Cardiovascular Diabetology

RESEARCH

Open Access

Predicting cardiovascular outcomes in Chinese patients with type 2 diabetes by combining risk factor trajectories and machine learning algorithm: a cohort study

Qi Huang[1†], Xiantong Zou[1†], Zhouhui Lian[2], Xianghai Zhou[1], Xueyao Han[1], Yingying Luo[1†], Shuohua Chen[3], Yanxiu Wang[3], Shouling Wu[3*] and Linong Ji[1*]

Abstract

Background Cardiovascular complications are major concerns for Chinese patients with type 2 diabetes. Accurately predicting these risks remains challenging due to limitations in traditional risk models. We aimed to develop a dynamic prediction model using machine learning and longitudinal trajectories of cardiovascular risk factors to improve prediction accuracy.

Methods We included 16,378 patients from the Kailuan cohort, splitting them into training and testing datasets. Using baseline characteristics and changes over a four-year observation period, we developed the ML-CVD-C (Machine Learning Cardiovascular Disease in Chinese) score to predict 10-year cardiovascular risk, including cardiovascular death, nonfatal myocardial infarction, and stroke. We compared the discrimination and calibration of ML-CVD-C with models using only baseline variables (ML-CVD-C [base]), China-PAR (Prediction for ASCVD Risk in China), and PREVENT (Predict Risk of cardiovascular disease EVENTs). Risk stratification improvements were assessed through net reclassification improvement (NRI) and integrated discrimination improvement (IDI). Transition analysis examined the changes in risk stratification over time.

Results The ML-CVD-C score achieved a C-index of 0.80 (95% CI: 0.78–0.82) in the testing cohort, significantly outperforming the ML-CVD-C (base) score, China-PAR, and PREVENT, which had C-index values of 0.62–0.65. ML-CVD-C also provided more accurate cardiovascular risk estimates, though all models tended to overestimate the prevalence of high-risk cases. Stratification by the ML-CVD-C score showed substantial improvement, with NRI

[†]Qi Huang and Xiantong Zou contributed equally to this work.

*Correspondence:
Shouling Wu
drwusl@163.com
Linong Ji
jiln@bjmu.edu.cn

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

gains of 57.7%, 44.1%, and 47.3%, and IDI gains of 10.1%, 7.9%, and 8.4% compared to the other three scores. Both the trajectory and machine learning algorithm contributed significantly to the enhancement of model performance. Transition analysis revealed that participants who remained in the same risk category or were reclassified to a lower category exhibited 22% and 86% reductions in cardiovascular risk compared to those reclassified to a higher risk category during the observation period.

Conclusions The ML-CVD-C model, incorporating dynamic cardiovascular risk trajectories and a machine learning algorithm, significantly improves risk prediction accuracy for Chinese patients with diabetes. This model may serve as a valuable tool for more personalized cardiovascular risk management in type 2 diabetes.

Keywords Machine learning, Risk prediction, Type 2 diabetes, Trajectories

Background

Type 2 diabetes mellitus (T2DM) is increasingly prevalent in China, posing significant challenges for effective management [1]. Cardiovascular disease remains the leading cause of morbidity and mortality in patients with T2DM [2], and accurate risk assessment is crucial for targeted interventions to prevent cardiovascular events. Current guidelines use various prediction models to estimate the 10-year cardiovascular risk, such as the PREVENT (Predict Risk of cardiovascular disease EVENTs) [3] in the United States and the China-PAR (Prediction for ASCVD Risk in China) score in China [4]. For patients with T2DM, the European Society of Cardiology (ESC) guidelines suggest using tools like SCORE-2D [5] (Systematic Coronary Risk Evaluation 2 for diabetes) and ADVANCE [6] to estimate cardiovascular risk. However, these models often lack the specificity required for accurately predicting risk among Chinese patients with T2DM.

While recent advances in machine learning have led to the development of population-specific cardiovascular risk models, these models have shown only modest improvements in predictive performance [7]. One reason was that traditional models generally assess risk based on a snapshot of baseline factors, which often fails to capture the evolving risk profile. Emerging evidence indicates that longitudinal risk factor changes provide insights into cardiovascular outcomes in T2DM [8]. Previous studies have shown that cumulative systolic blood pressure (SBP) load is a more acceptable predictor of cardiovascular events compared to traditional BP measures among patients with T2DM [9]. Additionally, post-treatment changes in blood glucose mediate the cardiovascular protective effects of sodium-glucose co-transporter-2 (SGLT2) inhibitors and glucagon-like peptide-1 agonists (GLP-1 RA) [10, 11].

We hypothesize that cardiovascular risk evolves dynamically with the changes of mediators, so that models that incorporate both baseline variables and their trajectories would perform better than those using baseline variables alone. Building on this rationale, our study leverages machine learning to develop and validate a novel dynamic cardiovascular risk prediction model in Chinese patients with T2DM (ML-CVD-C). The ML-CVD-C model integrates longitudinal trajectories of multiple risk factors over time, providing a dynamic evaluation of cardiovascular risk in a real-world setting.

Method

Study design and population

The Kailuan Study is a prospective cohort study conducted in Tangshan, China, registered with the Chinese Clinical Trial Registry (ChiCTR-TNC-11001489) [12]. From June 2006 to December 2021, employees and retirees of the Kailuan Group underwent biennial health examinations and completed questionnaires at the Kailuan General Hospital and its affiliated hospitals. Detailed study design and methodology have been described in previous publications [12]. In this study, we included individuals with diabetes, defined based on self-reported diabetes history, hemoglobin A1c (HbA1c) levels ≥ 6.5%, or fasting glucose ≥ 7.0 mmol/L. A 4-year observational window was established for collecting longitudinal data on risk factor trajectories following cohort enrollment. Individuals who experienced cardiovascular events or were censored during the first four years were excluded to avoid immortal bias. Participants were also excluded if they were under 40 or over 79 years of age, or if they had missing data on age, sex, or glucose levels, and the final cohort comprised 16,378 participants (Fig. 1).

Outcomes assessment

Participants were followed until the occurrence of the primary outcome, the most recent clinical follow-up, or December 31, 2021, whichever came first. The primary outcome was defined as the first occurrence of cardiovascular disease, including nonfatal myocardial infarction, nonfatal stroke, or cardiovascular death (MACE). Outcome data were sourced from biennial questionnaires, municipal social insurance, the Kailuan group social security, and Tangshan medical insurance systems, with identification based on ICD-10 codes.

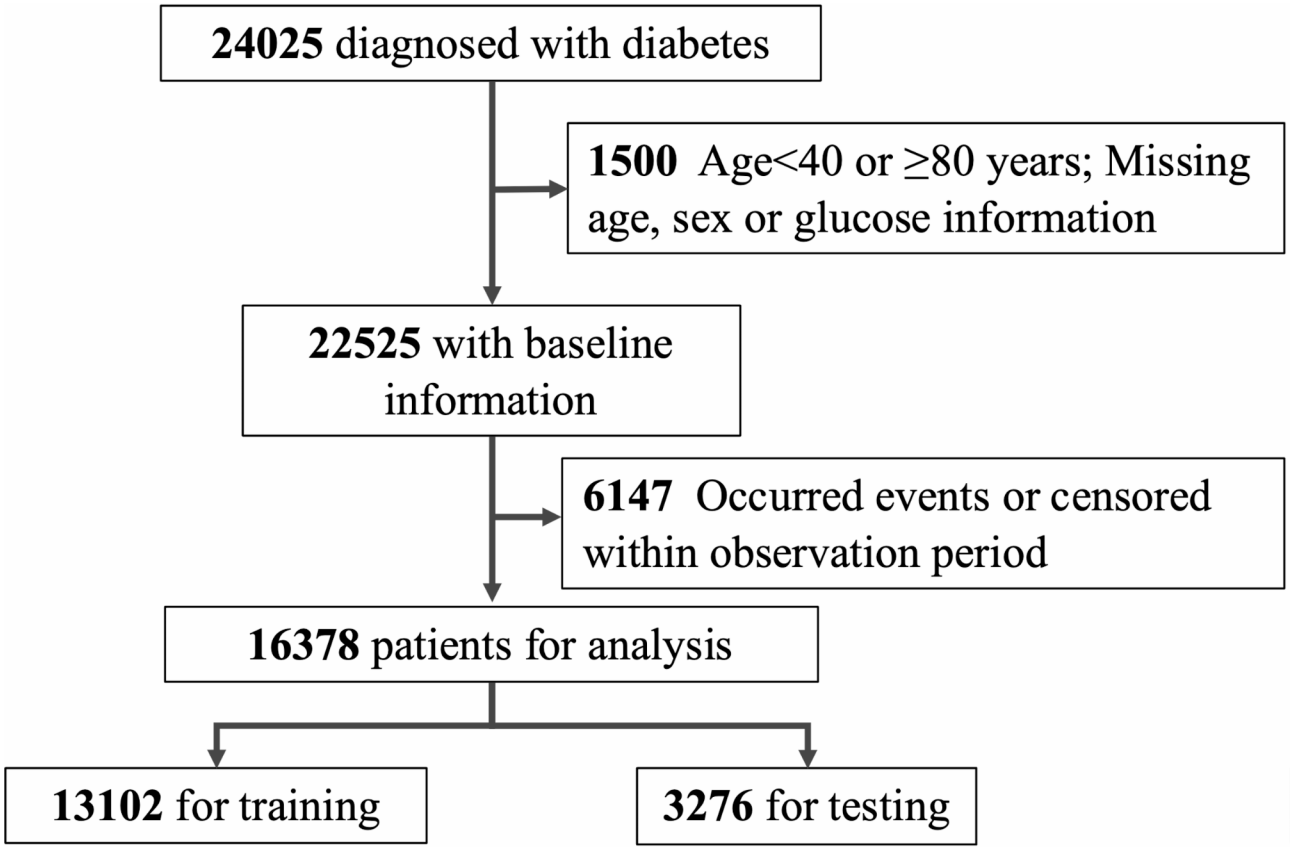

Fig. 1 Flow diagram of the training and test dataset. Participants missing baseline information, experienced cardiovascular events/censored within a 4-year observation window were excluded. The final population for analysis were randomly divided to training and testing sets at an 8:2 ratio

Data collection and variable measurement

Participants were asked to revisit the hospitals to have a face-to-face interview using a standardized questionnaire for collection of the demographic data and medication information. The physical measurement, such as height, weight, and blood pressure were also conducted in the Kailuan General Hospital and affiliated hospitals using the standardized criteria. BMI was calculated as weight (kg) divided by the square of height (m^2). Blood samples were collected following an 8-hour fast to assess fasting glucose, blood creatinine, alanine transaminase, triglycerides, total cholesterol, HDL cholesterol, and LDL cholesterol using the Hitachi 7600 autoanalyzer. The estimated glomerular filtration rate (eGFR) was calculated using the CKD-EPI equation. The left common carotid artery intima-media thickness (IMT) was measured by trained physicians using the VINNO M86 ultrasound machine.

Baseline clinical features related to cardiovascular disease included demographic data, physical measurements, disease history, and laboratory values (Table S1). To manage repeated measurements at visits, we utilized functional principal component (FPC) analysis to reduce dimensionality [13]. This method calculates individual FPC scores that align with the principal trajectories of the analyzed features. To achieve it, discrete data points at uneven time intervals for each feature were firstly smoothed to create continuous curves and decomposed into a mean function, representing the average value at each time point, and a Gaussian kernel function. Then the principal components and their variance contributions were derived using formula:

$$\int K(s,t)\varphi_k(s)\,ds = \lambda_k \varphi_k(t)$$

where $\varphi_k(t)$ is the k-th FPC. The number of FPCs K was chosen to explain over 95% of the total variance. Finally the FPC score of the feature for the individual i could be estimated as:

$$\xi_{ik} = \int [x_i(t) - \mu(t)]\varphi_k(t)\,dt$$

where $x_i(t)$ is the actual value of the individual i. FPC scores quantify how an individual's feature trajectory aligns with the principal components of that feature.

These scores were then integrated with baseline features for subsequent analysis and model development.

Model development

The dataset was randomly divided into derivation and testing cohorts in an 8:2 ratio, stratified by the occurrence of MACE events. Predictors with more than 20% missing data, high collinearity, or low variance were excluded. Missing baseline data were imputed using a random forest-based approach, which outperformed k-nearest neighbor and chained equations imputation methods in handling non-linear data with complex interactions [14]. Variables not normally distributed were log-transformed, and all continuous variables were normalized. Outliers beyond six standard deviations were removed (Fig. S1).

An XGBoost algorithm was developed utilizing the Cox negative log-likelihood as the loss function to predict the primary outcome. 5-fold cross-validation was performed within the derivation cohort to optimize the hyperparameters (Table S2). This final machine learning model for cardiovascular disease in the Chinese population is named ML-CVD-C.

Model assessment

To evaluate whether trajectory information could improve the discrimination of cardiovascular risk, we selected predictors including body mass index (BMI), systolic blood pressure (SBP), diastolic blood pressure (DBP), fasting plasma glucose (FPG), triglycerides (TG), total cholesterol (TC), high-density lipoprotein cholesterol (HDL), low-density lipoprotein cholesterol (LDL), alanine transaminase (ALT), and serum creatinine (SCr). We compared C-index values between a model using only baseline values and one that incorporated both baseline values and FPC scores. Wald's test was used to assess differences between models.

The performance of the ML-CVD-C model was evaluated in terms of discrimination and calibration. Discrimination was assessed using the C-index and the time-dependent area under the receiver operating characteristic curve (AUROC) in the training and testing datasets. Calibration was assessed using calibration plots, the expected–observed ratio, and Hosmer–Lemeshow $\chi 2$ (HL-$\chi 2$) tests. Feature importance within the model was quantified using SHapley Additive exPlanations (SHAP) values.

The ML-CVD-C model was used to calculate the ML-CVD-C score as a value between 0 and 1, which presents the probability of developing cardiovascular disease within the next 10 years. To provide benchmarks, China-PAR and PREVENT scores were used to assess cardiovascular risk, in alignment with the latest guidelines from Chinese and US authorities. The ML-CVD-C model was also employed to calculate the baseline risk using baseline values only, referred to as the ML-CVD-C (baseline) score. Participants were stratified into low (0-4.9%), medium (5-9.9%), high (10-19.9%), and very-high (≥ 20%) risk categories according to the 10-year cardiovascular risk scores calculated using ML-CVD-C, ML-CVD-C (base) score, China-PAR, and PREVENT, in accordance with thresholds commonly used in cardiovascular risk guidelines [15]. Cox regression was used to analyze the association between risk stratification and the primary outcome. Net reclassification improvement (NRI) and integrated discrimination improvement (IDI) for ML-CVD-C were also evaluated.

To assess longitudinal transitions in cardiovascular risk categories, we examined changes in ML-CVD-C risk stratification from baseline to the 4-year follow-up, estimating their association with cardiovascular outcomes, to provide insights into how changes in risk categorization over time correlated with cardiovascular outcomes.

Sensitivity analysis

To evaluate the impact of trajectory observation length on the predictive performance of the model, sensitivity analyses were conducted by adjusting the trajectory observation window in both the training and testing datasets. The observation windows varied from 0 to 6 years, and scores were recalibrated for each window to examine potential changes in prediction accuracy. The ML-CVD-C model was also compared to a traditional Cox regression model using the same variables to allow for a direct comparison between traditional statistical approaches and machine learning techniques within the same dataset and variable context.

Statistics analysis

Continuous variables were reported as either means ± standard deviations (SD) or medians [interquartile ranges (IQR)], while categorical variables were shown as counts (percentages). All statistical analyses were conducted using R software, version 4.2.2 (R Foundation for Statistical Computing, Vienna, Austria). The 'MissForest' package facilitated baseline data imputation, the 'fdapace' and 'longitudinalData' packages were employed for FPC analysis of the longitudinal data, and the 'XGBoost' package was utilized to build the XGBoost model. We performed all statistical tests two-sided, with a P-value of < 0.05 deemed indicative of statistical significance.

Results

There were 16,378 participants included in the final analysis, with a mean age of 55.3 years, and 4.2% ($N = 687$) had a history of cardiovascular disease at baseline. Over a mean follow-up of 8.5 years, there were 2,164 (13.2%) incident of primary outcome (Table 1).

Table 1 Characteristics of the training and testing sets

	Total	Training	Testing
Participants—no.	16,378	13,102	3276
MACE—no. (%)[#]	2164 (13.2)	1731 (13.2)	433 (13.2)
Age—yr	55.3 (8.7)	55. 3 (8.7)	55.2 (8.7)
Female—no. (%)	3078 (18.8)	2483 (19.0)	595 (18.2)
Current smoke—%	6633 (40.5)	5300 (40.5)	1333 (40.7)
Hypertension—no. (%)	6542 (39.9)	5193 (39.6)	1349 (41.2)
Hyperlipemia—no. (%)	4643 (28.3)	3657 (27.9)	986 (30.1)
History of cardiovascular disease—no. (%)	687 (4.2)	546 (4.2)	141 (4.3)
Antihypertensive drugs—no. (%)	4336 (26.5)	3425 (26.1)	911 (27.8)
Lipid-lowering drugs—no. (%)	2092 (12.8)	1610 (12.3)	482 (14.7)
Height—cm	167.9 (6.9)	167.8 (6.9)	168.0 (6.8)
Body weight—kg	74.0 (11.0)	74.0 (11.1)	74.2 (11.0)
BMI—kg/m^2	26.3 (3.4)	26.2 (3.4)	26.3 (3.3)
SBP—mmHg	138.7 (20.2)	138.7 (20.1)	138.9 (20.5)
DBP—mmHg	86.7 (11.3)	86.8 (11.3)	86.9 (11.5)
Heart rate—bpm	76.7 (10.9)	76.5 (10.9)	76.5 (10.7)
ALT—U/L	25.8 (23.3)	25.9 (23.3)	25.5 (23.3)
FPG—mmol/L	8.8 (4.1)	8.8 (4.1)	8.8 (4.2)
SCr -mmol/L	88.2 (41.2)	88.2 (42.5)	88.1 (35.6)
eGFR—ml/min/1.73m^2	83.7 (20.3)	83.8 (20.4)	83.6 (20.1)
Median TG—mmol/L	1.7 [1.1, 2.6]	1.7 [1.1, 2.6]	1.7 [1.1, 2.7]
Median TC—mmol/L	5.2 [4.6, 6.0]	5.2 [4.6, 6.0]	5.2 [4.6, 6.0]
Median HDL—mmol/L	1.4 [1.2, 1.7]	1.4 [1.2, 1.7]	1.4 [1.2, 1.7]
Median LDL—mmol/L	2.6 [2.1, 3.2]	2.6 [2.1, 3.2]	2.6 [2.0, 3.2]
Mean risk score—%			
ML-CVD-C	12.2 (10.9)	12.0 (10.7)	12.1 (10.9)
China-PAR	17.8 (11.5)	17.7 (11.5)	17.8 (11.5)
PREVENT	14.7 (8.0)	14.7 (8.1)	14.7 (8.0)
Base	18.3 (10.8)	18.2 (10.9)	18.3 (10.8)

Data are represented as mean (standard deviation) or median [interquartile range] for continuous variables or N (%) for categorical variables

[#] MACE included non-fatal myocardial infraction, non-fatal stroke, death from cardiovascular cause

BMI, body mass index. SBP, systolic blood pressure. DBP, diastolic blood pressure. eGFR, estimated glomerular filtration rate. ALT, alanine transaminase. FPG, fasting plasma glucose. TG, triglycerides. TC, total cholesterol. HDL, high-density lipoprotein cholesterol. LDL, low-density lipoprotein cholesterol. SCr, serum creatinine

The performance of cardiovascular risk scores

In the overall analysis, we observed that incorporating trajectory information into the model yielded moderate improvements in predictive performance compared to models relying solely on baseline data (Fig. S2). The ML-CVD-C model, developed using machine learning to incorporate multiple risk factor trajectories, demonstrated consistently strong predictive performance for cardiovascular events across various time frames in both the training and testing datasets (Figs. S3, 2A). The C-index of the ML-CVD-C risk score was 0.86 (0.85, 0.87) in the training set and 0.80 (0.78, 0.82) in the testing set (Table S3). In contrast, the China-PAR, PREVENT, ML-CVD-C (base) risk scores had C-index values of 0.62 (0.59, 0.65), 0.64 (0.61, 0.67) and 0.65 (0.62, 0.68), respectively, in the testing cohort. In terms of calibration (Fig. 2B–D), while the calibration plot indicated that participants at higher CVD risk were more likely to be overestimated by models, the ML-CVD-C, China-PAR, and PREVENT, ML-CVD-C (base) risk scores exhibited expected-observed ratios of 1.30, 1.92, 1.59 and 1.97, respectively. The ML-CVD-C risk score performed slightly better than the others.

Analysis of SHAP values indicated that among the top ten contributing variables, age and sex were the most influential baseline factors. For trajectory risk factors, the first principal component (FPC1) represents the primary component, capturing sustained exposure to the factors. The second (FPC2) and third (FPC3) components mainly reflect fluctuations in these factors or trend changes of a variable. FPC3 of ALT and FPC1 of fasting glucose and systolic blood pressure were the most significant contributors (Fig. 3).

Risk stratification with cardiovascular scores

The ML-CVD-C score was used to estimate the 10-year cardiovascular risk, allowing participants to be stratified into low, medium, high, and very high-risk categories. As expected, participants in higher risk categories were generally older, had a higher proportion of males, a greater prevalence of cardiovascular disease (CVD) history, and exhibited less favorable profiles in blood pressure, glucose, and lipid levels (Table S4). MACE incidence was highest in the very high-risk group, with an occurrence rate of 18.2%. In a subset of 1,437 participants who underwent follow-up measurements for common carotid artery intima-media thickness (IMT), we observed a significant increase in IMT in categories with higher risks ($P < 0.001$, Table S4).

The hazard ratios for MACE in participants in the medium-risk, high-risk, and very high-risk groups defined by ML-CVD-C were 3.97 [2.60, 6.07], 9.47 [6.32, 14.18], and 21.3 [14.2, 32.1], respectively, in reference to participants with low risk (Table S5). In contrast, the risk groups defined by China-PAR showed significantly elevated HRs for MACE only in the very high-risk group (HR 2.62 [1.66, 4.12] versus low-risk group). Although risk groups defined by both PREVENT and ML-CVD-C (base) scores exhibited a linear increase in hazard ratios for MACE, the effect sizes were smaller. Reclassification analyses demonstrated that ML-CVD-C risk stratification notably improved classification accuracy, with a net reclassification improvement (NRI) of approximately 50%

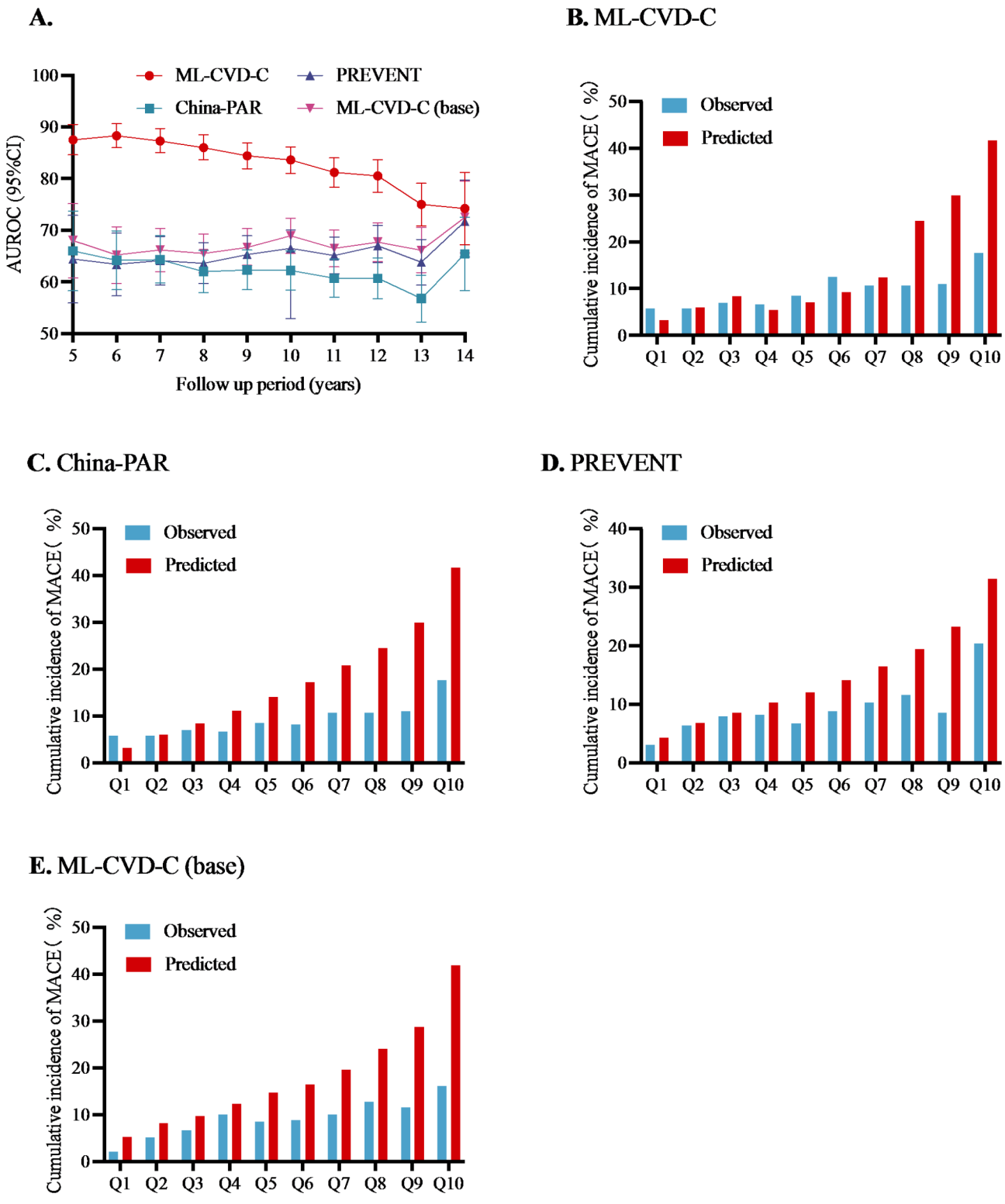

Fig. 2 Time-dependent AUROC and calibration curves for cardiovascular scores in the testing set. **A** The area under the receiver operating characteristic curve (AUROC) with 95% confidence intervals (CI) of cardiovascular score for predicting MACE in various follow-up time. **B**–**E** The calibration curve of the ML-CVD-C, China-PAR, PREVENT and ML-CVD-C (base) scores for 10-year cardiovascular risk. The plot shows the predicted vs. observed incidence of MACE based on deciles of predicted risk. Q1 to Q10 refer to deciles of risk score. MACE, major adverse cardiovascular outcome

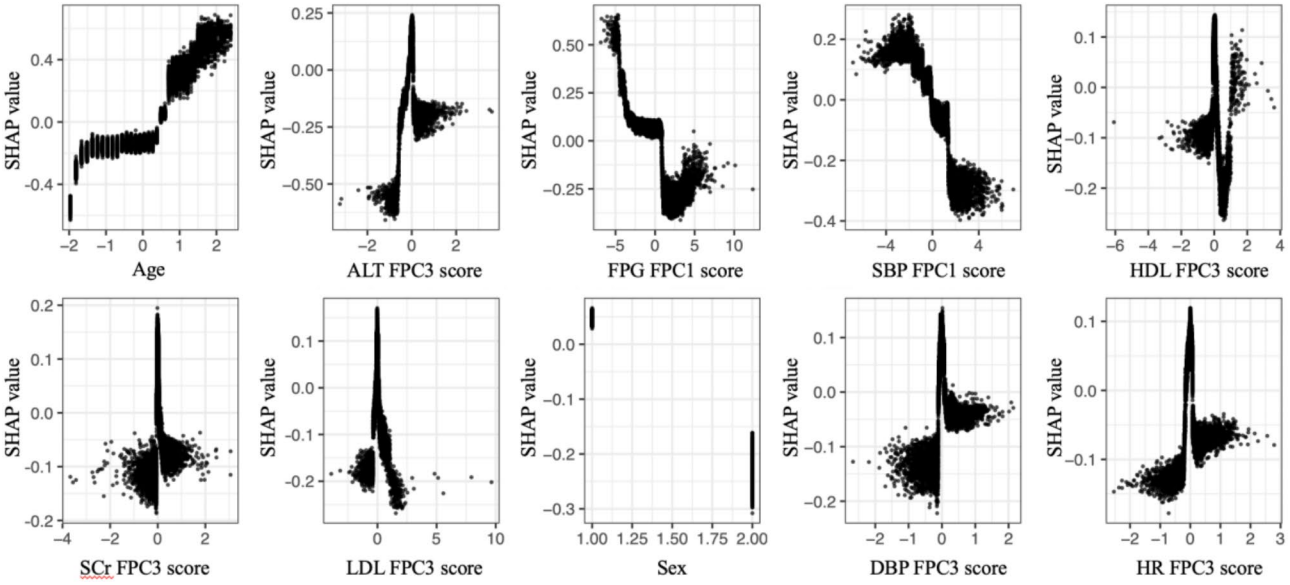

Fig. 3 Dependency plots of the top 10 important features in the ML-CVD-C quantified by SHapley Additive exPlanation (SHAP) values. These features are age; the third functional principal component (FPC3) score of alanine transaminase (ALT); the first FPC score (FPC1) of fasting plasma glucose (FPG); the FPC1 score of systolic blood pressure (SBP); the third FPC score of high-density lipoprotein cholesterol (HDL); the FPC3 score of serum creatinine (SCr); the FPC3 score of low-density lipoprotein cholesterol (LDL); sex; the FPC3 score of diastolic blood pressure (DBP); the FPC3 score of heart rate (HR). Clinically, the FPC1 score indicates the long-term average level of a variable and the FPC3 score represents the fluctuations, or trend changes of a variable. Categorical features are binary-encoded, while continuous features have been scaled and centered around the mean for the analysis. Each plot point corresponds to an individual data observation. The x-axis represents the deviation of feature values from the mean, and the y-axis depicts the SHAP value for each feature within the training dataset

Table 2 Reclassification improvement by ML-CVD-C

	ML-CVD-C versus China-PAR	ML-CVD-C versus ML-CVD-C base	ML-CVD-C versus PREVENT
NRI overall (95% CI)	57.7% (42.9, 77.1)	44.1% (29.9, 61.9)	47.3% (29.3, 61.2)
NRI cases (95% CI)	44.2% (33.7, 61.0)	35.9% (24.7, 52.8)	39.9% (27.6, 51.0)
NRI non-cases (95% CI)	13.5% (2.6, 24.9)	8.3% (0.3, 19.1)	7.4% (-1.4, 17.0)
IDI (95% CI)	10.1% (7.5, 12.7)	7.9% (5.6, 10.3)	8.4% (5.7, 10.9)

NRI, net reclassification improvement. IDI, integrated discrimination improvement

and an integrated discrimination improvement (IDI) of around 10% compared to other scores (Table 2).

The transition of cardiovascular risk assessed by ML-CVD-C

The ML-CVD-C dynamic prediction model allowed for cardiovascular risk estimation at both baseline and the 4-year follow-up, enabling the analysis of transitions in risk categories over the observation period. Participants initially categorized as low-risk or very high-risk were more likely to remain within their original categories after 4 years. Conversely, participants in the medium and high-risk groups demonstrated greater transitions across risk categories. Notably, nearly two-thirds of participants initially at high risk and about half of those at very high risk were reclassified into a lower category over time (Fig. 4). Participants who remained in the same risk category or were reclassified into a lower category exhibited significantly lower hazard ratios for cardiovascular events compared to those who were reclassified to a higher risk category over the observation period (HR: 0.78 [95% CI: 0.66, 0.92] and 0.14 [0.11, 0.16], respectively).

Sensitivity analysis

The sensitivity analysis assessed how varying the trajectory observation window impacted the model's performance, and the results showed that extending the observational window led to an increase in the C-index in both training and testing cohorts (Fig. S4). The application of machine learning techniques significantly improved predictive performance compared to traditional models, with the performance gap widening as the observation period increased.

Discussion

In this study, we developed a machine learning-based cardiovascular disease model (ML-CVD-C) for individuals with T2DM in the Chinese population. By incorporating trajectory analysis and machine learning techniques, our model demonstrated better discrimination and calibration compared to the currently recommended risk scores. Our stratification strategy notably enhanced

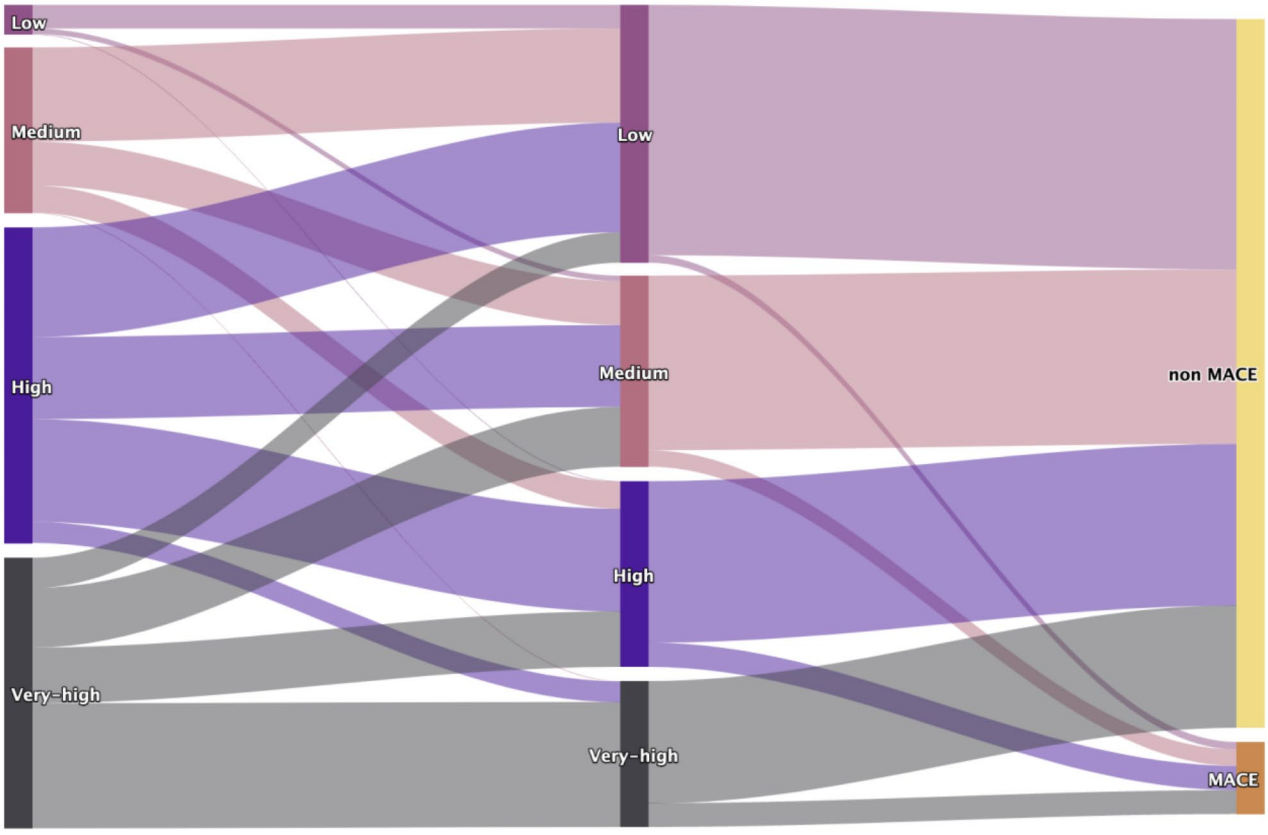

Fig. 4 Sankey diagram visualizing the progression of ML-CVD-C risk score and cardiovascular outcome. Patients used ML-CVD-C model to estimate risk at baseline and re-estimate at year 4. MACE, major adverse cardiovascular event

reclassification performance, providing a valuable tool for precision medicine in patients with T2DM in China.

Predicting cardiovascular risk in T2DM remains challenging. A previous study reported that the 22 existing cardiovascular scores only had moderate predictive ability over a 10-year follow-up period [16]. In our research, we observed that the widely recommended models, China-PAR and PREVENT, achieved a C-index of 0.62–0.65 in the Kailuan cohort, significantly lower than their performance in the general population [3, 4]. Similar findings were reported in a study conducted in Yinzhou district, where the C-index of the China-PAR score for cardiovascular events in participants with T2DM was 0.67 in men and 0.70 in women [17]. A systematic review reported that CVD prediction models for Chinese patients with T2DM had C-index values ranging from 0.70 to 0.85 [18]. A recent model using the XGBoost method based on 141,516 Chinese adults achieved a C-index of 0.74–0.76 in validation sample [19]. Notably, CVD risk prediction models tend to exhibit better discrimination for short-term predictions. One study achieved an AUROC exceeding 0.8 in 3-year predictions [20]. However, the predictive performance often declines over longer estimation periods, with some models showing reduced discrimination when extending the prediction horizon from 10 to 30 years [21]. In our study, the ML-CVD-C score demonstrated superior discrimination within the Kailuan cohort, highlighting its effectiveness for short-term CVD risk prediction. In our study, our ML-CVD-C score demonstrated better discrimination in the Kailuan cohort than other scores. The introduction of trajectory information contributed significantly to this improvement. While previous studies suggested that trajectory-based assessments could enhance diabetes management, these approaches typically relied on summary metrics like means, variability, or cumulative load of repeated measurements [9, 22]. Our approach utilized principal component analysis to generate overall scores that assess the alignment of individual and overall trajectory. Moreover, by leveraging machine learning, we effectively integrated the interplay of multiple complex variables. Our predictive results showed a C-index improvement to 0.80 in the testing cohort, also indicating a robust performance among existing machine-learning based cardiovascular prediction models [7]. Additionally, the ML-CVD-C model enabled better risk stratification in Chinese patients with T2DM. Our study suggested that this stratification significantly improved actual risk

estimation, NRI, and IDI. Notably, the improvement in NRI was more pronounced in populations with MACE, highlighting the enhanced specificity of the ML-CVD-C model.

The ML-CVD-C model explored the potential benefits of incorporating changes observed during patient revisits. The SHAP values indicated that revisited information provided additional predictive value independent of baseline values. The ADVANCE study highlighted that after adjustment for baseline SBP, sustained pressure elevation significantly increases the risk of major cardiovascular events, particularly cardiovascular death and myocardial infarction [9]. The Chinese Longitudinal Healthy Longevity Survey found a strong link between long-term systolic blood pressure load and cardiovascular mortality among the elderly [23]. Long-term glucose control also affected cardiovascular outcomes, with one study suggesting that personal glucose level control according to glucose variability was crucial for patients with T2DM [24]. This emphasizes the need for regular management for cardiovascular risk factors in patients with diabetes [25]. The ML-CVD-C model explanation indicates that fluctuations in ALT, creatinine, and LDL levels contribute to the model, while sustained elevations in SBP and FPG over four years were important factors predicting cardiovascular outcomes. This finding suggests that different monitoring strategies for these risk factors may improve cardiovascular disease management.

Follow-up of ML-CVD-C risk assessments revealed that many individuals initially classified needed reclassification upon subsequent evaluations. A recent study suggested the evolution of diabetes screening and management strategies has radically improved the cardiovascular risk profile of people with diabetes [26]. Reestimating cardiovascular risk over time in patients is crucial for understanding the progression of diabetes and for adjusting treatment strategies to mitigate cardiovascular complications effectively. High-risk individuals may not remain at elevated risk upon reevaluation with favorable control of cardiovascular risk factors [27]. Moreover, the cardiovascular protective effects of new-generation glucose-lowering medications such as SGLT2 inhibitors also reduced cardiovascular risk among patients at higher baseline risk [28, 29]. The ML-CVD-C scores, derived from multiple diabetes-related trajectory predictors, would be a better tool for capturing changes in cardiovascular risk profiles in an increasingly heterogeneous population. This model allowed for dynamic risk reclassification during follow-ups, paving the way for a shift in clinical practice from single-point predictions to continuous, multi-point risk surveillance.

Our study utilized a large population-based cohort to demonstrate that dynamic machine learning models that incorporate multiple interim factors during the disease course may lead to more accurate prognostic assessments in Chinese patients with T2DM. Despite the advancement, we observed that our ML-CVD-C score still tended to overestimate the actual cardiovascular risk, particularly in patients at higher risk, although it demonstrated better calibration than the other scores. Previous studies have reported inconsistent results regarding the calibration of cardiovascular risk models in the Chinese population [15, 30]: A study in Yinzhou suggested an underestimation of risk using cardiovascular scores, while another study in northern China indicated an overestimation. In our study, we observed a higher prevalence of medication use in individuals at higher risk. We postulated that a plausible explanation for the overestimation is that individuals at higher risk are more likely to receive intensive medical intervention, which subsequently reduces their risk. Information on diabetes medication and adherence to treatment regimens would be needed for better calibration, but such data were scarce in our study. Incorporating detailed treatment and medication data into future models could further enhance the accuracy and reliability of cardiovascular risk predictions in patients with T2DM. Interestingly, the low-risk group demonstrated a higher prevalence of self-reported hyperlipidemia but had lower levels of triglycerides, total cholesterol, and LDL cholesterol, with no significant differences in the use of lipid-lowering drugs. The self-reported diagnosis of hyperlipidemia in our study may explain this observation. In China, hypertriglyceridemia is more prevalent than hypercholesterolemia [31]. It is possible that low-risk individuals likely presenting primarily with hypertriglyceridemia, while high-risk individuals are more affected by hypercholesterolemia, which is more detrimental to cardiovascular health. Greater awareness and early detection of hyperlipidemia in the low-risk group may have facilitated timely lifestyle changes and improved lipid profiles in these patients. These observations underscore the limitations of the ML-CVD-C (base) model in capturing the dynamic progression of T2DM, which might reduce the reliability of risk stratification and the significance of transition analysis. Nonetheless, ML-CVD-C still provides a way to identify patients with prolonged high-risk exposure.

Limitations

This study has other limitations. Firstly, the validation of the models was conducted solely within the Kailuan cohort, necessitating further validation in other cohorts to confirm the findings. Additionally, the lack of regular check-ups and long-term follow-ups for T2DM in other domestic cohorts restricted our ability to validate the efficacy of our model more broadly. Secondly, the absence of detailed data on some advanced cardiovascular predictors, such as the urine albumin-creatinine ratio,

potentially impacted the predictive ability. Thirdly, the lack of follow-up data and more missing variables presented challenges in applying the ML-CVD-C model in real-world contexts. Finally, lack of direct data on treatment adherence and medication use limited the ability to assess their impact on calibration and prediction. The effect of new-generation glucose-lowering medications on cardiovascular risk reduction was not explicitly considered. Whether the application of the ML-CVD-C model can help physicians design personalized treatment strategies in current clinical practice also needs further investigation.

Future directions

To address these limitations, our future research on the machine learning model will focus on validating its performance across multiple cohorts with diverse demographic and clinical characteristics. We will explore whether integrating additional dynamic biomarkers can further improve predictive accuracy. Additionally, it is worthwhile to explore the model's potential for monitoring changes in cardiovascular disease risk in patients receiving medications such as SGLT2 inhibitors and GLP-1 receptor agonists, which are known to improve cardiovascular outcomes.

Conclusions

The utilization of the machine learning-based cardiovascular dynamic prediction model ML-CVD-C represents a significant advancement in the predictive capabilities for cardiovascular diseases in T2DM patients within the Chinese population. This model enables accurate assessment of cardiovascular outcomes, showcasing its potential for the precise stratification of patients with T2DM in China.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12933-025-02611-0 .

Supplementary Material 1

Author contributions

Q.H., X.Zo., L.J. and S.W. designed the study. Q.H. and X.Zo. wrote and revised the manuscript. Q.H. conducted the study and analysis; Z.L., X.Zh., X.H., Y.L.,S.C., Y.W, S.W., and L.J. were involved in the data collection, conception, and the interpretation of the results. All authors reviewed and approved the final version of the manuscript.

Funding

Our research was supported by the National Natural Science Foundation of China (T2341011 to X.Zo.), Beijing Nova Cross program (Z211100002121169) (to X.Zo.) and Special Fund for Young Scientists of National Science and Technology Major Project of China (2023ZD0508801) (to X. Zo).

Data availability

The data in our study could be made available upon request to the corresponding authors.

Declarations

Ethics approval and consent to participate

The study protocol was approved by the Kailuan General Hospital Ethics Committee, and all participants provided informed consent in adherence to the Declaration of Helsinki.

Conflict of interest

The authors declare no conflict of interest

Author details

[1]Department of Endocrinology and Metabolism, Peking University People's Hospital, Beijing 100044, China
[2]Wangxuan Institute of Computer Technology (WICT), Peking University, Beijing 100044, China
[3]Department of Cardiology, Kailuan General Hospital, Tangshan, China

Received: 12 November 2024 / Accepted: 23 January 2025
Published online: 07 February 2025

References

1. Wang L, Peng W, Zhao Z, Zhang M, Shi Z, Song Z et al. Prevalence and treatment of diabetes in China, 2013–2018. JAMA. 2021;326(24):2498–506. https://doi.org/10.1001/jama.2021.22208
2. Deshpande AD, Harris-Hayes M, Schootman M. Epidemiology of diabetes and diabetes-related complications. Phys Therapy. 2008;88(11):1254–64.
3. Khan SS, Matsushita K, Sang Y, Ballew SH, Grams ME, Surapaneni A et al. Development and Validation of the American Heart Association's PREVENT equations. Circulation. 2024;149(6):430–49. https://doi.org/10.1161/CIRCULATIONAHA.123.067626
4. Yang X, Li J, Hu D, Chen J, Li Y, Huang J, et al. Predicting the 10-year risks of atherosclerotic cardiovascular disease in Chinese population: the China-PAR project (prediction for ASCVD risk in China). Circulation. 2016;134(19):1430–40.
5. SCORE2-Diabetes Working Group and the ESC Cardiovascular Risk Collaboration, Pennells L, Kaptoge S, Østergaard HB, Read SH, Carinci F et al. SCORE2-Diabetes: 10-year cardiovascular risk estimation in type 2 diabetes in Europe. Eur Heart J. 2023;44(28):2544–56.
6. Kengne AP, Patel A, Marre M, Travert F, Lievre M, Zoungas S et al. Contemporary model for cardiovascular risk prediction in people with type 2 diabetes. Eur J Cardiovasc Prev Rehabil. 2011;18(3):393–8.
7. Wang M, Francis F, Kunz H, Zhang X, Wan C, Liu Y et al. Artificial intelligence models for predicting cardiovascular diseases in people with type 2 diabetes: a systematic review. Intell Based Med. 2022;6:100072.
8. Lindbohm JV, Sipilä PN, Mars N, Knüppel A, Pentti J, Nyberg ST et al. Association between change in cardiovascular risk scores and future cardiovascular disease: analyses of data from the Whitehall II longitudinal, prospective cohort study. Lancet Digital Health. 2021;3(7):e434–44.
9. Wang N, Harris K, Hamet P, Harrap S, Mancia G, Poulter N, et al. Cumulative systolic blood pressure load and cardiovascular risk in patients with diabetes. J Am Coll Cardiol. 2022;80(12):1147–55.
10. Buse JB, Bain SC, Mann JFE, Nauck MA, Nissen SE, Pocock S, et al. Cardiovascular risk reduction with liraglutide: an exploratory mediation analysis of the leader trial. Diabetes Care. 2020;43(7):1546–52.
11. Oshima M, Neuen BL, Li J, Perkovic V, Charytan DM, De Zeeuw D et al. Early change in albuminuria with canagliflozin predicts kidney and cardiovascular outcomes: a posthoc analysis from the CREDENCE trial. JASN. 2020;31(12):2925–36. https://doi.org/10.1681/ASN.2020050723
12. Wu S, Huang Z, Yang X, Zhou Y, Wang A, Chen L, et al. Prevalence of ideal cardiovascular health and its relationship with the 4-year cardiovascular events in a northern Chinese industrial city. Circ Cardiovasc Qual Outcomes. 2012;5(4):487–93.
13. Yao F, Müller HG, Wang JL. Functional data analysis for sparse longitudinal data. J Am Stat Assoc. 2005;100(470):577–90.
14. Stekhoven DJ, Bühlmann P. MissForest—non-parametric missing value imputation for mixed-type data. Bioinformatics. 2012;28(1):112–8.
15. Chen W, Liu X, Shen P, Chen Q, Sun Y, Wu J et al. Accuracy of the China-PAR and WHO risk models in predicting the ten-year risks of cardiovascular disease in the Chinese population. Zhonghua Liu Xing Bing Xue Za Zhi = Zhonghua liuxingbingxue zazhi. 2022;43(8):1275–81.

16. Dziopa K, Asselbergs FW, Gratton J, Chaturvedi N, Schmidt AF. Cardiovascular risk prediction in type 2 diabetes: a comparison of 22 risk scores in primary care settings. Diabetologia. 2022;65(4):644–56.
17. Liang J, Li Q, Fu Z, Liu X, Shen P, Sun Y et al. Validation and comparison of cardiovascular risk prediction equations in Chinese patients with type 2 diabetes. Eur J Prev Cardiol. 2023;zwad198.
18. Dong W, Wan E, Bedford L, Wu T, Wong C, Tang E, et al. Prediction models for the risk of cardiovascular diseases in Chinese patients with type 2 diabetes mellitus: a systematic review. Public Health. 2020;186:144–56.
19. Dong W, Wan EYF, Fong DYT, Tan KCB, Tsui WWS, Hui EMT, et al. Development and validation of 10-year risk prediction models of cardiovascular disease in Chinese type 2 diabetes mellitus patients in primary care using interpretable machine learning-based methods. Diabetes Obes Metab. 2024;26(9):3969–87.
20. Ding J, Luo Y, Shi H, Chen R, Luo S, Yang X, et al. Machine learning for the prediction of atherosclerotic cardiovascular disease during 3-year follow up in Chinese type 2 diabetes mellitus patients. J Diabetes Invest. 2023;14(11):1289–302.
21. Chen F, Wang J, Chen X, Yu L, An Y, Gong Q, et al. Development of models to predict 10-30-year cardiovascular disease risk using the Da Qing IGT and diabetes study. Diabetol Metab Syndr. 2023;15(1):62.
22. Martinez M, Santamarina J, Pavesi A, Musso C, Umpierrez GE. Glycemic variability and cardiovascular disease in patients with type 2 diabetes. BMJ Open Diabetes Res Care. 2021;9(1):e002032.
23. Zhang X, Yang J, Ruan H, Zheng Y, Zhao L, Li L, et al. Association of cumulative blood pressure with cardiovascular mortality in Chinese older people: a longitudinal prospective study. Exp Gerontol. 2022;168:111952.
24. Pei J, Wang X, Pei Z, Hu X. Glycemic control, HbA1c variability, and major cardiovascular adverse outcomes in type 2 diabetes patients with elevated cardiovascular risk: insights from the ACCORD study. Cardiovasc Diabetol. 2023;22(1):287.
25. Huang Q, Zou X, Gao P, Han X, Zhou X, Ji L. How does obesity affect mortality through blood pressure and blood glucose in Chinese and US citizens? Insights from a causal mediation analysis of two large cohorts. J Glob Health. 2023;13.
26. Pylypchuk R, Wells S, Kerr A, Poppe K, Harwood M, Mehta S, et al. Cardiovascular risk prediction in type 2 diabetes before and after widespread screening: a derivation and validation study. Lancet. 2021;397(10291):2264–74.
27. Arrieta F, Piñera M, Nogales P, Iglesias R, Tutor E, Salinero M, et al. P 30: Cardiovascular risk factors in type 2 diabetes along 3 years of follow-up. Diabetes Metab. 2012;38:S115.
28. Khan MS, Segar MW, Usman MS, Patel KV, Van Spall HGC, DeVore AD et al. Effect of canagliflozin on heart failure hospitalization in diabetes according to baseline heart failure risk. JACC Heart Fail. 2023;11(7):825–35.
29. Zou X, Huang Q, Luo Y, Ren Q, Han X, Zhou X, et al. The efficacy of canagliflozin in diabetes subgroups stratified by data-driven clustering or a supervised machine learning method: a post hoc analysis of canagliflozin clinical trial data. Diabetologia. 2022;65(9):1424–35.
30. Tang X, Zhang D, He L, Wu N, Si Y, Cao Y, et al. Performance of atherosclerotic cardiovascular risk prediction models in a rural Northern Chinese population: results from the Fangshan Cohort Study. Am Heart J. 2019;211:34–44.
31. Xia Q, Chen Y, Yu Z, Huang Z, Yang Y, Mao A, et al. Prevalence, awareness, treatment, and control of dyslipidemia in Chinese adults: a systematic review and meta-analysis. Front Cardiovasc Med. 2023;10:1186330.

Publisher's note