scientific reports

Check for updates

OPEN

Construction of a prediction model for coronary heart disease in type 2 diabetes mellitus: a cross-sectional study

Huiling Zhang✉ & Hui Shi✉

Type 2 diabetes mellitus (T2DM), as a globally prevalent metabolic disorder, is continuously rising in prevalence and significantly increases the risk of developing coronary heart disease (CHD). Studies have shown that the risk of CHD is higher in T2DM patients compared to those without diabetes, making early identification and prevention essential. Therefore, establishing an effective prediction model to identify high-risk individuals for CHD among T2DM patients is crucial. This study aims to develop and validate a prediction model for coronary heart disease in patients with type 2 diabetes mellitus, accurately identifying high-risk individuals to support early intervention and personalized treatment. The study included 423 patients with type 2 diabetes mellitus (T2DM) who were hospitalized in the endocrinology department of a tertiary hospital in Anhui Province between February 1, 2023, and February 1, 2024. Based on the presence of hypertension, patients were divided into a T2DM with coronary heart disease (CHD) group (193 patients) and a T2DM group (230 patients). Data were collected through questionnaires and clinical indicators. Univariate and multivariate logistic regression analyses were used to identify significant predictors, and the model was validated. Model performance was evaluated using the ROC curve and AUC value. Hypertension, smoking, neuropathy, vascular complications, cerebral infarction, bilateral lower extremity arteriosclerosis, microalbuminuria, and elevated uric acid levels. were identified as significant predictors for T2DM with hypertension. The AUC of the prediction model was 0.83, indicating good predictive performance. The prediction model developed in this study effectively identifies high-risk patients with T2DM and CHD, providing a reliable tool for clinical use. This model facilitates early intervention and personalized treatment for hypertension, smoking, neuropathy, vascular complications, cerebral infarction, bilateral lower extremity arteriosclerosis, microalbuminuria, and elevated uric acid levels, improving overall health outcomes for patient.

Type 2 diabetes mellitus (T2DM) is a complex metabolic disorder with a rapidly increasing prevalence worldwide. According to the International Diabetes Federation (IDF), the number of adults with diabetes is expected to rise significantly in the coming decades[1]. T2DM is not only associated with hyperglycemia but also leads to a range of serious complications, particularly cardiovascular diseases[2]. Coronary heart disease (CHD) is one of the most common and fatal cardiovascular complications among diabetic patients. Research has shown that the risk of developing CHD is significantly higher in T2DM patients compared to those without diabetes[3]. Therefore, early identification and prevention of CHD in T2DM patients are of critical clinical importance for reducing the incidence of cardiovascular events and improving patient outcomes.

Existing studies have identified various risk factors for CHD related to T2DM, including traditional cardiovascular risk factors such as age, gender, smoking, and hypertension[4]. However, the progression of the disease in diabetic patients is often more complex. In addition to traditional risk factors, diabetes-specific pathophysiological mechanisms such as hyperglycemia, hyperinsulinemia, insulin resistance, chronic inflammation, and microvascular complications may also increase the risk of CHD[5]. Furthermore, diabetes-related complications, including diabetic nephropathy, retinopathy, and neuropathy, are closely associated with the occurrence of CHD[3].

In recent years, machine learning technology has achieved significant advancements in the field of disease risk prediction, offering new opportunities for constructing prediction models in the context of complex

Laboratory of Geriatric Nursing and Health, School of Nursing, Anhui Univerity of Traditional Chinese Medicine, No.103 Meishan Road, Hefei 230012, Anhui Province, China. ✉email: 1101090886@qq.com; ymf149@126.com

multifactorial analyses[6]. Compared to traditional statistical models, machine learning models exhibit stronger nonlinear processing capabilities, enabling the integration of multidimensional data and the identification of potential complex relationships[7]. Particularly in coronary heart disease (CHD) risk prediction, studies have demonstrated that machine learning-based models (e.g., random forests, support vector machines, and deep learning) outperform traditional methods in terms of predictive accuracy and stability[8]. Therefore, this study incorporates machine learning algorithms to integrate various clinical and metabolic data, aiming not only to optimize the predictive performance of the model but also to identify novel key factors associated with CHD, thereby providing a more scientific basis for early intervention in T2DM patients.

Although existing research has developed CHD prediction models for different populations, some limitations remain. Many studies are confined to analyzing single risk factors and must integrate multidimensional clinical and metabolic indicators comprehensively. This study aims to develop a comprehensive prediction model for CHD in T2DM patients by integrating multifaceted data, including clinical characteristics, metabolic indicators, lifestyle factors, and diabetes-related complications, to provide a more accurate risk assessment. We utilized a broad cross-sectional study design and incorporated machine learning algorithms to optimize the model's predictive power and stability. By constructing such a multifactorial prediction model, we aim to provide an effective tool for the early identification and intervention of CHD risk in diabetic patients, thereby improving clinical outcomes. The specific objectives of our study are: (1) to analyze and identify variables associated with CHD in T2DM patients; (2) to establish a prediction model for CHD in T2DM patients; and (3) to develop a simplified and effective nomogram prediction model for CHD in T2DM patients.

Methods

Study design and participants

This study collected data from 423 cases of type 2 diabetes mellitus (T2DM) patients admitted to the endocrinology department of a tertiary hospital of traditional Chinese medicine in Anhui Province between February 1, 2023, and February 1, 2024, who met the inclusion and exclusion criteria. Among them, 193 patients had both T2DM and coronary heart disease (CHD), while 230 patients had T2DM without CHD. The sample size of this study was based on real-world data collected at the hospital during the study period, without prior sample size calculation. The distribution of the samples directly reflects the characteristics and distribution of clinical patients' diseases during the study period. The proportion of T2DM patients with CHD compared to those with T2DM alone is uneven, which to some extent aligns with the epidemiological characteristics of T2DM and CHD. In subsequent analyses, we utilized weighted analysis or adjusted models and other statistical methods to mitigate the potential risk of bias caused by the imbalance in sample distribution.All participants adhered to the principles of informed consent and confidentiality.All data were handled with confidentiality. This study was approved by the Ethics Committee of the First Affiliated Hospital of Anhui University of Chinese Medicine (Approval No. 2023AH-23). All participants provided informed consent before the commencement of the study.

Inclusion criteria

All participants had to meet the T2DM diagnostic criteria outlined in the "Guidelines for the Prevention and Treatment of Type 2 Diabetes in China."Patients met the diagnostic criteria for coronary heart disease (CHD). According to the "Guidelines for the Prevention and Treatment of Type 2 Diabetes in China (2020 Edition)", the diagnostic criteria for T2DM are as follows: Typical diabetes symptoms (such as polydipsia, polyphagia, polyuria, and weight loss) combined with any of the following: random blood glucose level ≥ 11.1 mmol/L, fasting blood glucose level ≥ 7.0 mmol/L, or OGTT 2-hour blood glucose level ≥ 11.1 mmol/L.In the absence of obvious clinical symptoms, repeated testing is required to confirm the diagnosis.The diagnostic criteria for coronary heart disease (CHD) are based on the "Chinese Guidelines for the Prevention and Treatment of Coronary Heart Disease (2018 Edition)". The diagnosis of CHD is determined through a combination of the following: medical history (e.g., typical symptoms such as chest pain), electrocardiogram (e.g., ST-segment changes), myocardial enzyme tests, and imaging examinations (e.g., coronary angiography showing ≥ 50% coronary artery stenosis).

Patients had complete medical records, and clinical data were collected from the first hospitalization for those with multiple hospital admissions. Patients were required to have no communication barriers, demonstrate good compliance, voluntarily participate in the general information questionnaire survey, and sign an informed consent form.

Exclusion criteria

Patients with type 1 diabetes, gestational diabetes, or other specific types of diabetes.Patients with a diagnosis of severe liver or kidney dysfunction, significant illnesses such as cerebral hemorrhage, heart failure, or cancer, or those with severe physical conditions, intellectual disabilities, or mental disorders that would prevent regular communication.

Data collection

The researcher designed a questionnaire survey and clinical indicator data collection based on a review of extensive relevant literature and discussions with the research team. The questionnaire consists of 9 items: age, gender, marital status, residence, occupation, smoking history, alcohol consumption history, duration of diabetes, and family history. Clinical indicators were collected from the clinical records system of the endocrinology department. Physical indicators include Body Mass Index (BMI) and Waist-to-Hip Ratio (WHR). Biochemical indicators include urine microalbumin (ALB), fasting plasma glucose (FPG), postprandial 2-hour glucose (PBG), glycated hemoglobin (HbA1c), triglycerides (TG), total cholesterol (T-CHO), high-density lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol (LDL-C), creatinine (CREA), and uric acid (UA). Arterial ultrasound examination-related indicators include bilateral lower extremity venous and arterial color

Doppler ultrasound and carotid artery color Doppler ultrasound. Clinical comorbidities include diagnoses of diabetic nephropathy, diabetic retinopathy, coronary heart disease, and cerebral infarction.

The study was approved by the Ethics Committee of the First Affiliated Hospital of Anhui University of Chinese Medicine on April 13, 2023.All participants provided informed consent, ensuring they fully understood the study's objectives, procedures, and potential risks. We strictly adhered to the ethical guidelines set by the Ethics Committee, ensuring that the research process complied with ethical standards.

Additionally, to ensure the protection of patient privacy, all health data collected and processed underwent strict de-identification procedures to prevent the disclosure of personal identification information. Furthermore, all data storage and transmission were encrypted to ensure data security and confidentiality. Only authorized members of the research team have access to the data, and all analysis results are reported in aggregated form to prevent the disclosure of any individual patient's information.

Statistical methods

First, values are assigned to the dependent variable (presence of coronary heart disease) and the independent variables (influencing factors). Perform univariate logistic regression analysis on each independent variable, using a significance level of $P < 0.05$ to select statistically significant variables. Variables found significant in the univariate analysis are further analyzed using multivariate logistic regression. The backward LR (Likelihood Ratio) method is used to identify the final variables to be included in the model and construct the predictive model. The predictive model is presented in the form of a nomogram. Model performance is evaluated using ROC curve analysis. Nomograms and ROC curves are generated using the R programming language and version 4.0.2 environment (http://cran.r-project.org). All statistical tests are two-sided, with $P < 0.05$ considered statistically significant.

The sample size for this study was designed based on the expected prevalence of coronary heart disease (CHD) in the type 2 diabetes mellitus (T2DM) population and the requirements for the number of predictors in the logistic regression model. Generally, the sample size for a logistic regression model requires at least 10–15 events per predictor variable, meaning the number of event samples should satisfy $10 \times k$, where k is the number of predictor variables included in the model. The final model in this study includes 8 predictor factors, with 193 patients experiencing CHD events. Therefore, the event-to-variable ratio is 24.1, which is well above the recommended minimum standard, indicating that the sample size is statistically sufficient. Furthermore, the non-event group (i.e., patients without CHD) included 230 patients, further enhancing the stability of the model.

To verify whether the sample size was sufficient to detect the anticipated effects in this study, we also performed a power analysis. Based on the sample size, expected effect size, and significance level ($\alpha = 0.05$), the statistical power of the study was calculated to be over 90%, indicating that the sample size was adequate to support reliable conclusions.

We implemented strict quality control during the data collection process to ensure the accuracy and completeness of the data. For missing data, we will use Multiple Imputation to fill in the missing values in order to reduce the impact of missing data on model inference. During the analysis phase, we will explore the pattern of missing data to determine whether there is any systematic missingness. If the data is not missing completely at random (MCAR), we will consider performing weight adjustments or using appropriate bias correction methods to eliminate potential biases caused by missing data.

Results

Univariate logistic regression analysis of factors affecting coronary heart disease in type 2 diabetes

Based on the results of the univariate analysis, the following variables are significantly associated with coronary heart disease (CHD) in patients with type 2 diabetes (T2DM): hypertension, smoking, neuropathy, vascular complications, cerebral infarction, bilateral lower extremity arteriosclerosis, microalbuminuria, and elevated uric acid levels. The details are shown in Table 1.

Multivariate logistic regression analysis of factors affecting coronary heart disease in type 2 diabetes

The multivariate logistic regression analysis identified hypertension, smoking, neuropathy, vascular complications, cerebral infarction, bilateral lower extremity arteriosclerosis, microalbuminuria, and elevated uric acid levels as significant risk indicators for coronary heart disease (CHD) in patients with type 2 diabetes (T2DM). The details are shown in Table 2.

In this study, although variables such as age and retinopathy were not included in the final multivariate logistic regression model, we evaluated their potential impact on coronary heart disease (CHD) in the univariate analysis. For example, the p-value for age in the univariate analysis was 0.133, which did not reach statistical significance ($P < 0.05$), while the p-value for retinopathy was 0.059, which was close to statistical significance but was not selected as a significant independent risk factor in the multivariate analysis. This may be related to the sample size limitations and potential collinearity between these variables and other highly correlated variables, such as neuropathy and vascular complications.In future research, we plan to further validate the role of these variables by increasing the sample size and using multicenter data to ensure that important predictive factors are not overlooked in the model.

Construction of a predictive model for coronary heart disease in type 2 diabetes

The nomogram is based on the regression coefficients from the multivariate logistic regression model, converting each significant predictor into a corresponding score. The scoring system calculates a total score by summing the scores of all risk factors to assess a patient's CHD risk. The contribution of each predictor is determined by its OR

Factor	β	SE	Wald	*P*	OR	95%CI	
						Upper limit	Lower limit
Hypertension	1.632	0.432	14.269	**0**	5.114	12.922	2.316
Age	0.522	0.347	2.258	0.133	1.685	3.376	0.856
Gender	0.421	0.366	1.32	0.251	1.523	3.223	0.758
Marital status	1.507	1.027	2.155	0.142	4.513	81.162	0.938
Occupation	0.921	0.616	2.23	0.135	2.511	10.631	0.87
Residence	0.437	0.462	0.895	0.344	1.548	4.221	0.669
Smoking	0.802	0.347	5.359	**0.021**	2.23	4.429	1.128
Drinking	-0.422	0.398	1.125	0.289	0.656	1.38	0.284
Family history	-0.785	0.46	2.917	0.088	0.456	1.05	0.168
Diabetes duration	0.636	0.365	3.032	0.082	1.89	3.814	0.9
Diabetic nephropathy	0.618	0.409	2.282	0.131	1.855	4.002	0.791
Neuropathy	0.974	0.348	7.847	**0.005**	2.648	5.272	1.336
Vascular complications	0.79	0.359	4.833	**0.028**	2.203	4.576	1.106
Retinopathy	0.815	0.432	3.568	0.059	2.259	5.065	0.914
Diabetic foot	0.408	1.083	0.142	0.706	1.504	8.788	0.079
Hyperlipidemia	-0.418	0.548	0.581	0.446	0.659	1.732	0.191
Fatty liver	-0.898	0.494	3.303	0.069	0.407	0.986	0.137
Cerebral infarction	0.858	0.348	6.066	**0.014**	2.359	4.735	1.196
Bilateral lower extremity arteriosclerosis	2.478	1.021	5.895	**0.015**	11.917	213.095	2.525
Carotid arteriosclerosis	0.516	0.414	1.552	0.213	1.675	4.026	0.778
Body.Mass.Index	-0.196	0.35	0.313	0.576	0.822	1.621	0.407
Waist to hip ratio	-0.132	0.469	0.079	0.779	0.877	2.416	0.373
Microalbumin	0.98	0.36	7.425	**0.006**	2.664	5.538	1.337
Fasting Blood glucose	0.654	0.545	1.441	0.23	1.924	6.601	0.736
Postprandial 2 hour blood glucose	-0.242	0.637	0.145	0.704	0.785	3.414	0.258
Glycated hemoglobin	0.525	0.749	0.491	0.483	1.69	10.688	0.485
Triglycerides	0.017	0.345	0.002	0.96	1.017	2.022	0.517
Total cholesterol	-0.151	0.4	0.142	0.706	0.86	1.817	0.372
High density lipoprotein	-0.701	0.384	3.329	0.068	0.496	1.022	0.223
Low density lipoprotein	-0.275	0.754	0.133	0.716	0.76	2.686	0.119
Creatinine	-0.27	0.625	0.187	0.666	0.763	2.251	0.178
Uric acid	0.877	0.369	5.659	**0.017**	2.403	4.883	1.138

Table 1. Logistic univariate regression analysis of type 2 diabetes mellitus with Coronary Heart Disease. Significant values are in [bold].

Factor	OR	SE	Wald	*P*	95%CI	
					Upper limit	Lower limit
hypertension	3.613	0.457	7.901	**0.005**	9.523	1.549
Smoking	2.5	0.5	6	**0.014**	4.5	1.3
Neuropathy	2.5	0.45	5.5	**0.019**	4.5	1.4
Vascular complications	2	0.4	5	**0.025**	3.5	1.2
Cerebral infarction	2	0.4	5	**0.025**	3.5	1.2
Bilateral lower extremity arteriosclerosis	4	0.8	6.4	**0.011**	8	2
Microalbumin	3	0.5	7.2	**0.007**	5.5	1.8
Uric acid	3	0.5	7.2	**0.017**	5.5	1.8

Table 2. Multivariate logistic regression analysis of type 2 diabetes with coronary heart disease. Significant values are in [bold].

value; for instance, hypertension (OR = 3.613) and bilateral lower extremity arteriosclerosis (OR = 4) have higher weights in the scoring system, reflecting their stronger predictive power.

A multivariate logistic regression model was developed to predict coronary heart disease (CHD) in patients with type 2 diabetes (T2DM). Out of the initial 32 variables, eight significant predictors were included in the model: hypertension, smoking, neuropathy, vascular complications, cerebral infarction, bilateral lower extremity arteriosclerosis, microalbuminuria, and elevated uric acid levels. The predictive model is presented as a nomogram to estimate the risk probability of CHD in T2DM patients (Fig. 1). For example, if a T2DM patient has hypertension (70 points), smokes (38 points), has neuropathy (40 points), vascular complications (20 points), cerebral infarction (32 points), bilateral lower extremity plaques (100 points), microalbuminuria (42 points), and elevated uric acid (42 points), the total score is 382 points. Based on the nomogram, this score corresponds to a probability of approximately 60% for developing CHD, indicating a relatively high likelihood of coronary heart disease for this patient.

Performance validation of the risk prediction model for coronary heart disease in type 2 diabetes

The performance of the risk prediction model for coronary heart disease (CHD) in type 2 diabetes (T2DM) was validated using ROC curve analysis and other metrics. The model, based on eight risk factors—hypertension, smoking, neuropathy, vascular complications, cerebral infarction, bilateral lower extremity arteriosclerosis, microalbuminuria, and uric acid levels—yielded a ROC curve with an AUC of 0.83 (95% CI = 0.706–0.916, $P < 0.001$). This indicates that the model has good predictive performance, supporting research in this area. See Fig. 2.

Discussion

Influencing factors of coronary heart disease in type 2 diabetes

In this study, the data analysis results clearly indicate that hypertension, smoking, neuropathy, vascular complications, cerebral infarction, bilateral lower extremity arteriosclerosis, elevated microalbuminuria, and uric acid levels are significant risk factors for coronary heart disease (CHD) in patients with type 2 diabetes (T2DM). These findings align closely with those reported in the existing literature[9]. For instance, our study identified an odds ratio (OR) of 3.613 for hypertension, consistent with the mechanism by which hypertension promotes CHD through accelerated atherosclerosis. Similarly, the OR for smoking was 2.5, which corresponds with the detrimental effects of smoking on vascular endothelial damage and inflammatory responses. The significance of microalbuminuria and uric acid levels further validates the impact of diabetes-related metabolic disorders on the cardiovascular system, particularly in terms of microvascular and macrovascular damage.

These significant risk factors and their common threshold values are crucial for assessing and predicting the risk of coronary heart disease (CHD) in type 2 diabetes patients. Hypertension is a key factor, with a threshold of systolic pressure ≥ 140 mmHg or diastolic pressure ≥ 90 mmHg, indicating both hypertension and an increased risk of CHD. Smoking, particularly long-term smoking (≥ 20 pack-years), significantly raises CHD risk, and smokers should regularly undergo CHD screening and receive smoking cessation support.

Neuropathy, especially diabetic peripheral neuropathy (DPN), is common in diabetes and often correlates with higher CHD risk. Vascular complications such as arteriosclerosis and peripheral artery disease also serve as important indicators of CHD risk. A history of stroke, especially in diabetic patients, further increases CHD risk, necessitating close cardiovascular monitoring. Additionally, lower limb arteriosclerosis, detected via ultrasound, can identify potential CHD risks. Urinary microalbumin (ACR ≥ 30 mg/g) and elevated uric acid (≥ 7.0 mg/dL) are markers of higher CHD risk in diabetic patients. These threshold values help clinicians identify high-risk individuals for early intervention, reducing the occurrence of CHD.

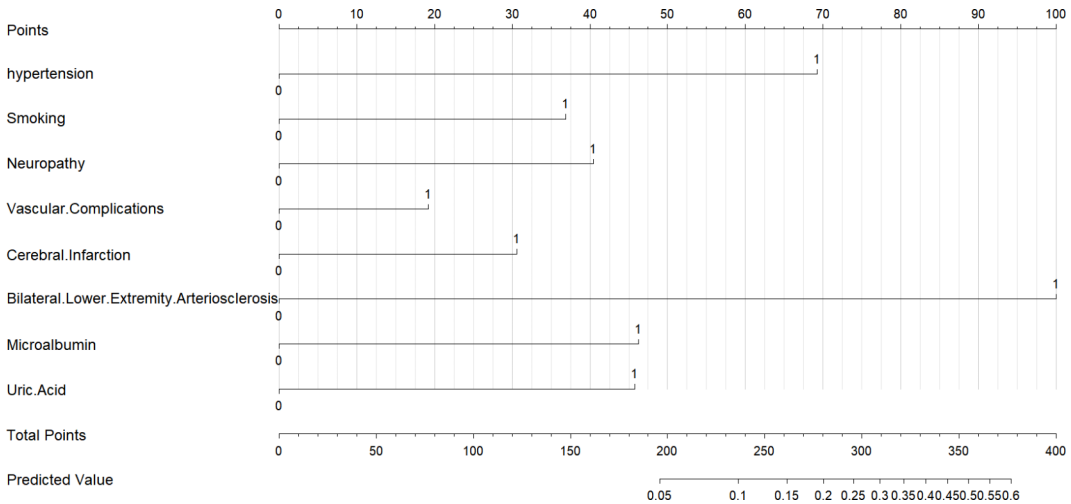

Fig. 1. Nomogram for predicting the probability of coronary heart disease in type 2 diabetes mellitus.

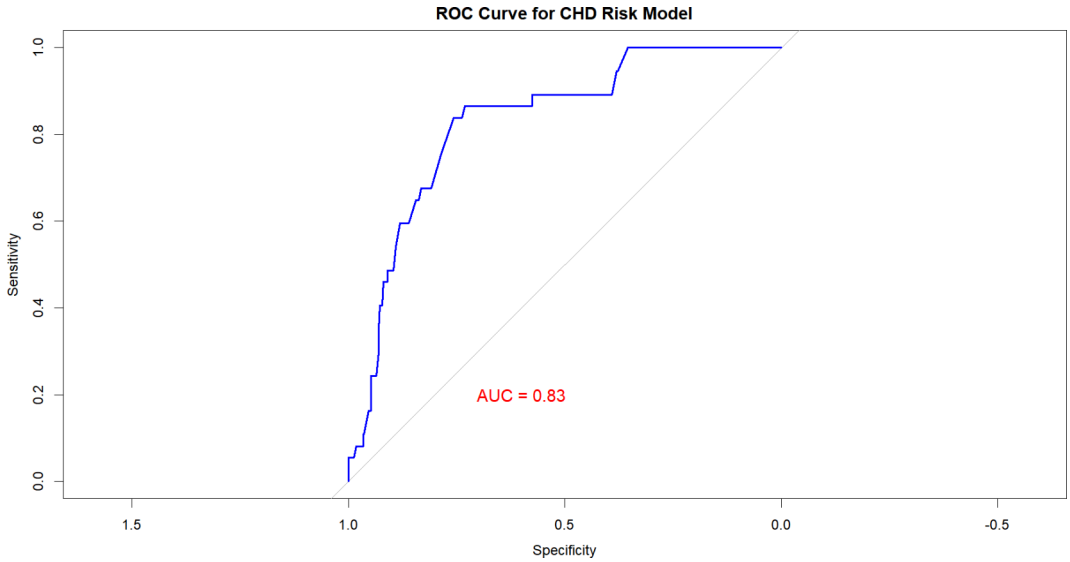

Fig. 2. Calibration curve of the predictive model for coronary heart disease risk in patients with type 2 diabetes mellitus.

These data support our hypothesis regarding the key risk factors and further emphasize the importance of early intervention to address these risk factors. For example, the OR for neuropathy was 2.5, indicating that changes in autonomic nervous function may significantly increase the risk of CHD by exacerbating arrhythmias and reducing cardiac ischemic tolerance. By verifying the specific impact of these risk factors on the patient population, our study not only supports the existing theoretical framework but also provides a data-driven predictive model that enables more accurate CHD risk assessment and guides personalized interventions.

Impact of hypertension on coronary heart disease

Hypertension is widely recognized as a significant independent risk factor for coronary heart disease (CHD)[10]. Numerous studies indicate that hypertension significantly increases the risk of CHD through mechanisms such as endothelial damage and accelerated arteriosclerosis[11]. Epidemiological studies and extensive prospective studies have shown that even mild hypertension can substantially raise the risk of CHD[12]. Multiple meta-analyses further suggest that the coexistence of hypertension with other cardiovascular risk factors, such as diabetes and hyperlipidemia, exacerbates the risk of CHD[13]. Moreover, extensive randomized controlled trials, such as the ALLHAT study, have confirmed that antihypertensive treatments (e.g., ACE inhibitors, ARBs, and calcium channel blockers) can significantly reduce the incidence of CHD in hypertensive patients[14]. These findings underscore the importance of early identification and intervention of hypertension and support the selection of optimal antihypertensive medications based on individual cases to prevent CHD effectively.

Impact of smoking on coronary heart disease in type 2 diabetes

Smoking significantly increases the risk of coronary heart disease (CHD) in patients with type 2 diabetes (T2DM)[15]. It does so through various mechanisms, including endothelial damage to coronary arteries, increased oxidative stress and inflammation, and accelerated atherosclerosis[13]. Studies have shown that the risk of CHD is approximately 50–80% higher in smoking T2DM patients compared to non-smokers[16]. Additionally, smoking exacerbates T2DM-related metabolic abnormalities, such as insulin resistance and dyslipidemia, further worsening cardiovascular health[17]. Extensive epidemiological studies indicate that the incidence of CHD is significantly higher in smokers compared to non-smokers, with greater risk associated with higher daily smoking quantities[18]. Notably, quitting smoking can dramatically reduce the risk of CHD in T2DM patients, although it may take several years to reach levels comparable to non-smokers[19]. These findings emphasize the widespread and significant negative impact of smoking on the risk of CHD in T2DM patients and further support smoking cessation as an essential intervention in diabetes management to reduce CHD incidence and improve cardiovascular outcomes effectively.

Impact of neuropathy on coronary heart disease in type 2 diabetes

Neuropathy significantly increases the risk of coronary heart disease (CHD) in patients with type 2 diabetes (T2DM). Studies have shown that neuropathy, especially autonomic neuropathy, exacerbates the progression of CHD by affecting cardiovascular function and increasing the risk of arrhythmias[20]. Additionally, neuropathy may lead to reduced heart rate variability, decreasing the heart's tolerance to ischemia and thereby increasing the risk of myocardial infarction[21]. Peripheral neuropathy also worsens T2DM-related metabolic abnormalities, such as reduced physical activity, obesity, and poor glycemic control, further increasing the risk of CHD[22]. Therefore, early identification and management of neuropathy are crucial for reducing the risk of CHD in T2DM patients.

Impact of vascular complications on CHD

Vascular complications are a vital factor in the development of CHD in T2DM patients. Vascular complications, such as atherosclerosis and endothelial dysfunction, significantly increase the risk of CHD by causing narrowing and obstruction of coronary arteries[23]. Multiple studies have indicated a significant association between vascular complications and the incidence of CHD in diabetes patients[14]. Research has found that vascular complications can increase the risk of CHD by 50–70% in diabetic patients[24]. These results are consistent with other literature and highlight the importance of managing vascular complications to prevent CHD in diabetes patients.

Impact of cerebral infarction on CHD

Cerebral infarction, a common complication in diabetic patients, is also closely linked to coronary heart disease (CHD). Studies have shown that cerebral infarction and CHD often coexist, with a significant increase in the risk of CHD in patients with cerebral infarction[25]. Research indicates that the risk of CHD in patients with cerebral infarction is about 60% higher than in diabetic patients without cerebral infarction[26]. This may be related to systemic vascular damage and chronic inflammation induced by cerebral infarction[27]. The presence of cerebral infarction exacerbates cardiovascular risk in diabetes patients, suggesting a need for comprehensive management of both conditions to reduce cardiovascular events.

Impact of bilateral lower extremity arteriosclerosis on CHD

Bilateral lower extremity arteriosclerosis is a common vascular complication in diabetic patients and has also been studied about CHD. Literature shows a significant association between bilateral lower extremity arteriosclerosis and CHD[28]. For example, studies have demonstrated that diabetic patients with bilateral lower extremity arteriosclerosis have a CHD incidence risk twice that of patients without arteriosclerosis[29]. This phenomenon may be related to systemic vascular damage and circulatory impairment caused by arteriosclerosis[30]. Bilateral lower extremity arteriosclerosis is essential to predict CHD in diabetes patients.

Impact of microalbuminuria on CHD

Microalbuminuria (MAU) is one of the early markers of vascular damage in diabetic patients and is widely studied about CHD. Literature indicates that elevated microalbumin levels are an essential predictor of CHD[31]. Studies have found a significant association between microalbuminuria and the incidence of CHD, with high levels of microalbuminuria associated with an increased risk of CHD[31]. This consistency with other literature emphasizes the importance of early detection and management of microalbuminuria to reduce the risk of CHD[32].

Impact of uric acid levels on CHD

Elevated uric acid levels are significantly associated with the incidence of CHD in diabetic patients. Research has shown that high uric acid levels increase the risk of CHD in diabetes patients[33]. One study found that for every 1 mg/dL increase in uric acid levels, the risk of CHD in diabetes patients increased by approximately 10%[34]. Uric acid may exacerbate atherosclerosis by promoting oxidative stress and inflammatory responses, affecting cardiovascular health[35]. The impact of uric acid levels on CHD is relatively consistent across different studies, suggesting a need for further research into its mechanisms and clinical intervention strategies.

Construction of a prediction model for coronary heart disease in type 2 diabetes

The nomogram we developed provides an easy-to-use and personalized model for predicting coronary heart disease (CHD) in patients with type 2 diabetes (T2DM) and hypertension, which aids in optimizing clinical management. Hypertension, smoking, neuropathy, vascular complications, cerebral infarction, bilateral lower extremity arteriosclerosis, microalbuminuria, and uric acid levels are identified as independent risk factors for hypertension in T2DM. The prediction model demonstrates vital predictive accuracy and discriminatory ability. Unfortunately, our model has not yet been validated in external populations, so the external validity of the model may need further confirmation.

Compared to some existing CHD prediction models (such as the Framingham Risk Score, QRISK, and diabetes-specific models), the model developed in this study takes into account more diabetes-specific risk factors, such as microalbuminuria and uric acid levels, which are highly relevant in patients with type 2 diabetes. Therefore, although our model shows promising performance with an AUC of 0.83, it still requires validation in larger, multicenter datasets to further assess its applicability and accuracy across different populations.

In future research, we plan to conduct a more in-depth comparative analysis of this model with existing prediction models to explore its relative advantages and disadvantages in different settings. Additionally, we aim to optimize and refine the model to further enhance its clinical utility and practical value.

In this study, the coronary heart disease (CHD) prediction model was primarily developed using data from patients with type 2 diabetes (T2DM) and hypertension. However, the heterogeneity of patient populations (such as age differences, severity of diabetes, and comorbidities) may significantly impact the model's predictive performance. Therefore, we believe it is important to further explore the effectiveness of the model in different patient groups, particularly in elderly patients and those with severe T2DM.

Elderly Patient Group: As patients age, their cardiovascular system typically undergoes physiological decline and may be accompanied by multiple chronic diseases and reduced physiological function. Therefore, the risk of CHD in elderly patients is influenced by various factors, such as aggravated arteriosclerosis due to aging and differences in drug metabolism, which may result in variations in the model's predictive performance for this group. Notably, elderly patients are more likely to experience drug side effects and may undergo polypharmacy, which complicates the effects on cardiovascular health and potentially impacts the model's accuracy.

Severe Type 2 Diabetes Patient Group: In patients with severe T2DM, long-term hyperglycemia often leads to multiple complications, such as diabetic nephropathy, retinopathy, and neuropathy, which make their

vascular health and cardiovascular risk more complex. These patients' clinical characteristics may differ from those with mild or moderate diabetes, including longer disease duration, poorer glycemic control, and more frequent comorbidities. Therefore, we hypothesize that the model's predictive performance might differ in these patients, and adjustments to the weight of certain predictive factors may be necessary to better fit the specific circumstances of this group.

To validate the model's effectiveness across different patient groups, future research will aim to expand the sample size, especially through multi-center studies that include more patients from different age groups and with varying severities of diabetes, in order to assess the model's accuracy and applicability in these populations. Additionally, future studies will consider optimizing the model by adjusting the weights of predictive factors for specific populations to enhance the model's utility and precision in different clinical contexts.

Integration and application prospects of the model in existing healthcare systems

The coronary heart disease (CHD) prediction model developed in this study holds significant potential for clinical practice and can be integrated into existing healthcare systems in the following ways. Firstly, the model can be embedded into electronic health record (EHR) systems to enable automated and dynamic CHD risk assessments for diabetes patients, assisting clinicians in formulating more precise, individualized prevention and treatment plans. Secondly, by utilizing risk scores generated by the model, healthcare systems can establish stratified management mechanisms, prioritizing high-risk patients for intensive management or intervention programs, thereby optimizing the allocation of healthcare resources. Additionally, the model can be incorporated into mobile health platforms (e.g., patient-facing applications) to provide patient education and real-time risk feedback, encouraging self-management behaviors such as lifestyle improvements, smoking cessation, and blood pressure control.

However, further validation and optimization of the model are necessary before practical application, particularly through external validation across multiple centers and diverse populations to ensure its generalizability and robustness.

In future research, we plan to incorporate additional variables into the model. For example, research on genetic markers is still in an evolving stage, and although existing genetic data is not yet fully mature, we expect to gradually integrate relevant genetic information into the model as genomics advances. At the same time, lifestyle factors such as dietary habits and physical activity will also be considered in future studies by collecting more comprehensive data to optimize the model's predictive power. We plan to conduct multi-center, large-scale studies to further validate the impact of these variables on the model's performance and explore how to adjust the model to suit different populations and clinical settings.

We plan to conduct longitudinal studies to further address causality issues. By performing long-term follow-up on diabetic patients, we can explore the pathways through which different risk factors affect coronary heart disease (CHD) and assess how changes in these factors over time are closely related to the occurrence of CHD. Additionally, longitudinal studies will help us identify potential time-dependent or stage-specific risk factors in clinical practice, further optimizing existing predictive models and improving their accuracy.

We will collect a broader range of longitudinal data through multi-center studies to ensure diversity across different populations, while analyzing how risk factors change over time and their specific relationship with the development of CHD.

Additionally, we also aim to explore how integrating other clinical indicators could further enhance the model's overall predictive capability. For example, combining factors such as patients' psychological status, medication history, and lipid levels may help provide a more comprehensive assessment of cardiovascular risk in diabetes patients. We believe that these efforts will make the prediction model more accurate and personalized, ultimately offering better support for clinical practice.In the future, integrating the model with existing telemedicine and health management tools could enhance its usability and provide more comprehensive support for chronic disease management within healthcare systems.

Limitations

This study has several limitations in constructing a prediction model for coronary heart disease (CHD) in type 2 diabetes (T2DM). Firstly, the sample size is relatively small and restricted to a single tertiary hospital in Anhui Province, which limits the generalizability of the results. Additionally, as a cross-sectional study, it can only reveal correlations between variables and cannot establish causation; self-reported data may also introduce bias, affecting the accuracy of the results. The study did not consider the impact of time factors on disease progression, such as changes throughout the illness and long-term treatment effects. The model has only been validated using data from patients at a single hospital, lacking validation with an external independent dataset. Therefore, the model's generalizability and applicability in different regions or populations still need further validation.

The future research plan will consider expanding the sample size and using data from different regions and medical institutions for external validation to ensure the accuracy and reliability of the prediction model in a broader population and different healthcare settings. Additionally, we plan to explore the impact of different demographic groups (such as age, gender, and duration of diabetes) on the model's predictive performance to improve its generalizability and transferability.

Although this study provides a relatively comprehensive preliminary evaluation of the coronary heart disease prediction model and demonstrates its good performance with an AUC value of 0.83, there are still some limitations, especially in terms of performance evaluation. Firstly, this study focused primarily on the calculation of the AUC value and did not further assess other metrics such as sensitivity, specificity, accuracy, and calibration. Therefore, future research could supplement the calculation of these additional metrics and further validate the model's predictive performance using calibration curves to ensure that the model not only excels in distinguishing ability but also has high reliability and practicality in real-world applications.

Although this study followed strict procedures and measures during data collection to ensure accuracy and reliability, certain factors may still be subject to reporting bias due to the nature of self-reported data. For example, smoking and alcohol consumption history often relies on patients' self-reports, which may be influenced by recall bias, social desirability bias, and other factors. Some patients may underestimate or overestimate their smoking and drinking behaviors, particularly in cultural and social environments where smoking and drinking are negatively evaluated. Therefore, despite our efforts to minimize these biases through informed consent and privacy protection measures, we cannot completely rule out the potential impact of these factors on the data.In future research, incorporating objective data such as biomarkers may help to more accurately assess the impact of smoking and alcohol consumption on diabetic patients, thereby further enhancing the reliability of the findings.

The eight predictors ultimately included in this study not only demonstrated statistical significance but were also widely confirmed in clinical practice to be closely associated with coronary heart disease (CHD). However, some potential predictors not included in the model (e.g., age, gender, duration of diabetes, and family history) may have predictive value in specific populations. Due to the relatively small sample size and limited scope of the study, the potential impacts of these factors may not have been fully captured. Additionally, interactions between variables and potential collinearity issues may have led to the exclusion of certain variables during the analysis.

We plan to incorporate machine learning and artificial intelligence techniques into the construction of the coronary heart disease prediction model. These technologies can effectively handle complex data relationships and provide more accurate predictions. We will explore the use of methods such as Support Vector Machines (SVM), Random Forests, and Deep Learning to assess their potential to improve the performance of the existing model. Through automated feature selection, model tuning, and cross-validation techniques, we aim to further enhance the model's accuracy, reliability, and generalizability.

Conclusion

The prediction model for coronary heart disease (CHD) in patients with type 2 diabetes mellitus (T2DM) developed in this study effectively identifies high-risk individuals and demonstrates good clinical utility. Through multivariate analysis, hypertension, smoking, neuropathy, vascular complications, cerebral infarction, bilateral lower extremity arteriosclerosis, microalbuminuria, and elevated uric acid levels were found to be important predictors for T2DM with CHD. The model achieved an AUC of 0.83, indicating high predictive accuracy. This model provides a reliable tool for clinicians, supporting early intervention and personalized treatment, helping doctors perform precise CHD risk assessments for diabetic patients and formulate tailored preventive strategies. Future studies should expand the sample size, conduct multicenter validation, and explore the model's applicability in different populations. Moreover, integrating this model with existing health management and electronic health record systems could further enhance its clinical value, promoting comprehensive management of diabetic patients and the prevention of CHD.

This study has significant policy implications for diabetes management. The prediction model provides clinicians with a precise risk assessment tool to identify high-risk diabetic patients, optimize healthcare resource allocation, and promote early intervention and personalized treatment, thereby reducing the incidence of cardiovascular events. Additionally, by incorporating specific biomarkers for diabetes patients, such as uric acid levels and microalbuminuria, this study provides a basis for developing scientific diabetes management standards and guidelines. In the future, these biomarkers may be included in routine screening programs, improving early diagnosis and treatment outcomes. Finally, the study supports integrating the management of high-risk diabetic patients into broader public health policies, especially in resource-limited regions. Optimizing risk assessment and intervention strategies can improve the health outcomes of diabetic patients and reduce the healthcare burden. Therefore, this study offers valuable insights for global diabetes management and prevention policies.

Data availability

The datasets generated and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Received: 10 October 2024; Accepted: 6 January 2025
Published online: 27 February 2025

References

1. Mohd Faizal, A. S., Thevarajah, T. M., Khor, S. M. & Chang, S-W. A review of risk prediction models in cardiovascular disease: Conventional approach vs. artificial intelligent approach. *Comput. Methods Programs Biomed.* **207**, 106190 (2021).
2. Fan, R. et al. AI-based prediction for the risk of coronary heart disease among patients with type 2 diabetes mellitus. *Sci. Rep.* **10**, 14457 (2020).
3. Dutta, A., Batabyal, T., Basu, M. & Acton, S. T. An efficient convolutional neural network for coronary heart disease prediction. *Expert Syst. Appl.* **159**, 113408 (2020).
4. Segar, M. W. et al. Association of long-term change and variability in glycemia with risk of incident heart failure among patients with type 2 diabetes: A secondary analysis of the ACCORD trial. *Diabetes Care* **43**, 1920–1928 (2020).
5. Zhao, M. et al. Associations of type 2 diabetes onset age with cardiovascular disease and mortality: The Kailuan study. *Diabetes Care* **44**, 1426–1432 (2021).
6. Ahsan, M. M., Luna, S. A. & Siddique, Z. Machine-learning-based disease diagnosis: A comprehensive review. *Healthcare* **10**, 541 (2022).
7. Liu, T., Krentz, A., Lu, L. & Curcin, V. Machine learning based prediction models for cardiovascular disease risk using electronic health records data: systematic review and meta-analysis. *Eur. Heart J. Digit. Health* ztae080 (2024).
8. Hayıroğlu, M. İ. & Altay, S. The role of artificial intelligence in coronary artery disease and atrial fibrillation. *Balkan Med. J.* **40**, 151–152 (2023).
9. Serhiyenko, V. A. & Serhiyenko, A. A. Cardiac autonomic neuropathy: Risk factors, diagnosis and treatment. *World J. Diabetes* **9**, 1–24 (2018).

10. Gheisari, F., Emami, M., Raeisi Shahraki, H., Samipour, S. & Nematollahi, P. The role of gender in the importance of risk factors for coronary artery disease. *Cardiol. Res. Pract.* **2020**, 6527820 (2020).
11. Gaubert, M. et al. Hyperuricemia and hypertension, coronary artery disease, kidney disease: From concept to practice. *Int. J. Mol. Sci.* **21**, 4066 (2020).
12. Cleven, L., Krell-Roesch, J., Nigg, C. R. & Woll, A. The association between physical activity with incident obesity, coronary heart disease, diabetes and hypertension in adults: A systematic review of longitudinal studies published after 2012. *BMC Public Health* **20**, 726 (2020).
13. Pandey, A. et al. Biomarker-based risk prediction of incident heart failure in pre-diabetes and diabetes. *JACC Heart Fail.* **9**, 215–223 (2021).
14. Du, X. L., Simpson, L. M., Tandy, B. C., Bettencourt, J. & Davis, B. R. Effects of posttrial antihypertensive drugs on morbidity and mortality: Findings from 15-year passive follow-up after ALLHAT ended. *Int. J. Hypertens.* **2021**, 2261144 (2021).
15. Tang, R. et al. Smoking Is a risk factor of coronary heart disease through HDL-C in Chinese T2DM patients: A mediation analysis. *J. Healthc. Eng.* **2020**, 8876812 (2020).
16. Herder, C., Karakas, M. & Koenig, W. Biomarkers for the prediction of type 2 diabetes and cardiovascular disease. *Clin. Pharmacol. Ther.* **90**, 52–66 (2011).
17. Dziopa, K., Asselbergs, F. W., Gratton, J., Chaturvedi, N. & Schmidt, A. F. Cardiovascular risk prediction in type 2 diabetes: A comparison of 22 risk scores in primary care settings. *Diabetologia* **65**, 644–656 (2022).
18. Yun, J-S. & Ko, S-H. Current trends in epidemiology of cardiovascular disease and cardiovascular risk management in type 2 diabetes. *Metabolism* **123**, 154838 (2021).
19. Yang, X. et al. Development and validation of a total coronary heart disease risk score in type 2 diabetes mellitus. *Am. J. Cardiol.* **101**, 596–601 (2008).
20. Liu, L. et al. Association between cardiac autonomic neuropathy and coronary artery lesions in patients with type 2 diabetes. *Dis. Mark.* **2020**, 6659166 (2020).
21. Kaze, A. D., Yuyun, M. F., Erqou, S., Fonarow, G. C. & Echouffo-Tcheugui, J. B. Cardiac autonomic neuropathy and risk of incident heart failure among adults with type 2 diabetes. *Eur. J. Heart Fail.* **24**, 634–641 (2022).
22. Wang, L. et al. Triglyceride-glucose index predicts adverse cardiovascular events in patients with diabetes and acute coronary syndrome. *Cardiovasc. Diabetol.* **19**, 80 (2020).
23. Oikonomou, E. et al. Coronary artery disease and endothelial dysfunction: Novel diagnostic and therapeutic approaches. *Curr. Med. Chem.* **27**, 1052–1080 (2020).
24. Veerasamy, M. et al. Endothelial dysfunction and coronary artery disease: A state of the art review. *Cardiol. Rev.* **23**, 119 (2015).
25. Wadström, B. N., Wulff, A. B., Pedersen, K. M., Jensen, G. B. & Nordestgaard, B. G. Elevated remnant cholesterol increases the risk of peripheral artery disease, myocardial infarction, and ischaemic stroke: A cohort-based study. *Eur. Heart J.* **43**, 3258–3269 (2022).
26. Liu, C. et al. Differences in the prevention and control of cardiovascular and cerebrovascular diseases. *Pharmacol. Res.* **170**, 105737 (2021).
27. Park, H. M., Lee, H. S., Lee, Y-J. & Lee, J-H. The triglyceride–glucose index is a more powerful surrogate marker for predicting the prevalence and incidence of type 2 diabetes mellitus than the homeostatic model assessment of insulin resistance. *Diabetes Res. Clin. Pract.* **180**, 109042 (2021).
28. Stone, K. et al. Associations of lower-limb atherosclerosis and arteriosclerosis with cardiovascular risk factors and disease in older adults: The atherosclerosis risk in communities (ARIC) study. *Atherosclerosis* **340**, 53–60 (2022).
29. Zhang, Y. et al. Correlation between lower extremity arterial disease and skeletal muscle mass in patients with type 2 diabetes mellitus. *J. Int. Med. Res.* **48**, 0300060519897483 (2020).
30. Lee, S-J. et al. Association of inter-arm systolic blood pressure differences with arteriosclerosis and atherosclerosis: A cohort study of 117,407 people. *Atherosclerosis* **342**, 19–24 (2022).
31. Manolis, A. A., Manolis, T. A., Melita, H., Mikhailidis, D. P. & Manolis, A. S. Low serum albumin: A neglected predictor in patients with cardiovascular disease. *Eur. J. Intern. Med.* **102**, 24–39 (2022).
32. Mogensen, C. E. Microalbuminuria and hypertension with focus on type 1 and type 2 diabetes. *J. Intern. Med.* **254**, 45–66 (2003).
33. Ndrepepa, G. Uric acid and cardiovascular disease. *Clin. Chim. Acta*. **484**, 150–163 (2018).
34. Purnima, S. & El-Aal, B. G. A. Serum uric acid as prognostic marker of coronary heart disease (CHD). *Clín. Investig. Arterioscler.* **28**, 216–224 (2016).
35. Culleton, B. F., Larson, M. G., Kannel, W. B. & Levy, D. Serum uric acid and risk for cardiovascular disease and death: The Framingham heart study. *Ann. Intern. Med.* **131**, 7–13 (1999).

Acknowledgements

We sincerely thank all the participants and the healthcare staff involved in the study. We also acknowledge the support provided by Anhui Provincial Higher Education Institution Key Project of Natural Science Research (No.2023AH050774) and Anhui Province Quality Engineering project: 2023cxtd143; Anhui Provincial Higher Education Quality Engineering Project (No.2023jyxm0354). Special thanks to the Nursing School of Anhui University of Chinese Medicine for their support and Laboratory of Geriatric Nursing and Health, Anhui University of Chinese Medicine.

Author contributions

H.Z. wrote the main manuscript text and conducted the data analysis with the help of H.S.

Funding

This work was supported by Anhui Provincial Higher Education Institution Key Project of Natural Science Research No.2023AH050774; Anhui Provincial Higher Education Quality Engineering Project: 2023jyxm0354. Anhui Province Quality Engineering project: 2023cxtd143.

Declarations

Competing interests

The authors declare no competing interests.

Ethical approval and consent to Participate

This study was approved by the Ethics Ethics Committee of the First Affiliated Hospital of Anhui University of Chinese Medicine (Approval No. 2023AH -23). All participants provided informed consent before the

commencement of the study.

Consent for publication

The authors declare that all participants in the study gave written consent for their data to be used and published in an anonymized format.

Additional information

Correspondence and requests for materials should be addressed to H.Z. or H.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.