Diabetology & Metabolic Syndrome

RESEARCH Open Access

T2D-LVDD: neural network-based predictive models for left ventricular diastolic dysfunction in type 2 diabetes

Yu Rong[1], Wei Liu[1], Ke Li[1], Jian Guo[2] and Xue-Ping Li[1*]

Abstract

Cardiovascular disease complications are the leading cause of morbidity and mortality in patients with Type 2 diabetes (T2DM). Left ventricular diastolic dysfunction (LVDD) is one of the earliest myocardial characteristics of diabetic cardiac dysfunction. Therefore, we aimed to develop an LVDD-risk predictive model to diagnose cardiac dysfunction before severe cardiovascular complications arise. We trained an artificial neural network model to predict LVDD risk with patients' clinical information. The model showed better performance than classical machine learning methods such as logistic regression, random forest and support vector machine. We further explored LVDD-risk/protective features with interpretability methods in neural network. Finally, we provided a freely accessible web server called LVDD-risk, where users can submit their clinical information to obtain their LVDD-risk probability and the most noteworthy risk indicators.

Key points

- Left ventricular diastolic dysfunction(LVDD) is the most common initial myocardial characteristic of diabetic cardiac dysfunction, We constructed an LVDD-risk predict model for earlier detection and diagnosis of serious cardiovascular disease.
- We built an LVDD-risk predict model using artificial neural network and then explored LVDD-risk/protective features using interpretability methods based on the built model.
- We have provided a web server tool for users to predict own LVDD-risk.

Keywords Diabetic complications, Left ventricular diastolic dysfunction, Cardiovascular disease, Machine learning, Neural network

*Correspondence:
Xue-Ping Li
1476117275@qq.com
[1]Xi'an Key Laboratory for Prevention and Treatment of Common Aging Diseases, Translational and Research Centre for Prevention and Therapy of Chronic Disease, Institute of Basic and Translational Medicine, Xi'an Medical University, Xi'an 710021, China
[2]Endocrinology Department of Shaanxi Provincial People's Hospital, Xi'an 710068, China

Introduction

Cardiovascular disease complications are the leading cause of morbidity and mortality in patients with type 2 diabetes mellitus (T2DM) [1]. T2DM can cause myocardial ischemia through microvascular [2] or macrovascular alterations [3], as well as directly impairing the myocardium [4]. Hyperglycemia, hyperinsulinemia, and insulin resistance can result in alterations in vascular homeostasis [5], reduced levels of nitric oxide, and increased levels of reactive oxygen species. These changes promote inflammation, leading to atherosclerosis and myocardial dysfunction [6]. Over time, the interplay of these factors frequently leads to myocardial fibrosis, hypertrophy, ischemia, diastolic and systolic dysfunction [7], and ultimately, heart failure [8] as the disease progresses. Currently, there are few therapies that can fully reverse the organ damage caused by the late cardiovascular complications of diabetes [9]. However, timely diagnosis and early intervention could possibly prevent or even reverse the occurrence and progression of diabetes complications [10, 11].

Numerous studies have reported that left ventricular diastolic dysfunction(LVDD) is the most common initial myocardial characteristic of diabetic cardiac dysfunction and LVDD is a easier impairment to diagnose from the wider range of early subclinical cardiac complications [12–14]. Several articles already talked about the application of LVDD as a diagnostic indicator for cardiovascular disease complications [12, 15]. Compared to current researches mainly focused on the cardiovascular events in diabetes itself [16, 17], such as the risk of heart failure or mortality, LVDD-targeted predictive model can advance the diagnostic time point of myocardial damage, leaving enough time and possibility for early intervention.

In this work, we investigated 805 Chinese T2DM patients with or without LVDD. We trained an artificial neural network to predict LVDD risk in T2DM patients based on the T2DM populations in the Chinese communities. In contrast to current studies that use Electrocardiogram [18–21] to predict cardiac dysfunction, we employ clinical indicators to model the prediction of LVDD. We then explored the risk/protective factors of LVDD using interpretability methods in neural network. Finally, we deployed the model online and provided a freely accessible web server named LVDD-risk (http://www.bigc.online/ExsgRNA) to make it more accessible for physicians and patients to use.

Materials and methods

Study population

805 patients with type 2 diabetes (T2DM) were investigated from June 2021 to December 2022 at endocrinology department of the Shaanxi People's Hospital. All patients were over 18 years old and diagnosed with T2DM based on the criteria for diabetes proposed by WHO diabetes expert committee in 2022. The following patients were exclused:

(1) Diagnosed with type 1 diabetes, gestational diabetes or any other special types of diabetes;
(2) Those with concomitant acute complications such as ketoacidosis or hyperosmolar coma;
(3) Those with concomitant acute and chronic infections, malignant tumors, or liver and kidney function damage;
(4) Those with concomitant hypertension, coronary atherosclerotic heart disease, rheumatic immune system disease, hypertensive heart disease, valvular heart disease, pericardial disease, or any other serious cardiovascular diseases.

Participants were diagnosed with LVDD according to internationally adopted standards set by the American Society of Echocardiography and European Association of Cardiovascular Imaging [22]. The early diastolic transmitral velocities(E) and late diastolic transmitral velocities(L) were measured by doppler echocardiography. The T2DM patients were classified into LVDD group($E/A < 1.0$, 597 patients) and control group($E/A \geq 1.0$, 208 patients). Ethical approval was obtained, and all participants provided informed consent before taking part.

Laboratory and clinical characteristics

The following clinical characteristics and laboratory examination results were collected: age, sex, body mass index(BMI), waist hip ratio(WHR), systolic blood pressure(SBP), diastolic blood pressure(DBP), smoking history, drinking history, diabetes family history. Diabetes complications: diabetes retinopathy(DR), diabetes peripheral neuropathy(DPN), diabetes nephropathy(DN), atherosclerosis(AS), fatty liver disease(FLD). Diabetes medication history: insulin, biguanides, αglycosidase inhibitor, dipeptide kallidinogenase inhibitor IV(DPPIV), sodium glucose transporter 2 inhibitor(SGLT2). Vitamin D(VitD), triglyceride(TG), high density lipoprotein(HDL), low density lipoprotein(LDL), white blood cell count(WBC), neutrophil/lymphocyte ratio(NLR), urine microalbumin/creatinine ratio(ACR), 24-hour urine protein(UTP), high-sensitive cardiac troponin(HSCTN).

Measured values less than lower quartile(QL) – $1.5 \times$ interquartile range(IQR) or greater than upper quartile(UQ) + $1.5 \times$ IQR were considered outliers and recognized as missing data. The missing data were replaced with the median. All measured data were normalized using Z-scores.

Statistical analysis and model development

All statistical analyses included in this study were conducted using Python software version 3.9.1. We built the artificial neural network(ANN) model using Keras (https://keras.io/) with the TensorFlow backend engine. ANN model had two hidden layers with 128 and 64 nodes. Three machine learning algorithms were used for comparison purposes and were implemented using the scikit-learn (https://scikit-learn.org/) library: logistic regression, random forest, and support vector machine (SVM).

ANN modeling

Artificial neural networks (ANNs) are computational models inspired by the human brain. They are comprised of a large number of connected nodes, each of which performs a simple mathematical operation. Each node's output is determined by this operation, as well as a set of parameters that are specific to that node. By connecting these nodes together and carefully setting their parameters, very complex functions can be learned and calculated.

a typical neural network consists of an input layer, an output layer, and one or more hidden layers in between. Each layer can have multiple nodes, where each node represents a neuron. Input signals from other neurons are transmitted to a neuron through links with associated weights. When the accumulated input reaches a certain threshold, the neuron outputs the received total input value. A fully connected layer exists between the hidden layer(s) and the output layer to assist in the final result output.

The mathematical expression for the input layer is as follows:

$$y = f\left(\sum_{i=1}^{n} \omega_i x_i - h\right)$$

The activation function f is commonly used in neural networks. Popular activation functions include the sigmoid, Tanh, and ReLU. When the sum of the products of all inputs and their corresponding connection weights exceeds a threshold, the output is 1; otherwise, the output is 0.

If the number of input layer nodes is n, the number of nodes in the hidden layer is l, and the number of nodes in the output layer is m. The weights from the input layer to the hidden layer are denoted as ω_{ij}, and the weights from the hidden layer to the output layer are denoted as ω_{jk}. The biases from the input layer to the hidden layer are denoted as a_j, and the biases from the hidden layer to the output layer are denoted as b_k. The learning rate is represented as η, and the activation function is denoted as f(x). All input variables should pass through the activation function, the output of the hidden layer, denoted as H_j, can be expressed as follows:

$$H_j = f\left(\sum_{i=1}^{n} \omega_{ij} x_i + a_j\right)$$

The predicted output of the neural network, denoted as O_k, can be calculated using the following formula:

$$O_k = \sum_{j=1}^{n} \omega_{jk} H_j + b_k$$

Logistic regression modeling

Logistic regression is used for binary classification where we use sigmoid function, that takes input as independent variables and produces a probability value between 0 and 1.The logistic regression model transforms the linear regression function continuous value output into categorical value output using a sigmoid function, which maps any real-valued set of independent variables input into a value between 0 and 1. This function is known as the logistic function.

Let the independent input features be:

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1m} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nm} \end{bmatrix}$$

And the dependent variable is Y having only binary value i.e. 0 or 1.

$$y = \begin{cases} 0 \, if \, class \, 1 \\ 1 \, if \, class \, 2 \end{cases}$$

Apply the multi-linear function to the input variables X.

$$z = \left(\sum\nolimits_{i=1}^{n} \omega_i x_i\right) + b$$

Then use the sigmoid function where the input will be z and find the predicted y between 0 and 1

$$\sigma\ (z) = \frac{1}{1 - e^{-z}}$$

Random forest modeling

The Random Forest (RF) algorithm is an ensemble classifier that utilizes classification or regression trees as base classifiers. It constructs an ensemble of decision trees by using the bootstrap sampling method to randomly select k samples with replacement from the original training

dataset N, creating k distinct training subsets. Each subset is then used to build a decision tree. By constructing these trees, the algorithm aims to address classification problems.

SVM modeling

Support Vector Machine (or SVM) is a machine learning technique used for classification tasks. Briefly, SVM works by identifying the optimal decision boundary that separates data points from different groups (or classes), and then predicts the class of new observations based on this separation boundary.Depending on the situations, the different groups might be separable by a linear straight line or by a non-linear boundary line.

ROC-AUC and PR-AUC

The efficacy of the models was evaluated using the receiver operating characteristic curve, area under the curve (ROC-AUC) and the precision-recall curve, area under the curve(PR-AUC).

ROC graphs are two-dimensional graphs where the true positive rate is plotted on the Y-axis and the false positive rate is plotted on the X-axis. An ROC graph illustrates the trade-offs between benefits (TPR, true positive rate) and costs (FPR, false positive rate).

A series of samples were sort in descending order based on the probability of being classified as positive. Starting from the highest probability and moving down, the "Score" value is used as the threshold. If the probability of a test sample belonging to the positive class is greater than this threshold, it is considered a positive sample; otherwise, it is considered a negative sample. Beginning with the first sample, set its score value as the threshold. Then, all subsequent samples, including the current one, are classified as negative. The TPR and FPR are calculated for each sample at the threshold, and the ROC curve is plotted. The area under curve (AUC) represents the area under the ROC curve, ranging between 0 and 1. A higher AUC value intuitively indicates better classifier performance.

Similarly, PR (Precision-Recall) graphs represent the graph which precision rate ($\frac{true\ positive}{true\ positive+false\ positive}$) is plotted on the Y axis and recall rate ($\frac{true\ positive}{true\ positive+false\ negative}$) is plotted on the X axis. The area under the PR curve (PR-AUC) represents the area under the PR curve, and a higher PR-AUC value also intuitively indicates better classifier performance.

The participants were randomly divided into training(5/6) and validation(1/6) sets. 5-fold cross-validation was performed on the training set, and the models were tested on the validation set. The details is shown in Fig. 1.

Feature identification by class-specific feature saliency map and lime explanation

We generated class-specific saliency maps to explore which types of patients were most and least likely to be predicted with LVDD. Feature identification was achieved by numerically generating a "synthetic patients data", which is representative of the class in terms of the above-mentioned classifier. Formally, let $S_c(p)$ be the score of the class c, computed by classifiers for a guide patient d; we would like to find an optimized patient, such that the $S_c(p)$ is highest [23, 24]. For the synthetic patients data that were most and least likely to be predicted with LVDD. We calculated the importance of features to predict results with lime(Local Interpretable Model-Agnostic Explanations) [25] and summed the weights of every factor of the 1000 generated patients.

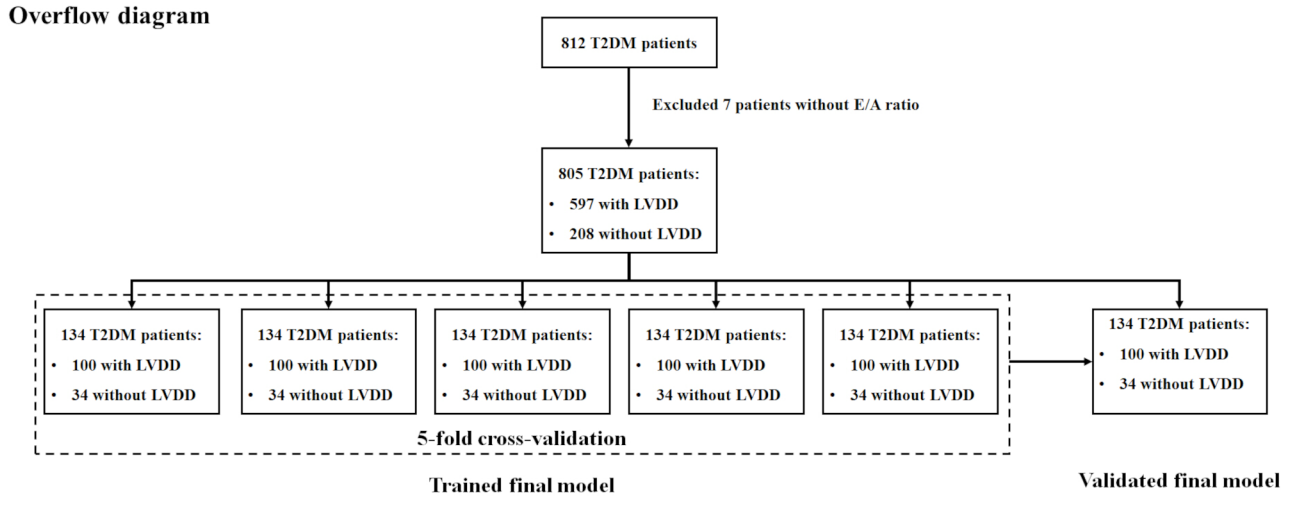

Fig. 1 The overflow diagram

Materials and code availability

The code and related datasets are freely available at https://github.com/ew314/LVDD_risk.

Results

Features characteristics and performance of ANN LVDD risk model

Among all 805 patients with type 2 diabetes mellitus (T2DM), 597 (74%) had LVDD. The laboratory test results and demographic characteristics of the participants are presented in Table 1. Significant differences were found in sex, smoking, drinking, diabetes peripheral neuropathy, the use of insulin and dipeptide kallidinogenase inhibitor IV, atherosclerosis, fatty liver disease, age, body mass index, systolic pressure, diastolic pressure, triglyceride, high density lipoprotein, low density lipoprotein, white blood cell count, neutrophil/lymphocyte ratio between the LVDD group and the control group($p < 0.05$). We further calculated the variance inflation factor(VIF) and performed spearman regression between features to check for multicollinearity (Table 2, Supplemental Figure. 1). These features showed less correlation. We remove the imbalanced age feature to ensure the generalizability of the model.

Compare to the logistic regression, random forest, and SVM algorithms, the ANN model showed the best performance in 5-fold-cross vaildation and the vaildation set(ROC-AUC = 0.741 ± 0.055/0.779, PR-AUC=0.881 ± 0.036/0.916), followed by logistic regression (ROC-AUC = 0.740 ± 0.052/0.753, PR-AUC = 0.882 ± 0.027/0.890), SVM(ROC-AUC = 0.723 ± 0.064/0.748, PR-AUC=0.863 ± 0.032/0.887) and random forest(ROC-AUC = 0.644 ± 0.069/0.661, PR-AUC=0.824 ± 0.045/0.867) (Fig. 2, Table 3). Therefore, the ANN model was chosen to build the LVDD risk prediction model.

The most risk/protective features of LVDD

To explored which features were the most risk/protective for the model's prediction. We generated 1000 fictional patients who were predicted to have LVDD(predicted

Table 1 Characteristics of the participants

	Control	LVDD	*P*-value
Male	151	374	$9.5*10^{-3}$
Famle	57	223	
24-Hour Urine Protein(mg/dL)	105.9(97.9 ~ 113.8)	107.8(102.5 ~ 113.0)	0.71
Age(year)	41.6(40.4 ~ 42.7)	59.2(58.6 ~ 59.9)	$2.5*10^{-112}$
Atherosclerosis(Y/N)	147/61	551/46	$2.6*10^{-15}$
Biguanides(Y/N)	115/93	375/222	0.06
Body Mass index(kg/m^2)	24.5(24.1 ~ 25.0)	23.7(23.5 ~ 23.9)	$1.5*10^{-4}$
Diabetes Nephropathy(Y/N)	88/120	266/331	0.57
Diabetes Peripheral Neuropathy(Y/N)	140/68	501/96	$3.0*10^{-7}$
Diabetes Retinopathy(Y/N)	30/178	110/487	0.19
Diastolic Pressure(mmHg)	76.5(75.5 ~ 77.5)	74.4(73.8 ~ 75.0)	$5.0*10^{-4}$
Dipeptide Kallidinogenase Inhibitor IV(Y/N)	14/194	95/502	$8.6*10^{-4}$
Drinking(Y/N)	52/156	109/488	0.04
Family History(Y/N)	99/109	250/347	0.15
Fatty Liver Disease(Y/N)	118/90	218/379	$3.6*10^{-7}$
High density lipoprotein(mmol/L)	1.0(1.0 ~ 1.0)	1.1(1.1 ~ 1.1)	$1.7*10^{-4}$
High-sensitive cardiac troponin(ng/mL)	7.1(6.7 ~ 7.5)	7.3(7.1 ~ 7.6)	0.3
Insulin(Y/N)	45/163	197/400	$2.1*10^{-3}$
Low density lipoprotein(mmol/L)	2.8(2.7 ~ 2.9)	2.6(2.6 ~ 2.7)	$3.6*10^{-3}$
Neutrophil/lymphocyte ratio(%)	1.6(1.5 ~ 1.6)	1.6(1.6 ~ 1.7)	0.04
Smoking(Y/N)	86/122	181/416	$3.6*10^{-3}$
Sodium glucose transporter 2 inhibitor(Y/N)	32/176	116/481	0.19
Systolic Pressure(mmHg)	116.5(115.2 ~ 117.8)	118.3(117.4 ~ 119.2)	0.03
Triglyceride(mmol/L)	1.6(1.5 ~ 1.6)	1.4(1.4 ~ 1.5)	0.04
Urine microalbumin/Creatinine ratio(%)	6.8(6.2 ~ 7.3)	7.0(6.5 ~ 7.4)	0.61
Vitamin D(ng/ml)	15.6(14.9 ~ 16.4)	16.3(15.8 ~ 16.8)	0.15
Waist-Hip Ratio	0.9(0.9 ~ 0.9)	0.9(0.9 ~ 0.9)	0.56
α Glycosidase inhibitor(Y/N)	38/170	202/395	$2.4*10^{-5}$
White blood cell count($10^9/L$)	6.2(6.0 ~ 6.4)	5.8(5.7 ~ 5.9)	$3.4*10^{-4}$

Control: Type 2 diabetes patients with no cardiac complications

LVDD: Type 2 diabetes patients with left ventricular diastolic dysfunction(LVDD)

P-value: Chi-square tests for category data and T-tests for numerical data

Table 2 VIF of all characteristics

	VIF
SEX	1.8
Smoking	1.8
Age	1.7
Drinking	1.5
Body Mass index	1.4
Diastolic Pressure	1.4
Fatty Liver Disease	1.4
High density lipoprotein	1.4
Systolic Pressure	1.4
Triglyceride	1.4
Waist-Hip Ratio	1.4
Low density lipoprotein	1.3
Atherosclerosis	1.2
High-sensitive cardiac troponin	1.2
White blood cell count	1.2
Biguanides	1.1
Diabetes Nephropathy	1.1
Diabetes Peripheral Neuropathy	1.1
Diabetes Retinopathy	1.1
Dipeptide Kallidinogenase Inhibitor IV	1.1
Insulin	1.1
Neutrophil/lymphocyte ratio	1.1
Sodium glucose transporter 2 inhibitor	1.1
Urine microalbumin/Creatinine ratio	1.1
Vitamin D	1.1
α Glycosidase inhibitor	1.1
24-Hour Urine Protein	1.0
Family History	1.0

VIF: Variance inflation factor

probability > 0.999) and 1000 fictional patients who were predicted not to have LVDD(predicted probability < 0.001). All features were normalized into Z-score, with the zero baseline representing the mean values(Fig. 3A). We then further calculated the feature contribution weights to the model's prediction results for each patient using the LIME explainer. The contribution weights were summed, and the weights with absolute values higher then 20 were marked as red(risk) or blue(protective). We found the 'DBP > 0.67', 'LDL > 0.68', 'SBP<=-0.75' were the most protective features, while 'VITD > 0.49', 'SBP > 0.68', 'DBP<=-0.64', 'HDL > 0.54' and 'WHR > 0.59' were most risky features.

The web server for LVDD risk prediction

Finally, we developed a web server (http://bigc.online:5000/) for users to predict their LVDD risk (Fig. 4). Patients could input their own information or choose it from a list. Considerating that most patients would not have all feature information, We performed leave-one-out feature selection to features required for model prediction (Fig. 5). These 7 most important features were required for web server, while other missing features were replaced with means. The result page displays a summary of LVDD risk probability, and two figures displayed the deviation of user information from the population means.

Discussion

In the present study, we created a neural network model based on 27 clinical information to predict LVDD risk among T2DM patients. The performance of the neural network were better than classical machine learning model such as logistic regression, random forest and SVM. We then explore the most risk/protective features with interpretability algorithm in neural networks based on the built neural network model. we developed a web server so that users could easily predict their LVDD risk probability and receive warnings about risk factors.

Cardiovascular complications could only be prevented or reversed at an early stage [26]. In comparison to current research, which primarily focuses on building models to predict cardiovascular events [16, 17], our aim was to build a model to predict LVDD risk for T2DM patients, which is an early sign of cardiovascular events [27, 28]. So that the alteration in myocardial structure could be detect at an earlier stage, leading to timely diagnosis and intervention. We built the model using an ANN model, which has the ability to extract high-level features from data through multi-level nonlinear mapping [29], giving it better performance than the three classical models. However, neural networks are black box models with complex internal operation processes, making it difficult to understand and explain the model's decision-making process. Therefore, we used two interpretability methods in neural networks, class-specific feature saliency maps [23, 24]and lime explainer [25], to find the key features for model prediction and the importance of risk/protective factors.

We find the higher HDL, VitD, SBP and WHR, and lower DBP, was the most risk factors for LVDD patients. These factors obtained in this study were similar to the results of previous studies. For example, the high SBP and low DBP have already been regenzed as cardiovascular risk factors [30], and their combination showed better performance. High WHR [31] and VitD [32] were also well-known risk factors for cardiovascular disease. However, our results showed that a high level of HDL was a risk factor, while high HDL level has been widely accepted as a protective factor [33]. These results seems contrary to common sense [34]. Many observational studies have demonstrated that low levels of HDL-C are associated with an increased risk of coronary heart disease [35]. However, the protective role of HDL cholesterol (HDL-C) has been seriously challenged by the evidence from recent genetic, epidemiologic and clinical trials. Nathalie Pamir found that lower HDL levels only predicted an increased risk of cardiovascular disease for

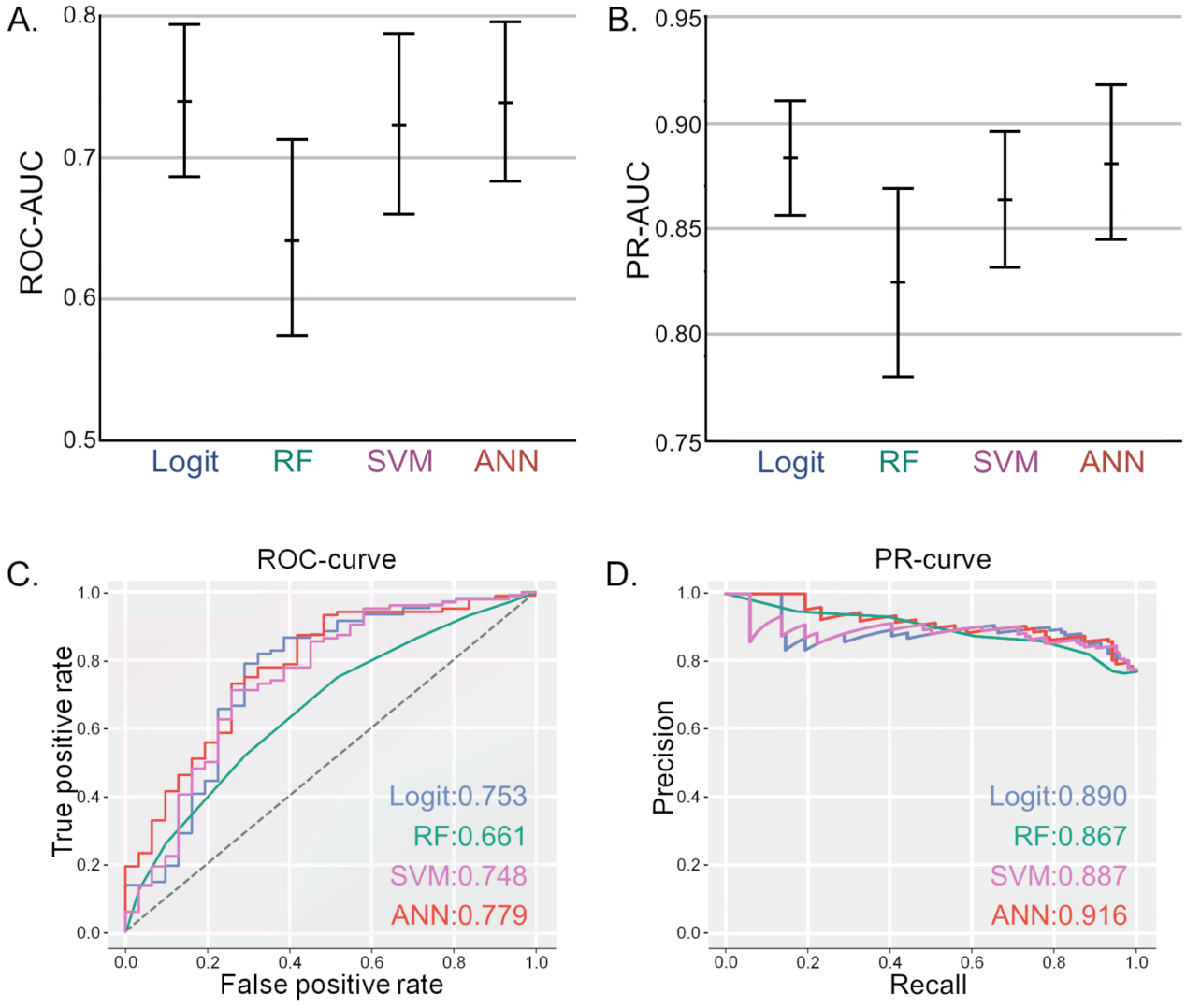

Fig. 2 **A&B**: Model performance (ROC-AUC&PR-AUC) of the logistic regression(Logit), random forest(RF), support vector machine(SVM) and artificial neural network(ANN) model in the 5-fold cross-validation process. **C&D**: Model performance (ROC-AUC&PR-AUC) in validation dataset

Table 3 ROC-AUC and PR-AUC of 5-cross-vaildated

Exp	Logistic Regression		Random forest		SVM		ANN	
	ROC-AUC	PR-AUC	ROC-AUC	PR-AUC	ROC-AUC	PR-AUC	ROC-AUC	PR-AUC
1	0.82	0.92	0.69	0.85	0.79	0.89	0.83	0.93
2	0.70	0.87	0.58	0.78	0.67	0.82	0.73	0.89
3	0.72	0.88	0.58	0.81	0.75	0.89	0.70	0.86
4	0.75	0.86	0.66	0.79	0.75	0.86	0.73	0.85
5	0.70	0.88	0.71	0.88	0.66	0.85	0.72	0.88
Fin	0.75	0.89	0.66	0.87	0.75	0.89	0.78	0.92

Exp: Every round experiment of 5-fold cross inspection

Fin: Finally vaildation

white adults and not for black adults [36]. Large-scale prospective cohort studies also contradict the previous finding of a linear inverse relationship between HDL and cardiovascular disease [7, 16–18]. Although it is a common finding that low levels of HDL predict increased cardiovascular risk, data from several cohorts have revealed a suggestion of increased cardiovascular outcomes in those with extremely high HDL levels [37]. High HDL-C was reported to be associated with increased cardiovascular risk in hypertensive patients [38], while

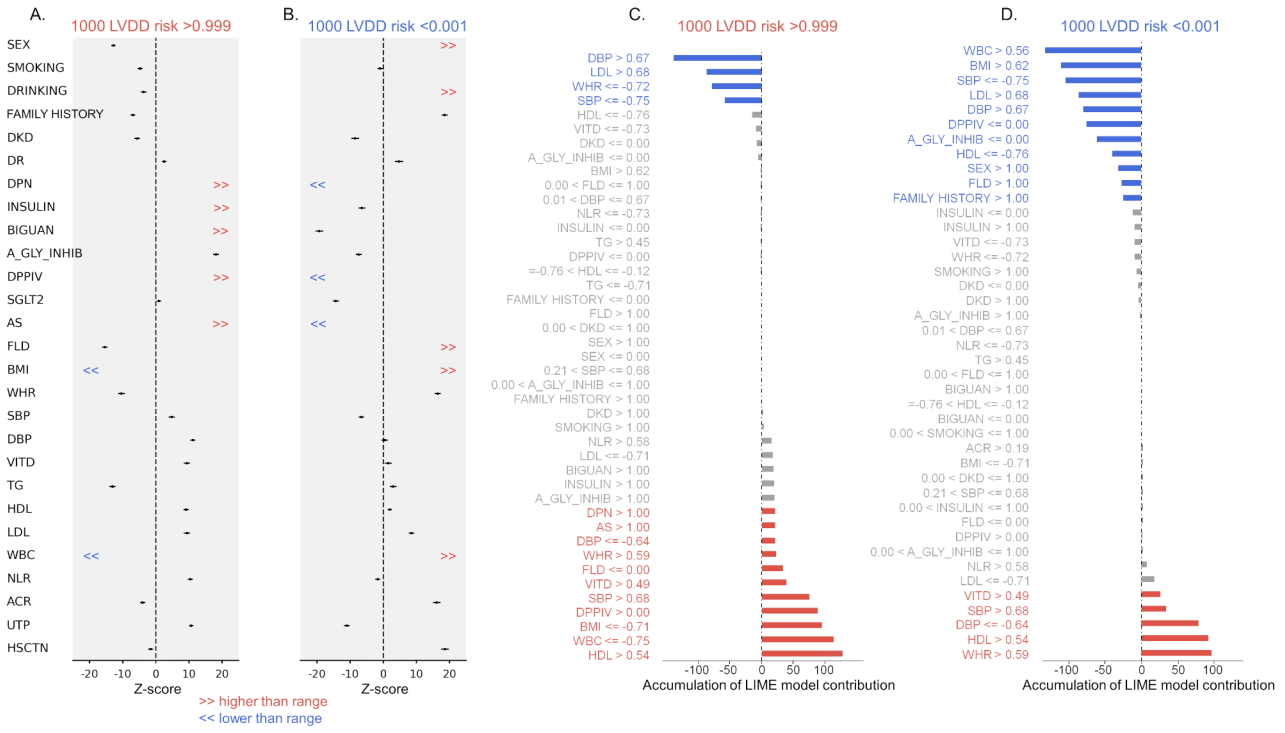

Fig. 3 **A&B**: Distribution of 1000 most/ least likely to have LVDD-risk patients' features. **C&D**: Accumulation of weights of features for the 1000 most/ least likely to have LVDD-risk patients. Absolute values higher then 20 were marked as red(risk)/protective(blue)

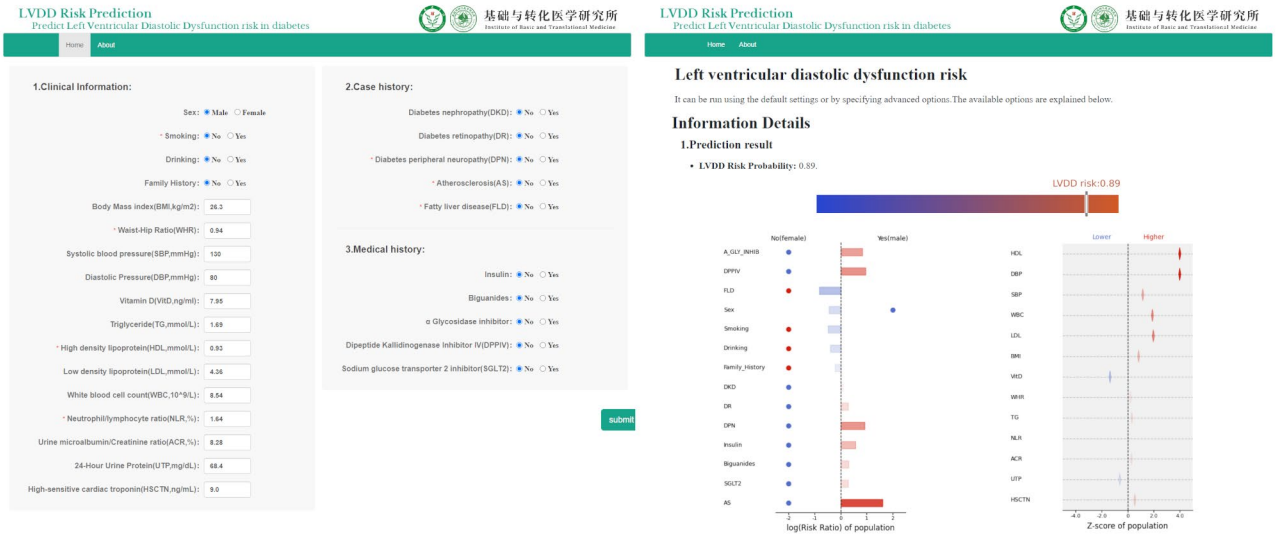

Fig. 4 Home page and result page of the web tools

hypertension is a common symptom of T2DM patients [39]. A study of the northern China population [40], detected that HDL-c levels showed a U-shaped relationship with all-cause mortality in younger participants(< 65 years old), and very high HDL-c levels(≥ 80 mg/dl) were independently associated with an increased total mortality risk. In our data, the HDL of LVDD patients were significantly higher than those without LVDD(ANOVA oneway test, P-value = 1.66×10^{-4}). We suggest that the association between HDL and LVDD risk in T2DM patients needs more research.

In Table 4, we summarize current studies that predict left ventricular dysfunction in T2DM patients. Studies of Soh, C.H., etc [41]., Chee KH, etc [42]. and Yan, Wf., etc [43]. primarily rely on instrumental measurements to predict LVDD in patients with diabetes. Halabi, A.,

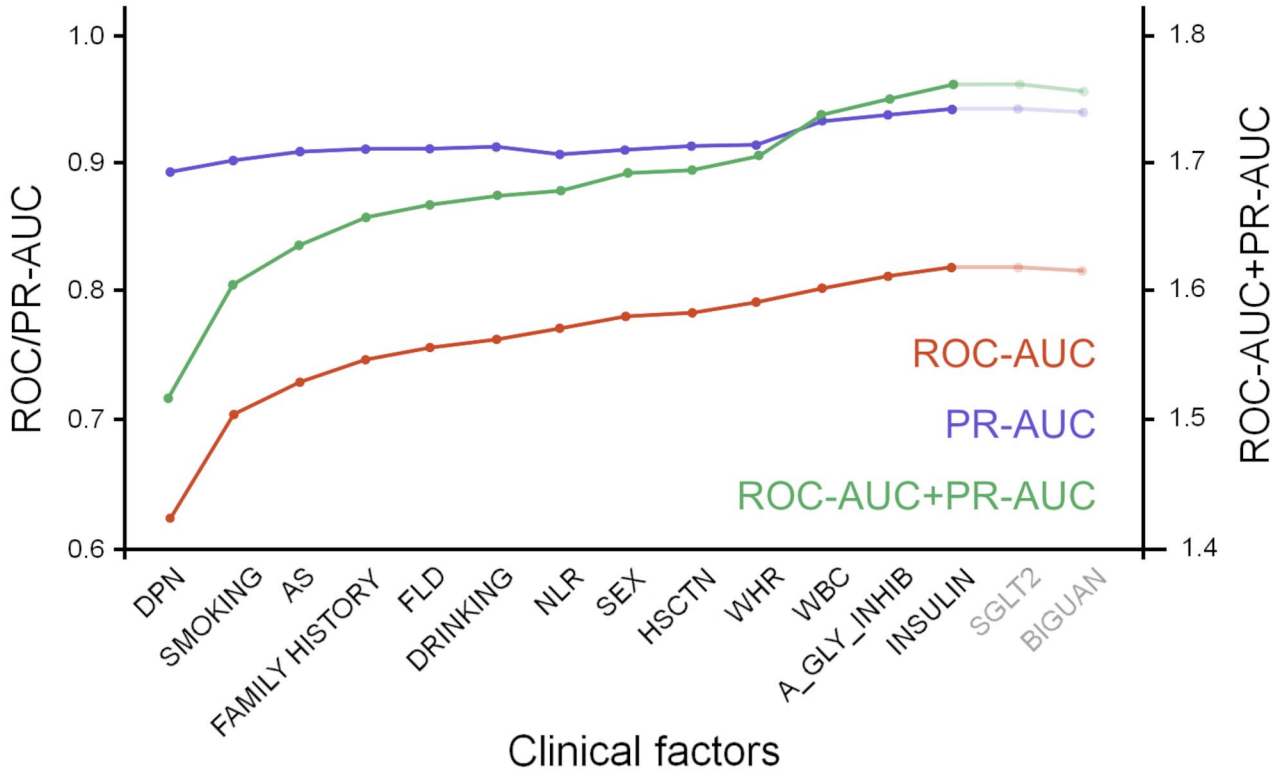

Fig. 5 Leave-one-out feature selection

Table 4 List of relevant studies on predicting left ventricular dysfunction in T2DM patients

Source	Input features	Output prediction	Participants	Year
Soh, C.H., etc. [41]	energy waveform electrocardiogram	subclinical left ventricular dysfunction	178 + 97	2024
Chee KH, etc. [42]	echocardiographic	left ventricular diastolic dysfunction	301	2021
Yan, Wf., etc. [43]	volume-time curve of cardiac magnetic resonance	left ventricular diastolic function	48	2021
Halabi, A., etc. [44]	36 clinical features	subclinical left ventricular dysfunction	804	2022
Hao, M., etc. [45]	41 clinical features	diastolic cardiac dysfunction	3030	2023
Chen Y, etc. [46]	10 clinical features	left ventricular diastolic function	84	2022

etc [44]. used 36 clinical indicators to model and found that natriuretic peptides and troponin are risk factors for subclinical LV dysfunction. Hao, M., etc [45]. checked the association between 41 clinical indicators and diastolic cardiac dysfunction and established a risk prediction model using six independent parameters: age, BMI, triglyceride, creatine phosphokinase isoenzyme, serum sodium, and urinary albumin/creatinine ratio. Chen Y, etc [46]. analyzed the association between 10 biochemical indicators and LVDD and found a significant increase in systolic blood pressure, cholesterol, LDL cholesterol, and fasting glucose in LVDD patients. However, in our results, while systolic blood pressure also emerged as a risk factor, LDL was found to be a protective factor for LVDD.

Limitations of the current study should be addressed. Our data only has outcome information without follow-time. Thus, we could only build classification model to predict LVDD outcome risk but without the time of onset. The LVDD diagnosis came from the echocardiography in medical records, which only provided classified labels, missing the detailed E/A ratio value. We could have further explored the features with linearly increasing LVDD risk if the E/A ratio value was available. We hope this issue will be addressed by collecting a larger sample size.

Conclusion

In summary, we contrust an ANN model to predict LVDD-risk for T2DM patients. This enables us to diagnose and detect myocardial damage in T2DM patients with early-stage heart failure.

Based on the built model, We explored the risk/protective factors of LVDD with interpretability methods in neural network. We find the higher HDL, VitD, SBP and WHR, and lower DBP, was the most risk factors for

LVDD patients. We also provided a freely accessible web server named LVDD-risk(http://www.bigc.online/ExsgRNA; http://www.xmuptcgd.top/LVDD) for users to predict LVDD risk.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13098-025-01714-8.

Supplementary Material 1: Figure S1: Spearman rank correlation among features.

Acknowledgements
We deeply appreciate the data support provided by the Shaanxi Provincial People's Hospital,.

Author contributions
Wei Liu collected the data. Yu Rong wrote the main manuscript text. All authors reviewed the manuscript.

Funding
This study is supported by the Research and Development Program of Innovation Chain for Key industries in Shaanxi Province(Program No. 2021ZDLSF02-09). Key Scientific Research Program of the Education Department in Shaanxi Province(Program No. 23JS054).

Declarations

Ethics approval and consent to participate
Ethical approval was obtained from the ethics committee of Xi'an Medical University, and all participants provided informed consent before taking part in the study.

Competing interests
The authors declare no competing interests.

Received: 18 August 2023 / Accepted: 25 April 2025
Published online: 17 May 2025

References

1. Ma C-X, Ma X-N, Guan C-H, Li Y-D, Mauricio D, Fu S-B. Cardiovascular disease in type 2 diabetes mellitus: progress toward personalized management. Cardiovasc Diabetol. 2022;21(1):74.
2. Faselis C, Katsimardou A, Imprialos K, Deligkaris P, Kallistratos M, Dimitriadis K. Microvascular complications of type 2 diabetes mellitus. Curr Vasc Pharmacol. 2020;18(2):117–24.
3. Viigimaa M, Sachinidis A, Toumpourleka M, Koutsampasopoulos K, Alliksoo S, Titma T. Macrovascular complications of type 2 diabetes mellitus. Curr Vasc Pharmacol. 2020;18(2):110–6.
4. Cui J, Liu Y, Li Y, Xu F, Liu Y. Type 2 diabetes and myocardial infarction: recent clinical evidence and perspective. Front Cardiovasc Med. 2021;8:644189.
5. Lee Y, Fluckey JD, Chakraborty S, Muthuchamy M. Hyperglycemia- and hyperinsulinemia-induced insulin resistance causes alterations in cellular bioenergetics and activation of inflammatory signaling in lymphatic muscle. Faseb J. 2017;31(7):2744–59.
6. Giacco F, Brownlee M. Oxidative stress and diabetic complications. Circ Res. 2010;107(9):1058–70.
7. Bahadoran Z, Mirmiran P, Ghasemi A. Role of nitric oxide in insulin secretion and glucose metabolism. Trends Endocrinol Metab. 2020;31(2):118–30.
8. Thomas MC. Type 2 diabetes and heart failure: challenges and solutions. Curr Cardiol Rev. 2016;12(3):249–55.
9. Liu R, Li L, Shao C, Cai H, Wang Z. The Impact of Diabetes on Vascular Disease: Progress from the Perspective of Epidemics and Treatments. *J Diabetes Res* 2022, 2022:1531289.
10. Herman WH, Ye W, Griffin SJ, Simmons RK, Davies MJ, Khunti K, Rutten GE, Sandbaek A, Lauritzen T, Borch-Johnsen K, et al. Early detection and treatment of type 2 diabetes reduce cardiovascular morbidity and mortality: A simulation of the results of the Anglo-Danish-Dutch study of intensive treatment in people with Screen-Detected diabetes in primary care (ADDITION-Europe). Diabetes Care. 2015;38(8):1449–55.
11. Galaviz KI, Narayan KMV, Lobelo F, Weber MB. Lifestyle and the prevention of type 2 diabetes: A status report. Am J Lifestyle Med. 2018;12(1):4–20.
12. Grigorescu ED, Lacatusu CM, Floria M, Mihai BM, Cretu I, Sorodoc L. Left ventricular diastolic dysfunction in type 2 Diabetes-Progress and perspectives. Diagnostics (Basel) 2019, 9(3).
13. von Bibra H, St John Sutton M. Diastolic dysfunction in diabetes and the metabolic syndrome: promising potential for diagnosis and prognosis. Diabetologia. 2010;53(6):1033–45.
14. Hassan Ayman KM, Abdallah Mahmoud A, Abdel-Mageed Eman A, Marwa S, Soliman Mona M, Kishk Yehia T. Correlation between left ventricular diastolic dysfunction and dyslipidaemia in asymptomatic patients with new-onset type 2 diabetes mellitus. Egypt J Intern Med. 2021;33(1):8.
15. Tanaka H, Tatsumi K, Matsuzoe H, Matsumoto K, Hirata K. -i: impact of diabetes mellitus on left ventricular longitudinal function of patients with non-ischemic dilated cardiomyopathy. Cardiovasc Diabetol. 2020;19(1):84.
16. Fan R, Zhang N, Yang L, Ke J, Zhao D, Cui Q. AI-based prediction for the risk of coronary heart disease among patients with type 2 diabetes mellitus. Sci Rep. 2020;10(1):14457.
17. Mansoori A, Sahranavard T, Hosseini ZS, Soflaei SS, Emrani N, Nazar E, Gharizadeh M, Khorasanchi Z, Effati S, Ghamsary M, et al. Prediction of type 2 diabetes mellitus using hematological factors based on machine learning approaches: a cohort study analysis. Sci Rep. 2023;13(1):663.
18. DESAI U, MARTIS RJ, ACHARYA UR, NAYAK CG, SESHIKALA G, K RS:, DIAGNOSIS OF MULTICLASS TACHYCARDIA BEATS USING RECURRENCE QUANTIFICATION ANALYSIS AND ENSEMBLE CLASSIFIERS. J Mech Med Biology. 2016;16(01):1640005.
19. Desai U, Nayak CG, Seshikala G, Martis RJ, Fernandes SL. Automated diagnosis of tachycardia beats. Smart computing and informatics: 2018// 2018; Singapore. Springer Singapore; 2018. pp. 421–9.
20. Desai U, Nayak CG, Seshikala G. Application of ensemble classifiers in accurate diagnosis of myocardial ischemia conditions. Progress Artif Intell. 2017;6(3):245–53.
21. Desai U, Nayak CG, Seshikala G, Martis RJ. Automated diagnosis of coronary artery disease using pattern recognition approach. Annu Int Conf IEEE Eng Med Biol Soc. 2017;2017:434–7.
22. Nagueh SF, Smiseth OA, Appleton CP, Byrd BF 3rd, Dokainish H, Edvardsen T, Flachskampf FA, Gillebert TC, Klein AL, Lancellotti P, et al. Recommendations for the evaluation of left ventricular diastolic function by echocardiography: an update from the American society of echocardiography and the European association of cardiovascular imaging. Eur Heart J Cardiovasc Imaging. 2016;17(12):1321–60.
23. Simonyan K, Vedaldi A, Zisserman A. Deep inside convolutional networks: visualising image classification models and saliency maps. Comput Sci 2013.
24. Chuai G, Ma H, Yan J, Chen M, Hong N, Xue D, Zhou C, Zhu C, Chen K, Duan B, et al. DeepCRISPR: optimized CRISPR guide RNA design by deep learning. Genome Biol. 2018;19(1):80.
25. Ribeiro MT, Singh S, Guestrin C. Why should I trust you?? Explaining the predictions of any classifier. In.; 2016: 97–101.
26. Chiesa ST, Marcovecchio ML. Preventing cardiovascular complications in type 1 diabetes: the need for a lifetime approach. Front Pediatr. 2021;9:696499.
27. Nagueh SF. Left ventricular diastolic function: Understanding pathophysiology, diagnosis, and prognosis with echocardiography. JACC Cardiovasc Imaging. 2020;13(1 Pt 2):228–44.
28. Zhou D, Yan M, Cheng Q, Feng X, Tang S, Feng Y. Prevalence and prognosis of left ventricular diastolic dysfunction in community hypertension patients. BMC Cardiovasc Disord. 2022;22(1):265.
29. Alzubaidi L, Zhang J, Humaidi AJ, Al-Dujaili A, Duan Y, Al-Shamma O, Santamaría J, Fadhel MA, Al-Amidie M, Farhan L. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. J Big Data. 2021;8(1):53.
30. Schillaci G, Pirro M, Mannarino E. Assessing cardiovascular risk: should we discard diastolic blood pressure? Circulation. 2009;119(2):210–2.
31. Gill D, Zuber V, Dawson J, Pearson-Stuttard J, Carter AR, Sanderson E, Karhunen V, Levin MG, Wootton RE, Klarin D, et al. Risk factors mediating the effect of body mass index and waist-to-hip ratio on cardiovascular outcomes: Mendelian randomization analysis. Int J Obes. 2021;45(7):1428–38.

32. Rai V, Agrawal DK. Role of vitamin D in cardiovascular diseases. Endocrinol Metab Clin North Am. 2017;46(4):1039–59.
33. Gordon T, Castelli WP, Hjortland MC, Kannel WB, Dawber TR. High density lipoprotein as a protective factor against coronary heart disease. The Framingham study. Am J Med. 1977;62(5):707–14.
34. Stone NJ, Robinson JG, Lichtenstein AH, Bairey Merz CN, Blum CB, Eckel RH, Goldberg AC, Gordon D, Levy D, Lloyd-Jones DM, et al. 2013 ACC/AHA guideline on the treatment of blood cholesterol to reduce atherosclerotic cardiovascular risk in adults: a report of the American college of cardiology/american heart association task force on practice guidelines. J Am Coll Cardiol. 2014;63(25 Pt B):2889–934.
35. Huxley RR, Barzi F, Lam TH, Czernichow S, Fang X, Welborn T, Shaw J, Ueshima H, Zimmet P, Jee SH, et al. Isolated low levels of high-density lipoprotein cholesterol are associated with an increased risk of coronary heart disease: an individual participant data meta-analysis of 23 studies in the Asia-Pacific region. Circulation. 2011;124(19):2056–64.
36. Zakai NA, Minnier J, Safford MM, Koh I, Irvin MR, Fazio S, Cushman M, Howard VJ, Pamir N. Race-Dependent association of High-Density lipoprotein cholesterol levels with incident coronary artery disease. J Am Coll Cardiol. 2022;80(22):2104–15.
37. Hirata A, Sugiyama D, Watanabe M, Tamakoshi A, Iso H, Kotani K, Kiyama M, Yamada M, Ishikawa S, Murakami Y, et al. Association of extremely high levels of high-density lipoprotein cholesterol with cardiovascular mortality in a pooled analysis of 9 cohort studies including 43,407 individuals: the EPOCH-JAPAN study. J Clin Lipidol. 2018;12(3):674–e684675.
38. Trimarco V, Izzo R, Morisco C, Mone P, Virginia Manzi M, Falco A, Pacella D, Gallo P, Lembo M, Santulli G, et al. High HDL (High-Density Lipoprotein) cholesterol increases cardiovascular risk in hypertensive patients. Hypertension. 2022;79(10):2355–63.
39. Petrie JR, Guzik TJ, Touyz RM. Diabetes, hypertension, and cardiovascular disease: clinical insights and vascular mechanisms. Can J Cardiol. 2018;34(5):575–84.
40. Li X, Guan B, Wang Y, Tse G, Zou F, Khalid BW, Xia Y, Wu S, Sun J. Association between high-density lipoprotein cholesterol and all-cause mortality in the general population of Northern China. Sci Rep. 2019;9(1):14426.
41. Soh CH, de Sá AGC, Potter E, Halabi A, Ascher DB, Marwick TH. Use of the energy waveform electrocardiogram to detect subclinical left ventricular dysfunction in patients with type 2 diabetes mellitus. Cardiovasc Diabetol. 2024;23(1):91.
42. Chee KH, Tan KL, Luqman I, Saiful SS, Chew YY, Chinna K, Tan ATB. Prevalence and predictors of left ventricular diastolic dysfunction in Malaysian patients with type 2 diabetes mellitus without prior known cardiovascular disease. Front Cardiovasc Med. 2021;8:676862.
43. Yan W-f, Gao Y, Zhang Y, Guo Y-k, Wang J, Jiang L, Li Y, Yang Z-g. Impact of type 2 diabetes mellitus on left ventricular diastolic function in patients with essential hypertension: evaluation by volume-time curve of cardiac magnetic resonance. Cardiovasc Diabetol. 2021;20(1):73.
44. Halabi A, Potter E, Yang H, Wright L, Sacre JW, Shaw JE, Marwick TH. Association of biomarkers and risk scores with subclinical left ventricular dysfunction in patients with type 2 diabetes mellitus. Cardiovasc Diabetol. 2022;21(1):278.
45. Hao M, Huang X, Liu X, Fang X, Li H, Lv L, Zhou L, Guo T, Yan D. Novel model predicts diastolic cardiac dysfunction in type 2 diabetes. Ann Med. 2023;55(1):766–77.
46. Chen Y, Yu M, Lan Y, Feng F, Jiang C. Development of a nomogram for predicting the risk of left ventricular diastolic function in subjects with type-2 diabetes mellitus. Int J Cardiovasc Imaging. 2022;38(1):15–23.

Publisher's note