OPEN

Machine Learning for the Prevalence and Severity of Coronary Artery Calcification in Nondialysis Chronic Kidney Disease Patients

A Chinese Large Cohort Study

Haitao Zhu, MD,† Changqing Yin, MD,† U. Joseph Schoepf, MD,*‡ Dongqing Wang, MD,† Changsheng Zhou, BD,* Guang Ming Lu, MD,* and Long Jiang Zhang, MD, PhD**

Purpose: This study sought to determine whether machine learning (ML) can be used to better identify the risk factors and establish the prediction models for the prevalence and severity of coronary artery calcification (CAC) in nondialysis chronic kidney disease (CKD) patients and compare the performance of distinctive ML models with conventional logistic regression (LR) model.

Materials and Methods: In all, 3701 Chinese nondialysis CKD patients undergoing noncontrast cardiac computed tomography (CT) scanning were enrolled from November 2013 to December 2017. CAC score derived from the cardiac CT was calculated with the calcium scoring software and was used to assess and stratify the prevalence and severity of CAC. Four ML models (LR, random forest, support vector machine, and k-nearest neighbor) and the corresponding feature ranks were conducted. The model that incorporated the independent predictors was shown as the receiver-operating characteristic (ROC) curve. Area under the curve (AUC) was used to present the prediction value. ML model performance was compared with the traditional LR model using pairwise comparisons of AUCs.

Results: Of the 3701 patients, 943 (25.5%) patients had CAC. Of the 943 patients with CAC, 764 patients (20.6%) and 179 patients (4.8%) had an Agatston CAC score of 1 to 300 and ≥ 300, respectively. The primary cohort and the independent validation cohort comprised 2957 patients and 744 patients, respectively. For the prevalence of CAC, the AUCs of ML models were from 0.78 to 0.82 in the training data set and the internal validation cohort. For the severity of CAC, the AUCs of the 4 ML models were from 0.67 to 0.70 in the training data set and from 0.53 to 0.70 in the internal validation cohort. For the prevalence of CAC, the AUC was 0.80 (95% confidence interval [CI]: 0.77-0.83) for ML (LR) versus 0.80 (95% CI: 0.77-0.83) for the traditional LR model ($P = 0.2533$). For the severity of CAC, the AUC was 0.70 (95% CI: 0.63-0.77) for ML (LR) versus 0.70 (95% CI: 0.63-0.77) for traditional LR model ($P = 0.982$).

Conclusions: This study constructed prediction models for the presence and severity of CAC based on Agatston scores derived from noncontrast cardiac CT scanning in nondialysis CKD patients using ML, and showed ML LR had the best performance.

Key Words: chronic kidney disease, coronary artery calcification, dialysis, machine learning

(*J Thorac Imaging* 2022;37:401–408)

Reports from international registries confirm that cardiovascular disease (CVD) accounts for ~45% of all-cause mortality among nondialysis chronic kidney disease (CKD) subjects.[1] Coronary artery calcification (CAC), independently predicting the risk of CVD, presents in 79% of nondialysis CKD patients and is more severe compared with the general population.[2,3] The prevalence and severity of CAC have prognostic significance and have been linked to all-cause mortality in nondialysis CKD patients.[4] Therefore, identifying the related risk factors and building the prediction model for the prevalence and severity of CAC in nondialysis CKD is necessarily required.

The prevalence and severity of CAC in CKD patients are highly associated with loss of renal function. In addition to the traditional CAC risk factors (age, hypertension, and hypercholesterolemia), CKD also confers nontraditional risk factors (abnormal mineral metabolism, increased oxidative stress, and inflammation) for the prevalence and severity of CAC in nondialysis CKD patients.[5] Due to the high CAC prevalence in nondialysis CKD patients, identifying the most relevant risk factors and further establishing the prediction model have a great clinic value to prevent the development and progression of CAC in this special subpopulation. Although many attempts have been made to identify risk factors, the findings were inconsistent and had limited success based on the conventional logistic regression (LR).[6,7] The development and progression of CAC in CKD patients are complex process, conventional LR methods had only the modest ability to work in this context.

Fundamental limitations of many conventional LR-based models include the selection of variables and the linearity of the model. Advances in machine learning (ML)

From the *Department of Medical Imaging, Jinling Hospital, Medical School of Nanjing University, Nanjing, Jiangsu; †Department of Medical Imaging, Affiliated Hospital of Jiangsu University, Jiangsu University, Zhenjiang, China; and ‡Division of Cardiovascular Imaging, Department of Radiology and Radiological Science, Medical University of South Carolina, Charleston, SC.

H. Z. and C. Y. contributed equally to this work.

Supported by The National Key Research and Development Program of China (2017YFC0113400 for L.J.Z.).

The authors declare no conflicts of interest.

Correspondence to: Long Jiang Zhang, MD, PhD, Department of Medical Imaging, Jinling Hospital, Medical School of Nanjing University, 305 Zhongshan East Road, Xuanwu District, Nanjing, Jiangsu Province 210002, China (e-mail: kevinzhlj@163.com).

DOI: 10.1097/RTI.0000000000000657

have resulted in the creation of algorithms capable of building models from large data sets with a multitude of variables, facilitating the construction of models for data-driven prediction or classification.[8] There are evidences suggesting that ML algorithms develop models from test inputs to derive predictions that can substantially outperform traditional regression prediction models in heart failure,[9,10] but the data in predicting the prevalence and severity of CAC are scant in CKD patients.

Therefore, based on the Agatston scores derived from the noncontrast cardiac computed tomography (CT) scanning, this study assessed the feasibility and accuracy of ML to predict the presence and severity of CAC, ascertained the potential risk factors in nondialysis CKD patients, and compared the performance with traditional LR.

MATERIALS AND METHODS

Study Participants

The local Institutional Review Board approved this retrospective study of the registry data (ChiCTR-OCH-14004447).[11] In all, 5102 CKD patients from Jinling Hospital, Medical School of Nanjing University were included between November 2013 and December 2017 in this study. Inclusion criteria were as follows: (a) CKD patients; CKD is defined as abnormalities of kidney structure or function, present for >3 months, with implications for health, according to Kidney Disease: Improving Global Outcomes (KDIGO) 2012 Clinical Practice Guideline for the Evaluation and Management of CKD and[12] (b) ability to provide informed consent for participation. The exclusion criteria were as follows: (a) CKD patients undergoing dialysis or renal transplantation; (b) patients with acute kidney injury, cirrhosis, polycystic kidney disease, or renal cell carcinoma, parathyroidectomy, or evident malignancies; (c) patients with history of invasive procedure for atherosclerotic CVD (coronary artery bypass graft, angioplasty, valve replacement, pacemaker placement or other vascular surgery) or with conditions making arterial calcification measurements technically impossible or unreliable, such as cardiac arrhythmias. The flowchart of this study is shown in Figure 1.

Coronary Artery Calcification Score (CACS) Measurements

All enrolled participants underwent cardiac CT scan without intravenous contrast administration using a dual-source CT scanner (Definition, Siemens Healthcare, Germany) with the following protocols: tube voltage 120 kVp, effective tube current 80 mAs, rotation time 330 ms, reconstructed slice thickness 3.0 mm, and pitch 1.5. The field of view was set to include the entire heart, and the z-axis direction included data from bifurcation of the pulmonary arteries to the apex of the heart during an expiratory breath holding with electrocardiography gating.

Imaging data sets were subsequently transferred to a workstation (Siemens Healthcare, Germany). A blinded reader with 3 years of experience in reading cardiac CT imaging (C.Y.) measured the CAC. Window width and level settings were adjusted to the investigator's discretion to optimally identify calcification. The CAC score was calculated with the commercially available calcium scoring software (CaScoring Software, Siemens Healthcare, Germany), which was used to identify and score any calcium in the 4 main coronary arteries (the left main, left anterior descending, left circumflex, and right coronary artery). Using a semiautomated threshold-dependent algorithm, calcifications within the coronary artery tree above a threshold of 130 HU were included and a minimum of 3 contiguous pixels were used for the identification of a calcific lesion. Each focus exceeding the minimum criteria was scored using the algorithm developed by Agatston et al,[13] calculated by multiplying the lesion area by a density factor derived from the maximal HU within this area.

A binary scale was used to classify the CT examination as to the presence or absence of CAC (Agatston score $=0$ vs. Agatston score >0). Otherwise, based on the distribution of Agatston scores, the CKD participants were divided into 3 groups as CAC score categories of 0 (no CAC), >1 to 300 (moderate CAC), and >300 (severe CAC).[14,15]

Measurements

Medical history, laboratory data, and anthropometric measurements were obtained for each enrolled patient. Age, sex, race/ethnicity, and medical history information were acquired by questionnaires.

Body mass index is calculated as weight in kilograms divided by height in meters squared. Smoking is defined as

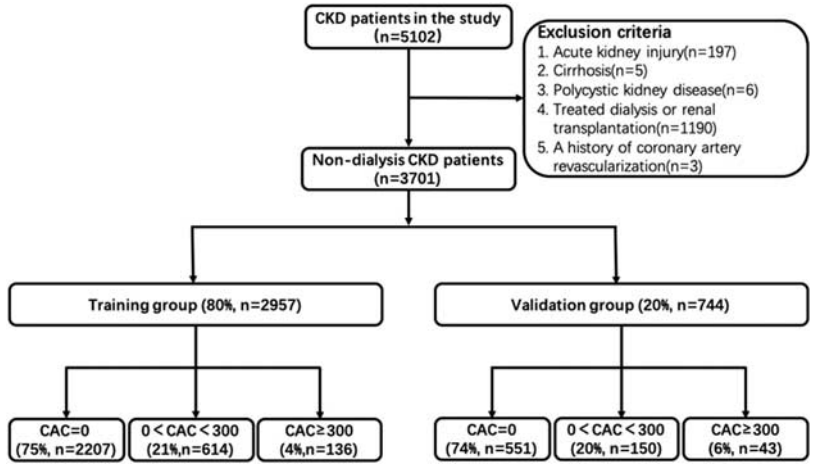

FIGURE 1. Flowchart of this study.

continuously smoking ≥1 cigarette a day for 6 months. Blood pressure in the seated position was measured using an aneroid sphygmomanometer after at least 5 minutes of quiet rest and the average of 3 measurements was used as the final result. Hypertension is defined as systolic pressure ≥140 mm Hg, diastolic pressure ≥90 mm Hg, or self-reported use of an antihypertensive medication. Diabetes is defined as a fasting plasma glucose level of 126 mg/dL or greater, a nonfasting plasma glucose level of 200 mg/dL or greater, or self-reported use of any antidiabetic medication.

Blood samples were drawn in a fasting state. Levels of glucose, total and high-density lipoprotein (HDL) cholesterol, low-density lipoprotein cholesterol, triglycerides, cystatin C, phosphorus, calcium, total parathyroid hormone (PTH), urea nitrogen, uric acid, creatinine, and alkaline phosphatase were measured from blood and urine samples using standard laboratory methods. Levels of C-reactive protein were measured using the particle-enhanced immunonephelometry method.

An overnight random morning urine sample was collected. Urine albumin was quantified using the turbidimetric method. Estimated glomerular filtration rate (eGFR) was calculated in the modified calculated using the 4-variable Modification of Diet in Renal Disease (MDRD) Study equation.[16]

ML Algorithms

ML methods were independently developed by in DeepWise Medical Research platform (https://research.deepwise.com/) and the detailed performance protocol as described previously.[17] Supervised ML algorithms with binary classification were used to build predictive models. Preprocessing of data included calculating complex parameters and manually marking each input variable as either numeric or categorical. To apply the data more efficiently and facilitate the construction of ML models, one-hot encoding was used for categorical data dimension. The primary steps involved the following: (1) splitting data into training (n = 2957 for the prevalence of CAC and n = 750 for the severity of CAC) and validation set (n = 744 for the prevalence of CAC and n = 193 for the severity of CAC; validation set is part of the training sets and only used for model super-parameter searching); (2) based on the input data and labels (patient demographic, clinical, and biochemical parameters information), we trained 4 ML models (LR, random forest [RF], support vector machine [SVM], and k-nearest neighbor [KNN]). LR model is a regression model. For each data sample, the LR model will output the probability of being positive. In the training stage, the goal of the model is to estimate a weight for each data dimension to minimize the differences between prediction and label. This weight matrix can tell us how each data dimension will influence the final prediction and this is why LR model has good interpretability. SVM is a binary classification model. For each data sample, the SVM model will output distance between the current data sample; the sign symbol of this distance indicates prediction is positive or negative. The goal of this model in training stage is to find a linear hyperplane separating positive data samples from negative data samples in the training set with maximum margin. Besides, we applied kernel function to the SVM model to map data into higher dimensions to make a more accurate hyperplane. RF model is an ensemble algorithm and consists of a collection of regression or classification decision trees. In the training stage, this model is to fit these trees to data. Besides, in the prediction stage, the model will output the average prediction of trees in the forest. KNN model is a regression model and consists of an input layer, a 64-unit hidden layer, and an output layer.

Given the sample size, performance evaluation was assessed using a 10-fold cross-validation. For the implementing procedure, the feature-selection method was used to reduce the overfitting problem. The best hyper-parameters and regularization parameters of each model would be searched automatically based on different metrics in a 10-fold cross-validation. After optimal hyperparameters and regularization parameters were chosen, the entire training cohort was used to train the model, and the performance was evaluated on internal validation cohorts.

Statistics Analysis

Statistical analyses were performed with SPSS 25.0 statistical software (IBM, Armonk, NY). The statistical significance levels were all 2-sided, with statistical significance set at 0.05. Continuous data were depicted as mean ± SD for normally distributed data, while median and interquartile range [interquartile range] was provided for non-normally distributed data. Categorical data were depicted as frequencies and percentages. For normally distributed data, independent sample t tests or analysis of variance tests were used if appropriate. For non-normally distributed data, independent samples nonparametric test (Mann-Whitney U test or Kruskal-Wallis H) was used for analysis. For categorical data, Pearson χ^2 tests or Fisher exact tests were used, if appropriate. The statistic value between the noncalcification and the calcification group was called P1, while the statistic value between moderate and severe groups was called P2.

The output of the classification model was calculated as the prediction probability of the prevalence or severity of CAC class. The performances of the models were shown as the receiver-operating characteristic curve (ROC) and area under the curve (AUC). The sensitivity and specificity were determined by the Youden index. DeLong test was used to compare AUCs of these models and Bonferroni correction was applied to determine the best performed ML model.[18] Feature importance was ranked according to the coefficient of each parameter provided by the corresponding ML algorithms.

RESULTS

Study Characteristics

Of the 5102 CKD patients, 1401 patients were excluded. The remaining 3701 patients were enrolled in the final analysis. Of the 3701 patients, 943 (25.5%) patients had CAC, while 2758 (74.5%) patients had no CAC. Of the 943 patients with CAC, 764 patients (20.6%) and 179 patients (4.8%) exhibited an Agatston CACS of 1 to 300 and ≥300, respectively. After random ranking, 2957 patients (Agatston score = 0: n = 2207; 1≤ Agatston score <300: n = 614; Agatston score ≥300: n = 136) and 744 patients (Agatston score = 0: n = 551; 1≤ Agatston score <300: n = 150; Agatston score ≥300: n = 43) comprised the primary training cohort and the independent validation cohort, respectively. Table 1 shows the demographical, clinical, and biochemical characteristics in the training and validation cohorts. No difference was found for all variables between the primary training cohort and the validation cohort ($P = 0.051$ to 0.924).

Risks Factors for the Prevalence and Severity of CAC

Baseline characteristics of participants according to CAC score categories (Agatston score = 0 and Agatston

TABLE 1. Characteristics of Demographic Information, Clinical Data, and Biochemical Parameters in the Primary and Validation Cohorts

	Grouping		
Indicators	**Primary Cohort (n = 2957)**	**Validation Cohort (n = 744)**	***P***
General information			
Number, n (%)	2957 (80.0)	744 (20.0)	
Sex (male), n (%)	1987 (67.2)	489 (65.7)	0.446
Age (y)	45.86 ± 13.61	46.32 ± 13.16	0.400
Diabetes (yes), n (%)	800 (27.1)	200 (26.9)	0.924
Hypertension (yes), n (%)	1466 (49.6)	381 (51.2)	0.426
Oral calcium (yes), n (%)	1113 (37.6)	251 (33.7)	0.049
Smoking (yes), n (%)	790 (26.7)	190 (25.5)	0.515
Body mass index (kg/m^2)	24.31 ± 3.67	24.44 ± 3.78	0.405
Biochemical indicators			
Triglycerides (mmol/L)	1.72 [1.24-2.48]	1.72 [1.25-2.41]	0.529
Total cholesterol (mmol/L)	5.10 [4.22-6.21]	5.09 [4.25-6.12]	0.815
Low-density lipoprotein (mmol/L)	3.10 [2.29-3.47]	3.10 [2.37-3.52]	0.647
High-density lipoprotein (mmol/L)	0.93 [0.69-1.06]	0.93 [0.69-1.07]	0.986
Calcium (mmol/L)	2.19 [2.04-2.30]	2.19 [2.06-2.30]	0.439
Phosphorus (mmol/L)	1.20 [1.04-1.38]	1.21 [1.06-1.38]	0.089
Alkaline phosphatase (U/L)	68 [54-76]	69 [52-76]	0.483
C-reactive protein (mg/L)	0.70 [0.1-3.5]	0.90 [0.1-4.0]	0.264
Urine protein (g/L)	1.91 [0.91-4.35]	2.09 [0.91-4.50]	0.298
Parathyroid hormone (ng/L)	46.76 [30.94-64.95]	48.76 [30.92-64.95]	0.324
25-$(OH)_2$-D_3 (μg/mL)	12.43 [6.17-16.40]	12.60 [6.35-16.77]	0.249
Glucose (mmol/L)	5.13 [4.65-5.71]	5.19 [4.70-5.76]	0.190
Uric acid (μmol/L)	428 [360-502]	425 [353-595]	0.375
Urea nitrogen (mmol/L)	23.4 [17.5-35.5]	23.6 [17.1-35.78]	0.659
Creatinine (mg/dL)	1.58 [1.20-2.48]	1.53 [1.16-2.63]	0.390
eGFR (mL/min/1.73 m2)	47.85 [27.39-67.47]	49.36 [25.97-68.98]	0.427
Cystatin C (mg/L)	1.87 [1.36-2.52]	1.79 [1.30-2.53]	0.287
Calcium phosphorus product (m^{2}mol^2/L^2)	2.59 [2.23-2.98]	2.62 [2.28-3.05]	0.032

Data are presented as n (%) or mean ± SD or median [25th, 75th percentile].

score > 0) are illustrated in Table 2. Compared with those without CAC, participants with CAC were to be older, male, more likely to have a history of diabetes, hypertension, on oral calcium agents and smoking, lower level of eGFR, greater level of body mass index, total cholesterol, HDL, calcium, phosphorus, alkaline phosphatase, urine protein, PTH, 25-$(OH)_2$-D_3, glucose, urea nitrogen, creatinine, cystatin C, and calcium phosphorus product (CPP) (all $P < 0.001$). A multivariable analysis identified older age, male, diabetes, hypertension, oral calcium agents, PTH, creatinine, HDL, glucose, and CPP as independent predictors for the presence of CAC in nondialysis CKD patients (age: 1.079 [1.070-1.088]; sex [male]: 1.904 [1.512-2.397]; diabetes: 2.177 [1.759-2.694]; hypertension: 1.393 [1.137-1.707]; oral calcium agent: 1.564 [1.283-1.07]; PTH: 1.002 [1.001-1.003]; creatinine: 0.922 [0.859-0.989]; HDL: 0.688 [0.520-0.911]; glucose: 1.107 [1.052-1.165]; and CPP: 1.582 [1.320-1.897], Table 3).

We next evaluated the risk factors and prediction model for the severity of CAC in all participants and 3 subgroups subdivided by CAC score categories (Agatston score = 0, 1 to 299, and ≥ 300). Older age and diabetes were associated with increasing CAC. There was a graded relationship between increasing CAC and higher levels of PTH, glucose, urea nitrogen, creatinine, cystatin C, CPP, and lower level of 25-$(OH)_2$-D_3 and eGFR (all $P < 0.001$) (Table 2). Due to the non-normal distribution of CAC score categories, negative log-log was first chosen as the connected function and then the hierarchical LR analysis was performed. A multivariable analysis identified older age, 25-$(OH)_2$-D_3, urea nitrogen, glucose as independent predictors for the severity of CAC in CKD patients (age: 1.049 [1.021-1.078]; 25-$(OH)_2$-D_3: 0.971 [0.943-0.999], urea nitrogen: 1.010 [1.000-1.019], and glucose: 1.140 [1.067-1.219] (Table 3).

ML Models and Performance

Figure 2 illustrates the ROC curves for training and internal validation sets, along with the corresponding AUCs of the LR, RF, SVM, and KNN. The AUCs for the training cohort are shown in Table 3. For the prevalence of CAC, the AUC was 0.82 (95% confidence interval [CI]: 0.80-0.83) for LR, 0.80 (95% CI: 0.79-0.82) for RF, 0.82 (95% CI: 0.80-0.83) for SVM, and 0.78 (95% CI: 0.76-0.79) for KNN, respectively. For the severity of CAC, the AUC was 0.67 (95% CI: 0.63-0.70) for LR, 0.70 (95% CI: 0.66-0.73) for RF, 0.69 (95% CI: 0.66-0.73) for SVM, and 0.67 (95% CI: 0.63-0.70) for KNN, respectively.

To further assess the models' performance, ROC curves and the corresponding AUC, sensitivity, and specificity of the 4 ML models and traditional LR prediction model were assessed in the internal validation data sets (Table 4). For the prevalence of CAC, the highest AUC was LR (0.80, 95% CI: 0.77-0.83) and SVM (0.80, 95% CI: 0.77-0.82), intermediary for RF (0.78, 95% CI: 0.75-0.81), and next for KNN (0.76, 95% CI: 0.73-0.79). LR had a good prediction ability for the prevalence of CAC that was better than the other three models ($P = 0.048$, 0.024, and 0.002 for RF, SVM, and KNN, respectively). There was no significant difference in performance when compared LR with the traditional LR prediction model (AUC: 0.80, 95% CI: 0.77-0.83; $P = 0.254$). For the severity of CAC, the

TABLE 2. Baseline Characteristics of Participants According to CAC Score Categories

Index	Noncalcification	Calcification			P_1	P_2
		All	1-299	≥300		
General information						
Number, n (%)	2207 (74.6)	750 (25.4)	614 (81.9)	136 (18.1)		
Sex (male), n (%)	1413 (64.0)	574 (76.5)	469 (76.4)	105 (77.2)	0.001	0.838
Age (y), n (%)	42.60 ± 12.82	55.45 ± 11.15	54.50 ± 11.08	59.74 ± 10.49	0.001	0.001
Diabetes (yes), n (%)	412 (18.7)	388 (51.7)	305 (49.7)	83 (61)	0.001	0.016
Hypertension (yes), n (%)	984 (44.6)	482 (64.3)	390 (63.5)	92 (67.5)	0.001	0.363
Oral calcium (yes), n (%)	727 (32.9)	386 (51.5)	308 (50.2)	78 (57.4)	0.001	0.129
Smoking (yes), n (%)	516 (23.4)	274 (36.5)	216 (35.2)	58 (42.6)	0.001	0.102
Body mass index	24.08 ± 3.71	24.98 ± 3.49	24.89 ± 3.45	25.34 ± 3.66	0.002	0.177
Biochemical indicators						
Triglycerides (mmol/L)	1.71 [1.25-2.47]	1.77 [1.21-2.51]	1.77 [1.19-2.54]	1.87 [1.33-2.34]	0.413	0.612
Total cholesterol (mmol/L)	5.05 [4.21-6.13]	5.20 [4.25-6.46]	5.19 [4.28-6.45]	5.34 [4.17-6.65]	0.020	0.977
Low-density lipoprotein (mmol/L)	3.09 [2.29-3.46]	3.14 [2.221-3.52]	3.14 [2.33-3.49]	3.03 [2.23-3.98]	0.489	0.840
High-density lipoprotein (mmol/L)	0.94 [0.70-1.07]	0.90 [0.68-1.00]	0.90 [0.69-0.98]	0.90 [0.66-1.09]	0.035	0.831
Calcium (mmol/L)	2.19 [2.06-2.30]	2.16 [2.01-2.28]	2.16 [2.00-2.28]	2.17 [2.01-2.28]	0.001	0.814
Phosphorus (mmol/L)	1.19 [1.02-1.36]	1.23 [1.08-1.41]	1.22 [1.08-1.40]	1.26 [1.09-1.45]	0.001	0.287
Alkaline phosphatase (U/L)	67 [53-74]	70 [58-80]	70 [58-80]	70 [58-84]	0.001	0.695
C-reactive protein (mg/L)	0.70 [0.10-3.30]	0.80 [0.10-4.00]	0.80 [0.10-4.00]	0.80 [0.10-4.00]	0.061	0.440
Urine protein (g/L)	1.79 [0.87-4.05]	2.25 [1.00-5.32]	2.16 [1.00-4.94]	2.92 [0.98-6.53]	0.001	0.081
Parathyroid hormone (ng/L)	44.96 [29.90-64.95]	53.66 [30.61-73.08]	49.26 [30.46-68.13]	64.95 [31.61-94.89]	0.003	0.004
25-$(OH)_2$-D_3 (μg/mL)	12.60 [6.53-17.02]	10.99 [5.17-14.68]	11.59 [5.33-15.19]	9.56 [4.29-12.60]	0.001	0.018
Glucose (mmol/L)	5.06 [4.62-5.54]	5.49 [4.80-6.46]	5.39 [4.77-6.31]	5.79 [5.16-7.65]	0.001	0.001
Uric acid (μmol/L)	427 [359-502]	432 [361-503]	431 [361-503]	436 [372-505]	0.496	0.668
Urea nitrogen (mmol/L)	22.30 [16.90-33.40]	27.00 [19.60-42.20]	26.10 [19.10-40.508]	31.50 [23.10-48.40]	0.001	0.001
Creatinine (mg/Dl)	1.53 [1.18-2.34]	1.78 [1.26-3.05]	1.72 [1.24-2.90]	1.99 [1.34-3.62]	0.001	0.018
eGFR (mL/min/1.73 m^2)	50.11 [29.75-69.94]	39.12 [20.38-59.42]	40.28 [21.88-61.40]	31.75 [15.14-55.43]	0.001	0.009
Cystatin C (mg/L)	1.79 [1.32-2.37]	2.14 [1.54-2.83]	2.08 [1.52-2.79]	2.14 [1.63-3.45]	0.001	0.019
Calcium phosphorus product (m^2mol^2/L^2)	2.57 [2.21-2.96]	2.62 [2.30-3.02]	2.62 [2.29-3.01]	2.69 [2.36-3.13]	0.003	0.191

Data are presented as n (%) or mean ± SD or median [25th, 75th percentile]. P_1 refers to the statistic value between noncalcification and calcification groups, while P_2 refers to the statistic value between moderate and severe groups.

highest AUC was LR (0.70, 95% CI: 0.63-0.77), the second highest for SVM (0.65, 95% CI: 0.57-0.71), then for RF (0.62, 95% CI: 0.55-0.69), and KNN (0.53, 95% CI: 0.45-0.60). There was no significant difference in performance when compared the reference LR with SWM ($P = 0.2544$) and traditional LR prediction model (AUC: 0.70, 95% CI: 0.63-0.77; $P = 0.982$), but significantly outperformed the RF ($P = 0.021$) and KNN ($P = 0.023$).

TABLE 3. Predictors for Presence and Severity of CAC in Multivariable Analysis

Predictors	Odds Ratio	95% CI	*P*
Presence			
Age	1.079	1.070-1.088	< 0.001
Sex (male)	1.904	1.512-2.397	< 0.001
Diabetes	2.177	1.759-2.694	< 0.001
Hypertension	1.393	1.137-1.707	0.001
Oral calcium agents	1.564	1.283-1.907	< 0.001
PTH	1.002	1.001-1.003	0.004
Creatinine	0.922	0.859-0.989	0.023
HDL	0.688	0.520-0.911	0.009
Glucose	1.107	1.052-1.165	< 0.001
CPP	1.582	1.320-1.897	< 0.001
Severity			
Age	1.049	1.021-1.078	< 0.001
25-$(OH)_2$-D_3	0.971	0.943-0.999	0.044
Urea nitrogen	1.010	1.000-1.019	0.044
Glucose	1.140	1.067-1.219	< 0.001

The feature ranks of the corresponding top 10 variants derived by the corresponding 3 ML algorithms for the prevalence of CAC (feature ranks are not available for KNN) as shown in Figures 3A–C. Among the variables, nondiabetes, creatinine, nonhypertension, and HDL were protective factors, while age, CPP, male, oral calcium, glucose, and PTH were related to higher risk for the prevalence of CAC. The feature ranks of the corresponding top 5 variables derived by the corresponding 2 ML algorithms for the severity of CAC (feature ranks are not available for SVM and KNN). Among the variants, 25-$(OH)_2$-D_3 was protective, while age, glucose, creatinine, and urea nitrogen were related to higher risk for the severity of CAC (Figs. 3D, E).

DISCUSSION

This study identified some novel risk factors, such as glucose and creatinine, that were highly associated with the prevalence and severity of CAC in nondialysis CKD patients and built the prediction model using ML methods based on a large Chinese cohort. The established 4 ML models, especially LR had the best prediction ability in the internal validation cohort, indicating the robustness of our developed ML models. Moreover, ML-based methods (LR) showed similar accuracy in predicting the prevalence and severity of CAC versus conventional multivariable LR analysis. This is the first large, retrospective cohort study to investigate the predictive value of associated risk factors

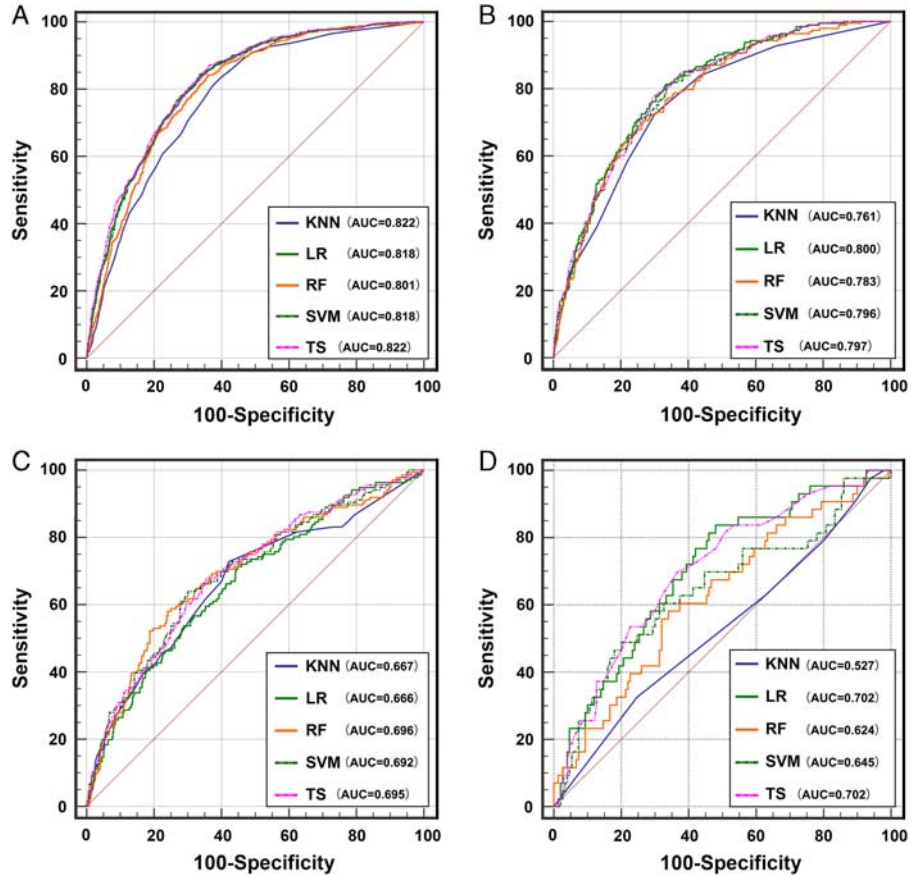

FIGURE 2. ROC curves and AUCs for the prevalence and severity of CAC in training and internal validation sets for LR, RF, SVM, KNN, and TS. A and B, ROC curves and AUCs for the prevalence of CAC in internal training (A) and internal validation (B) sets for LR, RF, SVM, and MLP, respectively. C and D, ROC curves and AUCs for the severity of CAC in internal training (C) and internal validation (D) sets for LR, RF, SVM, and MLP, respectively. TS indicates traditional logistic regression statistical. full color online

with the prevalence and severity of CAC based on ML methods in nondialysis Chinese CKD patients.

ML methods use computer algorithms to identify patterns in large data sets with a multitude of variables, and can be used to build the prediction models. ML has shown the potential to improve diagnostic accuracy and prognostic outcomes compared with conventional statistical methods.[19–21] The role of ML in building the prediction models for CVD has

TABLE 4. Performance of 4 ML and Traditional Logistic Regression Models to Predict the Prevalence and Severity of CAC in the Training and Internal Validation Data Sets

	Training Set, n = 2957					Internal Validation Set, n = 744				
Prevalence	**LR**	**RF**	**SVM**	**KNN**	**TS**	**LR**	**RF**	**SVM**	**KNN**	**TS**
AUC	0.82	0.80	0.82	0.78	0.82	0.80	0.78	0.80	0.76	0.80
95% CI	0.80-0.83	0.79-0.82	0.80-0.83	0.76-0.79	0.81-0.84	0.77-0.83	0.75-0.81	0.77-0.82	0.73-0.79	0.77-0.83
Sensitivity (%)	83.87	84.00	84.40	83.33	87.07	81.35	70.47	80.83	72.54	78.76
Specificity (%)	66.24	64.11	66.47	60.31	63.62	66.61	73.32	66.24	69.87	68.97
Delong test#	—	—	—	—	—	—	0.0476	0.0241	0.0020	0.2544
Severity	Training Set, n = 750					Internal Validation Set, n = 193				
AUC	0.67	0.70	0.69	0.67	0.70	0.70	0.62	0.65	0.53	0.70
95% CI	0.63-0.70	0.66-0.73	0.66-0.73	0.63-0.70	0.66-0.73	0.63-0.77	0.55-0.69	0.57-0.71	0.45-0.60	0.63-0.77
Sensitivity (%)	71.32	58.09	63.97	72.79	67.65	83.72	58.14	48.84	32.56	69.77
Specificity (%)	55.21	75.90	69.87	57.82	64.50	52.00	66.00	80.00	75.33	63.33
Delong test#	—	—	—	—	—	—	0.0208	0.2614	0.0234	0.9823

#$P < 0.05$ means a significant difference in the highest/lowest AUC of the ML and TS applied in the prevalence and severity internal validation set.
TS indicates traditional logistic regression statistical.

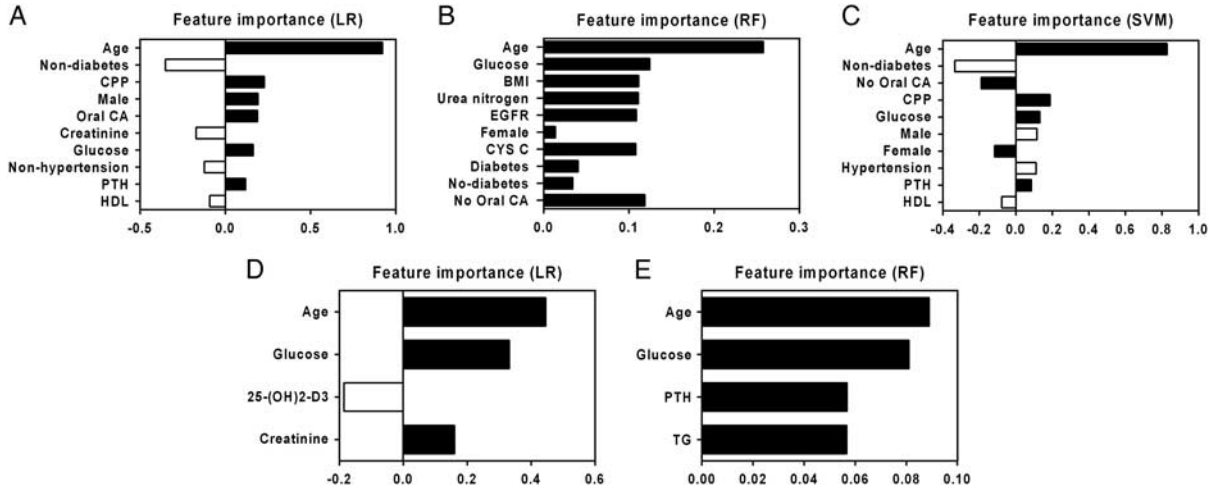

FIGURE 3. Feature ranks of the variables for the prevalence and severity of CAC derived from the corresponding ML algorithms. A–C, Feature ranks of top 10 variables (top to bottom) for the prevalence of CAC were derived by the corresponding LR, RF, and SVM. D and E, Feature ranks of top 10 variables (top to bottom) for the severity of CAC were derived from the corresponding LR and RF. The positive values were color-coded black, while negative values white. TS indicates traditional logistic regression statistical.

been studied and compared with the traditional LR method.[22–24] However, inconsistent performance between the ML methods and traditional LR method were obtained. For example, Motwani et al[25] reported that the prediction model based on ML methods was better than that of the traditional LR method in predicting 5-year all-cause mortality in patients undergoing coronary CT angiography. However, Chen et al[26] found that ML method-derived prediction model was ineffective in the prediction of a composite cardiovascular outcome in the Cardiovascular Outcomes in Renal Atherosclerotic Lesions (CORAL) population. Frizzell et al[27] found that ML algorithms did not improve prediction of 30-day heart failure readmissions compared with traditional prediction models. Christodoulou et al[28] found that the clinical prediction models based on ML lead to better AUCs than that based on LR. In this study, compared with the traditional multivariable LR analysis, the ML model (LR) showed a similar accuracy in predicting the presence and severity of CAC nondialysis CKD patients.

In this study, older age, sex (male), diabetes, and hypertension as independent predictors for the presence of CAC in nondialysis CKD patients are consistent with the Chronic Renal Insufficiency Cohort Study, while the findings were inconsistent with the Russo et al report.[29,30] In Chronic Renal Insufficiency Cohort Study, serum calcium and phosphate were also independent predictor factors for the severity of CAC in nondialysis CKD patients. In our prediction model, neither serum calcium level nor phosphorus level was an independent predictor factor for the severity of CAC. The difference for the risk factors between our study and previous studies can be interpreted as follows: (1) all previous risk factors for prediction of the prevalence and severity of CAC in nondialysis CKD patients were generated by traditional multivariable LR, which has several limitations, such as the nonlinear relationship between the factors and the outcome and the interactions among variables; (2) the cohort in this study was the Chinese population, and the former researches were based on American or European population. Race/ethnicity has been reported to be the independent risk factor for the prevalence and extent of coronary calcification, due to the lifestyle and economical variability. It was reported that in CKD patients the Asian population presented with the highest rates of CAC (64%, 95% CI: 54%-74%), followed by North America (61%, 95% CI: 51%-72%), European (59%, 95% CI: 42%-75%), and South America (53%, 95% CI: 3%-103%).[31] The Multi-Ethnic Study of Atherosclerosis (MESA) showed significant ethnic differences in CAC prevalence and severity.[32]

The present study had some limitations. First, this was a cross-sectional study which did not allow us to determine whether prevention or treatment of a risk factor could lead to an improvement in CAC prevalence and severity. Second, this was a single-center cohort and the absence of an external validation set. Although all samples were diagnosed in a Chinese population and from the signal center, which confirmed the examination used the standardized methods and minimized the bias, the absence of an external validation set may reduce the generalizability of our developed ML models. Third, our study was limited by the small number of moderate (Agatston score = 1 to 299) and severe (Agatston score $\geq$ 300) CAC patients. The low prevalence may reduce the prediction model performance in theory. For these subjects, a larger sample remains to be needed. Finally, further follow-up of the cohort will allow refinement of our risk estimates.

In conclusion, we developed the ML prediction models for the prevalence and severity of CAC in nondialysis Chinese CKD patients based on demographic information, clinical data and biochemical parameter characteristics, and we found that ML LR had the best predicting performance. The model can provide reliable quantitative individual risk assessments for the prevalence and severity of CAC in Chinese nondialysis CKD patients.

REFERENCES

1. Sarnak MJ, Amann K, Bangalore S, et al. Chronic kidney disease and coronary Artery disease: JACC state-of-the-art review. *J Am Coll Cardiol*. 2019;74:1823–1838.
2. Chen J, Budoff MJ, Reilly MP, et al. Coronary artery calcification and risk of cardiovascular disease and death among patients with chronic kidney disease. *JAMA Cardiol*. 2017;2:635–643.

3. Lai J, Akindavyi G, Fu Q, et al. Research progress on the relationship between coronary artery calcification and chronic renal failure. *Chin Med J (Engl)*. 2018;131:608–614.
4. Russo D, Corrao S, Miranda I, et al. Progression of coronary artery calcification in predialysis patients. *Am J Nephrol*. 2007;27:152–158.
5. Voelkl J, Cejka D, Alesutan I. An overview of the mechanisms in vascular calcification during chronic kidney disease. *Curr Opin Nephrol Hypertens*. 2019;28:289–296.
6. Lamarche MC, Hopman WM, Garland JS, et al. Relationship of coronary artery calcification with renal function decline and mortality in predialysis chronic kidney disease patients. *Nephrol Dial Transplant*. 2019;34:1715–1722.
7. Bundy JD, Chen J, Yang W, et al. Risk factors for progression of coronary artery calcification in patients with chronic kidney disease: the CRIC study. *Atherosclerosis*. 2018;271:53–60.
8. Obermeyer Z, Emanuel EJ. Predicting the future-big data, machine learning, and clinical medicine. *N Engl J Med*. 2016;375: 1216–1219.
9. Ahmad T, Lund LH, Rao P, et al. Machine learning methods improve prognostication, identify clinically distinct phenotypes, and detect heterogeneity in response to therapy in a large cohort of heart failure patients. *J Am Heart Assoc*. 2018;7:e008081.
10. Zack CJ, Senecal C, Kinar Y, et al. Leveraging machine learning techniques to forecast patient prognosis after percutaneous coronary intervention. *JACC Cardiovasc Interv*. 2019;12: 1304–1311.
11. Liu ZH, the China Dialysis Calcification Study (CDCS) Group. Vascular calcification burden of Chinese patients with chronic kidney disease: methodology of a cohort study. *BMC Nephrology*. 2015;16:1–7.
12. Stevens PE, Levin A. Kidney Disease: Improving Global Outcomes Chronic Kidney Disease Guideline Development Work Group Members. Evaluation and management of chronic kidney disease: synopsis of the kidney disease: improving global outcomes 2012 clinical practice guideline. *Ann Intern Med*. 2013;158:825–830.
13. Agatston AS, Janowitz WR, Hildner FJ, et al. Quantification of coronary artery calcium using ultrafast computed tomography. *J Am Coll Cardiol*. 1990;15:827–832.
14. Goff DC, Lloyd-Jones DM, Bennett G, et al. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *J Am Coll Cardiol*. 2014;63:2935–2959.
15. Hecht HS, Blaha MJ, Kazerooni EA, et al. CAC-DRS: Coronary Artery Calcium Data and Reporting System. An expert consensus document of the Society of Cardiovascular Computed Tomography (SCCT). *J Cardiovasc Comput Tomogr*. 2018;12:185–191.
16. Levey AS, Coresh J, Balk E, et al. National kidney foundation practice guidelines for chronic kidney disease: evaluation, classification, and stratification. *Ann Intern Med*. 2003;139:137–147.
17. Chen G, Lu M, Shi Z, et al. Development and validation of machine learning prediction model based on computed tomography angiography-derived hemodynamics for rupture status of intracranial aneurysms: a Chinese multicenter study. *Eur Radiol*. 2020;30:5170–5182.
18. Katz MH. *Multivariable Analysis: A Practical Guide for Clinicians and Public Health Researchers*. New York, NY: Cambridge University Press; 2011:1–233.
19. Beam AL, Kohane IS. Big data and machine learning in health care. *JAMA*. 2018;319:1317–1318.
20. Banerjee I, Sofela M, Yang J, et al. Development and performance of the pulmonary embolism result forecast model (PERFORM) for computed tomography clinical decision support. *JAMA Netw Open*. 2019;2:e198719.
21. van Velzen SGM, Lessmann N, Velthuis BK, et al. Deep learning for automatic calcium scoring in CT: validation using multiple cardiac CT and chest CT protocols. *Radiology*. 2020; 295:66–79.
22. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med*. 2019;25:44–56.
23. Sengupta PP, Kulkarni H, Narula J. Prediction of abnormal myocardial relaxation from signal processed surface ECG. *J Am Coll Cardiol*. 2018;71:1650–1660.
24. Johnson KM, Johnson HE. Scoring of coronary artery disease characteristics on coronary CT angiograms by using machine learning. *Radiology*. 2019;292:354–362.
25. Motwani M, Dey D, Berman DS, et al. Machine learning for prediction of all-cause mortality in patients with suspected coronary artery disease: a 5-year multicentre prospective registry analysis. *Eur Heart J*. 2017;38:500–507.
26. Chen T, Brewster P, Tuttle KR, et al. Prediction of cardiovascular outcomes with machine learning techniques: application to the Cardiovascular Outcomes in Renal Atherosclerotic Lesions (CORAL) study. *Int J Nephrol Renovasc Dis*. 2019;12:49–58.
27. Frizzell JD, Liang L, Schulte PJ, et al. Prediction of 30-day all-cause readmissions in patients hospitalized for heart failure: comparison of machine learning and other statistical approaches. *JAMA Cardiol*. 2017;2:204–209.
28. Christodoulou E, Ma J, Collins GS, et al. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *J Clin Epidemiol*. 2019; 110:12–22.
29. Russo D, Palmiero G, De Blasio AP, et al. Coronary artery calcification in patients with CRF not undergoing dialysis. *Am J Kidney Dis*. 2004;44:1024–1030.
30. He J, Reilly M, Yang W, et al. Risk factors for coronary artery calcium among patients with chronic kidney disease (from the Chronic Renal Insufficiency Cohort Study). *Am J Cardiol*. 2012; 110:1735–1741.
31. Wang XR, Zhang JJ, Xu XX, et al. Prevalence of coronary artery calcification and its association with mortality, cardiovascular events in patients with chronic kidney disease: a systematic review and meta-analysis. *Ren Fail*. 2019;41:244–256.
32. Fox CS, Larson MG, Keyes MJ, et al. Kidney function is inversely associated with coronary artery calcification in men and women free of cardiovascular disease: the Framingham Heart Study. *Kidney Int*. 2004;66:2017–2021.