A Novel Laboratory-Based Model to Predict the Presence of Obstructive Coronary Artery Disease
Comparison to Coronary Artery Disease Consortium 1/2 Score, Duke Clinical Score and Diamond-Forrester Score in China

Ling-yun Zhou,1,2 MD, Wen-jun Yin,1,2 MD, Jiang-lin Wang,1,2 MD, Can Hu,1,2,3 MD, Kun Liu,1,2,3 MD, Juan Wen,4 PhD, Li-ping Peng,4 PhD and Xiao-cong Zuo,1,2 PhD

Summary
Commonly used tools to assess the probability of obstructive-coronary artery disease (CAD) were derived based on Caucasian cohorts, with their performance in China is still unknown. Furthermore, most were established based on non-laboratory variables, contributing to the limited predictive ability to some extent. Thus, we developed and internally validated a laboratory-based model with data from a Chinese cohort of 8963 patients, with suspected stable chest pain, referred to catheter-based coronary angiography (CAG) from September 2007 to April 2019, and then compared the present model’s performance with the four most commonly used prediction tools, Coronary Artery Disease Consortium 1/2 Score (CAD1/2), Duke clinical score (DCS), and Diamond-Forrester score (DF). The final model was developed by random forest method, including 8 predictors derived from 70 variables. Five-fold cross-validation was performed to evaluate the model's prediction accuracy. In the external validation set, the present model showed a superior area under the receiver-operating curve (0.816), followed by DCS (0.66), CAD2 (0.61), CAD1 (0.59) and at last DF (0.58), respectively. Furthermore, the present model correctly classified 74.4% of obstructive-CAD patients as high-risk, and correctly classified more than one third of non-obstructive-CAD patients as low-risk. The present model’s net reclassification improvement (NRI) showed a significant positive reclassification over CAD1 (NRI = 0.60, P < 0.001), DF (NRI = 0.59, P < 0.001), CAD2 (NRI = 0.57, P < 0.001), and DCS (NRI = 0.43, P < 0.001). Decision curve analysis demonstrated that the present model provided a larger net benefit compared with CAD1/2, DCS, and DF. In conclusion, the novel model, using 8 laboratory and non-laboratory variables, performed well in risk stratifying patients with suspected chest pain regarding the presence of obstructive-CAD in the present Chinese cohort.

Key words: Risk stratification, Coronary artery stenosis, Prediction, Chest pain, Chinese Han population

Coronary artery disease (CAD) is a common cardiovascular disease with a high frequency of hospitalization, health care costs, and increased morbidity and mortality.1-3 Regularly, patients presenting with chest pain are told to undergo further diagnostic strategies to make an early and definite diagnosis of obstructive-CAD.4-6 The guideline-recommended, invasive, catheter-based coronary angiography (CAG) is the reference standard to diagnose obstructive-CAD because of its high sensitivity and specificity.6,7 However, CAG is expensive and associated with a risk of contrast-induced nephropathy, which increases in-hospital morbidity and mortality.6,7 Therefore, identifying patients at high risk of obstructive-CAD may benefit those at low risk of obstructive-CAD the most, helping them avoid excessive invasive testing. Many researchers made extensive efforts in this field,8-15 such as the Diamond-Forrester score (DF),16 the Coronary Artery Disease Consortium 1/2 score (CAD1/CAD2),17,18 and the Duke clinical score (DCS).19 However, all these models were developed based on age, sex, type of chest pain, and other non-laboratory variables, such as history of myocardial infarction, smoking, dyslipidemia, and diabetes mellitus. This may contribute to their limited predictive ability to some extent. Most recently, many researchers have made efforts to develop novel models by adding laboratory testing or biomarkers, such as genetic testing and several circulating proteins, to conventional risk factors to improve these tools’ predictive ability.16-26 Although
these models were conducted based on a much smaller cohort of patients, the results provided proof of the concept. Furthermore, prediction models perform well in populations similar to the ones for whom they were derived. As we know, all the aforementioned models were derived from Western populations. Moreover, these tools’ predictability remains unknown in the world’s developing regions, especially in China, where health resources are limited and risk stratification is urgently needed to avoid unnecessary health costs, for the lack of validation. Accordingly, in an attempt to provide further insight in this area, we undertook the present study to derive a laboratory-based prediction model from a large Chinese cohort of patients with suspected chest pain, and then assessed whether the present model enhances the predictability of obstructive-CAD compared to four most commonly used prediction tools, CAD1/2, DCS, and DF.

Methods

Ethics statement and study subjects: We performed a retrospective study included hospitalized patients with suspected chest pain who were referred for invasive CAG at the Third Xiangya Hospital of Central South University from September 2007 to April 2019. Patients undergoing CAG for acute processes, such as myocardial infarction and unstable angina pectoris, with a history of known CAD, and with more than 30% missing variables were excluded for the study. The study protocol was approved by the Medical Ethical Committee in the Xiangya Hospital of Central South University (No: 2017-S293). Patient written consent was waived as the data used in this study were anonymized. This study conformed to the ethical guidelines of the 1975 Declaration of Helsinki and Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD) guidelines.

Data acquisition and definitions: The index date was defined as the date of the CAG procedure. Detailed demographic and clinical characteristics were recorded from the electronic medical record system (EMRs) at the time of the procedure. The highest percent stenosis within any vessel on the CAG was recorded. Obstructive-CAD was defined as having two of above criteria. Non-specific chest pain was defined as having one or none of the criteria. Cigarette smoking was defined as current smoking or cessation of smoking within three months of the procedure.

Co-morbidities were identified by both the International Classification of Diseases, 10th Revision, coding (ICD 10) and detailed clinical information. Besides the diagnostic information, type 2 diabetes mellitus was also defined as fasting blood glucose (FBG) > 126 mg/dL, postprandial plasma glucose > 200 mg/dL, glycated hemoglobin > 6.5%, or the use of antidiabetic medications; hypertension was defined as repeated measurements of blood pressure > 140/90 mmHg or on treatment with antihypertensive medications; hyperlipidemia was defined as total cholesterol > 220 mg/dL, low-density lipoprotein cholesterol (LDL-C) > 140 mg/dL, fasting triglycerides > 150 mg/dL, high-density lipoprotein cholesterol (HDL-C) < 40 mg/dL, or the use of lipid-lowering medications; and chronic heart failure was defined as reduced left ventricular ejection fraction (LVEF) < 40%.

Statistical analysis: All statistical analyses were carried out using R software, V.3.3. The baseline continuous variable was presented as means ± standard deviation. The categorical variable was presented as absolute value and proportion. Baseline characteristics were compared between those with and without obstructive-CAD. A t test was used to compare the normally distributed continuous variables; otherwise, the Mann-Whiney U test was used. Categorical variables were compared by a chi-squared test. A 2-tailed value of P < 0.05 was established as the threshold of statistical significance.

The multiple imputation by chained equations (MICE) was used to deal with the missing CAG results. In present study, CAG was used to diagnose the presence of obstructive-CAD. However, an analysis restricted to patients undergoing invasive CAG could have been influenced by sample selection bias. Therefore, we imputed the missing CAG results for those patients who didn’t undergo CAG for numerous reasons by using the computed tomography (CT) based procedure, in addition to all other predictors. Consequently, the CT based procedure results were not included in the final predictive model. As a retrospective study, we hypothesized that the CAG results were missing at random. Therefore, the MICE package in R was used to perform multiple imputations based on the variables, including demographic information, co-morbidities, medications, and laboratory values.

All the enrolled patients were randomly split into a training set and a validation set in 3:1 manner. The method we used to develop and validate the prediction model was similar to that of our previous research, the random forests (RF) method, an ensemble of decision trees of high flexibility and sufficient accuracy, based on machine learning. All analyses for independent variable selection and model development were conducted on the training set. Mtry, the number of input variables randomly chosen at each split, and ntree, the number of trees in the forest, are two important parameters in RF. In the present study, the mtry is 3. The training data set is used to form the algorithm composed of 1000 trees, each of which is constructed using bootstrap samples from the training data and random feature selection. Each node is best split from a random selected set. When the RF algorithm best separates all instances, and this tree is able to classify all in-
Patient characteristics: Of a total of 21365 inpatients with suspected chest pain and referred to CAG, we finally included 8963 inpatients (5925 men, 3038 women) after exclusion criteria had been applied (Figure 1). Of them, 6499 patients underwent CAG, which revealed obstructive-CAD in 4720 (72.6%). Of the 2464 patients for whom we used multiple imputations to fill in the missing CAG results, 1327 (53.9%) underwent CT based procedure with negative results in 956 (72.1%). Baseline variables, including demographic information, comorbidities, medications, and laboratory values that differed between those with obstructive-CAD and those without obstructive-CAD are detailed in Table I and Supplemental Table.

Development and internal validation of present model: To develop and validate the prediction model, we randomly allocated three quarters \((n = 6722)\) of patients to the derivation dataset, with the remaining quarter \((n = 2241)\) assigned to the validation cohort. In the derivation cohort, those with obstructive-CAD \((4873 \text{ of } 6722, 72.5\%)\) tended to be older and male. Of the patients with obstructive-CAD, 3454 patients \((70.6\%)\) had evidence of ≥ 50% coronary stenosis in ≥ 2 vessels and 2105 patients \((43.1\%)\) had evidence of ≥ 50% coronary stenosis in ≥ 3 vessels. There were no statistical differences in the baseline characteristics between the training and validation sets (Table II).

Figure 2A shows the relationship between the cross-validation error and the number of variables. When there were eight variables, the error had a sharp decrease to 0.06. When the variables increased gradually to 70, the error remained at a similar level. Thus, our final model included eight indispensable features for obstructive-CAD prediction: age, gender, type of chest pain, hypertension, type 2 diabetes mellitus, smoking, LDL-C, and creatine. The eight variables’ importance is demonstrated in Figure 2B. The larger the importance number, the more important the variable. In addition, the change of AUC value is also used to assess the importance of the eight selected vari-
Table 1. Demographic, Clinical and Laboratory Data in Patients with or without Obstructive CAD in the Training Set

| Characteristics                        | Non-obstructive CAD | Obstructive CAD | P value |
|----------------------------------------|---------------------|-----------------|---------|
| Demographic                            |                     |                 |         |
| Age, years                             | 58.1 ± 10.2         | 62.9 ± 10.3     | 0.000   |
| Gender, female                         | 820 (44.8%)         | 1412 (28.9%)    | 0.000   |
| Vital Signs                            |                     |                 |         |
| Typical chest pain                     | 872 (47.6%)         | 3749 (76.6%)    | 0.000   |
| Atypical chest pain                    | 552 (30.18%)        | 614 (12.5%)     | 0.000   |
| Non-specific chest pain                | 405 (22.14%)        | 529 (10.81%)    | 0.000   |
| Medical history                        |                     |                 |         |
| Smoking                                | 457 (25.0%)         | 2089 (42.7%)    | 0.000   |
| Drinking                               | 472 (25.3%)         | 1766 (36.1%)    | 0.000   |
| Hypertension                           | 944 (51.6%)         | 3155 (64.5%)    | 0.000   |
| Type 2 diabetes mellitus               | 485 (26.5%)         | 2371 (48.5%)    | 0.000   |
| Laboratory values                      |                     |                 |         |
| White blood cell, × 10^9/L             | 6.67 ± 2.12         | 8.05 ± 2.17     | 0.000   |
| Red blood cell, × 10^12/L              | 4.45 ± 0.62         | 4.23 ± 0.59     | 0.000   |
| Hematocrit, %                          | 41.13 ± 3.89        | 40.58 ± 4.02    | 0.013   |
| Hemoglobin, g/L                        | 135.12 ± 15.47      | 133.27 ± 16.44  | 0.007   |
| Monocytes, × 10^9/L                    | 0.43 ± 0.15         | 0.48 ± 0.22     | 0.000   |
| Lymphocytes %                          | 28.12 ± 7.79        | 22.31 ± 8.86    | 0.000   |
| Neutrophil, × 10^9/L                   | 4.37 ± 2.11         | 5.62 ± 2.46     | 0.000   |
| Neutrophil %                           | 63.69 ± 8.71        | 68.12 ± 10.12   | 0.000   |
| Basophil %                             | 0.41 ± 0.26         | 0.35 ± 0.24     | 0.000   |
| Albumin, g/L                           | 40.13 ± 3.81        | 38.12 ± 4.02    | 0.000   |
| ALT, U/L                               | 30.66 ± 27.79       | 38.93 ± 98.34   | 0.003   |
| AST, U/L                               | 34.98 ± 59.02       | 77.32 ± 137.49  | 0.000   |
| Total bilirubin, μmol/L                | 15.04 ± 5.78        | 14.27 ± 5.81    | 0.012   |
| Total bile acids, μmol/L               | 5.05 ± 4.78         | 4.72 ± 4.93     | 0.001   |
| Globulin, g/L                          | 23.12 ± 4.08        | 27.66 ± 4.11    | 0.000   |
| Urea, mmol/L                           | 5.31 ± 1.76         | 5.76 ± 2.01     | 0.000   |
| Blood uric acid, mmol/L                | 319.62 ± 109.93     | 339.77 ± 112.38 | 0.000   |
| Creatinine, μmol/L                     | 74.66 ± 24.41       | 86.32 ± 41.66   | 0.000   |
| eGFR, mL/minute/1.73 m²                | 88.12 ± 19.64       | 80.12 ± 20.17   | 0.000   |
| Triglycerides, mmol/L                  | 1.61 ± 1.26         | 1.77 ± 1.46     | 0.005   |
| Total cholesterol, mmol/L              | 4.32 ± 0.73         | 4.66 ± 0.98     | 0.005   |
| HDL-C, mmol/L                          | 1.24 ± 0.30         | 1.14 ± 0.26     | 0.000   |
| LDL-C, mmol/L                          | 2.32 ± 0.71         | 2.58 ± 0.73     | 0.000   |
| FBG, mmol/L                            | 5.23 ± 1.64         | 6.12 ± 2.33     | 0.000   |
| Ca²⁺, mmol/L                           | 2.32 ± 0.19         | 2.26 ± 0.16     | 0.017   |
| INR                                    | 1.02 ± 0.12         | 1.04 ± 0.19     | 0.000   |
| LVEF, %                                | 61.69 ± 11.12       | 58.93 ± 10.38   | 0.019   |
| Medications                            |                     |                 |         |
| ACEI/ARB                               | 1029 (56.3%)        | 3876 (79.2%)    | 0.000   |
| β blocker                              | 1177 (64.4%)        | 4129 (84.4%)    | 0.000   |
| Statins                                | 1352 (73.9%)        | 4780 (97.7%)    | 0.000   |
| PPIs                                   | 616 (33.7%)         | 2678 (54.7%)    | 0.000   |
| Aspirin                                | 1269 (69.4%)        | 4804 (98.2%)    | 0.000   |
| P2Y12 receptor antagonists             | 446 (24.4%)         | 4355 (88.6%)    | 0.000   |
| Angiography results                    |                     |                 |         |
| ≥ 50% coronary stenosis in ≥ 2 vessels | 3454 (70.6%)        | 2105 (43.1%)    |         |
| ≥ 50% coronary stenosis in ≥ 3 vessels | 486 (26.6%)         | 3412 (69.7%)    | 0.000   |
| With invasive CAG only                 | 821 (44.9%)         | 155 (3.2%)      | 0.000   |
| With CT based procedure only           | 121 (6.6%)          | 945 (19.3%)     | 0.000   |
| With both invasive CAG and CT based procedure | 486 (26.6%) | 3412 (69.7%) | 0.000 |

ALT indicates alanine aminotransferase; AST, glutamic-oxalacetic transaminase; eGFR, estimated glomerular filtration rate; HDL-C, high density lipoprotein cholesterol; LDL-C, low density lipoprotein cholesterol; FBG, fasting blood glucose; INR, international normalized ratio; LVEF, left ventricular ejection fractions; ACEI, angiotension converting enzyme inhibitors; ARB, angiotensin receptor blockers; PPIs, proton pump inhibitors; CAG, coronary angiography; and CT, computed tomography.
I n t H e a r t J

May 2020 441

A NOVEL LABORATORY-BASED MODEL TO PREDICT CAD

Figure 2. A: The relationship between the cross-validation error and the number of variables. B: The importance of the eight obstructive-CAD variables. CAD indicates coronary artery disease; HP, hypertension; T2DM, type 2 diabetes mellitus; and Cr, creatine.

Table II. The Change of AUC Value When Each Variable Is Excluded in the Model and the Variables in the Training Set and the Validation Set

| Variables                  | The value of AUC after excluding a variable | The change value of AUC after excluding a variable | Training set ($n = 6722$) | Validation set ($n = 2241$) | $P$ value |
|----------------------------|--------------------------------------------|--------------------------------------------------|-----------------------------|----------------------------|-----------|
| All                        | 0.927                                      | 0.000                                            |                             |                            |           |
| Typical chest pain         | 0.847                                      | 0.080                                            | 4621 (68.7%)                | 1466 (65.4%)               | 0.251     |
| Type 2 diabetes mellitus   | 0.771                                      | 0.076                                            | 2856 (42.5%)                | 989 (44.1%)                | 0.143     |
| Smoking                    | 0.695                                      | 0.061                                            | 2546 (37.9%)                | 889 (39.7%)                | 0.212     |
| Gender, Female             | 0.649                                      | 0.046                                            | 2232 (33.2%)                | 803 (35.8%)                | 0.233     |
| Age, year                  | 0.627                                      | 0.022                                            | 60.2 ± 10.4                 | 60.9 ± 10.3                | 0.612     |
| LDL-C, mmol/L              | 0.612                                      | 0.015                                            | 2.44 ± 0.72                 | 2.46 ± 0.66                | 0.739     |
| Creatine, mmol/L           | 0.603                                      | 0.009                                            | 82.3 ± 33.4                 | 83.4 ± 41.2                | 0.677     |
| Hypertension               | 0.596                                      | 0.007                                            | 4099 (60.9%)                | 1414 (63.1%)               | 0.114     |

LDL-C indicates low density lipoprotein cholesterol.
Figure 3. Receiver operating characteristic analysis of the present predictive model, CAD1, DCS, and DF for the prediction of obstructive-CAD. CAD indicates coronary artery disease; CAD1, coronary artery disease consortium 1 score; DCS, Duke clinical score; DF, Diamond-Forrester score.

ables (Table II). There is no difference between the method based on AUC or out-of-bag validation in terms of sorting on the variables’ importance.

The present model obtained an AUC of 0.939 in the training set (Supplemental Figure A) and 0.927 in the internal validation set (Supplemental Figure B). Such results sufficiently indicated that a big separation for obstructive-CAD and non-obstructive CAD patients was indeed obtained from this prediction model. The present model’s prediction accuracy was 87.5%, with a sensitivity of 94.6%, specificity of 67.8%, and MCC of 0.668, respectively. The validation achieved 87.3% for accuracy, 94% for sensitivity, 69.3% for specificity, and the MCC for 0.669, respectively. The result of high prediction accuracy and successful prediction suggested that the new model was effective in predicting obstructive-CAD.

Comparison of predictive ability of present model to CAD1/2, DCS, and DF model in external validation cohort: Furthermore, we evaluated the present model’s predictive ability compared with the CAD1/2, DCS and DF in a cohort containing 1605 patients (1170 patients with obstructive-CAD and 435 patients without obstructive-CAD) for external validation. For the prediction of obstructive-CAD, the present predictive model had a superior area under the ROC (AUC = 0.816), followed by DCS (AUC = 0.66), CAD2 (AUC = 0.61), CAD1 (AUC = 0.59) and at last DF (AUC = 0.58), as shown in Figure 3.

To assess the present model’s improved risk category classification compared with each of the other models, we constructed a reclassification table, as shown in Table III. In those patients with obstructive-CAD, the present prediction model correctly classified 74.4% (871/1170) patients as high risk, and no patients were classified as low risk. On the other hand, in patients without obstructive-CAD, our model correctly classified more than one third of patients (33.2% (389/1170)), and 4.4% (165/1170) patients from a low or intermediate risk with CAD2, DF, CAD1, and DCS to the high-risk category, respectively. On the other hand, in patients without CAD, present model re-stratified 43.9% (176/435), 39.1% (170/435), 36.1% (157/435), and 26.2% (114/435) patients from intermediate or high risk with DCS, CAD1, DF, and CAD 2 to the low risk category, respectively. The total NRI of present model showed a significant positive reclassification over CAD1 (NRI = 60.1%, P < 0.001), DF (NRI = 59.4%, P < 0.001), CAD2 (NRI = 56.8%, P < 0.001), and DCS (NRI = 43.2%, P < 0.001).

Furthermore, a decision curve analysis was used to compare the different pre-test tools. As seen in Figure 5, the decision curve analysis graphically showed the net benefit of using each model to risk stratify patients. In present analysis, the novel tool provided a larger net benefit compared with the other four commonly used models.

Discussion

This study is the first to develop and validate both laboratory and non-laboratory variables-based prediction tools to estimate obstructive-CAD risk in a large Chinese cohort. Our results may assist clinicians in risk stratifying patients with suspected chest pain regarding the presence of obstructive-CAD in similar populations.

In the present cohort, 72.6% patients undergoing CAG were diagnosed as obstructive-CAD, while 72.1% of the patients undergoing CT based procedures showed negative results. It appears that the clinicians are prone to ordering invasive CAG, instead of CT based procedures, for patients considered more likely to have obstructive-CAD according to their clinical assessments. On the contrary, patients who are considered to be less likely to have obstructive coronary stenosis may undergo CT based procedures instead of invasive CAG when the physician can confidently rule out CAD with CT based procedures. Al-
though CAG is used as the “gold standard” for diagnosing obstructive-CAD, the high costs and invasiveness limit its clinical use, especially in developing countries, such as China. A CT based procedure is non-invasive, but the cost of CT angiography is still high in China. Using pre-test tools to identify patients who really need further diagnostic tests may reduce the healthcare costs for low-risk patients. Furthermore, contrast medium is necessary for both contrast-enhanced CT based procedures and invasive CAG. As reported in our previous study, the rate of acute kidney injury resulting from intravenous contrast-enhanced CT and CAG was 13.8% and 15.9%, respectively, which may contribute to continued loss of kidney function, chronic kidney disease, mortality, and increased costs. Using tools to screen patients who need further diagnostic tests may decrease the risk of contrast induced acute kidney injury. Moreover, CT based procedures or CAG may not be available in local Chinese medical centers; therefore, the present model may help clinicians determine which high-risk patients should undergo further diagnostic tests in higher levels of hospitals. Accordingly, developing this pre-test tool may be meaningful in those regions for identifying high-risk obstructive-CAD patients who need undergo further diagnostic strategies to reduce the disease

| Present predictive model | Low risk | Intermediate risk | High risk | Total |
|--------------------------|----------|-------------------|-----------|-------|
| With obstructive-CAD     |          |                   |           |       |
| CAD1                     |          |                   |           |       |
| Low risk                 | 0        | 0                 | 23        | 23    |
| Intermediate risk        | 0        | 245               | 366*      | 611   |
| High risk                | 0        | 54*               | 482       | 536   |
| Total                    | 0        | 299               | 871       | 1170  |
| CAD2                     |          |                   |           |       |
| Low risk                 | 0        | 23                | 105†      | 128   |
| Intermediate risk        | 0        | 241               | 471†      | 712   |
| High risk                | 0        | 35*               | 295       | 330   |
| Total                    | 0        | 299               | 871       | 1170  |
| DCS                      |          |                   |           |       |
| Low risk                 | 0        | 26                | 33†       | 59    |
| Intermediate risk        | 0        | 197               | 132†      | 329   |
| High risk                | 0        | 76*               | 706       | 782   |
| Total                    | 0        | 299               | 871       | 1170  |
| DF                       |          |                   |           |       |
| Low risk                 | 0        | 13                | 34†       | 47    |
| Intermediate risk        | 0        | 177               | 377†      | 554   |
| High risk                | 0        | 109*              | 460       | 569   |
| Total                    | 0        | 299               | 871       | 1170  |
| Without obstructive-CAD  |          |                   |           |       |
| CAD1                     |          |                   |           |       |
| Low risk                 | 21       | 33*               | 0         | 54    |
| Intermediate risk        | 170†     | 211               | 0         | 381   |
| High risk                | 0        | 0                 | 0         | 0     |
| Total                    | 191      | 244               | 0         | 435   |
| CAD2                     |          |                   |           |       |
| Low risk                 | 77       | 68*               | 0         | 145   |
| Intermediate risk        | 114†     | 176               | 0         | 290   |
| High risk                | 0        | 0                 | 0         | 0     |
| Total                    | 191      | 244               | 0         | 435   |
| DCS                      |          |                   |           |       |
| Low risk                 | 15       | 21*               | 0         | 36    |
| Intermediate risk        | 54†      | 102               | 0         | 156   |
| High risk                | 122†     | 121               | 0         | 243   |
| Total                    | 191      | 244               | 0         | 435   |
| DF                       |          |                   |           |       |
| Low risk                 | 34       | 11*               | 0         | 45    |
| Intermediate risk        | 96†      | 177               | 0         | 273   |
| High risk                | 61†      | 56                | 0         | 117   |
| Total                    | 191      | 244               | 0         | 435   |

*Inadequate reclassification. †Correct reclassification. CAD indicates coronary artery disease; CAD1, coronary artery disease consortium 1 score; DCS, Duke clinical score; and DF, Diamond-Forrester score.
Prediction tools perform best when used in populations similar to the cohorts from which they were derived. Hence, each prediction model has its own limitation in external validation. For example, the DF model overestimated the probability of obstructive-CAD in women since there were fewer female patients in the study cohort (6983 women versus 17013 men). On the contrary, CAD2, which added modifiable cardiovascular risk factors to the model, seemed to underestimate the actual prevalence of obstructive-CAD because it was created to estimate the pre-test probability of CAD in low-prevalence populations. Furthermore, almost all the prediction tools were developed for Caucasian patients and have undergone limited external validation in other regions of the world. Besides that, no prediction model has been developed and validated based on a Chinese population until now. As we know, some differences exist in the genetic or environmental factors, as well as lifestyle patterns, associated with the different frequencies of obstructive-CAD among the different races of the world. South Asian populations are believed to have a higher risk of obstructive-CAD than other ethnic cohorts. Consequently, a Caucasian population and Chinese population may have different risk factors associated with obstructive-CAD. In fact, all the above-mentioned models were not used regularly in Chinese clinical practice because of the well-known poor risk stratification capacity. As a result, it may be of clinical importance to establish a prediction model based on the Chinese population when assessing obstructive-CAD probability.

In the present study, we derived the prediction model
from 8963 Chinese patients. In general, the more variables chosen, the better the predictive discriminative ability. However, increasing the number of variables may make it difficult for clinical practice, so it is important to balance the prediction ability and the number of predictors in a model. However, previous prediction tools mostly focused on the convenience of clinical practice, resulting in limited enrolled predictors. None of the published models were established based on both non-laboratory and laboratory values, limiting their discriminative ability in general patients to some extent. The external validation of these models in the present Chinese cohort showed the present predictive model had superior AUC (0.816), followed by DCS (AUC = 0.66), CAD2 (AUC = 0.61), CAD1 (AUC = 0.59) and at last DF (AUC = 0.58). In clinical practice, laboratory testing can provide direct evidence to help physicians assess the patient’s current status. Thus, we recorded all the clinical information, including non-laboratory and laboratory variables, upon admission. After investigating the independent risk factors associated with the obstructive-CAD, eight predictors were finally confirmed. Although some variables enrolled in our model are similar to published prediction tools,24-31 adding two laboratory variables to the model seems to have incremental effects on its discriminative ability in present study. Both of the laboratory values (LDL-C and Cr) are tested regularly in different health resource settings with no additional costs in routine clinical practice, especially in the low- or middle-income countries, making the model more likely to be used in these regions.

To assess the present model’s performance in allocating participants to the proper risk category compared with the CAD1/2, DCS, and DF, which are recommended by recently updated guidelines, we applied those risk equations to an external validation cohort. As the results show, our prediction tool correctly classified as much as 74.4% of patients as high risk in those diagnosed with obstructive-CAD by invasive CAG. In patients without obstructive-CAD, the present prediction model correctly classified more than one third of patients (43.9%) as low risk. The most important finding is that no obstructive-CAD patients were classified as low risk by this model, and no non-obstructive CAD patients were classified as high risk. The decision curve analysis confirmed that this novel tool provided a larger net benefit compared with the other four commonly used models. Accordingly, this model performed well in risk stratifying patients with suspected chest pain regarding the presence of obstructive-CAD.

**Limitations:** Our study has several limitations. First, selection bias; to avoid this, we enrolled all the eligible patients with or without CAG results. Furthermore, we collected almost all the data recorded in the EMRs, including demographic, information, co-morbidities, concomitant drugs, and laboratory values, for each patient. MICE based on CT results and other clinical characteristics was used to impute missing CAG results. However, a large prospective study is needed to validate our data and shed additional light on this area. Second, in the present study, the eight variables used to predict obstructive-CAD were conventional CAD risk factors. Although a pre-test tool including both conventional factors and new biomarkers associated with CAD may improve predictability, we did not include the novel variables, such as genetic testing results. This is because of the limited information abstracted from the hospital information system, the inconvenience for clinical practice, and the willingness to reduce medical care costs, which is an urgent need, especially in low- or middle-income countries. We hope that present model, based on conventional CAD risk factors, may be a good option for cardiologists to identify patients who really need further diagnostic tests and reduce the disease burden in China. Third, although we abstracted the data as completely as we could, several clinical variables and complications that were not documented in the EMR could be missing and may contribute to unmeasured confounding.

**Conclusion**

The present model using eight laboratory and non-laboratory variables was developed, internally and externally validated in a large Chinese cohort, and showed superior predictive ability over CAD1/2, DCS, and DF models. This novel pre-test tool may be a choice for clinicians to risk stratify patients with chest pain for obstructive-CAD in China. This model’s utility in clinical practice should be confirmed in further multi-center prospective studies.

**Disclosure**

**Conflicts of interest:** The authors have no conflicts of interest to disclose.

**References**

1. Musunuru K, Kathiresan S. Genetics of common, complex coronary artery disease. Cell 2019; 177: 132-45.
2. Mintz GS, Gargiulini G. Intravascular imaging in coronary artery disease. Lancet 2017; 390: 793-809.
3. Task Force Members, Montalescot G, Sechtem U, et al. 2013 ESC guidelines on the management of stable coronary artery disease: the Task Force on the management of stable coronary artery disease of the European Society of Cardiology. Eur Heart J 2013; 34: 2949-3003.
4. Roffi M, Patrono C, Collet JP, et al. 2015 ESC Guidelines for the management of acute coronary syndromes in patients presenting without persistent ST-segment elevation: Task Force for the Management of Acute Coronary Syndromes in Patients Presenting without Persistent ST-Segment Elevation of the European Society of Cardiology (ESC). Eur Heart J 2016; 37: 267-315.
5. Ibanez B, James S, Agewall S, et al. 2017 ESC Guidelines for the management of acute myocardial infarction in patients presenting with ST-segment elevation: the Task Force for the management of acute myocardial infarction in patients presenting with ST-segment elevation of the European Society of Cardiology (ESC). Eur Heart J 2018; 39: 119-77.
6. Nash K, Hafeez A, Hou S. Hospital-acquired renal insufficiency.
Am J Kidney Dis 2002; 39: 930-6.
7. Piccolo R, Giustino G, Mehran R, Winnacker S. Stable coronary artery disease: revascularisation and invasive strategies. Lancet 2015; 386: 702-13.
8. Diamond GA, Forrester JS. Analysis of probability as an aid in the clinical diagnosis of coronary-artery disease. N Engl J Med 1979; 300: 1350-8.
9. Almeida J, Fonseca P, Dias T, et al. Comparison of coronary artery disease consortium 1 and 2 scores and dukes clinical score to predict obstructive coronary disease by invasive coronary angiography. Clin Cardiol 2016; 39: 223-8.
10. Genders TSS, Steyerberg EW, Alkadhi H, et al. A clinical prediction rule for the diagnosis of coronary artery disease: validation, updating, and extension. Eur Heart J 2011; 32: 1316-30.
11. Genders TSS, Steyerberg EW, Hunink MGM, et al. Prediction model to estimate presence of coronary artery disease: retrospective pooled analysis of existing cohorts. BMJ 2012; 344: e3485.
12. Pryor DB, Jr HFE, Lee KL, Calif RM, Rosati RA. Estimating the likelihood of significant coronary artery disease. Am J Med 1983; 75: 771-80.
13. Fihn SD, Gardin JM, Abrams J, et al. 2012 ACCF/AHA/ACP/AATS/PCNA/SCAI/STS guideline for the diagnosis and management of patients with stable ischemic heart disease: A report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines, and the American College of Physicians, American Association for Thoracic Surgery, Preventive Cardiovascular Nurses Association, Society for Cardiovascular Angiography and Interventions, and Society of Thoracic Surgeons. Circulation 2012; 126: e354-471.
14. LaFramboise WA, Dhir R, Kelly LA, et al. Serum protein profiles predict coronary artery disease in symptomatic patients referred for coronary angiography. BMC Med 2012; 10: 157.
15. Bolton JL, Stewart MCW, Wilson NF, Anderson N, Price JF. Improvement in prediction of coronary heart disease risk over conventional risk factors using SNPs identified in genome-wide association studies. PLOS ONE 2013; 8: e57310.
16. Ibrahim NE, Jr JJJ, Magaret CA, et al. A clinical and biomarker scoring system to predict the presence of obstructive coronary artery disease. J Am Coll Cardiol 2017; 69: 1147-56.
17. Yin WJ, Yi YH, Guan XF, et al. Preprocedural prediction model for contrast-induced nephropathy patients. J Am Heart Assoc 2017; 6: e004498.
18. Gong R, Chen MH, Peng LS, Wei SL. Common genes in coronary artery disease from Europe, Asia and North America regardless of race and lifestyle. Eur Rev Med Pharmacol Sci 2013; 19: 1092-100.
19. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. Circulation 2007; 115: 928-35.
20. Leening MIG, Vedder MM, Wittman JCM, Pencina MJ, Steyerberg EW. Net reclassification improvement: computation, interpretation, and controversies: a literature review and clinician’s guide. Ann Intern Med 2014; 160: 122-31.
21. Vickers AJ, Pepe M. Does the net reclassification improvement help us evaluate models and markers? Ann Intern Med 2014; 160: 136-7.
22. Calster BV, Wynants L, Verbeek JFM, et al. Reporting and interpreting decision curve analysis: A guide for investigators. Eur Urol 2018; 74: 796-804.
23. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making 2006; 26: 565-74.
24. Williams MA, Fleg JL, Ades PA, et al. Secondary prevention of coronary heart disease in the elderly (with emphasis on patients > 75 years of age): an American Heart Association scientific statement from the Council on Clinical Cardiology Subcommittee on Exercise, Cardiac Rehabilitation, and Prevention. Circulation 2002; 105: 1735-43.
25. Nakazato R, Arsanjani R, Achenbach S, et al. Age-related risk of major adverse cardiac event risk and coronary artery disease extent and severity by coronary CT angiography: results from 15187 patients from the International Multisite CONFIRM Study. Eur Heart J Cardiovasc Imaging 2014; 15: 586-94.
26. Naito R, Miyauchi K. Coronary artery disease and type 2 diabetes mellitus. Int Heart J 2017; 58: 475-80.
27. Ham mond T, T anguy F, Bourassa MG. Management of coronary artery disease: therapeutic options in patients with diabetes. J Am Coll Cardiol 2000; 36: 355-65.
28. Ali SF, Smith EE, Reeves MJ, et al. Smoking paradox in patients hospitalized with coronary artery disease or acute ischemic stroke: findings from get with the guidelines. Circ Cardiovasc Qual Outcomes 2015; 8: S73-80.
29. Nakarnishi R, Baskaran L, Gransar H, et al. Relationship of hypertension to coronary atherosclerosis and cardiac events in patients with coronary computed tomographic angiography. Hypertension 2017; 70: 293-9.
30. Lee TM, Lin YJ, Su SF, et al. Relation of systemic arterial pulse pressure to coronary atherosclerosis in patients with mitral stenosis. Am J Cardiol 1997; 80: 1035-9.
31. Frohlich ED. State of the Art lecture. Risk mechanisms in hypertensive heart disease. Hypertension 1999; 34: 782-9.

**Supplemental Files**

Supplemental Table
Supplemental Figure
Please see supplemental files: https://doi.org/10.1536/ihj.19-532