Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Association and Comparative Predictive Performance of TyG-BMI for Incident Cardiovascular Disease in Adults With Hypertension: A Prospective CHARLS Cohort Study

Yinglun Wei
Shenzhen Hospital (Futian) of Guangzhou University of Chinese Medicine, Guangzhou University of Chinese Medicine

Xiaoer Liu
Acupuncture Anesthesia Clinical Research Institute, Yueyang Hospital of Integrated Traditional Chinese and Western Medicine, Shanghai University of Traditional Chinese Medicine

Zhihao Wang
Shanxi Jiaoke New Materials Science and Technology Co., Ltd Shanxi Transportation Research Institute Group Co., Ltd

Yunqing Luo
Shenzhen Hospital (Futian) of Guangzhou University of Chinese Medicine, Guangzhou University of Chinese Medicine

Can Yun
Heilongjiang University of Chinese Medicine

Kecheng Liu
Guangdong Provincial Hospital of Chinese Medicine, Zhuhai Hospital

Xiaoyan Zhai
School of Basic Medicine, Shanxi University of Chinese Medicine, / Key Research Laboratory for Treating Multiple Sclerosis by Strengthening Qi and Promoting Blood Circulation

Xinzheng Tang
yayatom@163.com

Shenzhen Hospital (Futian) of Guangzhou University of Chinese Medicine, Guangzhou University of Chinese Medicine

---

Article

Keywords: TyG-BMI, Hypertension, Cardiovascular disease, Insulin resistance, Machine learning, CHARLS

Posted Date: November 23rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11455132/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Association and Comparative Predictive Performance of TyG-BMI for Incident Cardiovascular Disease in Adults With Hypertension: A Prospective CHARLS Cohort Study

Yinglun Wei[1†], Xiaoer Liu[2†], Zhihao Wang[3†], Yunqing Luo[1], Can Yun[4], Kecheng Liu[5], Xiaoyan Zhai[6*] and Xinzheng Tang[1*]

[1]Shenzhen Hospital (Futian) of Guangzhou University of Chinese Medicine, Guangzhou University of Chinese Medicine,Shenzhen 518034, China

[2]Acupuncture Anesthesia Clinical Research Institute, Yueyang Hospital of Integrated Traditional Chinese and Western Medicine, Shanghai University of Traditional Chinese Medicine, Shanghai, 200437, China

[3]Shanxi Jiaoke New Materials Science and Technology Co., Ltd，Shanxi Transportation Research Institute Group Co., Ltd, Taiyuan,030032, China

[4]Heilongjiang University of Chinese Medicine, Harbin 150040, China

[5]Guangdong Provincial Hospital of Chinese Medicine, Zhuhai Hospital, Zhuhai, 519000, China

[6]School of Basic Medicine, Shanxi University of Chinese Medicine, / Key Research Laboratory for Treating Multiple Sclerosis by Strengthening Qi and Promoting Blood Circulation, National Administration of Traditional Chinese Medicine, Jinzhong 030619, China

[†]Yinglun Wei, Xiaoer Liu, and Zhihao Wang contributed equally to this work

*Correspondence:

Xinzheng Tang

yayatom@163.com

Xiaoyan Zhai

zxy2005019@126.com

Full list of author information is available at the end of the article.

Abstract

The triglyceride-glucose body mass index (TyG-BMI) integrates insulin

resistance and obesity and may provide additional information for cardiovascular risk assessment. However, its association with cardiovascular disease (CVD) and its predictive value relative to other TyG-related indices remain uncertain in adults with hypertension. We included 3,092 participants aged ⩾45 years with hypertension and without CVD at baseline from the China Health and Retirement Longitudinal Study and followed them through 2018. During a median follow-up of 8 years, 886 participants developed CVD, including 659 cases of heart disease and 329 cases of stroke. Higher TyG-BMI was associated with greater cardiovascular risk, particularly for heart disease, while associations with stroke were weaker. Dose-response analyses suggested largely linear relationships with CVD and heart disease. Using five machine learning algorithms, we further assessed the incremental predictive value of TyG-BMI and other TyG-related indices. Adding TyG-BMI to the basic clinical model improved the area under the curve from 0.593 to 0.636 and yielded greater improvement in risk classification than the other indices. TyG-BMI may therefore provide complementary information for cardiovascular risk assessment in middle-aged and older adults with hypertension.

Keywords：TyG-BMI; Hypertension; Cardiovascular disease; Insulin resistance; Machine learning；CHARLS

1．Introduction

With the acceleration of global population aging and the rising prevalence of obesity, diabetes, and other metabolic abnormalities, cardiovascular disease (CVD) has become one of the most important public health challenges worldwide [1]. The Global Burden of Disease (GBD) study indicates that CVD remains the leading cause of global mortality and disability-adjusted life years (DALYs), with its disease burden being particularly pronounced among middle-aged and older populations. Although

considerable progress has been made in the prevention and treatment of CVD in recent years, the global prevalence and mortality of CVD continue to increase due to demographic shifts and the persistent accumulation of cardiovascular risk factors [2]. As one of the countries with the heaviest cardiovascular disease burden worldwide, China currently has more than 330 million people affected by various types of CVD [3]. CVD has become a major public health issue that threatens population health and increases the socioeconomic burden in China. As one of the most important modifiable risk factors for CVD, hypertension is closely associated with the occurrence of cardiovascular events, including coronary heart disease and stroke [4], with this association being particularly pronounced among middle-aged and older adults [5]. Therefore, identifying reliable, convenient, and clinically accessible risk assessment indicators in individuals with hypertension is of great significance for the early recognition and precise stratification of cardiovascular risk, as well as for the prevention of cardiovascular events and improvement of patient prognosis. Insulin resistance (IR) is considered an important pathophysiological mechanism underlying hypertension, metabolic disorders, and cardiovascular disease [6,7]. Previous studies have shown that IR can contribute to endothelial dysfunction, vascular inflammation, oxidative stress, arterial stiffness, sympathetic nervous system activation, and the progression of atherosclerosis, thereby increasing cardiovascular risk[8,9]. Although the hyperinsulinemic-euglycemic clamp remains the gold standard for assessing IR, its invasive nature, technical complexity, and high cost limit its widespread application in large-scale epidemiological studies and routine clinical practice [10]. Therefore, simple and cost-effective surrogate indicators of IR have received increasing attention in recent years.

Among these surrogate indicators, the triglyceride-glucose (TyG) index, calculated from fasting triglycerides and fasting plasma glucose, has emerged as a practical and reliable marker of IR [11]. Numerous studies have shown that an elevated TyG index is associated with hypertension, diabetes, metabolic syndrome, chronic kidney disease, and cardiovascular events [12]. However, growing evidence suggests that

combining the TyG index with obesity-related parameters may better reflect the complex interactions among lipid metabolism, glucose metabolism, and fat accumulation [7]. Accordingly, several TyG-related indices have been proposed as enhanced IR markers, including the triglyceride-glucose body mass index (TyG-BMI) [14], triglyceride-glucose waist circumference (TyG-WC) [15], and triglyceride-glucose waist-to-height ratio (TyG-WHtR) [16]. Among these indices, TyG-BMI integrates triglycerides, fasting plasma glucose, and BMI, thereby simultaneously reflecting metabolic dysfunction and obesity status [7,17].

Recent studies have shown that TyG-BMI is closely associated with various cardiometabolic diseases, including hypertension, coronary heart disease, stroke, and cardiovascular mortality [7]. In addition, prospective cohort studies based on the China Health and Retirement Longitudinal Study (CHARLS) have found that TyG-BMI and other TyG-related indices are independently associated with incident cardiovascular disease and heart disease [18,19]. Furthermore, restricted cubic spline (RCS) analyses in previous studies have indicated that the associations between TyG-related indices and cardiovascular outcomes are predominantly linear [20,21]. However, current evidence regarding the predictive performance of TyG-BMI compared with other TyG-related indices remains inconsistent. Some studies have suggested that TyG-WC or TyG-WHtR may have superior predictive ability for cardiovascular outcomes [20,22], whereas others have indicated that TyG-BMI may be more strongly associated with cardiovascular risk, particularly among individuals without diabetes [18]. Therefore, further research is needed to clarify the predictive value of TyG-BMI in specific high-risk populations.

Meanwhile, because machine learning (ML) techniques can capture complex nonlinear relationships and interactions among multiple variables, their application in cardiovascular risk prediction has increased in recent years. Compared with traditional statistical methods, ML algorithms may provide more accurate risk stratification. Recent studies have shown that incorporating IR-related indicators into ML models can improve the predictive performance for cardiovascular outcomes

[23–25]. However, evidence regarding the incremental predictive value of TyG-BMI in machine learning-based cardiovascular prediction models remains limited, particularly among middle-aged and older adults with hypertension.

Therefore, based on the nationally representative CHARLS cohort, this study aimed to investigate the association between TyG-BMI and the future risk of cardiovascular disease among middle-aged and older adults with hypertension. We compared the predictive performance of TyG-BMI with that of other TyG-related indices and further evaluated its incremental value in machine learning-based cardiovascular risk prediction models.

2.Methods

2. 1 Research Methods

2.1.1 Study Design and Population

This study was a nationwide prospective cohort study based on data from the China Health and Retirement Longitudinal Study (CHARLS). CHARLS is a nationally representative longitudinal survey of Chinese adults aged 45 years and older. It used a multistage stratified probability sampling method and covered 28 provinces, 150 counties or districts, and 450 communities across China, ensuring good national representativeness. The baseline survey was conducted in 2011–2012, with follow-up surveys carried out in 2013, 2015, and 2018. All data were collected by professionally trained interviewers using standardized questionnaires through face-to-face interviews. The CHARLS study was approved by the Biomedical Ethics Review Committee of Peking University (IRB00001052–11015), and all participants provided informed consent in accordance with the Declaration of Helsinki. More detailed information about CHARLS is available on its official website.（http：

//charls.pku.edu.cn/en）。

In the present study, we used data from the CHARLS surveys conducted in 2011, 2013, 2015, and 2018, with participants who completed the initial survey in 2011–2012 defined as the baseline population. Study participants were selected from the CHARLS baseline cohort. To construct the final analytic sample, the inclusion criteria were as follows:

The inclusion criteria were as follows:

1.aged ⩾45 years at baseline;

2.diagnosed with hypertension at baseline;

3.available data required for calculating TyG-related indices, with at least one valid follow-up record.

The exclusion criteria were as follows:

1.presence of cardiovascular disease at baseline, including heart disease or stroke;

2.missing key variables required for calculating TyG-related indices, such as triglycerides, fasting plasma glucose, body mass index, or waist circumference;

3.lack of follow-up information on cardiovascular event outcomes;

4.extreme or implausible values for exposure variables.

After the above screening process, a total of 3,092 participants were included in the final analysis from the 17,707 participants in the 2011 baseline survey (Fig. 1).

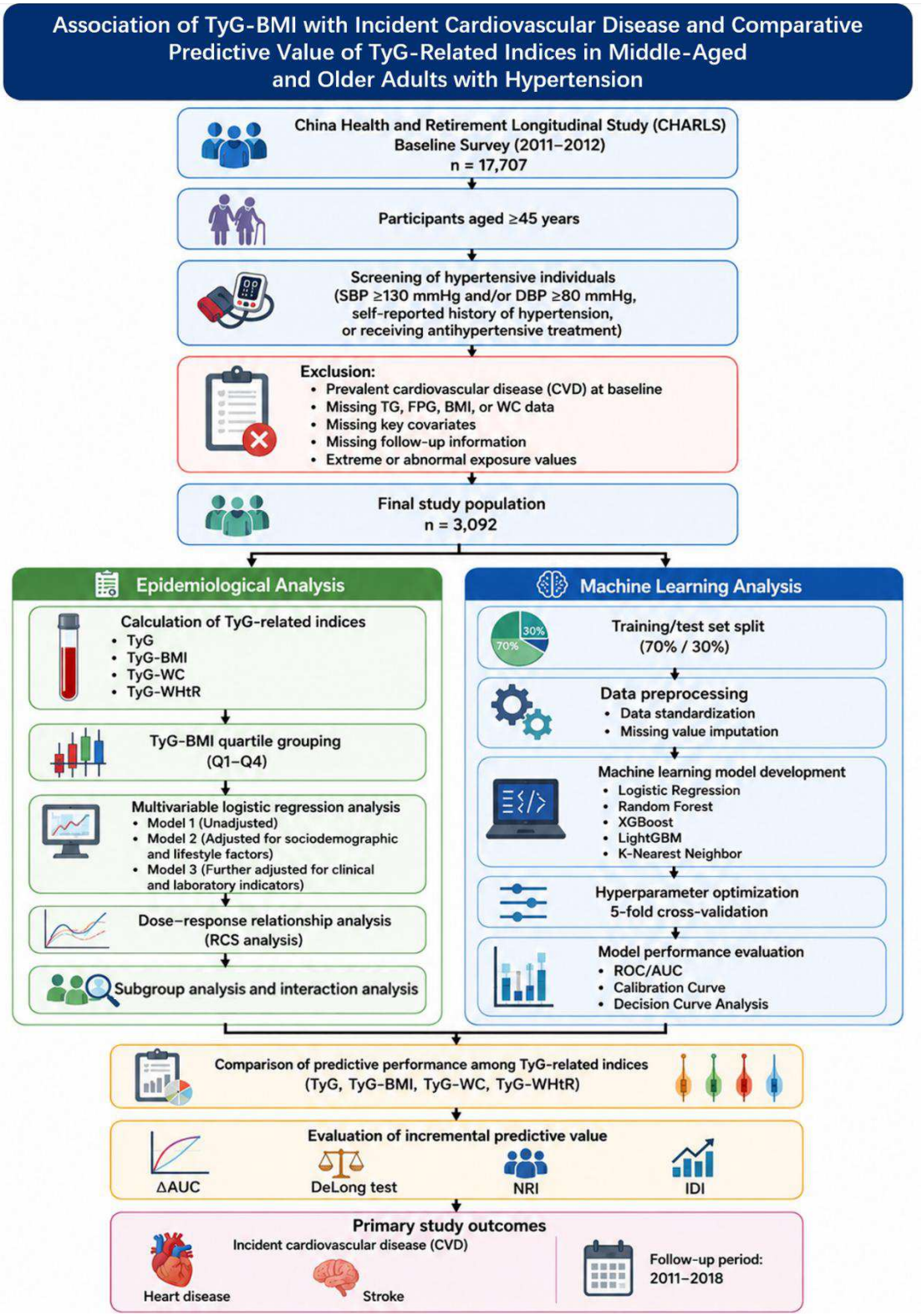

Fig. 1 Flowchart of the study

2.1.2 Calculation and evaluation of TyG-related indices

Fasting venous blood samples were collected from all participants following

standardized procedures. Fasting plasma glucose (FPG) and triglycerides (TG) were measured in certified laboratories using enzymatic colorimetric assays[25]. Anthropometric indicators included height, weight, and waist circumference (WC), which were measured three times by trained professionals according to standardized protocols, with the average value used as the final measurement. Height and weight were measured using a vertical stadiometer and a calibrated scale, respectively. Waist circumference was measured with a flexible tape at the level of the umbilicus while participants were standing [26]. TyG, TyG-BMI, TyG-WC, and TyG-WHtR were calculated using the following formulas [27–30]:

$$TyG = \ln[TG\ (mg/dL) \times FPG\ (mg/dL) / 2]$$

- $TyG\text{-}BMI = TyG \times BMI$
- $TyG\text{-}WC = TyG \times WC$
- $TyG\text{-}WHtR = TyG \times WHtR$

Body mass index (BMI) was defined as body weight in kilograms divided by height in meters squared, and waist-to-height ratio (WHtR) was defined as waist circumference divided by height. To further evaluate the potential dose–response relationship between TyG-BMI and the risk of incident CVD, TyG-BMI was categorized into quartiles from Q1 to Q4.

2.1.3 Diagnosis of Hypertension

Blood pressure was measured by uniformly trained professionals according to standardized procedures. Participants were required to remain seated and rest for at least 5 minutes before measurement. Systolic blood pressure (SBP) and diastolic blood

pressure (DBP) were then measured using an automated electronic blood pressure monitor. Each participant underwent three measurements, with an interval of 45–60 seconds between measurements, and the average of the three readings was used for subsequent analyses.

According to current international standards, hypertension was defined as meeting any of the following criteria [31]:

1. mean systolic blood pressure (SBP) ≥130 mmHg or diastolic blood pressure (DBP) ≥80 mmHg;
2. self-reported previous diagnosis of hypertension by a physician;
3. current use of antihypertensive medications.

2.1.4 Ascertainment of outcomes

The primary outcome of this study was incident cardiovascular disease (CVD) during follow-up from Wave 2 to Wave 4, including heart disease and stroke. CVD occurrence was determined based on self-reported physician diagnoses collected through standardized questionnaires in the CHARLS follow-up surveys. Participants were considered to have developed a CVD event if they reported being diagnosed by a physician with heart disease or stroke at any follow-up wave.

Heart disease was defined as an affirmative response to the question, “Have you ever been diagnosed by a doctor with heart disease, such as coronary heart disease, angina, myocardial infarction, or heart failure?” Stroke was defined as an affirmative response to the question, “Have you ever been diagnosed by a doctor with stroke?” [10,19]. Participants were assessed for incident cardiovascular outcomes during the

available follow-up surveys conducted in 2013, 2015, and 2018. Participants were considered to have developed incident CVD if they reported a new physician diagnosis of heart disease or stroke at any follow-up wave. Participants who did not report either outcome during their available follow-up assessments were classified as not having incident CVD.

2.1.5 Covariates

Covariates were selected based on previous medical studies, CHARLS-related research, sociodemographic factors, and the availability of variables in the CHARLS database. They mainly included sociodemographic characteristics, lifestyle factors, clinical disease status, and biochemical indicators. Sociodemographic variables included age, sex, educational level, and marital status. Lifestyle factors included smoking status and alcohol consumption. Anthropometric and clinical indicators included systolic blood pressure (SBP), diastolic blood pressure (DBP), and body mass index (BMI). In addition, chronic disease status was obtained based on self-reported physician diagnoses, including diabetes, dyslipidemia, and hypertension. Biochemical indicators were derived from fasting venous blood samples, including fasting plasma glucose (FPG), triglycerides (TG), total cholesterol (TC), high-density lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol (LDL-C), uric acid (UA), serum creatinine (SCR), and glycated hemoglobin (HbA1c).

All covariates were measured at baseline and adjusted for in multivariable analyses to control for the potential influence of confounding factors.

2.2 Model Development and Validation

To further evaluate the predictive value of composite metabolic indicators for CVD risk among individuals with hypertension, we used the baseline characteristics of the included population to develop machine learning models. Within the machine learning framework, we systematically compared the predictive performance of TyG-BMI with that of other TyG-related indices for CVD events.

First, participants were randomly divided into a training set (70%) and a testing set (30%) using stratified sampling to ensure comparable distributions of outcome events between the two sets. All data preprocessing procedures were performed based on the training set and then synchronously applied to the testing set to avoid data leakage. Continuous variables were standardized, and missing data were handled using appropriate imputation methods. Second, a basic prediction model was constructed using traditional cardiovascular risk factors, including demographic characteristics, lifestyle factors, clinical indicators, and biochemical parameters. On this basis, TyG and each TyG-related index, including TyG-BMI, TyG-WC, and TyG-WHtR, were separately added to the basic model to construct corresponding extended models. This approach enabled a direct comparison of the predictive performance of different TyG-related indices within the same machine learning framework.Third, to evaluate potential linear and complex nonlinear relationships between predictors and outcomes, multiple machine learning algorithms were used to develop and compare prediction models, including logistic regression, random forest, eXtreme Gradient Boosting (XGBoost), LightGBM, and k-nearest neighbors (KNN).Model hyperparameters were tuned using a grid search approach, and model performance was optimized with five-

fold cross-validation in the training set. The discriminative ability of the models was evaluated using receiver operating characteristic (ROC) curves and the area under the curve (AUC), with a higher AUC indicating better ability to distinguish individuals who developed CVD events from those who did not.The clinical utility of the models was assessed using decision curve analysis (DCA), which quantified the net benefit and potential clinical applicability of the models across different risk thresholds. When appropriate, the DeLong test was used to compare differences in AUCs between models, thereby evaluating statistical differences in model discrimination. Finally, models were constructed and evaluated separately for incident cardiovascular disease (CVD), heart disease, and stroke to explore potential heterogeneity in predictive performance across different cardiovascular outcomes.

2.3 Statistical analysis

2.3.1 Association Between TyG-BMI and CVD

Participants in this study were divided into four groups according to the quartiles of TyG-BMI (Q1–Q4). The distributions of continuous and categorical variables were compared across groups, and the potential dose–response relationship between TyG-BMI and study outcomes was preliminarily explored.

According to the distribution characteristics of the variables, different descriptive statistical methods were used for data analysis. Normally distributed continuous variables were expressed as mean ± standard deviation (SD), and between-group comparisons were performed using the independent-samples t test or one-way analysis of variance (ANOVA). Non-normally distributed continuous variables were expressed

as median and interquartile range (IQR), and between-group comparisons were performed using the Mann–Whitney U test. Categorical variables were presented as frequencies and percentages, and between-group differences were compared using the chi-square test. Missing covariate data were handled using multiple imputation to reduce potential bias and improve the robustness of the statistical analyses. Because the primary outcome was defined as the occurrence of incident CVD during the available follow-up period rather than the exact time to event, multivariable logistic regression was used to evaluate the association between baseline TyG-BMI and subsequent cardiovascular outcomes. Model 1 was an unadjusted model. Model 2 was adjusted for sociodemographic and lifestyle factors, including sex, marital status, educational level, smoking status, and alcohol consumption. Model 3 was the fully adjusted model, which further included clinical indicators and biochemical covariates based on Model 2.To avoid potential collinearity, variables used to construct the exposure indices were not simultaneously included in the corresponding models. For example, because the exposure variable in this study included TyG-BMI, BMI and fasting plasma glucose were not included as covariates in the regression models. The association between the TyG-BMI index and cardiovascular disease was expressed as odds ratios (ORs) and 95% confidence intervals (CIs). Based on the fully adjusted model, restricted cubic spline (RCS) regression analysis was further performed to explore the dose‑response relationship and potential nonlinear association between TyG-BMI and CVD risk. The optimal number of knots for the RCS model was selected according to the Akaike information criterion (AIC).Subgroup and sensitivity analyses were subsequently

conducted based on baseline characteristics. Interaction tests were performed to evaluate the stability and heterogeneity of the association between TyG-BMI and CVD risk across different subgroups. All statistical analyses were performed using Python Scikit-learn library (version 1.1.3, https://github.com/scikit-learn/scikit-learn). A two-sided P value <0.05 was considered statistically significant.

2.3.2 Model Performance Comparison Among TyG-Related Indices

To evaluate the incremental predictive value of TyG and TyG-related indices for CVD risk prediction models, a basic model including only traditional risk factors was first constructed. Subsequently, TyG, TyG-BMI, TyG-WC, and TyG-WHtR were separately added to the basic model to construct corresponding extended models.The discriminative ability of different models was compared by plotting receiver operating characteristic (ROC) curves and calculating the area under the curve (AUC). The incremental predictive performance of each index relative to the basic model was evaluated using the change in AUC (ΔAUC), and differences in AUCs between models were statistically compared using the DeLong test.

In addition, net reclassification improvement (NRI) and integrated discrimination improvement (IDI) were further calculated to assess improvements in risk reclassification ability and overall discriminative performance after adding TyG and its related indices to the model.

2.3.3 Evaluation of the Basic Performance of Machine Learning Models

In both the training and testing sets, multiple complementary metrics were used to comprehensively evaluate the discriminative ability, calibration performance, and

clinical utility of the machine learning models. Discriminative ability was assessed using receiver operating characteristic (ROC) curves and the area under the curve (AUC), which quantified the ability of the models to distinguish individuals who developed cardiovascular events from those who did not.Classification performance was evaluated using accuracy, precision, recall, and F1-score. Accuracy reflects the overall proportion of correctly classified samples. Precision represents the proportion of true positive cases among all samples predicted as positive. Recall reflects the sensitivity of the model in identifying positive cases. The F1-score is the harmonic mean of precision and recall and was used to comprehensively evaluate model performance.Model calibration was assessed using calibration curves, which compare the agreement between predicted probabilities and observed event rates. Greater consistency between predicted and observed risks indicates better calibration performance. Clinical utility was evaluated using decision curve analysis (DCA), which calculates the net benefit of the model across a range of threshold probabilities, thereby assessing its potential value in clinical decision-making.

All evaluation metrics were calculated for each machine learning model, and comparative analyses were performed to identify the model with the best overall predictive performance.

3.Results

3. 1 Baseline Characteristics

A total of 3,092 participants with hypertension were included in the final analysis. The mean age was 59.79 ± 9.14 years, and 53.33% were women. During follow-up

through 2018, 886 participants (28.65%) developed CVD, including 659 cases of heart disease and 329 cases of stroke .

Baseline characteristics of the study population differed across TyG-BMI quartiles (Table 1). Significant differences across TyG-BMI quartiles were observed for most demographic, clinical, metabolic, and cardiovascular characteristics, whereas serum creatinine did not differ significantly across groups ($P = 0.923$). Compared with participants who did not develop CVD, those who developed CVD were generally older and had worse cardiometabolic profiles. Specifically, participants with incident CVD had higher body mass index (BMI), triglyceride (TG), and fasting plasma glucose (FPG) levels, but lower high-density lipoprotein cholesterol (HDL-C) levels than those without CVD. In addition, TyG-BMI was higher among participants who developed CVD, suggesting a closer association with cardiovascular risk. Participants in the higher TyG-BMI quartiles were generally younger and had a higher proportion of women. Regarding socioeconomic characteristics, the proportion of married participants was higher in the higher TyG-BMI quartiles, and educational level also differed significantly across groups.As TyG-BMI levels increased, BMI, TG, LDL-C, FPG, HbA1c, WC, WHtR, the prevalence of diabetes, the prevalence of dyslipidemia, and the use of antihypertensive medications all gradually increased, whereas HDL-C levels decreased significantly (all $P < 0.05$). In addition, the prevalence of CVD, heart disease, and stroke increased progressively across TyG-BMI quartiles. Notably, the proportions of smoking and alcohol consumption showed a decreasing trend with increasing TyG-BMI levels.

360 Table 1 Baseline characteristics and cardiovascular outcomes by TyG-BMI quartiles

Characteristics	Total	Quartiles of TyG-BMI				P value
	(n=3092)	Quartile1(n=774)	Quartile 2(n=771)	Quartile 3 (n=774)	Quartile 4 (n=773)	
Age (years)	59.79±9.14	62.77±9.63	59.65±9.17	59.24±8.86	57.47±8.02	<0.001
Gender, n (%)						
Male	1443(46.67)	437(56.46)	358(46.43)	359(46.38)	289(37.38)	<0.001
Female	1649(53.33)	337(43.54)	413(53.57)	415(53.62)	484(62.61)	<0.001
Education level, n (%)						<0.001
Below middle school	2821(91.24)	737(95.21)	689(89.48)	694(89.66)	701(90.69)	
High school	246(7.96)	31(4.01)	76(9.87)	75(9.69)	64(8.28)	
College or above	24(0.71)	6(0.78)	5(0.64)	5(0.65)	8(1.03)	
Marital status, n (%)						<0.001
Married	2206(83.94)	632(81.65)	656(85.08)	688(88.89)	705(91.20)	
Others	422(16.06)	142(18.35)	115(14.92)	86(11.11)	68(8.80)	
BMI (kg/m^2)	24.16 ± 3.95	19.83 ± 1.75	22.79 ± 1.39	25.13 ± 1.61	28.89 ± 3.31	<0.001
TG (mg/dL)	111.51 (79.65–160.18)	77.88 (58.41–102.66)	99.12 (74.56–135.40)	124.34 (95.58–165.49)	170.80 (126.56–243.38)	<0.001
HDL-C (mg/dL)	50.77 ± 14.89	59.45 ± 15.47	53.55 ± 13.67	47.71 ± 12.89	42.38 ± 11.47	<0.001
LDL-C (mg/dL)	119.87 ± 35.12	112.12 ± 32.04	122.03 ± 33.03	124.03 ± 34.81	121.33 ± 39.04	<0.001
FPG (mg/dL)	103.32 (95.58–114.66)	99.72 (92.16–107.59)	100.98 (94.18–109.80)	104.22 (97.38–114.30)	110.88 (100.80–131.22)	<0.001
SBP (mmHg)	141.76 ± 18.21	141.60 ± 18.75	140.12 ± 17.54	142.90 ± 17.55	142.41 ± 18.86	0.0166
DBP (mmHg)	81.91 ± 10.81	79.94 ± 10.62	81.30 ± 10.71	82.75 ± 10.34	83.65 ± 11.20	<0.001
Alcohol consumption,n(%)	1270(41.07)	368(47.55)	314(40.73)	323(41.73)	265(34.28)	<0.001
Smoking, n(%)	1205(38.97)	400(51.68)	295(38.26)	286(36.95)	224(28.98)	<0.001
WC（cm）	87.25 ± 10.08	77.34 ± 6.65	84.09 ± 6.43	90.18 ± 6.70	97.37 ± 7.52	<0.001
UA（mg/dl）	4.39 (3.64–5.26)	4.18 (3.50–5.06)	4.35 (3.58–5.11)	4.42 (3.72–5.26)	4.66 (3.81–5.59)	<0.001
Scr（mg/dl）	0.76 (0.66–0.89)	0.76 (0.66–0.88)	0.77 (0.67–0.88)	0.76 (0.66–0.90)	0.76 (0.67–0.89)	0.923
HbA1c（%）	5.33 ± 0.89	5.12 ± 0.61	5.19 ± 0.60	5.34 ± 0.90	5.65 ± 1.20	<0.001
Diabetes, n(%)	216(6.99)	24(3.10)	21(2.72)	60 (7.75)	111(14.36)	<0.001
Dyslipidemia, n (%)	346(11.19)	34(4.39)	54(7.00)	82(10.59)	176(22.77)	<0.001
Antihypertensive therapy, n (%)	939(30.37)	145(18.73)	193(25.03)	262(33.85)	339(43.85)	<0.001
Heart disease, n(%)	659(21.31)	150(19.38)	152(19.71)	161(20.80)	196(25.36)	0.014
Stroke, n(%)	329(10.64)	61 (7.88)	75(9.72)	94(12.14)	99(12.81)	0.006
CVD, n(%)	886(28.65)	195(25.19)	207(26.85)	221(28.55)	263(34.02)	<0.001
WHtR	0.55 ± 0.07	0.49 ± 0.05	0.53 ± 0.05	0.57 ± 0.05	0.62 ± 0.05	<0.001
TyG-BMI	207.92 (181.75–237.60)	167.22 (155.84–174.79)	194.49 (188.16–201.01)	221.15 (214.49–229.09)	260.54 (247.69–277.77)	<0.001
TyG	8.68 (8.29–9.10)	8.27 (7.99–8.54)	8.54 (8.23–8.86)	8.79 (8.50–9.14)	9.19 (8.86–9.67)	<0.001
TyG-WC	756.53 (674.98–844.15)	637.25 (599.84–675.12)	716.21 (679.71–752.25)	796.50 (757.23–834.02)	896.48 (848.35–951.41)	<0.001

TyG-WHtR	4.80 (4.28–5.37)	4.02 (3.78–4.32)	4.53 (4.30–4.81)	5.01 (4.76–5.27)	5.68 (5.38–6.06)	<0.001

TyG-BMI, Triglyceride glucose-body mass index; SBP, systolic blood pressure; DBP, diastolic blood pressure; BMI, body mass index; WC, waist circumference; TG, triglyceride; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; Scr, serum creatinine; FPG, fasting plasma glucose; HbA1c, glycosylated hemoglobin A1c; UA, uric acid; IR, insulin resistance; TyG, triglyceride–glucose; TyG-WC, TyG-waist circumference; TyG-WHtR, TyG-waist-to-height ratio; CVD,Cardiovascular diseases

3.2 Association Between Baseline TyG-BMI and CVD Events

To evaluate the association between baseline TyG-BMI and incident cardiovascular outcomes, multivariable logistic regression models were used for analysis. The association of TyG-BMI was assessed both as a continuous variable and as a categorical variable based on quartiles. Three progressively adjusted models were constructed: Model 1 was the unadjusted model; Model 2 was further adjusted for sex, marital status, educational level, smoking status, and alcohol consumption based on Model 1; and Model 3 was the fully adjusted model, which further included clinical and biochemical indicators based on Model 2, including systolic blood pressure (SBP), diastolic blood pressure (DBP), high-density lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol (LDL-C), uric acid (UA), serum creatinine (Scr), glycated hemoglobin (HbA1c), diabetes, dyslipidemia, and antihypertensive treatment (Table 2).

Table 2. Multivariable logistic regression analyses of TyG-BMI and cardiovascular outcomes

TyG-BMI	**Model 1**		**Model 2**		**Model 3**	
	OR (95%CI)	**P value**	**OR (95%CI)**	**P value**	**OR (95%CI)**	**P value**
Cardiovascular diseases			Cardiovascular diseases		Cardiovascular diseases	

Continuous	1.004 (1.002-1.006)	<0.001	1.006(1.004- 1.008)	<0.001	1.003 (1.001-1.006)	0.006
Quartile 1	**Reference**		**Reference**		**Reference**	
Quartile 2	1.09(0.87- 1.37)	0.450	1.19(0.95 -1.51)	0.137	1.15 (0.90-1.46)	0.270
Quartile 3	1.19 (0.95 -1.49)	0.136	1.32(1.05 -1.66)	0.019	1.14 (0.89-1.46)	0.301
Quartile 4	1.53 (1.23 -1.91)	<0.001	1.80(1.43 – 2.28)	<0.001	1.36 (1.03-1.79)	0.028
P for trend	<0.001		<0.001		0.040	
Heart disease			**Heart disease**		**Heart disease**	
Continuous	1.004 (1.002 – 1.006)	<0.001	1.004 (1.002 – 1.006)	<0.001	1.003 (1.001-1.006)	0.012
Quartile 1	Reference		Reference		Reference	
Quartile 2	1.02(0.80 – 1.32)	0.86	1.08 (0.83 – 1.39)	0.580	1.08 (0.83-1.41)	0.574
Quartile 3	1.09(0.85– 1.40)	0.49	1.16 (0.90 – 1.50)	0.259	1.13(0.86-1.50)	0.388
Quartile 4	1.41(1.11 – 1.80)	0.005	1.55(1.20 – 1.99)	<0.001	1.37(1.01-1.85)	0.043
P for trend	0.004		<0.001		0.045	
Stroke			Stroke		Stroke	
Continuous	1.005 (1.002 – 1.007)	<0.001	1.007 (1.004 – 1.009)	<0.001	1.002 (0.999-1.006)	0.194
Quartile 1	Reference		Reference		Reference	
Quartile 2	1.26(0.89 – 1.80)	0.198	1.45 (1.01–2.08)	0.044	1.28(0.88-1.86)	0.190
Quartile 3	1.61 (1.15 – 2.27)	0.006	1.91 (1.35–2.70)	<0.001	1.35(0.92-1.96)	0.122
Quartile 4	1.72 (1.23– 2.40)	0.002	2.23 (1.57–3.18)	<0.001	1.29(0.85-1.95)	0.226
P for trend	<0.001		<0.001		0.257	

381 OR, odds ratio; CI, confidence interval

382 In the unadjusted model (Model 1), higher TyG-BMI levels were significantly
383 associated with increased risks of cardiovascular disease, heart disease, and stroke.
384 Compared with participants in the lowest TyG-BMI quartile (Q1), those in the highest
385 quartile (Q4) had a 53% higher risk of CVD (OR = 1.53, 95% CI: 1.23–1.91, P < 0.001),
386 a 41% higher risk of heart disease (OR = 1.41, 95% CI: 1.11–1.80, P = 0.005), and a
387 72% higher risk of stroke (OR = 1.72, 95% CI: 1.23–2.40, P = 0.002). Significant dose–

response relationships across TyG-BMI quartiles were observed for all cardiovascular outcomes (all P for trend < 0.01).

The P values for trend tests were calculated and found to be all less than 0.01, indicating significant dose–response relationships across TyG-BMI quartiles for all cardiovascular outcomes. In Model 2, after further adjustment for sociodemographic and lifestyle variables, these associations were slightly strengthened. Compared with participants in Q1, those in Q4 had an 80% higher risk of CVD (OR = 1.80, 95% CI: 1.43–2.28, P < 0.001), a 55% higher risk of heart disease (OR = 1.55, 95% CI: 1.20–1.99, P < 0.001), and more than a twofold higher risk of stroke (OR = 2.23, 95% CI: 1.57–3.18, P < 0.001). The trend associations across quartiles were statistically significant for all outcomes (all P for trend < 0.001). In the fully adjusted model (Model 3), TyG-BMI remained independently associated with CVD and heart disease. Compared with participants in Q1, those in Q4 had a 36% higher risk of CVD (OR = 1.36, 95% CI: 1.03–1.79, P = 0.028) and a 37% higher risk of heart disease (OR = 1.37, 95% CI: 1.01–1.85, P = 0.043). Significant dose–response relationships remained for CVD and heart disease, with P for trend values of 0.040 and 0.045, respectively.When TyG-BMI was analyzed as a continuous variable, each one-unit increase was associated with a 0.3% higher risk of CVD (OR = 1.003, 95% CI: 1.001–1.006, P = 0.006) and a 0.3% higher risk of heart disease (OR = 1.003, 95% CI: 1.001–1.006, P = 0.012). However, after full adjustment for confounding factors, the association between TyG-BMI and stroke was attenuated and no longer statistically significant (OR = 1.002, 95% CI: 0.999–1.006, P = 0.194).

Restricted cubic spline (RCS) analysis was used to evaluate the dose–response relationships between TyG-BMI and the risks of CVD, heart disease, and stroke (Fig. 2). In Model 1 and Model 2, TyG-BMI showed significant overall associations with the risks of CVD, heart disease, and stroke (all P-overall < 0.05), whereas no significant nonlinear associations were observed (all P-nonlinear > 0.05), suggesting that these associations were predominantly linear.In the fully adjusted model, the overall association between TyG-BMI and CVD remained statistically significant (P-overall = 0.037), and a marginally significant association was observed for heart disease (P-overall = 0.094), whereas the association with stroke was attenuated and no longer statistically significant (P-overall = 0.400). After full adjustment, no significant nonlinear associations were observed for any outcome (all P-nonlinear > 0.05), further indicating that the associations between TyG-BMI and cardiovascular outcomes were mainly linear.

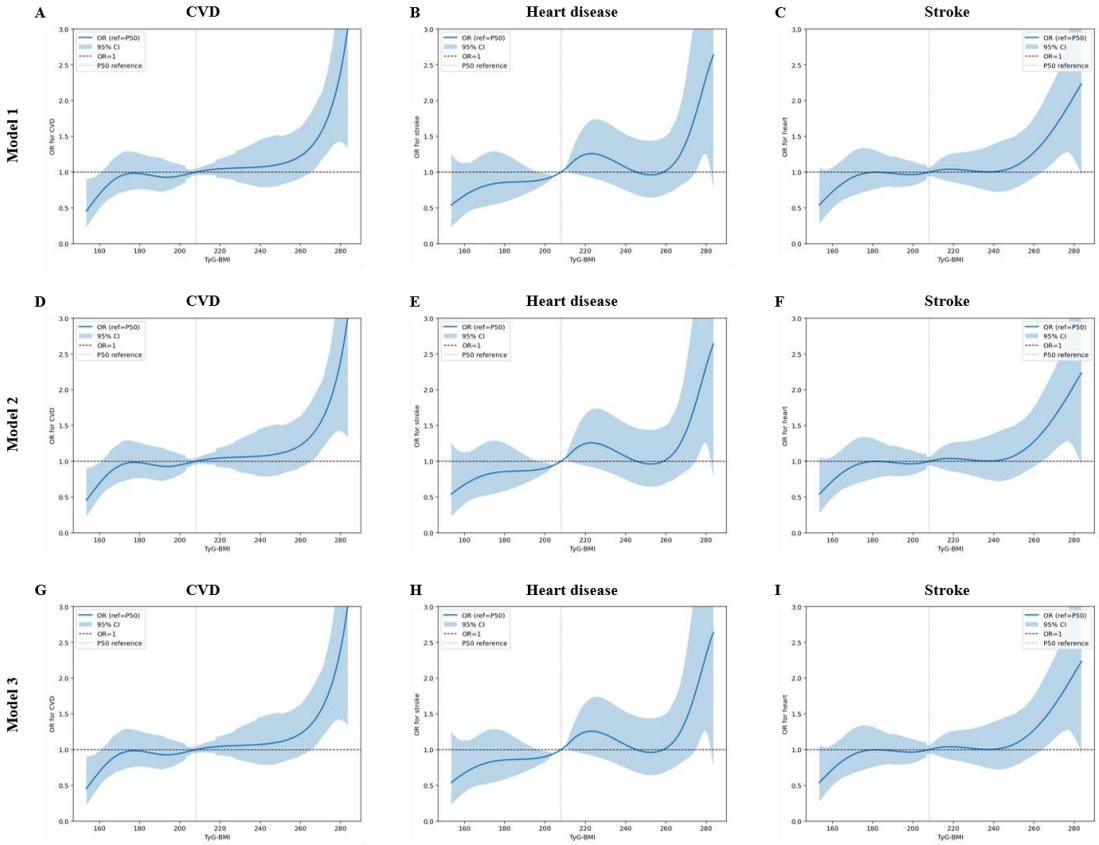

Figure 2. Restricted cubic spline analyses of TyG-BMI and cardiovascular outcomes. Restricted cubic spline analyses of the associations between TyG-BMI and cardiovascular disease (A, D, G), heart disease (B, E, H), and stroke (C, F, I) in Models 1–3. Model 1 was unadjusted. Model 2 was adjusted for sociodemographic and lifestyle factors, including sex, marital status, educational level, smoking status, and alcohol consumption. Model 3 was further adjusted for clinical indicators and biochemical covariates based on Model 2.

These results indicate that higher TyG-BMI levels are associated with an increased risk of cardiovascular outcomes, particularly CVD and heart disease, and that these associations are predominantly linear rather than nonlinear.

3.3 Comparison of model performance across TyG-related indices

To further compare the predictive performance of different TyG-related indices within the same machine learning framework, TyG, TyG-WC, TyG-WHtR, and TyG-BMI were separately incorporated into the basic prediction model, and their

incremental predictive performance for incident CVD was further evaluated.

Among all TyG-related indices, the model incorporating TyG-BMI showed the most pronounced improvement in predictive performance. The basic model without any TyG-related index had an AUC of 0.593 (95% CI: 0.56–0.62). After adding TyG-BMI, the AUC increased to 0.636 (95% CI: 0.61–0.66), corresponding to a ΔAUC of 0.043 (Fig. 3).

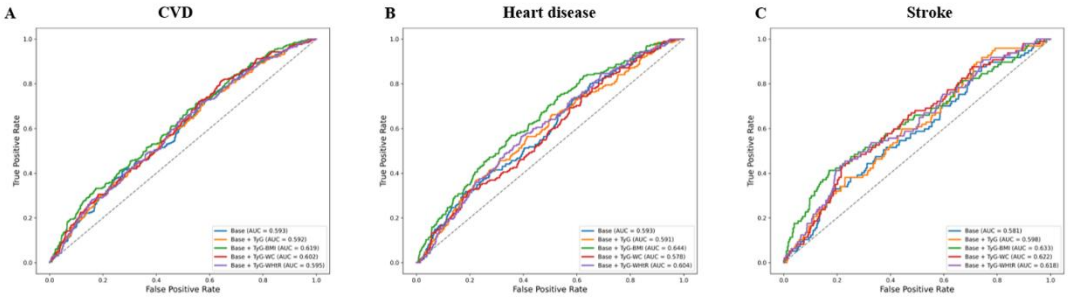

Figure 3. Comparison of the predictive performance of TyG-related indices for incident cardiovascular disease. CVD, cardiovascular disease; TyG, triglyceride–glucose; TyG-WC, TyG-waist circumference; TyG-BMI, TyG-body mass index; TyG-WHtR, TyG-waist-to-height ratio;

Table 3. Comparison of the incremental predictive value of TyG-related indices for incident cardiovascular disease

Model	AUC	Δ AUC	DeLong Z	P value	NRI	IDI
Base model	0.593（0.56-0.62）	-	-	-	-	-
Base + TyG	0.592（0.56-0.62）	-0.001	0.42	0.67	0.012	0.003
Base + TyG-WC	0.602（0.57-0.63）	0.009	1.12	0.26	0.045	0.008
Base + TyG-WHtR	0.595（0.56-0.62）	0.002	0.63	0.52	0.018	0.004
Base + TyG-BMI	0.636（0.61-0.66）	0.043	3.586	<0.001	0.192	0.036

Base model included age, sex, marital status, educational level, smoking status, alcohol consumption, systolic blood pressure (SBP), high-density lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol (LDL-C), diabetes, dyslipidemia, and antihypertensive therapy. TyG, TyG-BMI, TyG-WC, and TyG-WHtR were added individually to the base model. ΔAUC represents

the change in the area under the receiver operating characteristic curve compared with the base model. Model discrimination was compared using the DeLong test. NRI, net reclassification improvement; IDI, integrated discrimination improvement.

The DeLong test showed that this improvement was statistically significant ($Z = 3.586$, $P < 0.001$). In addition, the model incorporating TyG-BMI demonstrated the highest net reclassification improvement (NRI = 0.192) and integrated discrimination improvement (IDI = 0.036), suggesting significant improvements in both risk stratification and model discrimination.

Compared with TyG-BMI, the other TyG-related indices provided smaller improvements in predictive performance. Adding TyG alone did not improve model discrimination (AUC = 0.592, $\Delta AUC = -0.001$, $P = 0.67$). The improvements observed after incorporating TyG-WC (AUC = 0.602, $\Delta AUC = 0.009$, $P = 0.26$) and TyG-WHtR (AUC = 0.595, $\Delta AUC = 0.002$, $P = 0.52$) were also modest and did not reach statistical significance in the DeLong tests. Although these indices yielded positive NRI and IDI values, their incremental predictive improvements were smaller than those observed after adding TyG-BMI (Table 3).

In summary, compared with other TyG-related indices, TyG-BMI demonstrated superior incremental predictive value for incident CVD and may serve as a more effective metabolic marker for cardiovascular risk prediction among individuals with hypertension.

3. 4 Performance Evaluation of Machine Learning Models Based on TyG-BMI

To further evaluate the predictive performance of TyG-BMI within a machine learning framework, five models were developed using the training set and evaluated

in the independent testing set: Logistic Regression (LR), Random Forest (RF), eXtreme Gradient Boosting (XGB), LightGBM (LGBM), and K-Nearest Neighbors (KNN). Model performance was assessed using accuracy, precision, recall, F1-score, and AUC. Among the five models, LR achieved the highest AUC (0.632) and accuracy (0.713), whereas LGBM showed the highest recall (0.179) and F1-score (0.258) (Table 4).

Table 4. Predictive performance of machine learning models for incident cardiovascular disease

Model	Accuracy	Precision	Recall	F1-score	AUC
LR	0.713043	0.484375	0.118321	0.190184	0.631656
XGB	0.702174	0.428571	0.137405	0.208092	0.619024
RF	0.700000	0.393939	0.099237	0.158537	0.590521
LGB	0.706522	0.460784	0.179389	0.258242	0.587815
KNN	0.698913	0.325581	0.053435	0.091803	0.539200

Performance metrics were evaluated using the testing set. Accuracy, precision, recall, F1-score, and area under the receiver operating characteristic curve (AUC) were used to assess model performance.

All models demonstrated some predictive ability for cardiovascular disease, but their discriminative performance varied. The Logistic Regression model achieved the highest AUC value (0.632) and the highest accuracy (0.713), suggesting that it performed best in distinguishing individuals who would develop CVD from those who would not. The XGBoost model had an AUC of 0.619, slightly lower than that of the LR model, whereas the AUCs of the Random Forest, LightGBM, and KNN models were 0.591, 0.588, and 0.539, respectively. Although the LightGBM model achieved the highest recall (0.179) and F1-score (0.258), its overall discriminative ability remained lower than that of the LR model.

Notably, contrary to expectations, tree-based ensemble learning algorithms, including XGB, RF, and LGBM, did not demonstrate clearly superior predictive performance compared with traditional Logistic Regression. This may suggest that, in the present study cohort, the relationships between cardiovascular disease risk and the included variables were mainly relatively stable and linear, with limited complex nonlinear patterns; therefore, the traditional Logistic Regression model was able to capture the major risk information effectively.In addition, the incidence of outcome events in this study was relatively low, and there was a certain degree of class imbalance, which resulted in generally low recall and F1-score values across the models. This is also a relatively common phenomenon in long-term follow-up cohort studies.

Decision curve analysis (DCA) was performed to evaluate the potential clinical utility of the prediction models. For CVD, heart disease, and stroke, the models demonstrated higher net benefits than the “treat-all” and “treat-none” strategies across a range of threshold probabilities (Fig. 4). Calibration curves showed relatively good agreement between predicted probabilities and observed event rates in both the training and testing sets, with the curves generally approaching the ideal reference line (Fig. 4).

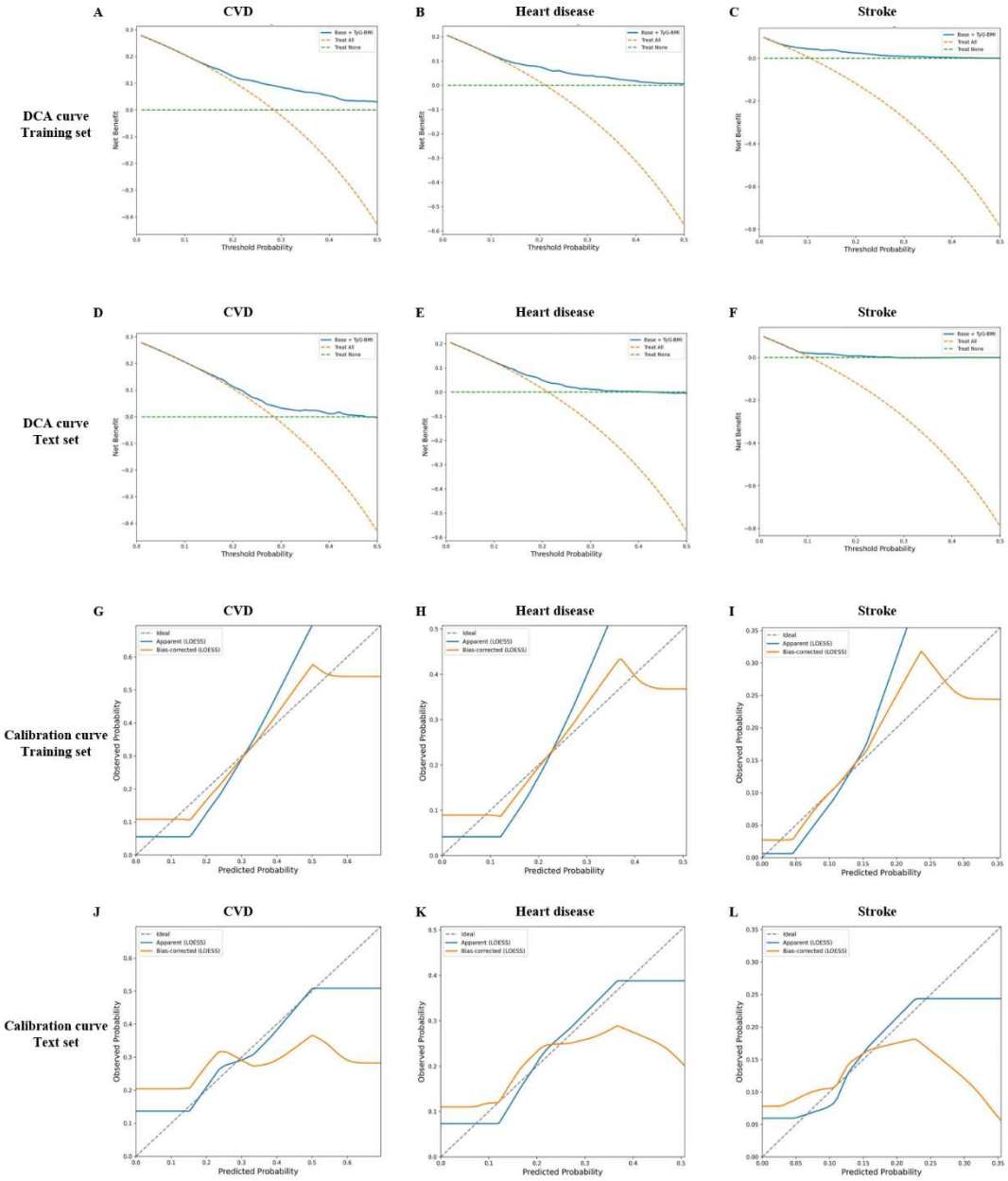

Figure 4. Decision curve and calibration analyses of the TyG-BMI prediction model. Decision curve and calibration analyses of the TyG-BMI-based machine learning model for predicting incident cardiovascular outcomes. Decision curve analyses for cardiovascular disease (A, D), heart disease (B, E), and stroke (C, F) in the training (A–C) and testing (D–F) sets are shown in the upper panels. Calibration curves for cardiovascular disease (G, J), heart disease (H, K), and stroke (I, L) in the training (G–I) and testing (J–L) sets are shown in the lower panels.

Considering all evaluation metrics, the Logistic Regression model demonstrated

the best overall predictive performance. In addition, the constructed models showed good clinical utility, supporting the potential application of machine learning approaches in cardiovascular risk assessment among individuals with hypertension.

3.6 Subgroup Analysis and Interaction Analysis

To further evaluate the robustness and potential heterogeneity of the association between TyG-BMI and the risk of incident CVD, subgroup analyses were performed according to age (<60 or ≥60 years), sex, BMI (<24 or ≥24 kg/m²), and diabetes status. Interaction tests were conducted to assess whether the association differed across subgroups.

Overall, the association between TyG-BMI and an increased risk of incident CVD was generally consistent in direction across different subgroups (Fig. 5A). The association between TyG-BMI and increased CVD risk reached statistical significance among younger participants (<60 years) and women. Similar trends toward increased risk were also observed among participants aged ≥60 years, men, those with BMI ≥24 kg/m², and patients with diabetes, although some subgroup associations did not reach statistical significance, possibly due to the limited number of events and insufficient statistical power.

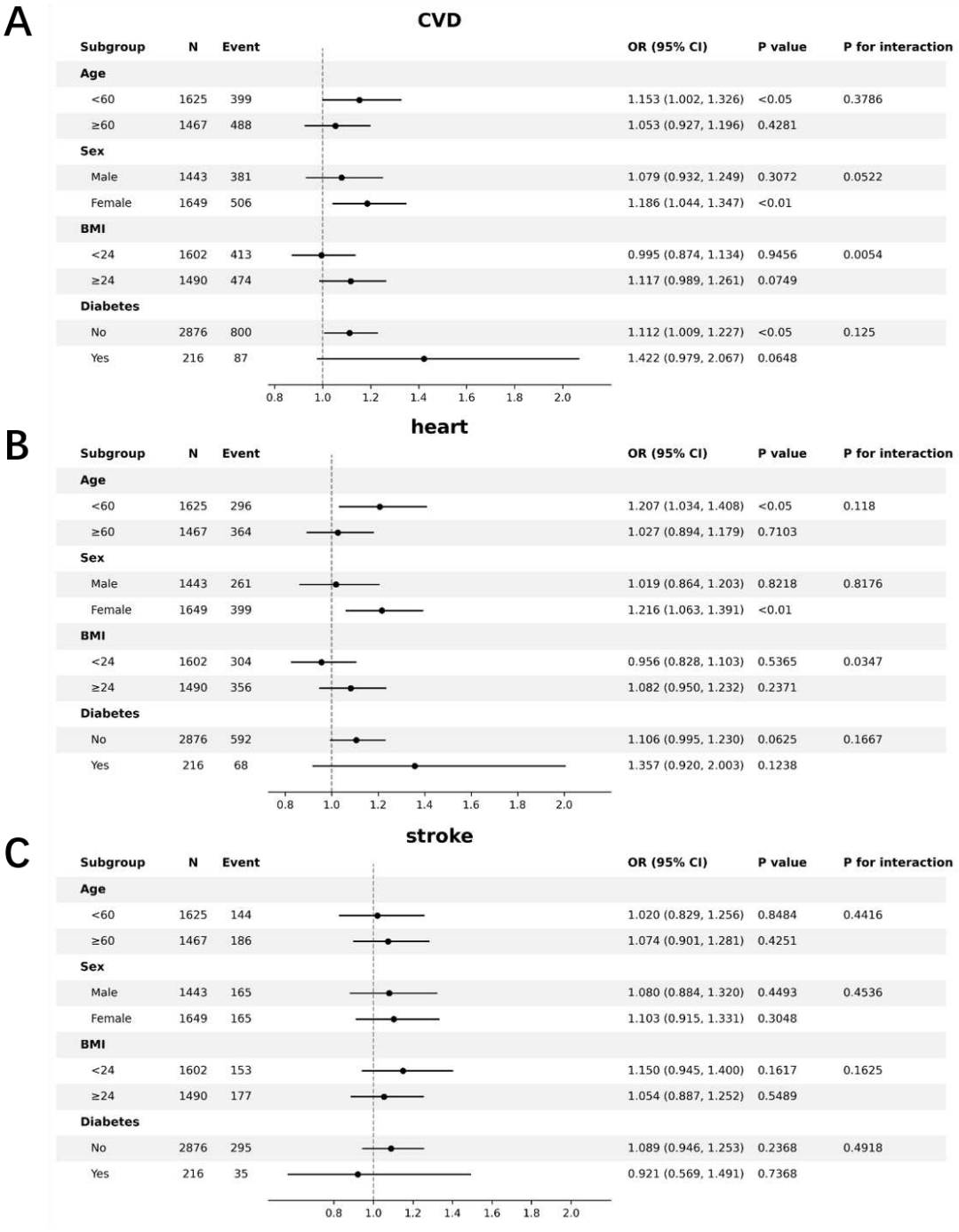

Fig. 5 Subgroup analysis of the association between estimated TyG-BMI and A CVD, B heart disease, and C stroke. CVD, cardiovascular disease; OR, odds ratio; BMI, body mass index

Interaction analysis showed that only BMI significantly modified the association between TyG-BMI and incident CVD (P for interaction = 0.005), suggesting a certain degree of heterogeneity in the association between TyG-BMI and CVD risk across

different BMI levels. In contrast, the interactions for age (P for interaction = 0.379), sex (P for interaction = 0.052), and diabetes status (P for interaction = 0.125) did not reach statistical significance, indicating that the association between TyG-BMI and CVD risk was generally stable across different age, sex, and diabetes-status subgroups.

For the heart disease outcome, TyG-BMI was significantly associated with an increased risk of heart disease among participants aged <60 years and among women, while a significant interaction was observed across BMI strata (P for interaction = 0.035; Fig. 5B). No significant interactions were observed for age, sex, or diabetes status. For the stroke outcome, no statistically significant association between TyG-BMI and stroke risk was observed in any subgroup, and no significant interactions were found for age, sex, BMI, or diabetes status (all P for interaction > 0.05; Fig. 5C), suggesting no evident subgroup heterogeneity in the association between TyG-BMI and stroke risk.

4.Discussion

In this nationwide prospective cohort study based on the CHARLS database, we systematically evaluated the association between TyG-BMI and future cardiovascular risk among middle-aged and older adults with hypertension, and further explored its predictive value using machine learning methods. The main findings of this study were as follows: (1) higher TyG-BMI levels were independently associated with increased risks of CVD and heart disease; (2) restricted cubic spline (RCS) analysis showed that the associations between TyG-BMI and cardiovascular outcomes were predominantly linear; (3) compared with other TyG-related indices, incorporating TyG-BMI into the basic model provided the greatest incremental predictive value; and (4) machine learning models incorporating TyG-BMI demonstrated good discriminative ability,

calibration performance, and clinical utility in predicting cardiovascular outcomes.Overall, this study suggests that TyG-BMI is not only an important surrogate indicator reflecting insulin resistance status, but may also serve as an effective tool for cardiovascular risk stratification and early identification of high-risk individuals among middle-aged and older adults with hypertension.

Hypertension is one of the most important modifiable risk factors for cardiovascular disease and is often accompanied by multiple cardiometabolic risk factors, such as obesity, dysglycemia, lipid metabolism disorders, and chronic low-grade inflammation [32]. In recent years, increasing evidence has indicated that insulin resistance plays a critical role in the development and progression of hypertension and cardiovascular diseas[e] [33,34]. As a composite index composed of triglycerides, fasting plasma glucose, and body mass index, TyG-BMI can simultaneously reflect abnormalities in glucose metabolism, lipid metabolism disorders, and obesity status. Therefore, it is considered a relatively comprehensive and practical surrogate indicator for assessing insulin resistance [7].

From a pathophysiological perspective, higher TyG-BMI levels may promote the development of cardiovascular disease through multiple pathways. First, insulin resistance can impair insulin-mediated nitric oxide production and endothelium-dependent vasodilation, leading to vascular endothelial dysfunction and impaired microvascular perfusion [33]. Second, insulin resistance can promote oxidative stress and chronic inflammatory responses, induce inflammatory cell infiltration and the release of inflammatory cytokines, thereby accelerating the development of atherosclerosis [9].

In addition, insulin resistance can activate the sympathetic nervous system and the renin–angiotensin–aldosterone system, further contributing to elevated blood pressure, vascular remodeling, and increased arterial stiffness[35].Meanwhile, obesity, particularly visceral fat accumulation, can secrete large amounts of pro-inflammatory adipocytokines, aggravating systemic inflammation and metabolic disorders and further promoting cardiovascular injury [36]. Therefore, the metabolic abnormalities and obesity-related information integrated by TyG-BMI may more comprehensively reflect the pathological processes underlying the development and progression of cardiovascular disease than glucose or lipid indicators alone, thereby demonstrating stronger risk prediction ability.

The findings of the present study are generally consistent with previous studies. Tang et al. [18] found in a nationwide cohort study that TyG-BMI was significantly and positively associated with the risk of incident heart disease. Li et al. [19], in a study based on the CHARLS database among individuals with cardiovascular-kidney-metabolic (CKM) syndrome, also reported that higher TyG-BMI levels were independently associated with an increased risk of future CVD. In addition, a meta-analysis conducted by Rao et al. [37], which included multiple prospective cohort studies, showed that higher TyG-BMI levels were significantly associated with an increased risk of future CVD, and this association remained stable across different study designs and populations.The present study further validates these findings in a specific high-risk population of middle-aged and older adults with hypertension, suggesting that the association between TyG-BMI and cardiovascular risk has good stability and

generalizability.

Notably, this study found a significant association between TyG-BMI and the risk of heart disease, whereas the association between TyG-BMI and stroke risk did not reach statistical significance in the fully adjusted model. Several factors may explain this finding. First, the relatively limited number of stroke events in this study may have reduced the statistical power to detect a significant association. Second, stroke outcomes in the CHARLS database were based on self-reported physician diagnoses, and ischemic stroke and hemorrhagic stroke could not be further distinguished. Different stroke subtypes may have distinct pathological mechanisms and risk factor profiles, which may weaken the overall association.In addition, among individuals with hypertension, blood pressure itself may be the predominant determinant of stroke occurrence, whereas the independent contribution of metabolic abnormalities reflected by TyG-BMI to stroke risk may be relatively weaker. Therefore, although the association with stroke did not reach statistical significance in this study, the overall direction of the effect remained consistent. Future studies with larger sample sizes and more refined outcome classifications are still needed for further validation.

Restricted cubic spline analysis in this study showed that the associations of TyG-BMI with the risks of CVD and heart disease were predominantly linear, with no obvious nonlinear association observed. This finding is generally consistent with previous studies on TyG-related indices [25,38]. The linear dose–response relationship suggests that cardiovascular risk gradually increases with continuously elevated TyG-BMI levels, without an apparent threshold effect.This means that even among

individuals who have not yet met the traditional diagnostic criteria for obesity or diabetes, a mild increase in TyG-BMI may indicate an increased cardiovascular risk. Therefore, continuous monitoring of changes in TyG-BMI levels in clinical practice may help identify high-risk individuals earlier and facilitate timely intervention.

In addition to traditional epidemiological analyses, this study further compared the predictive performance of several TyG-related indices, including TyG, TyG-BMI, TyG-WC, and TyG-WHtR. The results showed that TyG-BMI had the highest predictive performance. Compared with the basic model, the inclusion of TyG-BMI increased the model AUC from 0.593 to 0.636, with an incremental AUC of 0.043, and the DeLong test indicated that this difference was statistically significant. Meanwhile, NRI and IDI analyses further confirmed that TyG-BMI effectively improved individual risk reclassification and overall model discrimination.In contrast, the improvements in predictive performance provided by TyG, TyG-WC, and TyG-WHtR were relatively limited. This finding suggests that there may be a synergistic effect between the overall obesity information represented by BMI and the metabolic abnormalities reflected by TyG, enabling TyG-BMI to more comprehensively reflect an individual's cardiovascular risk status.

In recent years, machine learning techniques have been widely applied in the field of cardiovascular risk prediction because of their ability to handle complex nonlinear relationships and high-dimensional data [39]. In this study, TyG-BMI was further incorporated into machine learning models for validation. The results showed that models including TyG-BMI demonstrated better discriminative ability, calibration

performance, and clinical net benefit in the testing set.In particular, decision curve analysis showed that models incorporating TyG-BMI achieved higher net benefits across a wide range of threshold probabilities, suggesting good potential for clinical application. Compared with traditional risk assessment models, machine learning approaches can more fully capture complex interactions among risk factors. As an easily accessible and low-cost metabolic indicator, TyG-BMI may provide an important supplement for the future development of individualized cardiovascular risk prediction models.

This study has several strengths. First, it was based on the nationally representative prospective CHARLS cohort, with participants from multiple regions of China, thereby providing good generalizability. Second, this study focused on middle-aged and older adults with hypertension, a high-risk population for cardiovascular disease, which has important practical implications for clinical risk management. Third, this study simultaneously compared the predictive performance of multiple TyG-related indices and systematically evaluated the relative advantages of TyG-BMI.Fourth, by integrating traditional regression analysis, restricted cubic spline analysis, incremental prediction assessment, and machine learning approaches, this study validated the predictive value of TyG-BMI from multiple dimensions, thereby improving the reliability of the study conclusions.

Nevertheless, this study still has several limitations. First, as an observational cohort study based on CHARLS data, causal relationships cannot be definitively inferred. Second, cardiovascular outcomes mainly relied on self-reported physician

diagnoses, which may be subject to recall bias or outcome misclassification. Third, this study only analyzed TyG-BMI measured at baseline and was unable to evaluate the association between dynamic changes in TyG-BMI and cardiovascular risk. Fourth, cardiovascular outcomes were assessed at discrete CHARLS follow-up waves rather than according to exact clinical onset dates. Therefore, the present study focused on cumulative incident cardiovascular events during the available follow-up period using logistic regression and did not explicitly account for individual differences in time to event or follow-up duration. Future studies with more precise event-time information are needed to further validate these findings using time-to-event models. Fifth, the CHARLS database lacks information on stroke subtypes and some imaging data, which limited in-depth analyses of different types of cardiovascular events. Finally, the machine learning models in this study were validated only within the CHARLS cohort. Further validation in independent external populations is still needed before these models can be generalized to broader clinical applications. Future multicenter studies with larger sample sizes and longer follow-up durations should be conducted, incorporating dynamic metabolic indicators and external validation cohorts to further confirm the clinical utility of TyG-BMI in cardiovascular risk prediction.

In conclusion, among middle-aged and older adults with hypertension, higher TyG-BMI levels were independently associated with increased risks of CVD and heart disease. Compared with other TyG-related indices, TyG-BMI showed superior incremental predictive value and improved the performance of machine learning-based cardiovascular risk prediction.These findings suggest that TyG-BMI may serve as a

simple, economical, and clinically valuable biomarker for early cardiovascular risk stratification in individuals with hypertension, providing new insights and a potential tool for early risk identification and precise intervention in this population.

5.Conclusion

This study, based on the nationally representative prospective CHARLS cohort, systematically evaluated the association between TyG-BMI and future cardiovascular disease risk among middle-aged and older adults with hypertension, and further compared the predictive performance of different TyG-related indices in machine learning models. The results showed that higher TyG-BMI levels were independently associated with an increased risk of incident cardiovascular disease, particularly heart disease, and that this association was predominantly linear.Compared with other TyG-related indices, TyG-BMI demonstrated superior incremental predictive value in cardiovascular risk prediction and significantly improved model discrimination and risk stratification performance.

In addition, the TyG-BMI-based machine learning models constructed in this study, particularly tree-based models such as LightGBM, random forest, and XGBoost, demonstrated good discriminative ability, calibration performance, and clinical net benefit in predicting cardiovascular outcomes. This indicates that combining machine learning methods with metabolic indicators may further improve the efficiency of identifying high-risk individuals.These findings support TyG-BMI as a simple, economical, easily accessible metabolic biomarker with high clinical utility, which may be used for early cardiovascular risk assessment and precise risk stratification among

middle-aged and older adults with hypertension.

In summary, this study further confirmed the close association between TyG-BMI and cardiovascular risk. By incorporating TyG-BMI into a machine learning-based prediction framework, this study extends the traditional risk assessment paradigm and provides new insights and a potential tool for early screening and precise prevention of cardiovascular disease among individuals with hypertension. Future studies with larger sample sizes, multicenter designs, and external validation are still needed. Further incorporation of dynamic trajectories of TyG-BMI may help optimize model stability and enhance its clinical generalizability.

References

[1] Update A H A S. 2025 Heart disease and stroke statistics: a report of US and global data from the American Heart Association[J]. Circulation, 2025, 151(8): e41-e660.

[2] Global Burden of Cardiovascular Diseases and Risks 2023 Collaborators. Global, regional, and national burden of cardiovascular diseases and risk factors in 204 countries and territories, 1990-2023[J]. Journal of the American College of Cardiology, 2025, 86(22): 2167-2243.

[3] Mingbo L I U, Xinye H E, Xiaohong Y, et al. Interpretation of report on cardiovascular health and diseases in China 2023[J]. J Clin Cardiol, 2024, 40(8): 599-616.

[4] Chinese Society of Cardiology of Chinese Medical Association, Cardiovascular Disease Prevention and Rehabilitation Committee of Chinese Association of Rehabilitation Medicine, Cardiovascular Disease Committee of Chinese Association of

Gerontology and Geriatrics, et al. Chinese guideline on the primary prevention of cardiovascular diseases[J]. Cardiology Discovery, 2021, 1(02): 70-104.

[5] Zhang Z Y, Xiang G U, Zhe T, et al. Homocysteine, hypertension, and risks of cardiovascular events and all-cause death in the Chinese elderly population: a prospective study[J]. Journal of geriatric cardiology: JGC, 2021, 18(10): 796.

[6] Zhang B, Jiang D, Ma H, et al. Association between triglyceride-glucose index and its obesity indicators with hypertension in postmenopausal women: a cross-sectional study[J]. Frontiers in Nutrition, 2025, 12: 1623697.

[7] Song K, Xu Y, Wu S, et al. Research status of triglyceride glucose-body mass index (TyG-BMI index)[J]. Frontiers in Cardiovascular Medicine, 2025, 12: 1597112.

[8] Horton W B, Love K M, Gregory J M, et al. Metabolic and vascular insulin resistance: partners in the pathogenesis of cardiovascular disease in diabetes[J]. American Journal of Physiology-Heart and Circulatory Physiology, 2025, 328(6): H1218-H1236.

[9] Brie A D, Christodorescu R M, Popescu R, et al. Atherosclerosis and insulin resistance: is there a link between them?[J]. Biomedicines, 2025, 13(6): 1291.

[10] Zhou H, Shi Y, Zhou X. Comparison of nine insulin resistance surrogates for predicting cardiovascular disease: a cohort study[J]. Diabetology & Metabolic Syndrome, 2025, 17(1): 365.

[11] Du T, Yuan G, Zhang M, et al. Clinical usefulness of lipid ratios, visceral adiposity indicators, and the triglycerides and glucose index as risk markers of insulin resistance[J]. Cardiovascular diabetology, 2014, 13(1): 146.

765 [12] Sun Y, Ji H, Sun W, et al. Triglyceride glucose (TyG) index: A promising biomarker
766 for diagnosis and treatment of different diseases[J]. European journal of internal
767 medicine, 2025, 131: 3-14.
768 [13] Er L K, Wu S, Chou H H, et al. Triglyceride glucose-body mass index is a simple
769 and clinically useful surrogate marker for insulin resistance in nondiabetic
770 individuals[J]. PloS one, 2016, 11(3): e0149731.
771 [14] Zheng S, Shi S, Ren X, et al. Triglyceride glucose-waist circumference, a novel
772 and effective predictor of diabetes in first-degree relatives of type 2 diabetes patients:
773 cross-sectional and prospective cohort study[J]. Journal of translational medicine, 2016,
774 14(1): 260.
775 [15] Lim J, Kim J, Koo S H, et al. Comparison of triglyceride glucose index, and related
776 parameters to predict insulin resistance in Korean adults: An analysis of the 2007-2010
777 Korean National Health and Nutrition Examination Survey[J]. PloS one, 2019, 14(3):
778 e0212963.
779 [16] Xie J, Zhang X, Shao H, et al. An affordable approach to classifying type 2 diabetes
780 based on fasting plasma glucose, TyG index and BMI: a retrospective cohort study of
781 NHANES Data from 1988 to 2014[J]. Diabetology & metabolic syndrome, 2022, 14(1):
782 113.
783 [17] Tang X, Zhang K, He R. The association of triglyceride-glucose and triglyceride-
784 glucose related indices with the risk of heart disease in a national cohort study[J].
785 Cardiovascular diabetology, 2025, 24(1): 54.
786 [18] Li W, Shen C, Kong W, et al. Association between the triglyceride glucose-body

mass index and future cardiovascular disease risk in a population with Cardiovascular-Kidney-Metabolic syndrome stage 0–3: a nationwide prospective cohort study[J]. Cardiovascular diabetology, 2024, 23(1): 292.

[19] Qiao Y, Wang Y, Chen C, et al. Association between triglyceride-glucose (TyG) related indices and cardiovascular diseases and mortality among individuals with metabolic dysfunction-associated steatotic liver disease: a cohort study of UK Biobank[J]. Cardiovascular Diabetology, 2025, 24(1): 12.

[20] Zeng L, Zhao L, Wan J, et al. Cumulative average triglyceride glucose-waist height index and incident cardiovascular disease in middle-aged and older adults: A nationwide cohort study from the china health and retirement longitudinal study[J]. Plos one, 2026, 21(2): e0333827.

[21] Nye J. Triglyceride-Glucose Index-Related Parameters Predict Hypertension, CVD Risk[J]. Medical Bag, 2024: NA-NA.

[22] Huang Q, Zou X, Lian Z, et al. Predicting cardiovascular outcomes in Chinese patients with type 2 diabetes by combining risk factor trajectories and machine learning algorithm: a cohort study[J]. Cardiovascular Diabetology, 2025, 24(1): 61.

[23] Oikonomou E K, Khera R. Machine learning in precision diabetes care and cardiovascular risk prediction[J]. Cardiovascular Diabetology, 2023, 22(1): 259.

[24] Dong B, Chen Y, Yang X, et al. Estimated glucose disposal rate outperforms other insulin resistance surrogates in predicting incident cardiovascular diseases in cardiovascular-kidney-metabolic syndrome stages 0–3 and the development of a machine learning prediction model: a nationwide prospective cohort study[J].

Cardiovascular Diabetology, 2025, 24(1): 163.

[25] Zhao Y, Hu Y, Smith J P, et al. Cohort profile: the China health and retirement longitudinal study (CHARLS)[J]. International journal of epidemiology, 2014, 43(1): 61-68.

[26] Liu R, Yang G, Liu Y, et al. Which insulin resistance indices best predict future frailty progression in a cardiovascular-kidney-metabolic syndrome stage 0–3 population: a national prospective cohort and machine learning study[J]. Diabetology & Metabolic Syndrome, 2025, 17(1): 1-16.

[27] Cui C, Qi Y, Song J, et al. Comparison of triglyceride glucose index and modified triglyceride glucose indices in prediction of cardiovascular diseases in middle aged and older Chinese adults[J]. Cardiovascular diabetology, 2024, 23(1): 185.

[28] Tao L C, Xu J, Wang T, et al. Triglyceride-glucose index as a marker in cardiovascular diseases: landscape and limitations[J]. Cardiovascular diabetology, 2022, 21(1): 68.

[29] Flegal K M. Body-mass index and all-cause mortality[J]. The Lancet, 2017, 389(10086): 2284-2285.

[30] D’Agostino Sr R B, Vasan R S, Pencina M J, et al. General cardiovascular risk profile for use in primary care: the Framingham Heart Study[J]. Circulation, 2008, 117(6): 743-753.

[31] Whelton P K, Carey R M, Aronow W S, et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA guideline for the prevention, detection, evaluation, and management of high blood pressure in adults:

a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines[J]. Journal of the American College of Cardiology, 2018, 71(19): e127-e248.

[32] Petramala L, Gigante A, Sarlo F, et al. Relevance of obesity-related organ damage and metabolic syndrome classification in cardiovascular and renal risk stratification in patients with essential hypertension[J]. Frontiers in cardiovascular medicine, 2024, 11: 1369090.

[33] Horton W B, Love K M, Gregory J M, et al. Metabolic and vascular insulin resistance: partners in the pathogenesis of cardiovascular disease in diabetes[J]. American Journal of Physiology-Heart and Circulatory Physiology, 2025, 328(6): H1218-H1236.

[34] Fazio S, Mercurio V, Tibullo L, et al. Insulin resistance/hyperinsulinemia: an important cardiovascular risk factor that has long been underestimated[J]. Frontiers in cardiovascular medicine, 2024, 11: 1380506.

[35] Okazaki R A, Rizvi S H, Lyons R, et al. Vascular dysfunction in diabetes and pharmacotherapeutic opportunities: a focus on endothelial cell health[J]. American Journal of Physiology-Heart and Circulatory Physiology, 2025, 329(3): H705-H718.

[36] Lee M J, Kim J. The pathophysiology of visceral adipose tissues in cardiometabolic diseases[J]. Biochemical pharmacology, 2024, 222: 116116.

[37] Rao X, Xin Z, Yu Q, et al. Triglyceride-glucose-body mass index and the incidence of cardiovascular diseases: a meta-analysis of cohort studies[J]. Cardiovascular diabetology, 2025, 24(1): 34.

[38] Mao Q, Liu N, Xiong T, et al. Comparison of triglyceride-glucose index and triglyceride-glucose related indexes in predicting cardiovascular disease incidence among populations with cardiovascular-kidney-metabolic syndrome stages 0-3: a nationwide prospective cohort study[J]. Frontiers in Endocrinology, 2025, 16: 1562135.

[39] Kasartzian D I, Tsiampalis T. Transforming cardiovascular risk prediction: a review of machine learning and artificial intelligence innovations[J]. Life, 2025, 15(1): 94.

Abbreviations

CVD	Cardiovascular diseases
IR	Insulin resistance
WC	Waist circumference
HbA1c	Glycosylated hemoglobin A1c
TyG	Triglyceride–glucose
TyG-WC	TyG-waist circumference
TyG-BMI	TyG-body mass index
TyG-WHtR	TyG-waist-to-height ratio
CHARLS	China Health and Retirement Longitudinal Study
SBP	Systolic blood pressure
DBP	Diastolic blood pressure
LDL-C	Low density lipoprotein cholesterol
HDL-C	High-density lipoprotein cholesterol
Scr	Serum creatinine
FPG	Fasting plasma glucose
UA	Uric acid
TG	Triglyceride
KNN	K-Nearest Neighbor

LGB	Light Gradient Boosting Machine
RF	Random Forest
XGB	EXtreme Gradient Boosting
LR	Logistic Regression
RCS	Restricted cubic splines
AUC	Area under the ROC curve
DCA	Decision curve analysis
OR	Odds ratio

Declarations

Ethics approval and consent to participate

This study protocol was reviewed and approved by the Ethical Review Committee of Peking University (IRB 00001052-11015), and all participants provided written informed consent at the time of participation.

Clinical trial number

This research was funded by the Natural Science Foundation of Shanxi (grant No. 202203021221196).

Consent for publication

Not applicable.

Data availability

The data supporting the findings of this study are available the CHARLS website (http://charls.pku.edu.cn/en).

Competing interests

The authors declare no competing interests.

Funding

This research was funded by the Natural Science Foundation of Shanxi (grant No. 202203021221196).

Author contributions

Yinglun Wei, Xiaoer Liu, and Zhihao Wang contributed to the study design, data

analysis and interpretation, and drafting of the manuscript.Yunqing Luo, Can Yun, Kecheng Liu, Xiaoyan Zhai and Xinzheng Tang revised the manuscript critically for important intellectual content. All authors approved the final manuscript.

Acknowledgements

We would like to thank the CHARLS project members and every respondent for the time and effort that they have devoted to the CHARLS project.

Author details

[1]Shenzhen Hospital (Futian) of Guangzhou University of Chinese Medicine, Guangzhou University of Chinese Medicine,Shenzhen 518034, China

[2] Acupuncture Anesthesia Clinical Research Institute, Yueyang Hospital of Integrated Traditional Chinese and Western Medicine, Shanghai University of Traditional Chinese Medicine, Shanghai, 200437, China

[3]Shanxi Jiaoke New Materials Science and Technology Co., Ltd，Shanxi Transportation Research Institute Group Co., Ltd, Taiyuan,030032, China

[4]Heilongjiang University of Chinese Medicine, Harbin 150040, China

[5]Guangdong Provincial Hospital of Chinese Medicine, Zhuhai Hospital, Zhuhai, 519000, China

[6]School of Basic Medicine, Shanxi University of Chinese Medicine, / Key Research Laboratory for Treating Multiple Sclerosis by Strengthening Qi and Promoting Blood Circulation, National Administration of Traditional Chinese Medicine, Jinzhong 030619, China