Vascular Health and Risk Management

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

Machine Learning-Based Prediction Model for Predicting the Effect of the Serum γKlotho Level on Susceptibility to Coronary Heart Disease

Zi-Tong Guo[1], Xiao-Lin Yu[2], Hui Cheng[2], Tuersunjiang Naman[2]

[1]Department of Cardiology, First Affiliated Hospital of Xinjiang Medical University, Urumqi, Xinjiang, People's Republic of China; [2]Department of Cardiology, People's Hospital of Xinjiang Uygur Autonomous Region, Urumqi, Xinjiang, People's Republic of China

Correspondence: Tuersunjiang Naman, Email tursunjan1016@163.com

Objective: This study investigates the relationship between serum γKlotho levels and coronary heart disease (CHD) risk and develops a machine learning model for CHD prediction.

Methods: A total of 1435 subjects were enrolled for analysis and randomized as training (n = 969, 70%) or validation (n = 466, 30%) group. The training group was used for univariate regression. Thereafter, least absolute shrinkage and selection operator (LASSO) regression was conducted for selecting independent risk factors for CHD. Using independent risk factors for CHD, nine machine learning models were developed, the best model was selected by evaluating them, and the model was validated by decision curve analysis (DCA).

Results: The factors independently associated with CHD risk were age, the serum level of γKlotho, LDL-C, sex, diabetes, hypertension, and smoking status. We used these risk factors to construct nine popular machine-learning models. Among all models, the RF model was better appropriate; thus, we visualized and validated this model, which showed promising clinical application.

Conclusion: Serum γKlotho levels are novel biomarker which positively related to CHD risk. Additionally, the RF model can better predict the risk of CHD, and RF model is better appropriate to predicting the CHD risk in clinics.

Keywords: coronary heart disease, γKlotho, random forest, RF, prediction model

Introduction

Coronary heart disease (CHD) is the main factor associated with morbidity and mortality globally.[1] There are various risk factors related to CHD, including clinical and biological risk factors.[1] The known CAD-related risk factors include diabetes, hypertension, renal dysfunction, dyslipidemia, and a history of smoking cigarettes.[2,3] Coronary atherosclerosis and calcium accumulation into coronary artery to different degrees are some of biological risk factors for CHD;[4] the clinical and biological causes of CAD are not limited to those described above. There are various unknown risk factors for CHD. A recent study[5] reported that: serum Klotho level is a risk factor for CHD. According to a population-based study, serum klotho level is negatively associated with risk of CHD.[6] Understanding the interplay between these clinical as well as biological risk and novel biomarkers like γKlotho is crucial for comprehensive risk assessment of CHD.

Encoded by Klotho gene, Klotho protein is present in parathyroid glands, brain, kidney, skeletal muscle, reproductive organs, vasculature, reproductive glands, and urinary bladder.[7,8] Klotho exists in membranous and soluble forms. The membranous klotho targets phosphatonin fibroblast growth factor (FGF)-23. Moreover, secreted klotho (70 kDa) can be produced through alternative RNA splicing, which can release the whole cleaved extracellular domain.[9] Its soluble type is distributed in blood, urine, and cerebrospinal fluid due to transmembrane-type shedding via different cellular surface metalloproteases.

Received: 3 December 2024
Accepted: 17 May 2025
Published: 27 May 2025

The klotho family includes αKlotho, βKlotho, and γKlotho, and the serum levels of αKlotho[10] and β-Klotho[11] are correlated with the risk of CHD. However, no study has revealed the association of serum γKlotho level with CHD risk or the specific function of γKlotho.

Machine learning (ML) algorithms are particularly advantageous for analyzing complex datasets, as they can identify nonlinear relationships and interactions among multiple variables, and it can provide researchers with powerful tools to accurately predict the risk of diseases. Thus, it can be used for various medical applications, like diagnosis, treatment, medical image explanation, and outcome estimation.[12,13] It makes them ideal for predicting multifactorial diseases like CHD, but no study has developed an ML-based prediction model to predict the relation of serum γKlotho level with CHD risk.

Thus, the present work focused on investigating the correlation of the serum γKlotho level with CHD risk and develop an ML-based prediction model for predicting CHD risk.

Materials and Methods

Study Design and Subjects

There were altogether 1435 cases included from the People's Hospital of Xinjiang Uyghur Autonomous Region between January 2013 and December 2018. All patients provided written informed consent for participation. Among all participants, 921 were CHD patients and 514 were healthy individuals, according to our preset eligibility criteria (Figure 1).

All patients received coronary angiography before or during hospitalization. CAD patients whose coronary angiography suggested >50% luminal stenosis in at least one coronary artery, who previously underwent coronary artery

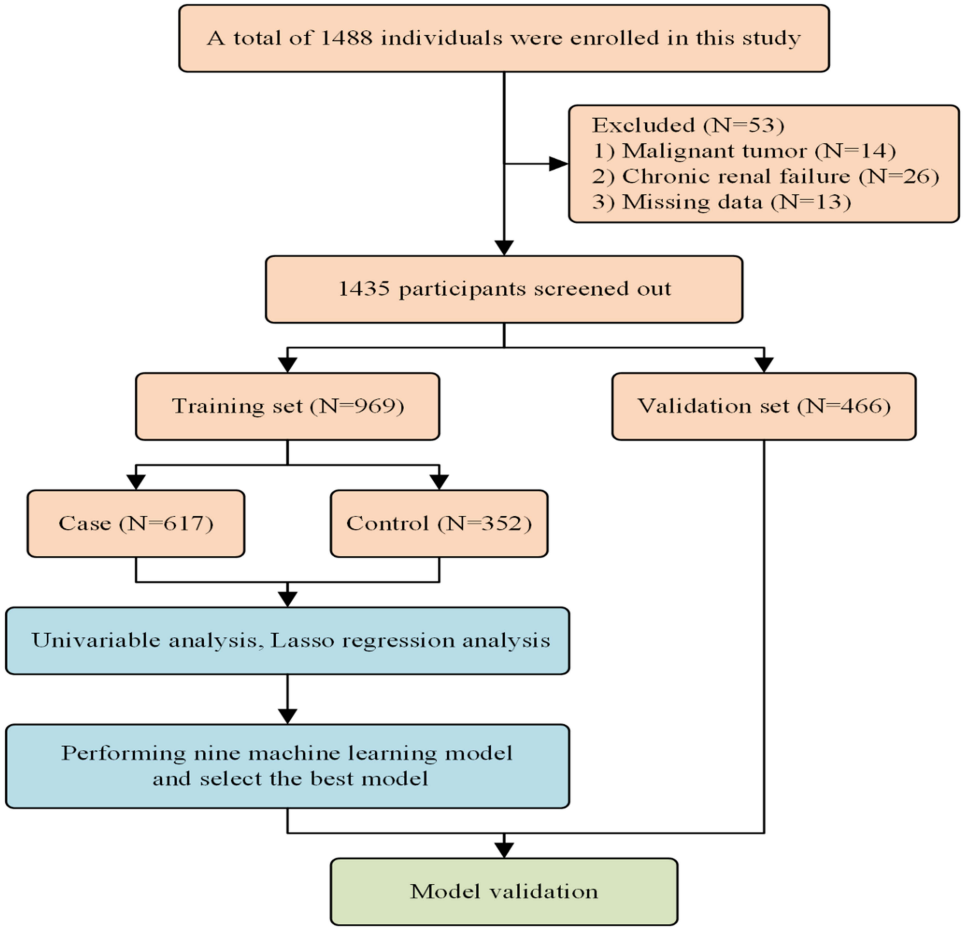

Figure 1 A study flowchart.

bypass grafting or percutaneous coronary intervention, and whose angina symptoms improved following rest or nitroglycerin administration were enrolled. CHD cases meeting criteria below were eliminated: unstable hemodynamics, acute decompensated HF, liver/kidney/autoimmune/hematological disorders, noncardiac disorders with life expectancy below one year, cachexia, and unwillingness to take part in the present work.

Healthy controls included in the study had the following characteristics: those having <50% coronary artery luminal stenosis validated through coronary angiography performed by at least two senior cardiologists, with no additional diseases, or with angina on exertions.

Blood Sample Collection and Laboratory Tests

Blood was sampled in every CAD patient and healthy control one day post-admission and to examine blood parameters like white blood cell (WBC) count and platelet (PLT) count via electrical impedance analysis (EIA), hemoglobin concentration, chemical colorimetry, total cholesterol (TC), triglyceride (TG), creatinine (CR), blood urea nitrogen (BUN), and high-density and low-density lipoprotein-cholesterol (HDL-C and LDL-C, respectively) using a chemiluminescence method at the laboratory of the People's Hospital of Xinjiang Uyghur Autonomous Region.

Measurement of the Plasma Level of γKlotho

For separating the plasma and blood cells, the blood samples underwent 10 min of centrifugation at 1500 rpm with anticoagulant ethylenediaminetetraacetic acid (EDTA) via an Eppendorf high-speed centrifuge. Additionally, blood plasma was maintained under −80°C before testing serum γKlotho level.

The serum γKlotho content was analyzed with a γKlotho ELISA kit (mlbio, Shanghai China); the examiners were blinded to clinical data of participants. For monitoring genotyping quality, 10% of the samples were duplicated.

Definitions of Cardiovascular Risk Factors

Body mass index (BMI) can be computed through dividing body weight (kg) by body height square (m). Smokers were people who smoked for at least six months or within the last six months. Drinkers were people consuming >100 g alcoholic beverages weekly during the previous month. The definition of hypertension was made in line with the 2011 New NICE guidelines for hypertension,[14] which include systolic blood pressure (SBP) ≥140 mmHg, diastolic blood pressure (DBP) ≥90 mmHg, or the antihypertensive medications over the last two weeks. The definition of diabetes mellitus (DM) was made as a glucose level of 11.1 mmol/L (200 mg/dL) 2 h following oral glucose use (75 g), a fasting plasma glucose level ≥7.0 mmol/L (126 mg/dL), antidiabetic medications, or a history of diabetes.

Statistical Analysis

Normally distributed continuous data were represented by mean ± SD and compared by Student's unpaired *t*-tests, whereas non-normally-distributed ones were represented by medians with interquartile ranges and analyzed by Mann–Whitney *U*-tests. Categorical data were presented as numbers with proportions and examined by Chi-square tests.

Patients were randomized into the training (n = 969, 70%) or validation (n = 466, 30%) group. Patients in training group were further classified in case (n = 617) or control (n = 352) subgroup.

In the training group, we performed univariate analysis and then selected variables satisfying $P < 0.2$ for LASSO regression analysis. Feature selection was conducted for dimensionality reduction,[15] with those satisfying nonzero coefficients upon LASSO regression being selected[16] and regarded as factors independently predicting CHD risk. These variables were used in nine popular machine learning models (support vector machine (SVM), random forest (RF), naïve Bayes algorithm (NB), K-nearest neighbor algorithm (KNN), linear discriminant analysis algorithm (LDA), generalized linear model (GLM), adaptive boosting algorithm (ADA), Recursive partitioning and regression trees (RPART), and gradient booster (GBM)). The best model among these was selected and validated via assessment of the model clinical utility through decision curve analysis (DCA). DCA determines net benefits acquired by subtracting the true-positive patient proportion from the false-positive patient proportion and weighted by relative risk of quitting interventions in comparison with negative results of needless interventions, under different threshold probabilities.[17]

Results were examined using SPSS 27.0, R 4.3.1, and Anaconda 1.10.0 software.

Results

Altogether, 1435 subjects were recruited into the present work; 921 (64.2%) participants were in the case group, 514 (35.8%) participants were in the control group, 819 (57.1%) participants were male, and 616 (42.9%) participants were female. We randomly divided all participants into a training group with 969 (68%) participants and a validation group with 466 (32%) participants. We compared the basic patient characteristics between training and validation groups and found that all variables were not significantly different between them (Table 1).

Univariate regression was also conducted between case and control subgroups in training group. Age, sex, smoking status, alcohol consumption status, hypertension, diabetes, red blood cell count, total bilirubin level, LDL-C level, and serum levels of γKlotho were significantly different ($P < 0.05$) (Table 2).

Table 1 Comparison of Basic Patient Features Between Training and Validation Groups

Variables		Training Group (n = 969)	Validation Group (n = 466)	Z/χ^2	*P*
Age, years		52(43,61.5)	54(45,64)	−1.647	0.099
Sex	Male, n(%)	559(57.7)	260(55.8)	0.461	0.497
	Female, n(%)	410(42.3)	206(44.2)		
BMI, kg/m^2		25.82(23.24,28.73)	25.96(23.58,28.7)	−0.474	0.636
Smoking		236(24.4)	100(21.5)	1.471	0.225
Alcohol		122(12.6)	57(12.2)	0.037	0.847
Hypertension		498(51.4)	231(49.6)	0.418	0.518
Diabetes		219(22.6)	118(25.3)	1.297	0.255
WBC,10^9/L		6.69(5.7,7.62)	6.54(5.55,7.76)	−1.108	0.268
RBC, 10^{12}/L		4.89(4.49,5.32)	4.83(4.49,5.26)	−0.701	0.483
PLT, 10^9/L		232(201.5,257)	229.5(192,260)	−0.521	0.602
Hemoglobin, g/L		139(128,148)	139(129,148)	−0.257	0.797
Serum sodium, mmol/L		141.5(139.46,143.34)	141.51(139.76,143.35)	−0.712	0.476
Chlorine, mmol/L		106.03(103.85,108.02)	106(103.76,107.91)	−0.241	0.809
Serum potassium, mmol/L		3.98(3.7,4.25)	3.99(3.75,4.31)	−1.449	0.147
Serum calcium, mmol/L		2.41(2.3,2.52)	2.4(2.3,2.52)	−0.374	0.709
AST, μμ/L		28.17(22.09,43.67)	28.54(22.26,44.21)	−0.421	0.674
ALT, μμ/L		26.42(18.49,44.54)	25.07(17.77,38.28)	−1.759	0.079
BUN, mmol/L		6.22(5.13,7.38)	6.08(4.98,7.33)	−0.812	0.417
CR, mmol/L		72.56(61.39,87.2)	70.88(60.44,86.68)	−1.069	0.285
Albumin, g/L		40.91(38.68,43.53)	40.96(38.78,43.62)	−0.907	0.365
Total bilirubin, μmol/L		11.3(7.5,15)	10.8(7.78,14.4)	−0.897	0.37
Total cholesterol, mmol/L		3.59(2.79,4.25)	3.53(2.72,4.25)	−0.778	0.437
TG, mmol/L		1.29(1.11,1.85)	1.31(1.1,1.94)	−0.339	0.735
HDL-C, mmol/L		1.11(0.94,1.28)	1.14(0.96,1.29)	−1.622	0.105
LDL-C, mmol/L		2.4(1.9,3)	2.5(2,3.1)	−1.567	0.117
CRP, mg/L		5.17(5.01,6.51)	5.17(5.01,6.75)	−0.056	0.955
γKlotho, ng/mL		59.52(49.08,69.05)	59.34(46.5,68.36)	−0.986	0.324

Table 2 Univariate Regression of Training Set

Variables		Control (n = 352)	Case (n = 617)	Z/χ2	*P*
Age, years		44(34,50)	58(50,67)	−16.778	<0.001
Sex	Male, n(%)	171(48.6)	388(62.9)	18.791	<0.001
	Female, n(%)	181(51.4)	229(37.1)		

(*Continued*)

Table 2 (Continued).

Variables	Control (n = 352)	Case (n = 617)	Z/χ2	P
BMI, kg/m^2	25.56(23.02,28.72)	25.96(23.33,29.01)	−1.094	0.274
Smoking	65(18.5)	171(27.7)	10.407	0.001
Alcohol	31(8.8)	91(14.7)	7.191	0.007
Hypertension	159(45.2)	339(54.9)	8.569	0.003
Diabetes	56(15.9)	163(26.4)	14.151	<0.001
WBC,10^9/L	6.51(5.54,7.78)	6.77(5.8,7.55)	−1.289	0.197
RBC,10^{12}/L	4.8(4.48,5.2)	4.92(4.5,5.34)	−2.553	0.011
PLT, 10^9/L	235.5(206.25,257)	229(196,258)	−1.948	0.051
Hemoglobin, g/L	138(130,145)	139(127,149.5)	−0.121	0.903
Serum sodium, mmol/L	141.64(140.05,143.32)	141.41(139.14,143.36)	−1.045	0.296
Chlorine, mmol/L	105.96(103.74,107.58)	106.11(103.95,108.22)	−1.185	0.236
Serum potassium, mmol/L	3.94(3.69,4.23)	4.01(3.7,4.26)	−1.161	0.246
Serum calcium, mmol/L	2.4(2.32,2.51)	2.41(2.29,2.52)	−0.133	0.894
AST, μμ/L	28.23(22.79,40.45)	28.17(21.82,44.5)	−0.168	0.866
ALT, μμ/L	26.75(18.52,43.3)	26.38(18.47,44.99)	−0.205	0.837
BUN, mmol/L	6.07(5.12,7.33)	6.28(5.13,7.39)	−0.663	0.507
CR, mmol/L	71.21(59.98,90.95)	72.96(62.55,86.22)	−0.765	0.444
Albumin, g/L	40.46(38.73,43.23)	41.22(38.6,43.7)	−0.886	0.376
Total bilirubin, μmol/L	11.65(8.53,14.9)	11.2(7.2,15.15)	−2.002	0.045
Total cholesterol, mmol/L	3.72(2.81,4.25)	3.51(2.75,4.24)	−1.092	0.275
TG, mmol/L	1.3(1.12,1.88)	1.28(1.11,1.84)	−0.456	0.649
HDL-C, mmol/L	1.11(0.96,1.3)	1.11(0.94,1.27)	−1.031	0.303
LDL-C, mmol/L	2.4(2,2.8)	2.5(1.9,3.1)	−2.047	0.041
CRP, mg/L	5.17(5.04,6.46)	5.17(2.26,6.55)	−0.687	0.492
γKlotho, ng/mL	52.26(43.73,61.33)	66.01(53.64,71.13)	−11.341	<0.001

We selected 12 variables with $P > 0.2$, which may be risk factors for CAD, and performed LASSO regression for optimizing variable selection. Based on these findings, 12 variables were reduced to seven variables (age, sex, hypertension, smoking status, diabetes, and serum levels of γKlotho and LDL-C) (Figure 2), and we regarded these seven variables as risk factors for CAD.

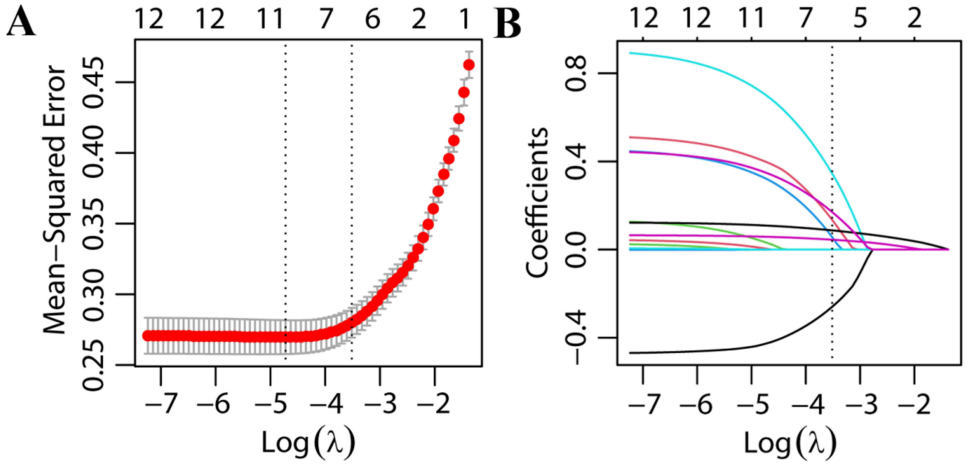

Figure 2 Feature selection based on least absolute shrinkage and selection operator (LASSO) algorithm. (**A**) Selection of tuning parameter (λ) into LASSO model with 10-fold cross-validation according to the minimum criteria. A plot showing squared error versus log (λ) was drawn. Each vertical line was made by optimum values with 1 standard error of minimum (1-SE criteria) and minimum criteria; λ = 0.030 was selected (1-SE criteria) through 10-fold cross-validation. (**B**) LASSO coefficients of 12 features. A plot of the coefficient versus log (λ) sequence was drawn. Each vertical line is made by the 10-fold cross-validation-selected value, in which optimal λ gave seven nonzero coefficients.

https://doi.org/10.2147/VHRM.S508351

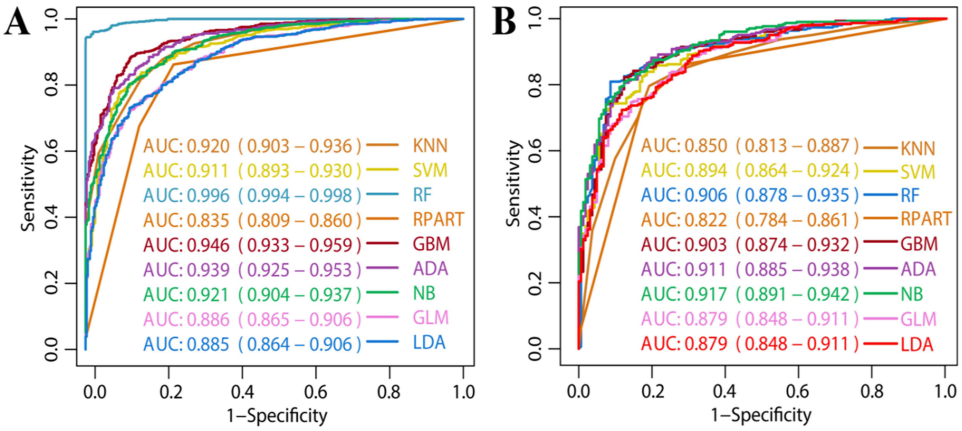

Figure 3 ROC and AUC value of nine machine learning models inTraining set (**A**) and Validation set (**B**).
Abbreviations: KNN, K-nearest neighbor algorithm; SVM, Support vector machine; RF, Random forest; RPART, Recursive partitioning and regression trees; GBM, Gradient boosted machine; ADA, Adaptive boosting algorithm; NB, Naïve Bayes algorithm; GLM, Generalized linear model; LDA, Linear discriminant analysis algorithm.

Model Construction

We used these seven variables to construct nine popular machine learning models and assessed them via area under the receiver operating characteristic (ROC) curve (AUC). These results showed that RF model had greatest AUC (0.996) in the training set, which was greater (0.906) than that of most of the models for validation set (but not the highest) (Figure 3).

We calculated the related parameters (sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV), accuracy, recall, F1 score, and precision) of the nine models in both training and validation sets. Consequently, all those above parameters for RF model were greater in training set and that those above parameters were relatively greater than those of the other nine models in the validation set (Table 3).

An integrated analysis of training and validation sets revealed that RF model was the best for predicting CAD risk.

Independent risk factors for CAD were used to visualize RF the model via a variable importance bar plot (Figure 4A) and Shapley additive explanations (SHAP) (Figure 4B). In the variable importance bar plot, the feature significance was

Table 3 ML Models Used to Predict the Risk of CHD

Classifier		Sensitivity	Specificity	PPV	NPV	Accuracy	Recall	F1-Score	Precision
Train set	KNN	0.878	0.795	0.883	0.789	0.848	0.878	0.881	0.883
	SVM	0.895	0.756	0.865	0.804	0.844	0.895	0.880	0.865
	RF	0.968	0.943	0.968	0943	0.959	0.968	0.968	0.968
	RPART	0.862	0.767	0.866	0.761	0.828	0.862	0.864	0.866
	GBM	0.909	0.815	0.896	0.837	0.875	0.909	0.903	0.896
	ADA	0.891	0.821	0.897	0.812	0.866	0.891	0.894	0.897
	NB	0.906	0.741	0.860	0.818	0.846	0.906	0.882	0.860
	GLM	0.872	0.707	0.839	0.759	0.812	0.872	0.855	0.839
	LDA	0.869	0.710	0.840	0.755	0.811	0.869	0.854	0.840
Validation set	KNN	0.849	0.716	0.849	0.716	0.803	0.849	0.849	0.849
	SVM	0.862	0.735	0.859	0.739	0.818	0.862	0.860	0.859
	RF	0.891	0.735	0.863	0.783	0.837	0.891	0.877	0.863
	RPART	0.862	0.710	0.848	0.732	0.809	0.862	0.855	0.848
	CBM	0.888	0.753	0.871	0.782	0.841	0.888	0.879	0.871
	ADA	0.888	0.784	0.885	0.789	0.852	0.888	0.887	0.885
	BN	0.901	0.728	0.862	0.797	0.841	0.901	0.881	0.862
	GLM	0.885	0.679	0.838	0.759	0.813	0.885	0.861	0.838
	LDA	0.872	0.704	0.847	0.745	0.813	0.872	0.859	0.847

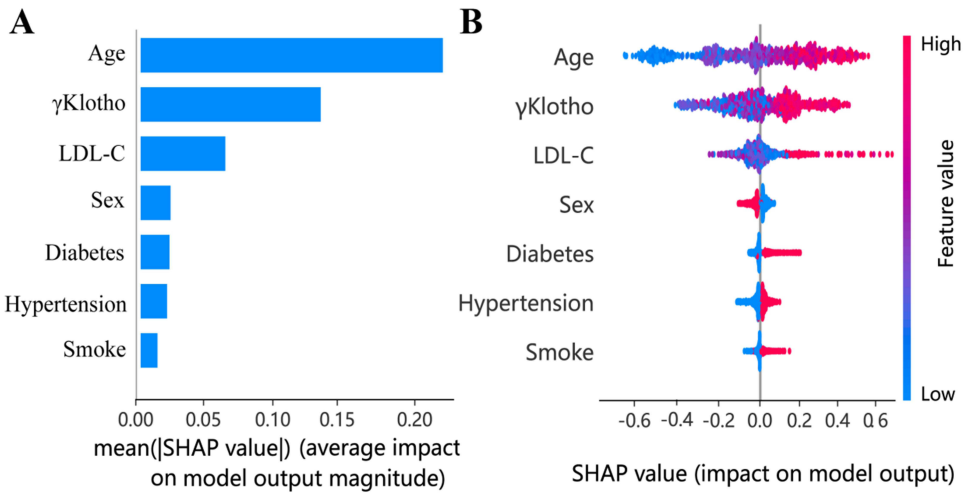

Figure 4 Variable importance bar plot represents feature ranking (**A**) and SHAP value for model output (**B**), with one point representing one variable in each observation. Then, point color was analyzed based on the relative variable height, with red and blue representing high and low levels, respectively.

represented by the bar length.[18] SHAP, as a game theory approach, contributes to the intuitive and accurate explanation of ML model output.[19] Its base value was defined as the output obtained by averaging all variables from training group, and it reflects the sample average value. Those independent risk factors for CHD according to degree of importance (from age to smoking status) are presented in Figure 4. Regarding the dichotomous classifier, a greater SHAP prediction score was related to an increased CHD risk.

A web calculator based on this model was developed for clinicians to predict patient's CHD progression (https://xiaoaipredictmodel.shinyapps.io/workrun11/), after inputting corresponding data of the patient, the prediction can be made automatically (Figure 5), and the cutoff value of RF model in training set is 0.599.

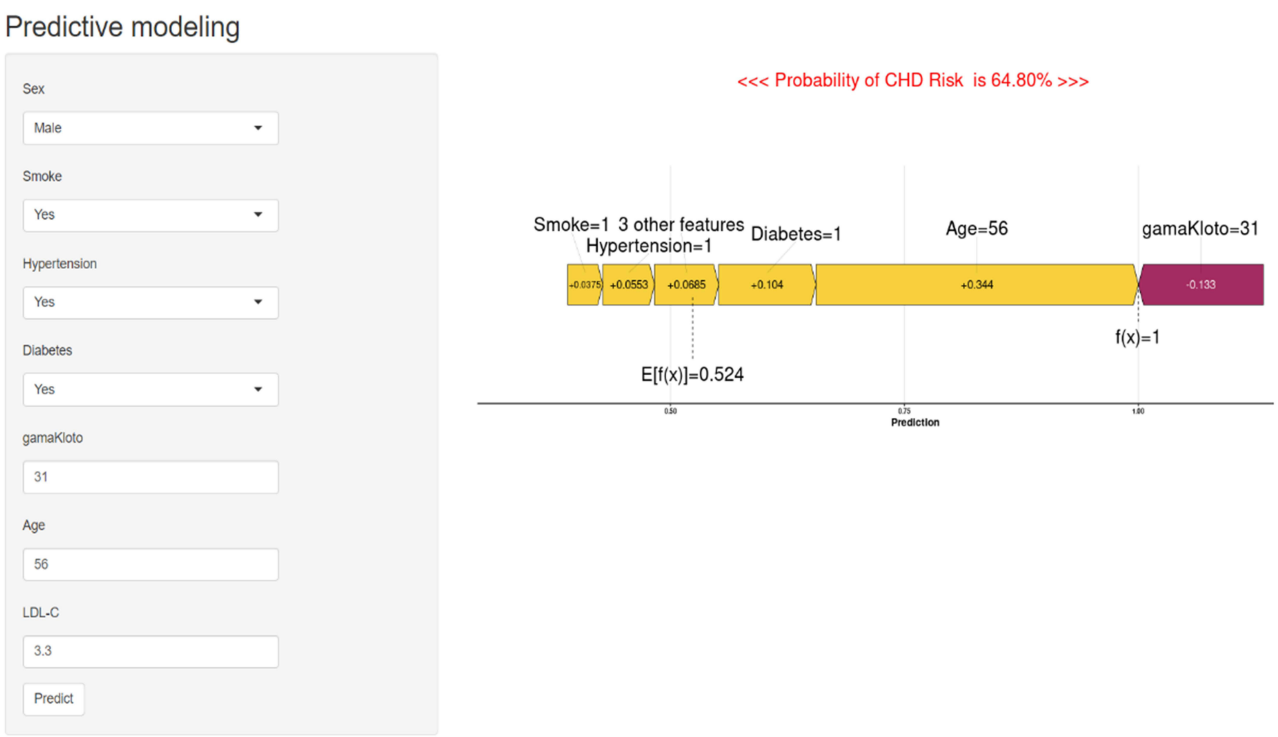

Figure 5 The calculator for predicting CHD progression in real time. Feature importance was calculated by the mean contribution of every observation, which is equal to the traditional method.

https://doi.org/10.2147/VHRM.S508351

The model interpretability was examined using two representative samples: a man with serum γ-Klotho level of 34.17 ng/mL (Figure 6A and B) and another man with that of 74.13 ng/mL (Figure 6C and D). The number on each variables name (Figure 6B and D) expressed the degree of influence on risk, and the red and blue colors suggested whether the factor increased or decreased CHD risk, respectively. The factors were all pooled, which offered a prediction probability corresponding to the SHAP value. For men who had a serum γKlotho level of 34.17 ng/mL, there was a low prediction probability (0.02) of CHD; for those who had a serum γKlotho level of 74.13 ng/mL, there was a high prediction probability (0.98) of ICM.

We performed model validation via DCA; if we used our model to predict CHD risk, greater net benefits were obtained than with "treat all" or "treat none" strategies, according to 0.0–1.0 threshold probabilities (Figure 7A and B), which indicated that the established RF model had greater clinical usefulness.

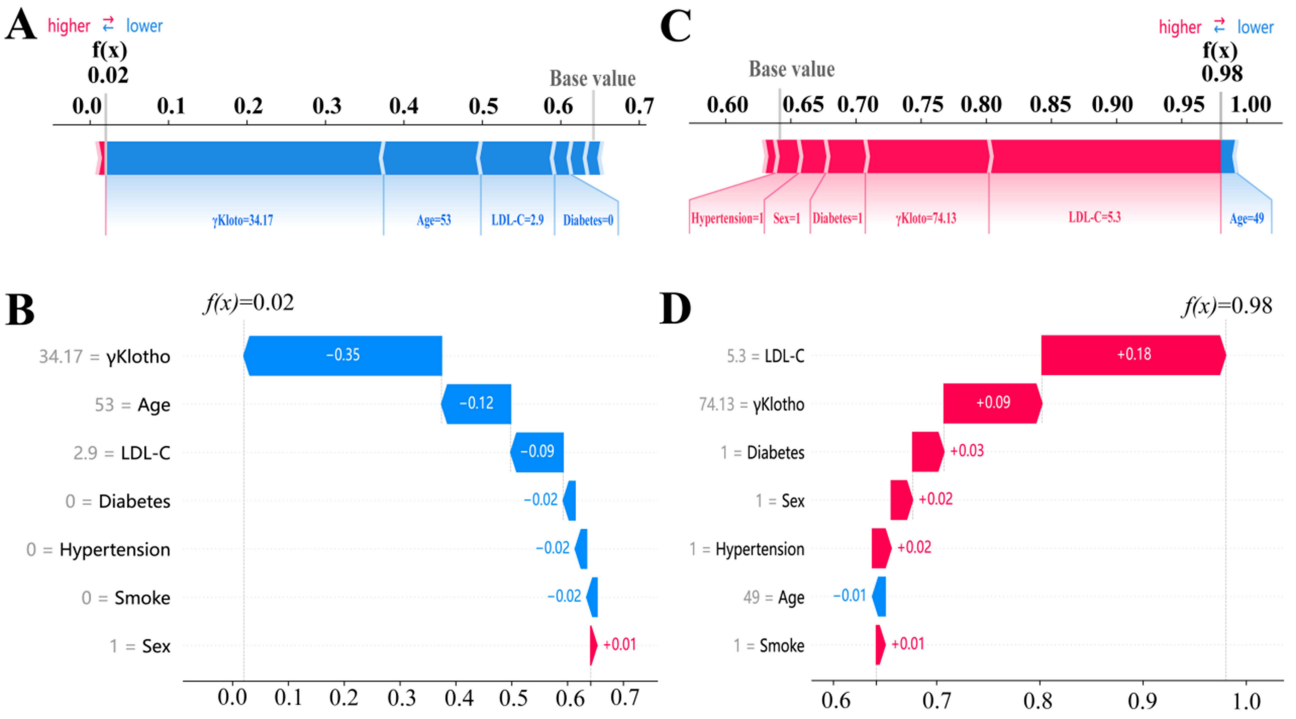

Figure 6 The model prediction results were interpreted using two samples. (**A** and **B**) Personalized prediction (**A**) and Waterfall plot for evolution the risks contributed by each feature for individual at low risk (**B**). (**C** and **D**) Personalized prediction (**C**) and Waterfall plot for evolution the risks contributed by each feature for individual at high risk (**D**).

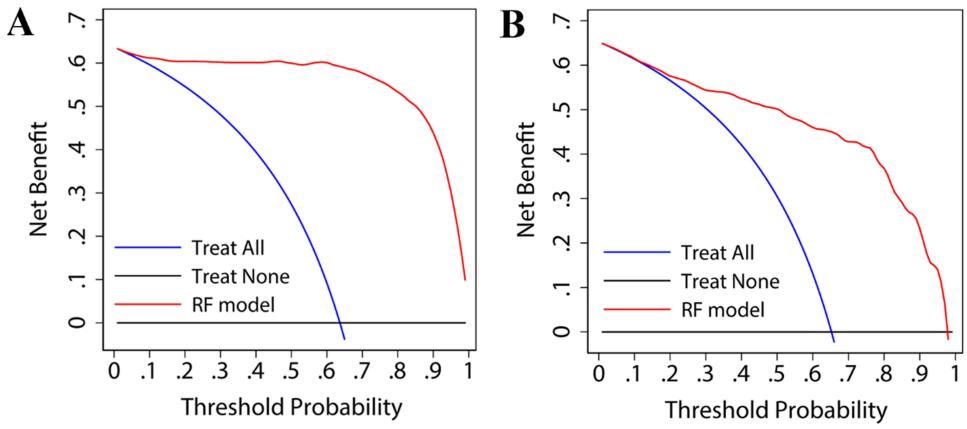

Figure 7 DCA of RF model using training set (**A**) and validation set (**B**).

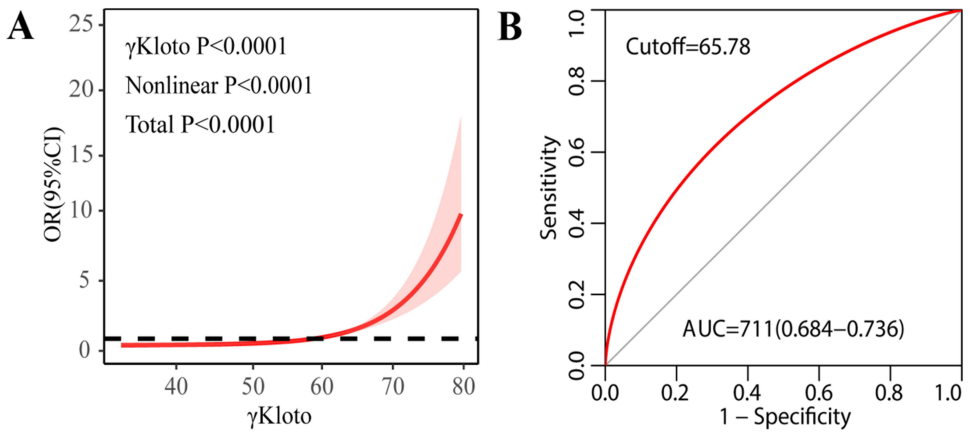

Figure 8 Restricted cubic spline (RCS) of serum level of γKlotho and risk of CHD (**A**), Receiver operating characteristic curves (ROC) and cutoff value of serum level of γKlotho (**B**).

In the present work, serum γKlotho level is a novel biomarker that is related to CHD risk. To identify the correlation of serum γKlotho level with CHD risk, restricted cubic spline (RCS) analysis was conducted (Figure 8A). Our results revealed that the incidence of CHD was positively ($P_{overall} < 0.001$) and non-linearly ($P_{non\text{-}linearity} = 0.7396$) correlated with the serum level of γKlotho.

The serum level of γKlotho, a new biomarker, is positively related to CHD risk and can be used to diagnose CHD clinically. To evaluate its diagnostic and discriminative performance for CHD, an ROC curve (Figure 8B) was drawn for γKlotho, and the AUC was 0.711 (0.684–0.736), whereas the threshold γKlotho level was 65.78 ng/mL, which indicated that the serum level of γKlotho exhibited the favorable prediction value for CHD risk.

Discussion

Based on our single-center, retrospective cohort study, age, serum γ-Klotho level, LDL-C, sex, diabetes, hypertension, and smoking status are risk factors for CHD, and a predictive model was constructed using a machine learning method. Among the CHD-related risk factors, age, LDL-C, sex, diabetes, hypertension, and smoking status were previously suggested to be the CAD-related risk factors. In this work, we identified a novel biomarker, the serum level of γKlotho, as a risk factor for CHD. The results indicated that the serum level of γKlotho is a risk factor for CHD, which shows positive relation to CHD risk.

Klotho, first discovered in 1997, can encode a new antiaging protein that is related to various biological processes mostly associated with human longevity.[20]

Since the Klotho gene was discovered, interest in this gene has increased, which has dramatically improved our understanding of the aging process. Serum Klotho expression in the human body decreases after individuals turn 40 years old[21] and can be detected in patients who develop aging-related disorders including hypertension, cancer, or kidney disease.[22]

Three Klotho forms have been discovered in mice and humans, including full-length transmembrane, secreted, and shed forms.[21] Secreted Klotho is dominant in the human body, particularly in the kidney.[23] The Klotho family includes 3 members encoded by three genes:[9,24] α-Klotho, β-Klotho, and γ-Klotho.

Of them, α-Klotho gene is on chromosome 13q12 and contains six exons, β-Klotho gene is on chromosome 4p14 and contains five exons, while γ Klotho gene is on chromosome 15q22.31 and contains 14 exons. The first two genes have high expression within enterohepatic tissue and the kidney, whereas the γKlotho gene shows high and selective expression within mouse eyes and brown adipose tissue.[25]

The αKlotho (KL) gene, which was first discovered to be an antiaging protein, may protect the vascular wall and is associated with cardiovascular disease,[26] In the progression of CHD, αKlotho plays a critical role in Klotho/FGF23 axis and regulating ERK/MAPK pathway in Cardiomyocyte,[27] the lower the serum αKlotho level the higher the risk of CHD.[5] βKlotho (known as KLB) is another subtype of Klotho protein whose molecular weight is around 130 kDa that acts

specifically as a membrane coreceptor and was identified after α-Klotho;[28] it has 41% amino acid sequence similarity to α-Klotho.[28,29] The lower serum β-klotho expression is related to the increased CVD risk.[11] It can also optimize the clinical outcomes for patients with T2DM. By regulating the homeostasis of glucose metabolism.[30] γKlotho, also referred to as lactase-like protein, is a member of glycosyl hydrolase 1 family, which serves as the single-pass membrane protein that contains 567 amino acids. KLγ is a member of the Klotho subfamily.[31] Research on KLγ is limited. In a previous study, γKlotho was investigated in three diseases, including breast,[32] prostate,[33] and bladder cancers.[33] Among these factors, increased γKlotho levels are related to aggressiveness and dismal prognosis of cancer patients; however, the correlation between serum γKlotho expression and cardiovascular disorders and specific function of it has not been previously reported. The present work is the first to reveal the relation of the serum level of γKlotho with CHD susceptibility and that the serum level of γKlotho is a risk factor for CHD. The higher the serum level of γKlotho, the higher the risk of CHD. Although the mechanism by which serum γKlotho leads to CHD is unknown, the relationship is confirmed. No study has described the function of γKlotho, possibly because the level of γKlotho results in CHD directly or indirectly. Additionally, the serum γKlotho level increases after CHD to prevent CHD and reverse it from CHD to healthy status. As the specific mechanisms are unclear, maybe it plays a critical role same with αKlotho and βKlotho, thus the causal relationship between CAD and γKlotho needs to be determined by performing functional research.

An interpretable ML-based risk stratification model for predicting the effect of serum γKlotho level on CHD and predicting CHD risk was constructed and validated in this study. By comparing nine machine learning models, we found that RF model had a superior prediction ability to others, and AUCs were 0.996 and 0.906 in training and validation sets, separately. We used the visualization function in the variable importance bar plot and SHAP[34] to determine how one specific variable affected the model output. We found that age and the serum level of γKlotho contribute more to the risk of CHD. The probability of SHAP prediction increases with increasing serum levels of γKlotho. We validated this RF model via DCA and found that when this model was used for predicting CHD risk, the net benefits over "treat all" or "treat none" strategies were obtained at the 0.0–1.0 threshold probabilities, indicating that this model had good clinical usefulness.

There are some limitations in the present work. Firstly, the single-center, retrospective study had a small sample size, while results from other populations were unknown. Secondly, serum γKlotho level showed positive relation to CHD risk, but the specific mechanism by which high serum levels of γKlotho can result in CHD is not known. Third, the MF model more accurately predicted the association of serum γKlotho level with CHD risk, but several confounding factors affect the accuracy of the association between the serum γKlotho level and CHD risk; thus, a matched study is needed.

In conclusion, we found that the serum level of γKlotho is a risk factor for CHD and is positively correlated with the susceptibility to CHD. RF was the best model in our study; it can accurately predict the risk of CHD and has good clinical usefulness.

Data Sharing Statement

Data used in this study are available from author (Tuersunjiang Naman) Email:tursunjan1016@163.com.

Ethics Approval and Consent to Participate

This study complies with the Declaration of Helsinki, and the Ethical Committee of people's hospital of xinjiang uygur autonomous region approved this study (SYDW2024032906). Every participant in this study signed the informed consent forms.

Funding

This work was supported by the Natural science project of Xinjiang Uyghur autonomous region (Grant Number: 2023D01C70).

Disclosure

The authors declare no conflicts of interest in this work.

References

1. Madhavan MV, Gersh BJ, Alexander KP, Granger CB, Stone GW. Coronary artery disease in patients ≥80 years of age. *J Am Coll Cardiol*. 2018;71(18):2015–2040. doi:10.1016/j.jacc.2017.12.068
2. Kannel WB. Coronary heart disease risk factors in the elderly. *Am J Geriatr Cardiol*. 2002;11(2):101–107. doi:10.1111/j.1076-7460.2002.00995.x
3. Fried LF, Shlipak MG, Crump C, et al. Renal insufficiency as a predictor of cardiovascular outcomes and mortality in elderly individuals. *J Am Coll Cardiol*. 2003;41(8):1364–1372. doi:10.1016/S0735-1097(03)00163-3
4. Meng H, Ruan J, Yan Z, et al. New progress in early diagnosis of atherosclerosis. *Int J Mol Sci*. 2022;23(16):8939. doi:10.3390/ijms23168939
5. Lee J, Kim D, Lee HJ, Choi JY, Min JY, Min KB. Association between serum klotho levels and cardiovascular disease risk factors in older adults. *BMC Cardiovasc Disord*. 2022;22(1):442. doi:10.1186/s12872-022-02885-2
6. Semba RD, Cappola AR, Sun K, et al. Plasma klotho and cardiovascular disease in adults. *J Am Geriatr Soc*. 2011;59(9):1596–1601. doi:10.1111/j.1532-5415.2011.03558.x
7. Lim K, Lu TS, Molostvov G, et al. Vascular Klotho deficiency potentiates the development of human artery calcification and mediates resistance to fibroblast growth factor 23. *Circulation*. 2012;125(18):2243–2255. doi:10.1161/CIRCULATIONAHA.111.053405
8. Donate-Correa J, Mora-Fernández C, Martínez-Sanz R, et al. Expression of FGF23/KLOTHO system in human vascular tissue. *Int J Cardiol*. 2013;165(1):179–183. doi:10.1016/j.ijcard.2011.08.850
9. Kuro OM. The Klotho proteins in health and disease. *Nat Rev Nephrol*. 2019;15(1):27–44. doi:10.1038/s41581-018-0078-3
10. Xu JP, Zeng RX, He MH, Lin SS, Guo LH, Zhang MZ. Associations between serum soluble α-Klotho and the prevalence of specific cardiovascular disease. *Front Cardiovasc Med*. 2022;9:899307. doi:10.3389/fcvm.2022.899307
11. Wu Y, Chen Z, Duan J, et al. Serum levels of FGF21, β-Klotho, and BDNF in stable coronary artery disease patients with depressive symptoms: a cross-sectional single-center study. *Front Psychiatry*. 2020;11:587492. doi:10.3389/fpsyt.2020.587492
12. Jiang F, Jiang Y, Zhi H, et al. Artificial intelligence in healthcare: past, present and future. *Stroke Vasc Neurol*. 2017;2(4):230–243. doi:10.1136/svn-2017-000101
13. Deo RC. Machine Learning in Medicine. *Circulation*. 2015;132(20):1920–1930. doi:10.1161/CIRCULATIONAHA.115.001593
14. Ritchie LD, Campbell NC, Murchie P. New NICE guidelines for hypertension. *BMJ*. 2011;343:d5644.
15. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw*. 2010;33(1):1–22. doi:10.18637/jss.v033.i01
16. Kidd A, McGettrick M, Tsim S, Halligan D, Bylesjo M, Blyth K. Survival prediction in mesothelioma using a scalable lasso regression model: instructions for use and initial performance using clinical predictors. *BMJ Open Respir Res*. 2018;5(1):e000240. doi:10.1136/bmjresp-2017-000240
17. Kerr K, Brown M, Zhu K, Janes H. Assessing the clinical impact of risk prediction models with decision curves: guidance for correct interpretation and appropriate use. *J Clin Oncol*. 2016;34(21):2534–2540. doi:10.1200/JCO.2015.65.5654
18. Rahman MS, Islam KR, Prithula J, et al. Machine learning-based prognostic model for 30-day mortality prediction in Sepsis-3. *BMC Med Inf Decis Making*. 2024;24(1):249. doi:10.1186/s12911-024-02655-4
19. Singh R, Lanchantin J, Sekhon A, Qi Y. Attend and predict: understanding gene regulation by selective attention on chromatin. *Adv Neural Inf Process Syst*. 2017;30:6785–6795.
20. Kuro-o M, Matsumura Y, Aizawa H, et al. Mutation of the mouse klotho gene leads to a syndrome resembling ageing. *Nature*. 1997;390(6655):45–51. doi:10.1038/36285
21. Xu Y, Sun Z. Molecular basis of Klotho: from gene to function in aging. *Endocrine Reviews*. 2015;36(2):174–193. doi:10.1210/er.2013-1079
22. Neyra JA, Hu MC, Moe OW. Klotho in clinical nephrology: diagnostic and therapeutic implications. *Clin J Am Soc Nephrol*. 2020;16(1):162–176. doi:10.2215/CJN.02840320
23. Olauson H, Mencke R, Hillebrands JL, Larsson TE. Tissue expression and source of circulating αKlotho. *Bone*. 2017;100:19–35. doi:10.1016/j.bone.2017.03.043
24. Zhang Y, Wang L, Wu Z, Yu X, Du X, Li X. The expressions of Klotho family genes in human ocular tissues and in anterior lens capsules of age-related cataract. *Current Eye Research*. 2017;42(6):871–875. doi:10.1080/02713683.2016.1259421
25. Fon Tacer K, Bookout AL, Ding X, et al. Research resource: comprehensive expression atlas of the fibroblast growth factor system in adult mouse. *Mol Endocrinol*. 2010;24(10):2050–2064. doi:10.1210/me.2010-0142
26. Keles N, Caliskan M, Dogan B, et al. Low serum level of Klotho is an early predictor of atherosclerosis. *Tohoku J Exp Med*. 2015;237(1):17–23. doi:10.1620/tjem.237.17
27. Jia Z, Liu Q, Xie Y, et al. Klotho/FGF23 axis regulates cardiomyocyte apoptosis and Cytokine release through ERK/MAPK pathway. *Cardiovasc Toxicol*. 2023;23(9–10):317–328. doi:10.1007/s12012-023-09805-6
28. Ito S, Kinoshita S, Shiraishi N, et al. Molecular cloning and expression analyses of mouse betaklotho, which encodes a novel Klotho family protein. *Mech Dev*. 2000;98(1–2):115–119. doi:10.1016/S0925-4773(00)00439-1
29. Ogawa Y, Kurosu H, Yamamoto M, et al. BetaKlotho is required for metabolic activity of fibroblast growth factor 21. *Proc Natl Acad Sci USA*. 2007;104(18):7432–7437. doi:10.1073/pnas.0701600104
30. Hua S, Liu Q, Li J, Fan M, Yan K, Ye D. Beta-klotho in type 2 diabetes mellitus: from pathophysiology to therapeutic strategies. *Rev Endocr Metab Disord*. 2021;22(4):1091–1109. doi:10.1007/s11154-021-09661-1
31. Ito S, Fujimori T, Hayashizaki Y, Nabeshima Y. Identification of a novel mouse membrane-bound family 1 glycosidase-like protein, which carries an atypical active site structure. *BBA*. 2002;1576(3):341–345. doi:10.1016/S0167-4781(02)00281-6
32. Trošt N, Peña-Llopis S, Koirala S, et al. γKlotho is a novel marker and cell survival factor in a subset of triple negative breast cancers. *Oncotarget*. 2016;7(3):2611–2628. doi:10.18632/oncotarget.6006
33. Onishi K, Miyake M, Hori S, et al. γ-Klotho is correlated with resistance to docetaxel in castration-resistant prostate cancer. *Oncol Lett*. 2020;19(3):2306–2316. doi:10.3892/ol.2020.11308
34. Lundberg SM, Erion G, Chen H, et al. From local explanations to global understanding with explainable AI for trees. *Nature Mach Intell*. 2020;2(1):56–67. doi:10.1038/s42256-019-0138-9