models constructed by Wu et al. and Yokota et al., the discrimination of the models constructed in this study is similar, but the former two studies used blood glucose and HbA1c levels as predictors (23, 24). Blood glucose and HbA1c levels are essential for diagnosing prediabetes and diabetes and are highly correlated with the risk of developing the diseases, so the rationality of using these two indicators as predictors for constructing prediction models is questionable. The Simplified.model constructed in this study by simplifying the variables has moderate discrimination for identifying prediabetes and diabetes, and the included indicators are commonly used in clinical physical examinations. Therefore, it can be considered that the model has certain applicability and is expected to contribute to the prevention and management of prediabetes and diabetes in the future.

In previous studies, several diabetes prediction models have been developed using statistical models such as logistic regression, Cox proportional hazards model, or Weibull distribution analysis.



The predictive accuracy of these traditional statistical methods, as measured by the C-index, ranged from approximately 0.74 to 0.94. In recent years, the development of artificial intelligence (AI) technology has presented new opportunities and challenges for diabetes prevention, diagnosis, and treatment. AI's main applications in diabetes include automated retinal screening, clinical diagnostic support, patient self-management tools, and risk stratification (51, 52). Currently, the aggregated AUC (Area Under the Curve) of artificial intelligence in diabetes prediction and risk stratification is approximately 0.86–0.88, indicating a high level of discrimination ability (53, 54). However, it is premature to conclude that machine learning surpasses traditional statistical analysis in predicting incident diabetes in specific populations. Furthermore, AI-generated models may suffer from overfitting, leading to highly accurate predictions for the training population but significantly reduced accuracy when applied to the validation population. Although there are still challenges in using machine learning models to predict incident diabetes in clinical practice, we firmly believe that in the future, more efficient machine learning models and the availability of larger omics databases will undoubtedly contribute to further improving prediction accuracy.

This study used logistic regression analysis to identify the independent risk factors for prediabetes and diabetes, most of which were independent risk factors. The results of this study help to strengthen people's awareness and control of the risk factors for diabetes, and take proactive measures to control the relevant factors for prediabetes and diabetes, which is of great significance for the prevention and control of diabetes. Furthermore, the predictive models constructed based on the results of this study can help identify high-risk individuals for prediabetes and diabetes.

This study still has several limitations. Firstly, all samples were obtained from a single hospital, and the sample size was relatively small. Additionally, the study subjects were all from the same region, which may lead to biases and lack of representativeness, making it challenging to generalize the research findings. Secondly, due to inadequate data collection, some crucial variables (such as medications taken by participants, ethnicity, lifestyle patterns, sleep habits, and smoking status) were not investigated. Including these variables could significantly enhance the impact of the research results. Furthermore, this study was cross-sectional and employed logistic regression analysis, which can identify independent influencing factors but cannot establish causal relationships. Lastly, the predictive model constructed in this study was only internally validated and lacked external validation and real-world application research.

Hence, future research could consider enlarging the sample size and adopting a multi-center research design to improve the reliability of the conclusions. Additionally, more scientifically rigorous study designs, such as cohort studies and randomized controlled trials, could be employed to further assess causality and explore comprehensive diabetes risk factors. Ultimately, more advanced algorithms should be considered to develop predictive models with better discrimination and evaluation, which can be applied to the identification of high-risk populations in the real world.

## 5 Conclusion

This study found that Age, BMI, SBP, U.GLU, PRO, TP, and ALT were independent risk factors, while GLB and HDL.C were independent protective factors for developing prediabetes from the normal population. For those who progressed from prediabetes to diabetes, Age, BMI, SBP, U.GLU, PRO, and TG were independent risk factors, while HDL.C was an independent protective factor. By including these factors as predictors, a prediction model was developed that had moderate discriminative ability for identifying individuals at high risk for prediabetes and diabetes. In this study, blood glucose and HbA1c were used as the diagnostic gold standards, while other risk indicators were considered as predictive factors. This approach allows for the identification and early warning of individuals at risk of prediabetes and diabetes among the population undergoing health examinations.

## Data availability statement

The original contributions presented in the study are included in the article/[Supplementary Material](#). Further inquiries can be directed to the corresponding authors.

## Ethics statement

The studies involving humans were approved by Peking University Shenzhen Hospital Ethical Review Committee. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

## Author contributions

DG, XC and LY acquired, analyzed, discussed the data and drafted the manuscript. DG, YZ, JL and CY analyzed and discussed the data. YC, WY and JW designed the research and revised the manuscript. All authors contributed to the article and approved the submitted version.

## Funding

This work was funded by Shenzhen Fund for Guangdong Provincial High-level Clinical Key Specialties (SZGSP014), funded by National Nature Science Foundation of China (82070961), funded by Shenzhen Key Medical Discipline Construction Fund (No.SZ XK037), funded by Shenzhen Science and Technology Program (No.JCYJ20220818103207015), funded by SanMing Project of Medicine in Shenzhen (SZSM201812091).

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fendo.2023.1225696/full#supplementary-material>

## References

- Tabak AG, Herder C, Rathmann W, Brunner EJ, Kivimaki M. Prediabetes: a high-risk state for diabetes development. *Lancet* (2012) 379(9833):2279–90. doi: 10.1016/S0140-6736(12)60283-9
- Saeedi P, Petersohn I, Salpea P, Malanda B, Karuranga S, Unwin N, et al. Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9(th) edition. *Diabetes Res Clin Pract* (2019) 157:107843. doi: 10.1016/j.diabres.2019.107843
- Wang L, Gao P, Zhang M, Huang Z, Zhang D, Deng Q, et al. Prevalence and ethnic pattern of diabetes and prediabetes in China in 2013. *JAMA* (2017) 317(24):2515–23. doi: 10.1001/jama.2017.7596
- Bigna JJ, Nansseu JR, Katte JC, Noubiap JJ. Prevalence of prediabetes and diabetes mellitus among adults residing in Cameroon: A systematic review and meta-analysis. *Diabetes Res Clin Pract* (2018) 137:109–18. doi: 10.1016/j.diabres.2017.12.005
- Anjana RM, Deepa M, Pradeepa R, Mahanta J, Narain K, Das HK, et al. Prevalence of diabetes and prediabetes in 15 states of India: results from the ICMR-INDIAB population-based cross-sectional study. *Lancet Diabetes Endocrinol* (2017) 5(8):585–96. doi: 10.1016/S2213-8587(17)30174-2
- Hashemi SJ, Karandish M, Cheraghian B, Azhdari M. Prevalence of prediabetes and associated factors in southwest Iran: results from Hoveyze cohort study. *BMC Endocr Disord* (2022) 22(1):72. doi: 10.1186/s12902-022-00990-z
- Diabetes Prevention Program Research G, Knowler WC, Fowler SE, Hamman RF, Christophi CA, Hoffman HJ, et al. 10-year follow-up of diabetes incidence and weight loss in the Diabetes Prevention Program Outcomes Study. *Lancet* (2009) 374(9702):1677–86. doi: 10.1016/S0140-6736(09)61457-4
- Apolzan JW, Venditti EM, Edelstein SL, Knowler WC, Dabelea D, Boyko EJ, et al. Long-term weight loss with metformin or lifestyle intervention in the diabetes prevention program outcomes study. *Ann Intern Med* (2019) 170(10):682–90. doi: 10.7326/M18-1605
- Richter B, Hemmingsen B, Metzendorf MI, Takwoingi Y. Development of type 2 diabetes mellitus in people with intermediate hyperglycaemia. *Cochrane Database Syst Rev* (2018) 10(10):CD012661. doi: 10.1002/14651858.CD012661.pub2
- American Diabetes A. (2) Classification and diagnosis of diabetes. *Diabetes Care* (2015) 38(Suppl):S8–S16. doi: 10.2337/dc15-S005
- Zimmet PZ, Magliano DJ, Herman WH, Shaw JE. Diabetes: a 21st century challenge. *Lancet Diabetes Endocrinol* (2014) 2(1):56–64. doi: 10.1016/S2213-8587(13)70112-8
- Zheng Y, Ley SH, Hu FB. Global aetiology and epidemiology of type 2 diabetes mellitus and its complications. *Nat Rev Endocrinol* (2018) 14(2):88–98. doi: 10.1038/nrendo.2017.151
- Menke A, Casagrande S, Geiss L, Cowie CC. Prevalence of and trends in diabetes among adults in the United States, 1988–2012. *JAMA* (2015) 314(10):1021–9. doi: 10.1001/jama.2015.10029
- Huang Y, Cai X, Mai W, Li M, Hu Y. Association between prediabetes and risk of cardiovascular disease and all cause mortality: systematic review and meta-analysis. *BMJ* (2016) 355:i5953. doi: 10.1136/bmj.i5953
- Paschou SA, Bletsas E, Saltiki K, Kazakou P, Kantreva K, Katsaounou P, et al. Sleep apnea and cardiovascular risk in patients with prediabetes and type 2 diabetes. *Nutrients* (2022) 14(23). doi: 10.3390/nu14234989
- Strain WD, Paldanius PM. Diabetes, cardiovascular disease and the microcirculation. *Cardiovasc Diabetol* (2018) 17(1):57. doi: 10.1186/s12933-018-0703-2
- Forbes JM, Cooper ME. Mechanisms of diabetic complications. *Physiol Rev* (2013) 93(1):137–88. doi: 10.1152/physrev.00045.2011
- Dal Canto E, Ceriello A, Ryden L, Ferrini M, Hansen TB, Schnell O, et al. Diabetes as a cardiovascular risk factor: An overview of global trends of macro and micro vascular complications. *Eur J Prev Cardiol* (2019) 26(2\_suppl):25–32. doi: 10.1177/2047487319878371
- Saleh M, Kim JY, March C, Gebara N, Arslanian S. Youth prediabetes and type 2 diabetes: Risk factors and prevalence of dysglycaemia. *Pediatr Obes* (2022) 17(1):e12841. doi: 10.1111/ijpo.12841
- Rodriguez-Segade S, Rodriguez J, Camina F, Sanmartin-Portas L, Gerpe-Jamardo J, Pazos-Couselo M, et al. Prediabetes defined by HbA(1c) and by fasting glucose: differences in risk factors and prevalence. *Acta Diabetol* (2019) 56(9):1023–30. doi: 10.1007/s00592-019-01342-5
- Ceriello A, Prattichizzo F. Variability of risk factors and diabetes complications. *Cardiovasc Diabetol* (2021) 20(1):101. doi: 10.1186/s12933-021-01289-4
- Lamina C, Ward NC. Lipoprotein (a) and diabetes mellitus. *Atherosclerosis* (2022) 349:63–71. doi: 10.1016/j.atherosclerosis.2022.04.016
- Wu J, Zhou J, Yin X, Chen Y, Lin X, Xu Z, et al. A prediction model for prediabetes risk in middle-aged and elderly populations: A prospective cohort study in China. *Int J Endocrinol* (2021) 2021:2520806. doi: 10.1155/2021/2520806
- Yokota N, Miyakoshi T, Sato Y, Nakasone Y, Yamashita K, Imai T, et al. Predictive models for conversion of prediabetes to diabetes. *J Diabetes Complications* (2017) 31(8):1266–71. doi: 10.1016/j.jdiacomp.2017.01.005
- Echouffo-Tcheugui JB, Selvin E. Prediabetes and what it means: the epidemiological evidence. *Annu Rev Public Health* (2021) 42:59–77. doi: 10.1146/annurev-publhealth-090419-102644
- Zuo H, Shi Z, Hussain A. Prevalence, trends and risk factors for the diabetes epidemic in China: a systematic review and meta-analysis. *Diabetes Res Clin Pract* (2014) 104(1):63–72. doi: 10.1016/j.diabres.2014.01.002
- Liu C, Foti K, Grams ME, Shin JI, Selvin E. Trends in self-reported prediabetes and metformin use in the USA: NHANES 2005–2014. *J Gen Intern Med* (2020) 35(1):95–101. doi: 10.1007/s11606-019-05398-5
- Zhang Y, Nie J, Zhang Y, Li J, Liang M, Wang G, et al. Degree of blood pressure control and incident diabetes mellitus in chinese adults with hypertension. *J Am Heart Assoc* (2020) 9(16):e017015. doi: 10.1161/JAHA.120.017015
- Shahin Y, Khan JA, Samuel N, Chetter I. Angiotensin converting enzyme inhibitors effect on endothelial dysfunction: a meta-analysis of randomised controlled trials. *Atherosclerosis* (2011) 216(1):7–16. doi: 10.1016/j.atherosclerosis.2011.02.044
- Yamazaki T, Mimura I, Tanaka T, Nangaku M. Treatment of diabetic kidney disease: current and future. *Diabetes Metab J* (2021) 45(1):11–26. doi: 10.4093/dmj.2020.0217
- Dou L, Jourde-Chiche N. Endothelial Toxicity of High Glucose and its by-Products in Diabetic Kidney Disease. *Toxins (Basel)* (2019) 11(10). doi: 10.3390/toxins11100578
- Chen J, Guo HJ, Qiu SH, Li W, Wang XH, Cai M, et al. Identification of newly diagnosed diabetes and prediabetes using fasting plasma glucose and urinary glucose in a chinese population: A multicenter cross-sectional study. *Chin Med J (Engl)* (2018) 131(14):1652–7. doi: 10.4103/0366-6999.235884
- Okada R, Yasuda Y, Tsushita K, Wakai K, Hamajima N, Matsuo S. Trace proteinuria by dipstick screening is associated with metabolic syndrome, hypertension, and diabetes. *Clin Exp Nephrol* (2018) 22(6):1387–94. doi: 10.1007/s10157-018-1601-3
- Jun M, Ohkuma T, Zoungas S, Colagiuri S, Mancia G, Marre M, et al. Changes in albuminuria and the risk of major clinical outcomes in diabetes: results from ADVANCE-ON. *Diabetes Care* (2018) 41(1):163–70. doi: 10.2337/dc17-1467
- Sabanayagam C, Shankar A, Lim SC, Lee J, Tai ES, Wong TY. Serum C-reactive protein level and prediabetes in two Asian populations. *Diabetologia* (2011) 54(4):767–75. doi: 10.1007/s00125-011-2052-5

36. Tilves CM, Zmuda JM, Kuipers AL, Nestlerode CS, Evans RW, Bunker CH, et al. Association of lipopolysaccharide-binding protein with aging-related adiposity change and prediabetes among african ancestry men. *Diabetes Care* (2016) 39(3):385–91. doi: 10.2337/dc15-1777
37. Abro MUR, Butt A, Baqa K, Waris N, Khalid M, Fawwad A. Association of serum liver enzyme Alanine Aminotransferase (ALT) in patients with type 2 diabetes. *Pak J Med Sci* (2018) 34(4):839–43. doi: 10.12669/pjms.344.15206
38. Qian K, Zhong S, Xie K, Yu D, Yang R, Gong DW. Hepatic ALT isoenzymes are elevated in gluconeogenic conditions including diabetes and suppressed by insulin at the protein level. *Diabetes Metab Res Rev* (2015) 31(6):562–71. doi: 10.1002/dmrr.2655
39. An R, Ma S, Zhang N, Lin H, Xiang T, Chen M, et al. AST-to-ALT ratio in the first trimester and the risk of gestational diabetes mellitus. *Front Endocrinol (Lausanne)* (2022) 13:1017448. doi: 10.3389/fendo.2022.1017448
40. Zhou Y, Yang G, Qu C, Chen J, Qian Y, Yuan L, et al. Predictive performance of lipid parameters in identifying undiagnosed diabetes and prediabetes: a cross-sectional study in eastern China. *BMC Endocr Disord* (2022) 22(1):76. doi: 10.1186/s12902-022-00984-x
41. Gao YX, Man Q, Jia S, Li Y, Li L, Zhang J. The fasting serum triglyceride levels of elderly population with different progression stages of diabetes mellitus in China. *J Diabetes Complications* (2017) 31(12):1641–7. doi: 10.1016/j.jdiacomp.2017.08.011
42. Janghorbani M, Soltanian N, Amini M, Aminorroaya A. Low-density lipoprotein cholesterol and risk of type 2 diabetes: The Isfahan diabetes prevention study. *Diabetes Metab Syndr* (2018) 12(5):715–9. doi: 10.1016/j.dsx.2018.04.019
43. Hermans MP, Ahn SA, Rousseau MF. log(TG)/HDL-C is related to both residual cardiometabolic risk and beta-cell function loss in type 2 diabetes males. *Cardiovasc Diabetol* (2010) 9:88. doi: 10.1186/1475-2840-9-88
44. Gong R, Liu Y, Luo G, Liu W, Jin Z, Xu Z, et al. Associations of TG/HDL ratio with the risk of prediabetes and diabetes in chinese adults: A chinese population cohort study based on open data. *Int J Endocrinol* (2021) 2021:9949579. doi: 10.1155/2021/9949579
45. Campos Muniz C, Leon-Garcia PE, Serrato Diaz A, Hernandez-Perez E. Diabetes mellitus prediction based on the triglyceride and glucose index. *Med Clin (Barc)* (2023) 160(6):231–6. doi: 10.1016/j.medcli.2022.07.003
46. Betteridge J. Benefits of lipid-lowering therapy in patients with type 2 diabetes mellitus. *Am J Med* (2005) 118(Suppl 12A):10–5. doi: 10.1016/j.amjmed.2005.09.013
47. Pang XN, Hu Y, Yuan Y, Shen JP, Zha XY, Sun X. Lower levels sex hormone-binding globulin independently associated with metabolic syndrome in pre-elderly and elderly men in China. *J Geriatr Cardiol* (2013) 10(1):28–33. doi: 10.3969/j.issn.1671-5411.2013.01.006
48. Iyidir OT, Degertekin CK, Yilmaz BA, Altinova AE, Toruner FB, Bozkurt N, et al. Serum levels of fetuin A are increased in women with gestational diabetes mellitus. *Arch Gynecol Obstet* (2015) 291(4):933–7. doi: 10.1007/s00404-014-3490-3
49. von Eckardstein A, Widmann C. High-density lipoprotein, beta cells, and diabetes. *Cardiovasc Res* (2014) 103(3):384–94. doi: 10.1093/cvr/cvu143
50. Cochran BJ, Ong KL, Manandhar B, Rye KA. High density lipoproteins and diabetes. *Cells* (2021) 10(4). doi: 10.3390/cells10040850
51. Nomura A, Noguchi M, Kometani M, Furukawa K, Yoneda T. Artificial intelligence in current diabetes management and prediction. *Curr Diabetes Rep* (2021) 21(12):61. doi: 10.1007/s11892-021-01423-2
52. Ellahham S. Artificial intelligence: the future for diabetes care. *Am J Med* (2020) 133(8):895–900. doi: 10.1016/j.amjmed.2020.03.033
53. Olusanya MO, Ogunsakin RE, Ghai M, Adeleke MA. Accuracy of machine learning classification models for the prediction of type 2 diabetes mellitus: A systematic survey and meta-analysis approach. *Int J Environ Res Public Health* (2022) 19(21). doi: 10.3390/ijerph192114280
54. Kodama S, Fujihara K, Horikawa C, Kitazawa M, Iwanaga M, Kato K, et al. Predictive ability of current machine learning algorithms for type 2 diabetes mellitus: A meta-analysis. *J Diabetes Investig* (2022) 13(5):900–8. doi: 10.1111/jdi.13736