Development of various-type feature-based artificial intelligence gatekeeper solution for the screening of coronary artery disease

Hyung-Bok Park

Department of Medicine

The Graduate School, Yonsei University

Development of various-type feature-based artificial intelligence gatekeeper solution for the screening of coronary artery disease

Directed by Professor Hyuk-Jae Chang

The Doctoral Dissertation
submitted to the Department of Medicine,
the Graduate School of Yonsei University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Medical Science

Hyung-Bok Park

December 2022

This certifies that the Doctoral Dissertation
of Hyung-Bok Park is approved.

Thesis Supervisor: Hyuk-Jae Chang

Thesis Committee Member#1: Jung-Sun Kim

Thesis Committee Member#2: Bon-Kwon Koo

Thesis Committee Member#3: Woong Kook

Thesis Committee Member#4: Dong-Ho Shin

The Graduate School
Yonsei University

December 2022

ACKNOWLEDGEMENTS

My most profound appreciation should go to Professor Hyuk-Jae Chang, my Ph.D. advisor, and mentor, for his time, effort, and understanding in helping me succeed in my research. He gave me tremendous help and support throughout my entire cardiology fellowship and Master's degree. I cannot appreciate him enough.

I also would like to say special thank you to Dr. Yongtaek Hong, Yeonggul Jang, Hyunseok Jeong, and Jina Lee for their consistent guidance and support during this project. Their vast wisdom and wealth of experience have inspired me throughout my research.

Furthermore, completing this collaborative project was only possible with the participation and wise guidance of the Seoul National University Mathematical science research team, Professors Woong Kook and Otto van Koert, Drs. Wonse Kim and Sung-Tae Jin, and Byungchang So. I was extremely privileged to participate in this remarkable project with all the participants and present the results as my doctoral dissertation.

To conclude, I'd like to thank God, my wife, Youngmi Han, and my two sons, Euihyun Park and Euichan Park. It would have been impossible to finish my Master's degree without their unwavering support over the past years.

<TABLE OF CONTENTS>

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Development of various-type feature-based artificial intelligence gatekeeper solution for the screening of coronary artery disease

Hyung-Bok Park

Department of Medicine
The Graduate School, Yonsei University

(Directed by Professor Hyuk-Jae Chang)

Introduction: The "primary tests" consisted of cardiovascular (CV) risk factors, chest X-ray radiography (CXR), and electrocardiography (ECG) for screening of coronary artery disease (CAD) generally lead to further downstream tests, which are not readily accessible for primary physicians or cost-effective. Therefore, we aimed to develop the various-type feature-based artificial intelligence (AI) gatekeeper solutions for CAD prediction to improve the diagnostic accuracy of these "primary tests".

Methods: We analyzed four clinical trials of CONSERVE (NCT01810198), CREDENCE (NCT02173275), 3V FFR-FRIENDS (NCT01621438), and PARADIGM (NCT02803411). All four datasets (n=5,643) were used for machine-learning ridge-regression analysis to generate a risk factor-based CAD prediction model, and a serially followed up PARADIGM database (n=1,463) was used for Bayesian quantile regression (BQR) analysis to explore the comprehensive relationship between CV risk factors and coronary artery stenosis and its progression. Meanwhile, coronary artery calcium (CAC) score having patients' dataset (n=559) was used for developing radiomic score-based machine learning model for prediction of ≥100 CAC score from CXR and the deep convolutional neural network (CNN) based ischemia detection in ECG. The integrated AI-gatekeeper solution model for CAD prediction was developed by extracting significant features from

those three different CAD prediction models.

Results: (1) The CV risk factors-based CAD prediction model by ridge-regression method showed a good performance of AUC 0.75 (95% CI 0.69-0.81). In BQR analysis, the 90th percentiles of the DS of the three vessels and their maximum DS change were 41%–50% and 5.6%–7.3%, respectively. Typical anginal symptoms were associated with the highest quantile (90%) of DS in the LAD; diabetes was associated with higher quantiles (75% and 90%) of DS in the LCx; dyslipidemia was associated with the highest quantile (90%) of DS in the RCA, and other symptoms showed some association with the LCx and RCA. High-density lipoprotein cholesterol showed a dynamic association with DS change in the per-patient analysis. (2) The radiomic score was the most prominent factor for CAC score ≥100 prediction (Odds ratio = 2.33; 95% Confidence interval [CI] = 1.62-3.44; $p < 0.001$) compared to clinical information. The radiomic score-based machine learning model showed AUC 0.84; 95% CI = 0.79-0.87 in predicting CAC score ≥100. (3) The CNN-based ischemia detection ECG model showed a modest result of AUC 0.60 (95% CI 0.54-0.66) in the training cohort and AUC 0.60 (95% CI 0.50-0.69) in the validation cohort. However, when the clinical variables were added, the model performance significantly improved both in the training and validation cohort up to an AUC of 0.71 (95% CI 0.66-0.76) and 0.67 (95% CI 0.59-0.69). (4) The final AUC of the integrated model for CAD prediction based on clinical risk factors, CXR, and ECG was AUC 0.77 (Sensitivity 73%, Specificity 69%) in the external validation.

Conclusions: In this study, we developed the integrated various-type feature-based AI-gatekeeper solution for CAD prediction using "primary tests" composed of CV risk factors, CXR, and ECG. This novel method may be widely applicable to clinical practice and improve the pre-test probability of coronary artery disease, particularly in the primary care setting.

Key words : artificial intelligence, coronary artery disease, cardiovascular risk factors, chest X-ray radiography, electrocardiography

Development of various-type feature-based artificial intelligence gatekeeper solution for the screening of coronary artery disease

Hyung-Bok Park

Department of Medicine

The Graduate School, Yonsei University

(Directed by Professor Hyuk-Jae Chang)

I. INTRODUCTION

Cardiovascular disease (CVD) is the primary cause of morbidity and mortality worldwide, with a global burden of 17 million deaths annually.[1] Coronary artery disease (CAD) accounts for over 50% of deaths and continues to increase.[2] Patients with chest pain or dyspnea, which are common symptoms that force them to visit the hospital, would be asked about clinical risk factors of CAD and tested by chest x-ray (CXR), electrocardiogram (ECG), and routine laboratory tests, including lipid profile.

These so-called “primary tests” for screening CAD generally lead to further downstream functional tests such as treadmill test, stress echocardiography, and stress/rest myocardial single-photon emission computed tomography.[3] However, the diagnostic yield for obstructive CAD through the confirmative test of invasive coronary angiography (ICA) hasn’t exceeded 50%.[4] Given that more than half of the patients have been inevitably getting unnecessary ICAs, revealing non-obstructive CADs that do not require coronary

intervention, the cost of those unnecessary ICAs is approximately 2 billion dollars in the USA and 4.8 to 12 trillion won in Korea.[3] To overcome these issues, non-invasive coronary computed tomography angiography (CCTA) has emerged as the gatekeeper for ICA, enhancing the diagnostic yield for obstructive CAD to up to 70%.[5] However, CCTA also has the potential risk of radiation hazard and contrast medium use and overestimation of CAD severity or non-interpretable issues due to artifacts or poor image quality, typically requiring confirmatory ICA testing.[6,7] Most of all, these downstream tests are not readily accessible for primary physicians nor are they cost-effective.

Recently introduced artificial intelligence (AI) machine learning or deep learning technology provides a new opportunity to shed new light on “primary tests” to find creative, cost-effective solutions. We, therefore, using artificial intelligence methods, sought to improve the diagnostic accuracy of the “primary tests” such as clinical risk factors, including laboratory tests, CXR, and ECG. Our specific study aims are as follows: (1) to evaluate the comprehensive dynamic relationship between multiple cardiovascular (CV) risk factors and different stages of CAD by applying the Bayesian quantile regression model using a serially followed-up database in a vessel-specific manner. And then configure the clinical risk factor model using the complete database; (2) to develop the integrated framework combined machine learning and radiomic scores for identifying moderate-to-severe coronary artery calcium (CAC) using CXR; (3) to determine the ischemic changes in resting ECG applying deep convolutional neural network (CNN) method; (4) to develop the various-type feature-based integrated AI gatekeeper solution for CAD prediction.

II. MATERIALS AND METHODS

1. Study participants and pooled analysis of 4 clinical studies

We analyzed the previously published three randomized controlled trials: CONSERVE (clinicaltrials.gov: NCT01810198),[8] CREDENCE (clinicaltrials.gov: NCT02173275),[9] and 3V FFR-FRIENDS (clinicaltrials.gov: NCT01621438)[10] which are open-label, international, multicenter trials as well as open-label, international, multicenter dynamic observational PARADIGM (clinicaltrials.gov: NCT02803411) registry.[11] The Institutional Review Board of Yonsei University College of Medicine approved this study (IRB Number: Y-2017-0084). All procedures were performed by the ethical standards of the Declaration of Helsinki, revised in 2013. Due to the retrospective nature of this study design, sample size calculation was not needed, and the patient's informed consent can be waived since the subject selection will be made from previously published anonymized data.

2. AI-based clinical risk factor model for CAD prediction

A. Study design of Bayesian quantile regression model using PARADIGM dataset

We analyzed the data from PARADIGM registry designed to track coronary atherosclerosis in serially acquired coronary computed tomography angiography (CCTA). Between 2003 and 2015, 2,252 consecutive patients with suspected or known CAD who underwent serial CCTA at an interscan interval of ≥2 years were enrolled. After the exclusion of patients with non-interpretable scans at baseline or follow-up CCTA (n=492), documented CAD before baseline CCTA (n=227), and incomplete clinical information such as CV risk factors, symptom variables, and laboratory results at baseline or follow-up CCTA (n=70), 1,463 patients who underwent per-segment-based quantitative CCTA plaque analysis including lumen diameter stenosis (DS) were evaluated.

The baseline clinical characteristics and laboratory data were used as clinical variables and the per-segment-based quantitative CCTA findings were used for a set of outcomes. We performed a vessel-wise analysis with these data at all outcome-level settings using the Bayesian truncated quantile regression model. For the vessel-wise analysis, all 18 coronary segments were classified into the following three vessels: left anterior descending (LAD), left circumflex (LCx), and right coronary artery (RCA). The largest quantitative DS measurement in each vessel (LAD, LCx, or RCA) was regarded as the representative value for each vessel and the largest DS among the vessels was regarded as the representative value for each patient. Most often, the LAD was included (n=1,264) followed by the RCA (n=864) and the LCx (n=718).

We tested the following two models: the DS model (Model 1) and DS change model (Model 2). Multiple CV risk factors including the symptom variables were used to predict quantile DS values for the three vessels and each patient in Model 1 and also used to predict quantile DS changes in Model 2. The quantile regression model for DS prediction (Model 1) was defined as follows:

$$\begin{gathered}\textit{DS of}\ \text{LAD, LCx, RCA, and per-patient} \\ = \boldsymbol{\alpha} + \sum_{i=1}^{7} \boldsymbol{\beta_i} \cdot \boldsymbol{Baselines_i} + \sum_{i=1}^{4} \boldsymbol{\gamma_i} \cdot \boldsymbol{Symptom\ Types_i} + \sum_{i=1}^{3} \boldsymbol{\delta_i} \cdot \boldsymbol{Lab\ Exams_i} + \boldsymbol{\epsilon_\theta}\end{gathered} \quad (1)$$

where $\boldsymbol{Baselines_i}$ were baseline CV risk factors including age, sex, body mass index (BMI), smoking, diabetes, hypertension, and dyslipidemia; $\boldsymbol{Symptom\ Types_i}$ were categorical risk factors denoting the types of patients' symptoms comprised "typical angina, atypical angina, non-cardiac pain, and others" with "asymptotic" as the reference category; $\boldsymbol{Lab\ Exams_i}$ were continuous variables from laboratory examinations including high-

density lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol (LDL-C), and triglycerides (TG); ϵ_θ was the error term with its θth quantile equal to zero (in our study, θ were 10%, 25%, 50%, 75%, and 90%).

Model 2 used the changes in DS values as the outcome variable, and the quantile regression model was specified as follows:

DS change of LAD, LCx, RCA, and per-patient

$$= \alpha + \sum_{i=1}^{7} \beta_i \cdot \boldsymbol{Baselines}_i + \sum_{i=1}^{4} \gamma_i \cdot \boldsymbol{Symptom\ Types}_i + \sum_{i=1}^{3} \delta_i \cdot \boldsymbol{Lab\ Exams}_{.i} + \epsilon_\theta \quad (2)$$

B. Study design of machine-learning ridge-regression model using the complete dataset (CONSERVE, CREDENCE, 3V FFR-FRIENDS, and PARADIGM)

We used machine-learning ridge-regression model which is a logistic regression model with L-2 penalty to develop risk factors based CAD prediction model. A complete database combining CONSERVE, CREDENCE, 3V FFR-FRIENDS, and PARADIGM studies (n=5,643) was used for the training dataset after excluding patients with missing clinical variables (Table 1). An external validation dataset (n=249) was used from Bundang Seoul national university hospital (Table 2). Parameters from ridge-regression model were trained in the training dataset and pre-trained model was inferred to the validation dataset.

Table 1. Baseline characteristics of complete dataset for risk factor model

	Patients (n=5,643)
Age, years	59.63 ± 10.70
Male	1095 (64.8%)
Body mass index, kg/m^2	25.03 ± 3.36
Current smoker	866 (15.34%)
Diabetes mellitus	733 (13.02%)
Hypertension	2,993 (53.04%)
Dyslipidemia	2,396 (42.46%)
Laboratory data	
HDL cholesterol, mg/dL	49.94 ± 20.50
LDL cholesterol, mg/dL	113.80 ± 35.15
Triglycerides	139.34 ± 86.04
Symptoms	
Asymptomatic	1332 (23.60%)
Typical angina	820 (14.53%)
Atypical angina	1974 (34.98%)
Non-cardiac pain	1214 (21.51%)
Others	385 (6.82%)
Obstructive CAD DS ≥ 50%	851 (15.1%)

Values are presented as means ± SDs or n (%)

CAD, coronary artery disease; LDL, low-density lipoprotein; HDL, high-density lipoprotein

Table 2. Baseline characteristics of validation dataset for risk factor model

	Patients (n=249)
Age, years	67.93 ± 10.99
Male	158 (63.2%)
Body mass index, kg/m^2	25.37 ± 3.55
Current smoker	57 (22.8%)
Diabetes mellitus	82 (32.8%)
Hypertension	152 (60.8%)
Dyslipidemia	80 (32%)
Laboratory data	
HDL cholesterol, mg/dL	51.41 ± 21.95
LDL cholesterol, mg/dL	90.116 ± 30.63
Triglycerides	114.56 ± 57.11
Symptoms	
Asymptomatic	55 (22.0%)
Typical angina	100 (40%)
Atypical angina	62 (24.8%)
Non-cardiac pain	33 (13.2%)
Others	0 (0%)
Obstructive CAD DS ≥ 50%	129 (51.6%)

Values are presented as means ± SDs or n (%)

Same abbreviation was used as in Table 1

3. Radiomics scoring based ≥100 calcium score prediction model from CXR

A. Study populations

We analyzed the combined data from two independent clinical studies, PARADIGM and CREDENCE. We included 652 patients who underwent CAC scanning at Severance Hospital, Seoul, South Korea, between January 1, 2010, and June 31, 2016. The inclusion criteria were as follows: (1) CXR examination performed within six months before or after the CAC scan and (2) availability of a posterior-anterior CXR view. We excluded 93 patients who had poor CXR image quality and undesirable objects projected onto the cardiac contour over CXR, such as pacemaker, implantable cardioverter-defibrillator leads, or other foreign materials. A total of 559 patients (308 men and 251 women; mean age, 62.4 ± 9.4 years; range, 38–88 years) were included in this study, and 391 patients were allocated to the training cohort and 168 patients to the independent validation cohort (7:3 ratio). The study population flowchart is shown in Figure 1.

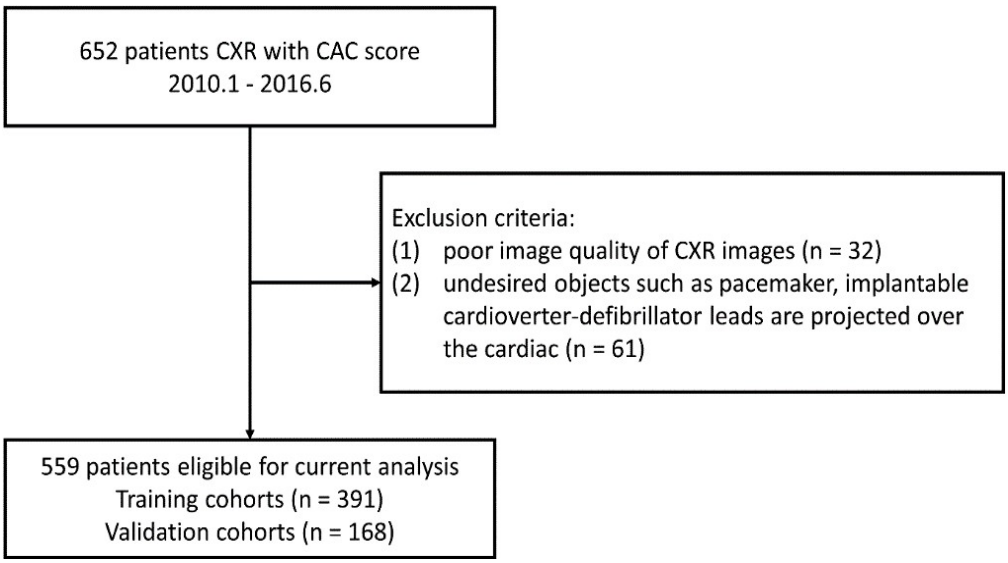

Figure 1. Consort diagram for machine learning based radiomics scoring model

B. Overall scheme

The input image was preprocessed, and radiomic features were extracted from the manually annotated cardiac contours. We performed feature selection in the training cohort through the least absolute shrinkage and selection operator (LASSO) regression, developed a radiomics score formula, and evaluated it in the validation cohort. For prediction modeling, the radiomic score-based machine learning model was trained and tested in 10-fold cross-validation. The overall flow is presented in Figure 2.

Figure 2. Flow chart of the overall scheme of this study.

C. Dataset curation and image preprocessing

Anonymized CXR images of 559 were obtained using the digital imaging and communications in medicine (DICOM) format and transferred to another standalone in-house annotation program. Two experienced cardiologists (H.B.P and R.H) manually delineated the cardiac contour over the CXR images using an in-house program. Annotated cardiac contours were corrected using the B-spline curve method. The aorta was not included in the annotated cardiac area to prevent including calcification of the aortic arch or other noncardiac structures. An example of manual contour annotation is shown in Figure 3. CXR images were normalized by referring to DICOM header information, and the annotated masks and images were resized to 512 × 512 pixels using the bilinear interpolation method.

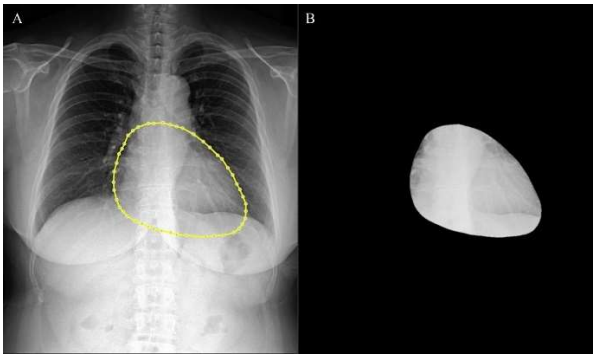

Figure 3. Example of region of interest on posteroanterior chest X -ray radiographs (CXR). (A) is an example of cardiac annotation by the expert. (B) is a cardiac segmented image.

D. CAC score assessment

The CAC score is typically assessed with a CAC scan. Figure 4 shows that calcification of main coronary arteries was observed in CXR compared to computed tomography (CT). The CAC score was also measured by two experienced cardiologists (H. B. P. and R. H.) using a commercially available calcium scoring software (Vitrea fx 6.4, Vital Images) according to the protocol described by Agatston et al.[12] The CAC score was calculated from the four main coronary arteries, including the left main (LM), left anterior descending (LAD), left circumflex (LCX), and right coronary artery (RCA), and an integrated score from all coronary lesions was used. CAC score was calculated by adding the scores of all coronary lesions, which were calculated by multiplying the area of the lesion by the density in Hounsfield units.[12] CAC score was defined with four categories: absent (0 AU), mild (1 to 99 AU), moderate (100 to 399 AU), and severe (≥400 AU).[13,14] A cutoff value of 100 AU was used to predict moderate-to-severe CAC scores. We defined a binary label for the machine learning model with CAC score as follows: zero was <100 AU (low to mild risk) and one was ≥100 AU (moderate to high risk).

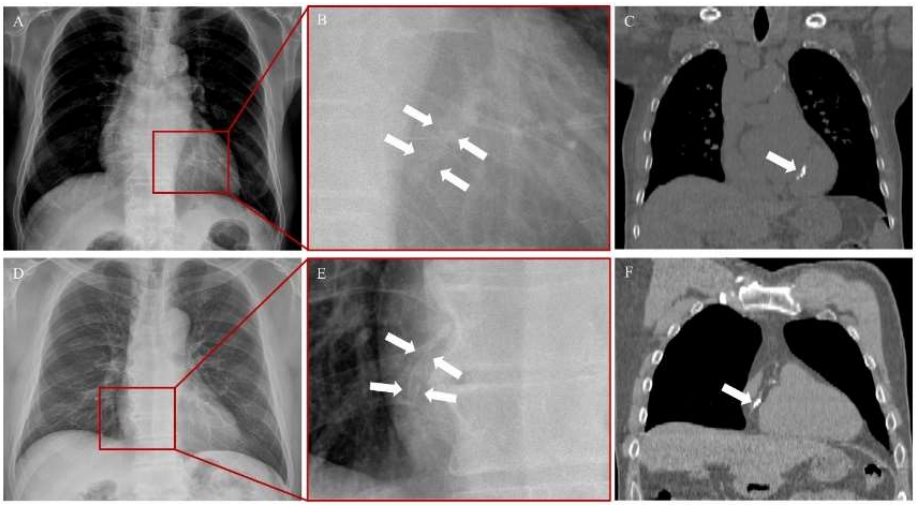

Figure 4. Representative chest X-ray radiographs (CXR) and computed tomography (CT) images of a patients with high coronary artery calcium (CAC) score.
(A, B) An 80-year-old male with an overall CAC score of 2525. The left anterior descending (LAD) calcium score observed on calcium scan was 522 (C). (D, E) A 52-year-old male with an overall CAC score of 1618. The right coronary artery (RCA) calcium score observed on calcium scan was 308 (F).

E. Radiomics feature extraction

We employed an open-source Python package, Pyradiomics,[15] to extract the radiomic features. The steps of radiomics feature extraction are as follows: (1) image segmentation with an annotated mask, (2) discretization and preprocessing of segmented images, and (3) application of first-, second-, or high-order (texture) statistics.[16] We extracted 455 radiomics features from the segmented cardiac area on CXR. The radiomic features consisted of 18 first-order statistics features and 24 gray-level co-occurrence matrix (GLCM), 16 gray-level run-length matrix (GLRLM), 16 gray-level size zone matrix (GLSZM), 14 gray-level dependence matrix (GLDM), and 5 neighboring gray tone

difference matrix (NGTDM) features. We applied four wavelet filters to the image and extracted the same set of radiomic features from the wavelet response image.

F. Radiomic score model

LASSO regression was used to select significant features in the training cohort.[17] The LASSO regression model was trained to classify moderate to severe CAC score (CAC score ≥100). A radiometric score was estimated with a linear combination of selected features weighted by their respective coefficients through LASSO. We compared two models: clinical information (CI) model and clinical information + radiomic score (CI-RS) model to evaluate the incremental value of radiomic score to the clinical information for CAC score prediction in the training cohort. For clinical information, age, gender, and BMI were used. We evaluated CI-RS model in the validation cohort to test reproducibility. The R software and "glmnet" package (R foundation for Statistical computing, Vienna, Austria, URL: http://www.R-project.org, 2016) were used for the LASSO logistics regression model analysis.

G. Radiomic score-based machine learning model

Linear regression is a statistical model, but it can be regarded as supervised machine learning because the parameters are learned by minimizing a loss of function from the given datasets. The random forest is an "ensemble learning" method consisting of aggregating a large number of decision trees. In most cases, random forest performs better than linear regression in binary classification task. Therefore, we developed a random forest based machine learning model for the prediction of CAC score ≥100. We tested two types of machine learning models: CI model and CI-RS model to evaluate the importance of radiomic score as an input of machine learning algorithm. The performance of prediction model was evaluated with 10-fold cross validation on the entire 559 patients.

4. AI-based ischemic change analysis in ECG

A. Study population and design

We analyzed the same database of Radiomics scoring-based CAC prediction model (n=559) from PARADIGM and CREDENCE datasets to generate an integrated AI gatekeeper solution combining clinical variables, CXR, and ECG. The ECG model for classifying coronary artery disease consisted of a combination of 1d CNN and Residual Unit. Because the data was partially synchronized, we used a total of 4 pairs which tied 12 leads by three as input to the model. After going through the same model for each pair, we extracted representative features through a pooling process, and the final classification was performed by combining these inputs. We used the Adam optimizer to learn the model for 50 epochs with a value of a learning rate of 0.0001, resulting in the performance of AUC 0.62 on the training set and AUC 0.56 on the test set. We extracted a feature map of size 64 × 4 from all data from the last convolutional layer of the corresponding model and combined three selected representative features with features from other modalities through principal component analysis to be used in a multimodal model.

5. Integrated AI-gatekeeper solution modeling for CAD prediction

We obtained the final probability from the clinical risk factor model, radiomics score from the CXR model, and three major features from the ECG model per each data. Using the obtained features as input, we constructed a simple single neural model to find the optimal weight between features and to learn to perform CAD prediction. At this time, the cost function used binary cross-entropy, and the model was updated using the Stochastic gradient descent optimizer.

6. Statistical analysis

All statistical analyses were performed using R software with package “ctqr” (version 4.1.0, R Foundation for Statistical Computing, Vienna, Austria). Continuous variables were compared using the t-test for variables with a normal distribution or the Mann-Whitney U test for variables with an abnormal or unknown distribution. Categorical variables were compared using the chi-square test or Fisher's exact test. The reported statistical significance levels were all two-sided, and p values less than 0.05 were considered statistically significant. R software was used to build and evaluate the prediction model. Prediction performance was evaluated using the area under the curve (AUC) values of the ROC curve in the training and validation cohorts.

Table 3. Baseline characteristics for Bayesian quantile regression model

	Patients (n=1,463)
Age, years	61.8± 9.1
Male	1095 (64.8)
Body mass index, kg/m^2	25.6 ± 3.4
Current smoker	320 (19.2)
Diabetes mellitus	404 (24.1)
Hypertension	993 (59.4)
Dyslipidemia	772 (46.3)
Laboratory data	
HDL cholesterol, mg/dL	49.6 ± 13.5
LDL cholesterol, mg/dL	112.5 ± 35.4
Triglycerides	145.5 ± 86.6
Symptoms	
Asymptomatic	370 (22.2)
Typical angina	109 (6.5)
Atypical angina	1038 (62.2)
Non-cardiac pain	133 (8.0)
Others	139 (8.3)

Values are presented as means ± SDs or n (%)

Same abbreviation was used as in Table 1

III. RESULTS

1. AI-based clinical risk factor model for CAD prediction

A. Bayesian quantile regression model using PARADIGM dataset

The baseline characteristics of the study population are presented in Table 3. The mean patient age was 62 years; 35.2% were women, 59.4% had hypertension, 46.3% had dyslipidemia, and 24.1% had diabetes mellitus. Most patients had atypical angina (62.2%) and typical anginal symptoms were observed in only 6.5% of the patients. The quantile estimates of 10%, 25%, 50%, 75%, and 90% for the three vessels and their per-patient values of the DS and DS changes are shown in Table 4. The mean measurements of the 90th percentiles were 41%–50% and 5.6%–7.3% in DS and DS change, respectively. Figures 4–7 show the error bar charts of the coefficient estimates with 95% confidence intervals for the selected risk factors for which at least one estimate was statistically significant among the five quantiles (10%, 25%, 50%, 75%, and 90%), respectively for regression Models 1 and 2. The y-axes were log-scaled for clear visibility of the error bar charts.

Table 4. The quantile estimates of 10%, 25%, 50%, 75%, and 90% for DS and DS change in the three vessels and per-patient

	Quantiles	**10%**	**25%**	**50%**	**75%**	**90%**
DS	LAD	7.39	14.03	23.68	35.40	47.57
	LCx	5.02	11.50	20.15	30.90	41.33
	RCA	6.14	13.04	22.80	32.93	44.49
	Per-patient	10.03	17.43	27.80	39.74	50.22
DS change	LAD	−2.18	−0.16	1.26	3.37	5.62
	LCx	−2.01	−0.19	1.22	3.15	5.65
	RCA	−1.57	0.14	1.64	3.84	7.02
	Per-patient	−0.77	0.67	2.37	4.49	7.31

In the per-vessel analysis of DS, the typical anginal symptom was associated with the highest quantile (90%) of DS in the LAD; diabetes was associated with higher quantiles (75% and 90%) of DS in the LCx; dyslipidemia was associated with the highest quantile (90%) of DS in the RCA, whereas other symptoms showed some association with the LCx and RCA (Figure 5). Overall, the per-patient analysis of DS, age, and hypertension was positively associated with all DS quantiles; in contrast, HDL-C was negatively associated with most DS quantiles (Figure 6).

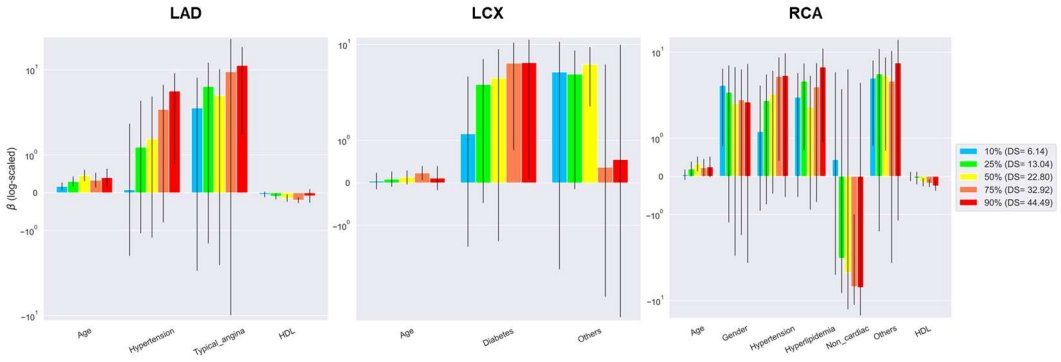

Figure 5. Bayesian quantile regression models for three vessels (LAD, LCx, and RCA) regarding DS

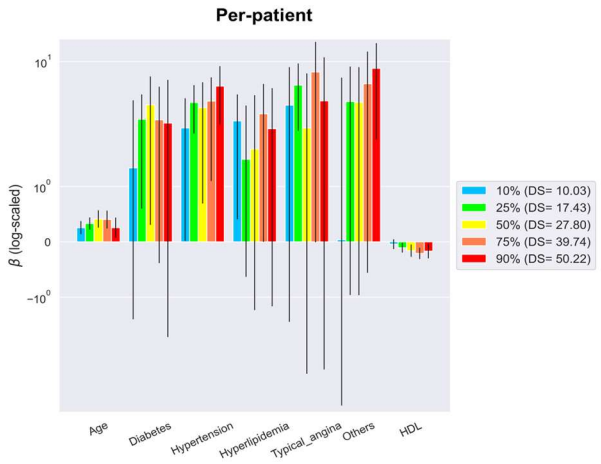

Figure 6. Bayesian quantile regression models for per-patient regarding DS

In the per-vessel analysis of DS change, HDL-C showed a clear and dynamic relationship with DS change in the LAD and RCA; hypertension also showed a dynamic relationship with DS change in the LCx and DS change severity (Figure 7). In the overall per-patient analysis of DS change, age, smoking, and hypertension showed a tendency to increase DS change, although no consistent associations were observed. However, unlike LDL-C, which showed no significant association with DS change, HDL-C showed a dynamic association with DS change which changed from positive to negative with DS severity (Figure 8).

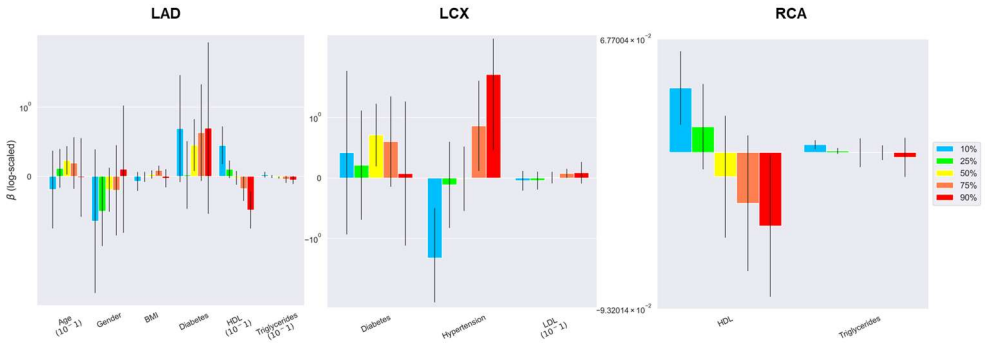

Figure 7. Bayesian quantile regression models for three vessels (LAD, LCX, and RCA) regarding DS change

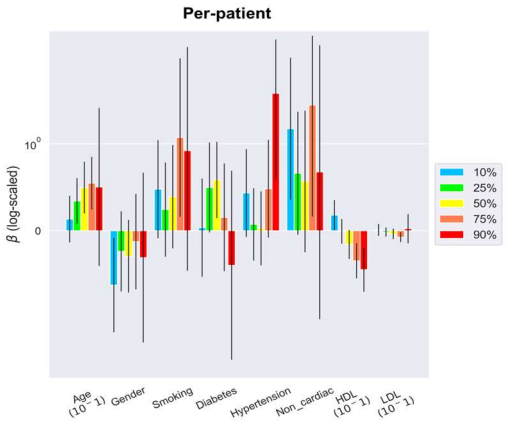

Figure 8. Bayesian quantile regression models for per-patient regarding DS change

B. Machine-learning ridge-regression model using complete dataset (CONSERVE, CREDENCE, 3V FFR-FRIENDS, and PARADIGM)

The baseline characteristics of the study population are presented in Table 1. The mean patient age was 59.6 years; 44.1% were women, 53% had hypertension, 42.5% had dyslipidemia, and 13% had diabetes mellitus. Most patients had atypical angina (35%) and typical anginal symptoms were observed in only 14.5% of the patients. Table 5 shows estimates of ridge-regression model. Complete dataset combined 4 independent studies (CONSERVE, CREDENCE, 3V FFR-FRIENDS, and PARADIGM) was used for training. The trained model were inferred to the external validation dataset (n=249) and the AUC was 0.753 (95% CI 0.69-0.81). (Figure 9).

Table 5. Estimates of ridge-regression model

Risk Factor	Criterion level	Coefficient Estimates
Sex		0.7978
BMI		-0.0188
Smoking		0.1587
Diabetes		0.2029
Hypertension		0.3015
Dyslipidemia		0.4247
Creatinine	(>1.4)	0.4734
HDL (logged)		-0.4747
LDL		-0.0009
Triglycerides (logged)		0.1695
Typical angina		0.8218
Atypical angina		-0.6226
Non cardiac		-0.1462
HbA1c level 1	(diabetes=1) & (7≤HbA1c<9)	0.3364
HbA1c level 2	(diabetes=1) & (HbA1c≥9)	0.6029
Age 0	(40,50)	0.4320
Age 1	(50,60)	0.9517
Age 2	(60,70)	1.3720
Age 3	(70,)	1.9340

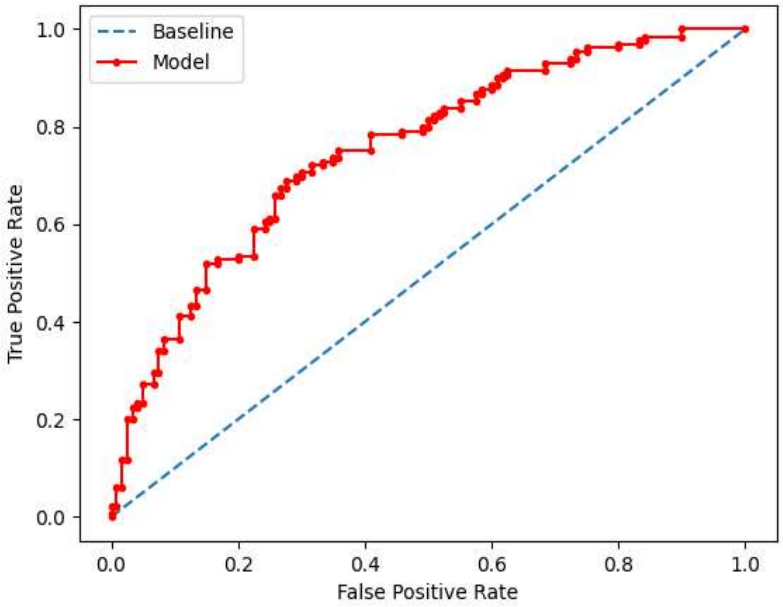

Figure 9. ROC curve of risk factor based ridge-regression model for CAD prediction

2. Radiomics scoring based ≥100 calcium score prediction model from CXR

A. Patient characteristics

The baseline characteristics of the study population are shown in Table 4. Of the 559 patients, 40.4% (n=226) had CAC scores ≥100 and 59.6% (n=333) had CAC scores <100. In addition, a more than 30-fold higher calcium score difference was observed in the CAC score ≥100 group (567.5 ± 662.5 vs. 16.4 ± 24.9, $p < 0.001$). There were no significant differences in sex and BMI between the two groups. The patients were divided into training and validation cohorts in a 7:3 ratio with all clinical characteristics being well matched, including similar calcium score differences between the CAC score ≥100 and CAC score <100 groups in each cohort (Table 6).

Table 6. Baseline characteristics for radiomics scoring based machine learning model

	CAC score < 100 (N = 333)	CAC score ≥ 100 (N = 226)	p-value
Clinical characteristics			
Age, years (±SD)	60.1 ± 9.06	65.9 ± 8.81	<0.001
Male, n (%)	176 (52.9%)	132 (58.4%)	0.227
BMI (kg/m^2)	24.8 ± 3.05	24.8 ± 3.38	0.999
Total cholesterol (ml/dL)	183 ± 37.0	173 ± 39.6	0.005
HDL (ml/dL)	49.0 ± 11.6	47.5 ± 11.5	0.150
SBP (mm Hg)	125 ± 14.4	127 ± 15.0	0.039
DBP (mm Hg)	76.7 ± 9.67	76.9 ± 9.69	0.823
Hypertension	162 (48.6%)	162 (71.7%)	<0.001
Hyperlipidemia	70 (21.0%)	74 (32.7%)	0.003
Diabetes mellitus	59 (17.7%)	82 (36.3%)	<0.001
Smoking history	48 (14.4%)	36 (15.9%)	0.710
CT Measurements			
CAC score (Agatston units)	16.4 ± 24.9	567.5 ± 662.5	<0.001

Values are mean ± SD, n (%).

CAC, coronary artery calcium; BMI, body mass index; HDL, high-density lipoprotein; SBP, systolic blood pressure; DBP, diastolic blood pressure; DS, diameter stenosis.

B. Radiomic score model

Based on the training cohort, 455 extracted radiomic features were reduced to 11 potential predictors using the LASSO regularization method (Figure 10). The optimal value of λ was obtained by performing 10-fold cross-validations to find the max AUC value and the max AUC + 1 standard error (SE) value.[18] The radiomic score was calculated using the following formula:

$$
\begin{aligned}
\text{Radiomics score} = {} & -0.4356 - 0.0204 \times \text{original\_firstorder\_InterquartileRange} \\
& - 0.1291 \times \text{original\_firstorder\_Skewness} \\
& + 0.0737 \times \text{original\_gldm\_DependenceVariance} \\
& + 0.3162 \times \text{original\_gldm\_LargeDependenceLowGrayLevelEmphasis} \\
& \times 0.2282 \times \text{original\_gldm\_SmallDependenceLowGrayLevelEmphasis} \\
& + 0.2390 \times \text{wavelet\_LH\_glcm\_Idn} \\
& - 0.0243 \times \text{wavelet\_LHglcm\_MaximumProbability} \\
& + 0.0176 \times \text{wavelet\_HL\_firstorder\_10Percentile} \\
& - 0.2769 \times \text{wavelet\_HL\_glcm\_ClusterShade} \\
& + 0.0346 \times \text{wavelet\_HH\_gldm\_DependenceVariance} \\
& - 0.1002 \times \text{wavelet\_LL\_glcm\_Imc2}
\end{aligned}
$$

The radiomics score in the training and validation cohorts was calculated using the formula and it was not significantly different between training and validation cohorts (-0.42 ± 0.68 vs. -0.47 ± 0.68, $p = 0.465$) (Table 7).

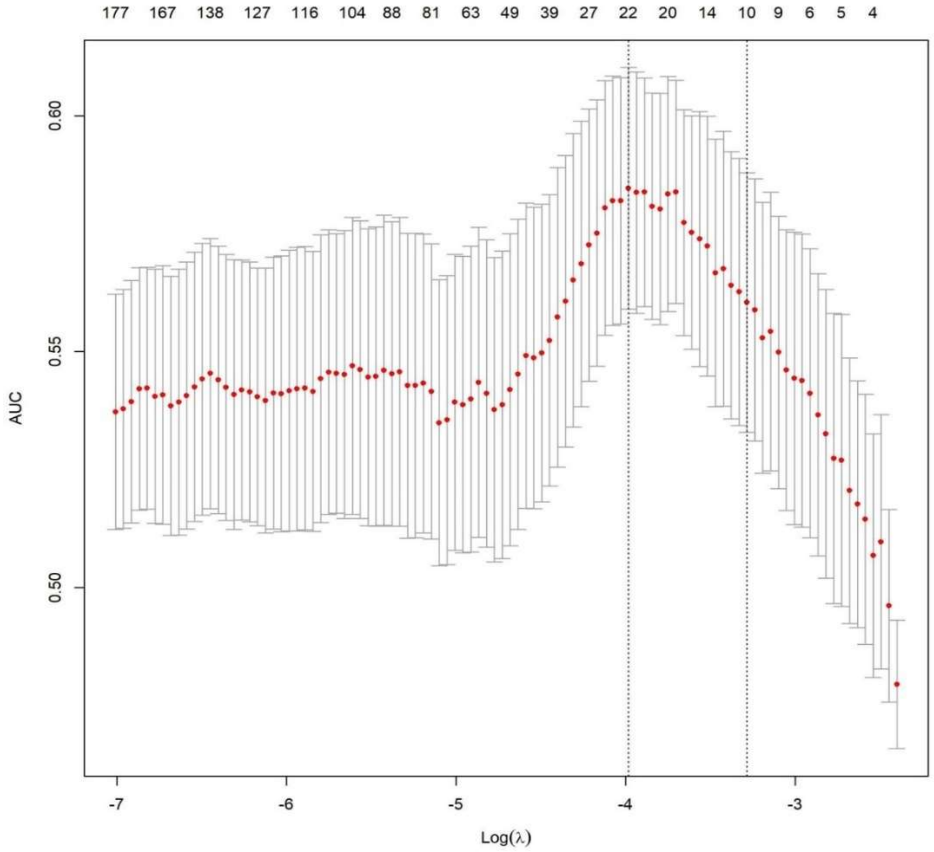

Figure 10. Radiomics feature selection using the least absolute shrinkage and selection operator (LASSO) regression model in the training cohort.

Dotted vertical lines were drawn at the optimal values using the maximum criteria and the 1 standard error (SE) of the maximum criteria (the 1-SE criteria). A λ value of 0.037 with log (λ) − 3.29 was chosen (the 1-SE criteria) according to 10-fold cross validation (CV), where optimal λ resulted in six nonzero coefficients.

Table 7. Patient characteristics in the training and validation cohorts of radiomics scoring based machine learning model

	Training cohort			p-value	Validation cohort			p-value	*P* between Training and Validation Set
	Total	CAC score <100 (N=233)	CAC score ≥100 (N=158)		Total	CAC score <100 (N=100)	CAC score ≥100 (N=68)		
Clinical characteristics									
Age, years (±SD)	62.6 ± 9.42	60.3 ± 9.03	66.1 ± 8.91	<0.001	61.9 ± 9.36	59.6 ± 9.15	65.3 ± 8.64	<0.001	0.412
Male, n (%)	210 (53.7%)	122 (52.4%)	88 (55.7%)	0.585	98 (58.3%)	54 (54.0%)	44 (64.7%)	0.222	0.360
BMI (kg/m^2)	24.9 ± 3.28	24.9 ± 3.10	24.8 ± 3.55	0.865	24.5 ± 2.94	24.5 ± 2.93	24.6 ± 2.98	0.768	0.251
Total cholesterol (ml/dL)	179 ± 38.1	183 ± 37.1	173 ± 38.8	0.012	178 ± 38.9	181 ±36.8	173 ± 41.6	0.194	0.808
HDL (ml/dL)	48.2 ± 11.6	49.1 ± 11.5	46.9 ± 11.6	0.066	48.8 ± 11.6	48.7 ± 11.9	49.0 ± 11.3	0.860	0.574
SBP (mm Hg)	126 ± 14.7	125 ± 14.5	128 ± 15.0	0.098	125 ± 14.7	124 ± 14.2	127 ± 15.3	0.215	0.434
DBP (mm Hg)	76.5 ± 9.61	76.8 ± 9.50	76.1 ± 9.78	0.528	77.3 ± 9.82	76.5 ± 10.1	78.5 ± 9.33	0.171	0.384
Hypertension	230 (58.8%)	115 (49.4%)	115 (72.8%)	<0.001	94 (56.0%)	47 (47.0%)	47 (69.1%)	0.007	0.591
Hyperlipidemia	100 (25.6%)	51 (21.9%)	49 (31.0%)	0.056	44 (26.2%)	19 (19.0%)	25 (36.8%)	0.017	0.963
Diabetes mellitus	100 (25.6%)	44 (18.9%)	56 (35.4%)	<0.001	41 (24.4%)	15 (15.0%)	26 (38.2%)	0.001	0.852
Smoking history	56 (14.3%)	33 (14.2%)	23 (14.6%)	0.999	28 (16.7%)	15 (15.0%)	13 (19.1%)	0.623	0.560
CT Measurements									
CAC score	227 ± 427	16.5 ± 25.1	537 ± 537	<0.001	268 ± 641	16.3 ± 24.6	638 ± 889	<0.001	0.449
Radiomics score	-0.42±0.68	-0.58±0.65	-0.18±0.65	<0.001	-0.47±0.68	-0.58±0.66	-0.31±0.69	0.012	0.465

C. Validation of radiomic score

Table 8 shows the incremental value of radiomic score to clinical information to predict moderate to severe CAC score. Among the used clinical information, age was the only predictive factor in CI model and CI-RS model (CI model; odds ratio [OR] = 1.09; 95% confidence interval [CI] = 1.06 – 1.12; p <0.001, CI-RS model; OR = 1.08; 95% CI = 1.05 – 1.11; p <0.001). However, the radiomics score was shown to be the most prominent factor for CAC score ≥100 prediction (CI-RS model; OR = 2.33; 95% CI = 1.62 – 3.44; p <0.001) and showed moderate correlation between radiomics score and CAC score (Spearman Correlation Coefficient 0.48, p <0.05).

Table 8. Incremental value of radiomic score to clinical information

	CI Model		CI-RS Model	
AUC	0.690 (0.636-0.744)		0.729 (0.677-0.781)	0.022
Variables	Odds radios (95% CI)	p value	Odds radios (95% CI)	p value
Age	1.085 (1.057-1.115)	<0.001	1.075 (1.046-1.106)	<0.001
Female	0.644 (0.411-1.000)	0.052	0.688 (0.433-1.086)	0.111
BMI	1.040 (0.971-1.114)	0.263	1.017 (0.948-1.093)	0.638
Radiomic score			2.330 (1.621-3.436)	<0.001

CI Model, clinical information Model; CI-RF, clinical information + radiomic score; AUC, area under the curve; BMI, body mass index; CI, confidence interval.

In the training cohorts, CI-RS model showed significantly higher AUC than CI model (CI model vs. CI-RS model: 0.69 vs. 0.73; p = 0.022). We further evaluated the performance of CI-RS model in the validation cohort and the performance of CI-RS was not significantly different between training and validation cohorts (training vs. validation: AUC 0.73, 95% CI 0.68 – 0.78 vs. AUC 0.72 95% CI 0.64 – 0.80). The model performance evaluation is

summarized in Table 9. Figure 11 shows the performance comparison of CI model and CI-RS model in the training and validation cohorts.

Table 9. Validation of radiomic score

	Training cohort	Validation cohort
AUC	0.729 (0.677-0.781)	0.717 (0.636-0.798)
Sensitivity	0.614	0.735
Specificity	0.751	0.720
NPV	0.742	0.800
PPV	0.626	0.641

AUC, area under the curve; NPV, negative predictive value; PPV, positive predictive value.

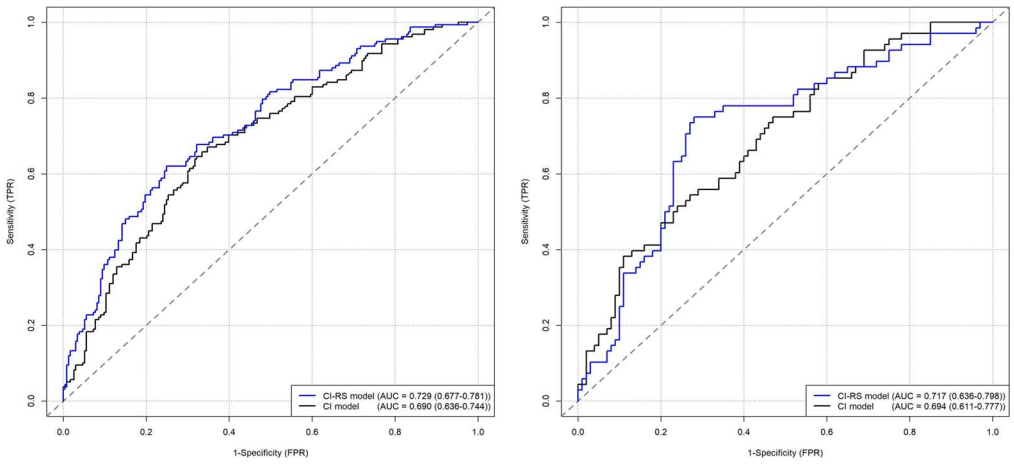

Figure 11. The receiver operating characteristic (ROC) curves of the CI model and CI-RS model derived from training (A) and validation cohorts (B)

D. Validation of radiomic score-based machine learning model

A machine learning based CI-RS model showed superior performance than CI model for the prediction of CAC score ≥100 (CI model vs. CI-RS model: AUC 0.701 vs. AUC 0.831)

in 10-fold cross validation (Figure 12). The sensitivity, specificity, negative predictive value, and positive predictive value of CI-RS model were superior to CI model at the best diagnostic decision point. Table 10 showed the 10-fold cross validation performance of CI model and CI-RS model.

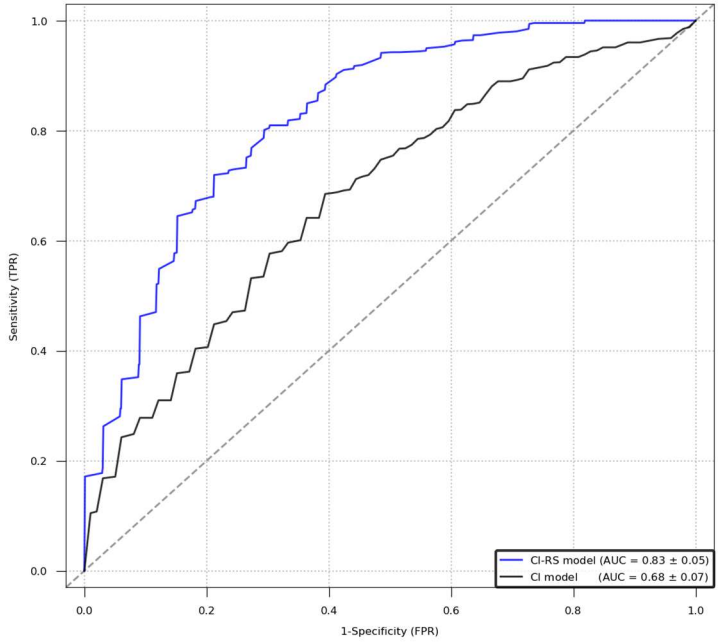

Figure 12. The receiver operating characteristic (ROC) curves of the CI model and CI-RS model derived from training (A) and validation cohorts (B)

Table 10. Performance of coronary artery calcium score prediction model

	CI Model	CI-RS Model
AUC	0.701 (0.63-0.73)	0.831 (0.79-0.87)
Sensitivity	0.761	0.806
Specificity	0.616	0.773
NPV	0.801	0.860
PPV	0.579	0.716

3. AI-based ischemic change analysis in ECG

The CNN-based ischemia detection ECG model showed a modest result of AUC 0.60 (95% CI 0.54-0.66) in the training cohort and AUC 0.60 (95% CI 0.50-0.69) in the validation cohort. When the clinical variables were added to this model, the model performance significantly improved up to an AUC of 0.71 (95% CI 0.66-0.76) in the training cohort as well as validation cohort (AUC 0.67, 95% CI 0.59-0.69).

4. Integrated AI-gatekeeper solution modeling for CAD prediction

In the training cohort, the final AUC of the random forest of an integrated model for CAD prediction based on clinical risk factors, CXR, and ECG was 0.82 (95% CI 0.72-0.91). However, relatively lower performance was shown in the validation cohort as 0.67 (95% CI 0.54-0.80). The performance of the clinical risk factors and CXR combined model showed a random forest AUC of 0.81 (95% CI 0.70-0.90), indicating that there was no additive benefit of ECG in the final integrated model. When we differentiated the prediction endpoint from 50% DS to 75% DS in the same 559 patients with ten-fold-cross validation, the integrated model showed an AUC of 0.74, relatively lower performance than 50% DS.

Lastly, in the external validation cohort (n=249), the final AUC of integrated AI-gatekeeper solution for CAD prediction was AUC 0.77 (Sensitivity 73%, Specificity 69%).

IV. DISCUSSION

In the present study, we developed the various-type feature-based AI-gatekeeper solution for CAD prediction. First, we explored the comprehensive relationship between CV risk factors and baseline-graded subclinical to clinical coronary artery stenosis and its progression to identify significant risk factors for CAD screening using a serially followed-

up database (PARADIGM; n=1,463). Based on this data, we then developed the general risk factor-based CAD prediction model using the complete database (CONSERVE, CREDENCE, 3V FFR-FRIENDS, and PARADIGM; n=5,643) and proved the reliability by external validation. This machine-learning based general risk factor model can be utilized as a web-based risk factor calculator. Second, we developed an integrated framework that combined machine learning and radiomic scores to predict moderate-to-severe CAC scores (≥100) using the database with CAC score information (PARADIGM and CREDENCE; n=559) and proved the reliability of this model by external validation. Third, we used the CNN method to perform AI-based ischemic change analysis in resting 12 leads ECGs. Although we could not show meaningful results, we identified the possibility of its additive benefit when combined with the clinical risk factor model. Finally, we configured the integrated AI model for CAD prediction based on clinical risk factors, CXR, and ECG and confirmed excellent performance of AUC 0.82.

1. AI-based clinical risk factor model for CAD prediction

We introduced the clinical utility of the Bayesian truncated quantile regression machine learning method to evaluate the comprehensive relationship between CV risk factors and different stages of CAD and its progression. HDL-C showed a consistent negative association with most DS levels and a dynamic relationship from positive to negative along with DS change severity, indicating that high HDL-C has a preventive effect on baseline CAD as well as CAD progression. Typical angina symptoms were only associated with a high quantile of stenosis in the LAD and not in the LCx or RCA. Likewise, diabetes was strongly associated with LCx, and dyslipidemia was associated with RCA. Other symptoms such as dyspnea, syncope, palpitation, dizziness, and epigastric pain showed a relationship

with stenosis in the LCx and RCA. Although it has been known by clinical experience that LAD lesions are associated with typical anginal symptoms owing to their considerable accountability in the entire coronary perfusion, no scientific evidence has been provided.[19,20] In this study, we suggested the use of the BQR model. Similarly, it is known by experience that LCx or RCA lesions are more likely to be associated with vague symptoms than LAD lesions.[19,21] This could also be demonstrated using the BQR model. In addition, HDL-C showed a dynamic interrelationship with graded coronary stenosis and stenosis progression, which was the most distinctive utility of the BQR model that could not be achieved in any other standard regression models.

Since Koenker and Bassett first introduced quantile regression models, they have been used in various research areas, such as investment, economics, and engineering, due to their multiple advantages over standard regression analysis.[22] Quantile regression has recently been regarded as an efficient analysis tool for income and wage studies in labor economics. The Bayesian Tobit quantile regression, an advanced version of the plain quantile regression model, has been utilized to estimate outage costs in the engineering field.[23-25] Although Wehby et al. firstly introduced the utility of the BQR model in the medical field by presenting the different risk factors for low and high birth weight,[26] it is not widely adopted probably because its interpretation seems somewhat unintuitive since the concept of quantile is less familiar than means.[27] However, with the increased interest in machine learning methods in medical research, quantile regression has recently attracted attention as a valuable data analysis tool in the medical research area.[28] Although clinical models for estimating the pretest probability of CAD based on age, sex, and symptom typicality in patients with stable angina have been developed,[29,30] recent studies raised the overestimation issue of these models, potentially due to the exclusion of other important

CV risk factors such as diabetes, dyslipidemia, hypertension, smoking, and obesity.[31,32] Novel imaging markers, including calcium score and multiple risk factor assessment using the machine learning method, have been evaluated to overcome this issue. However, most studies have shown modest performance for predicting obstructive CAD and are limited to a single outcome variable of 50% DS.

This is the first to apply BQR analysis to the CV area, particularly in predicting CAD, and comprehensively explore the association between CV risk factors with symptom characteristics and different stages of CAD and its progression. We believe that this pilot study can provide a framework for the cost-efficient utilization of previously overlooked clinical information, thereby facilitating the development of a more accurate CAD pretest probability model.

We then applied the machine-learning ridge-regression method to the complete database combining CONSERVE, CREDENCE, 3V FFR-FRIENDS, and PARADIGM studies, the most extensive database for machine-learning based risk factor model for CAD screening. We demonstrated excellent performance for CAD prediction with an AUC of 0.75 with external validation. This risk factor-based calculator could be utilized in primary physician care or even self-checkup.

2. Radiomics scoring based ≥100 calcium score prediction model from CXR

We also developed an integrated framework that combined machine learning and radiomic scores to identify the CAC score in CXR and experimentally validated the radiomics score as a predictive factor in CAC score prediction. We demonstrated the feasibility of a moderate-to-severe CAC score (≥100) prediction model using the proposed integrated framework. Radiomic features were extracted from the cardiac area, and a radiomic score

for predicting a CAC score of ≥100 was calculated. The model developed in the training cohort was subjected to a validation test. In the training cohort, our clinical variable-adjusted model (radiomics score with basic clinical variables) demonstrated a good prediction performance with an AUC of 0.73. The present model also showed consistent performance, with an AUC of 0.72 in the validation cohort.

Non-enhanced calcium scanning is a robust non-invasive tool for CAD screening as well as cardiovascular disease (CVD) risk assessment using a low radiation dose of 1mSv.[33] Numerous studies have demonstrated that CAC scores have superior risk stratification performance than traditional risk factor-based methods, showing higher net reclassification benefits, particularly in asymptomatic population [34-36]. Moderate CAC scores of 100 to 400 corresponded to 12.9 to 16.4% of 10-year event rate, and CAC scores >400 to >1000 corresponded to 22.5 to 28.6% and 37% of 10-year event rate, respectively.[33] Hence, a CAC score of ≥100 should require medical attention for aggressive statin therapy along with active CAD surveillance. In patients with stable chest pain, the CAC score also exhibited a clinical benefit in improving the pre-test probability for CAD.[37-39] Haberl et al. demonstrated the high diagnostic accuracy of a CAC score cutoff value of 100 for the detection of obstructive CAD with a sensitivity of 95% and a specificity of 79%.[37] Furthermore, the CAC score proved to have a better detection rate of obstructive CAD compared with conventional stress tests such as exercise treadmill electrocardiography or technetium-stress single photon emission computed tomography.[40-42] Although CAC scans can be performed using a low radiation dose, radiation hazards still exist, especially as they are cumulated by an ionizing radiation dose. Furthermore, recent technical advancements in coronary computed tomography angiography (CCTA), it can also be performed with 1 to 2mSv,[43] similar to a low-dose CAC scan, and recent studies have shown that statins

stabilize atherosclerotic plaques while calcium progression presents with an increased CAC score.[44] Hence, the clinical utility of CAC scan is currently relatively diminishing, particularly in patients with stable chest pain, as CCTA has become the initial diagnostic modality of choice regardless of CVD risks.[45,46] Furthermore, in asymptomatic patients, more than 11,000 population-based the Korea Initiatives on Coronary Artery Calcification (KOICA) registries have revealed that 55% of the study population had a CAC score of zero and 86% had a CAC score of less than 100.[47] In terms of CAD screening, it can be pointed out that majority of population had an unnecessary CAC scan. Therefore, the new imaging modality replacing CAC scan, to detect moderate to severe calcification corresponding to CAC score ≥100 using a lower radiation dose, would be utilized as the gatekeeper of CCTA and to improve the pre-test probability of CAD.

On the other hand, CXR is a rapid and cost-effective study widely performed from primary care clinics to tertiary hospitals with 0.14 mSv of a radiation dose,[48] but it is more specialized in detecting lung problems and thus has limited clinical utility for the diagnosis of CAD.[49-51] However, in the present study, we demonstrated that the combination of machine learning and radiomic scores can revitalize the use of CXR for CAD screening by predicting moderate-to-severe coronary calcification. This novel method can be easily applied to the asymptomatic population for CAD screening as well as to patients with stable chest pain as an effective gatekeeper for CCTA, and might be substituted for a CAC scan to reduce unnecessary radiation exposure.

Recently emerged radiomics is a computer-aided technique for extracting a large number of subtle or visually unidentifiable features, such as intensities, textures, or wavelets, as quantified values from a digitalized medical image.[52-54] In previous studies, radiomics features extracted from CT, ultrasound, positron emission tomography, and magnetic

resonance have shown good performance in differentiating tissue characteristics or for the prediction of disease progression, particularly in pre-cancerous lesions,[52] and are actively being applied to simple chest radiography.[55] The major clinical hurdle of deep neural networks is the ambiguity of the black box procedures; however, machine learning with radiomic features is more explainable and clinically applicable.[56] Therefore, the combination of machine learning and radiomics can show a synergistic effect in revealing generally overlooked features. Recent studies have also demonstrated the clinical utility of radiomics from CXR combined with machine learning algorithms for identifying patients with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2).[57-59] Kamel et al. first demonstrated the feasibility of convolutional neural networks (CNNs) for the prediction of CAC scores from CXR. They demonstrated the presence of coronary calcification predicting zero or non-zero Agatston scores on posterior anterior CXR, showing an AUC of 0.73.[60] In contrast, the present study introduced a novel machine learning-based radiomics feature analysis for moderate-to-severe CAC score prediction. To the best of our knowledge, this is the first study to use machine learning-based radiomics analysis applied to CXR for the detection of coronary calcification. The critical limitation of previous CNNs is that the algorithm takes an entire image as an input to the network, not focusing on the cardiac contour, and easily mistakenly detects undesired objects, such as cardiac devices, aortic calcifications, or other thoracic bones, as the coronary calcification.[60] In addition, because of the ambiguity of the convolutional feature extraction operation, it was impossible to discern which CXR features were relevant for predicting the presence of a CAC score. Furthermore, it is well known fact that CAC score has clear differential distribution patterns according to age, gender, and BMI. Adjusting these basic clinical variables is essential for generating a CAC score prediction model.[33] However, previous

studies by Kamel et al. did not reflect the impact of this basic information. In contrast, the present model targeted radiomic features from cardiac contouring in the CXR image and suggested the most relevant radiomics features to predict moderate to severe CAC scores.[60] Moreover, we adjusted the prediction model with basic clinical variables, but age was the only significant factor for moderate-to-severe CAC score prediction.

Although our prediction model showed moderate accuracy, the performance could be improved when the volume of the training dataset increased and traditional CVD risk factors were adjusted. This novel approach for predicting a CAC score ≥100 using a CXR image might save substantial healthcare costs and reduce radiation dose by reducing unnecessary CAC scans and playing a role as the gatekeeper of CCTA by increasing the pre-test probability of CAD. This method can also be easily applied as a routine practice to patients who visit primary health care facilities, irrespective of whether they are symptomatic or asymptomatic.

3. AI-based ischemic change analysis in ECG

We did not provide the CNN-based ischemia detection model in resting 12 leads-ECG. It is well known that resting ECG presents no ischemic change such as ST elevation/depression or T-wave inversion until significant hemodynamic limitation; in other words, ischemia happens. That is why various stress tests were developed for detecting myocardial ischemia. Although we hypothesized that the deep learning image analysis method per se could identify that the human eye cannot define ischemic change of resting ECG, no significant results were not achieved. However, we found that the CNN ECG analysis model could provide additional benefits when combined with other

prediction models or integrated models. Further investigation and advancement might be needed to improve the CNN-based ischemia detection algorithm in ECG.

4. Integrated AI-gatekeeper solution modeling for CAD prediction

We finally developed the various-type feature-based integrated AI gatekeeper solution based on clinical risk factors, CXR, and ECG for CAD screening and demonstrated the excellent performance of AUC 0.82. To the best of our knowledge, this is the first integrated model for CAD prediction using so-called “primary tests”. This novel method may be widely applicable to clinical practice and improve the pre-test probability of coronary artery disease, particularly in a primary physician care setting.

5. Study limitations

This study has several limitations. In the BQR analysis, we only included 1,463 patients with complete clinical information. Most had LAD lesions, and the LCx and RCA lesions were only on 465 and 340 vessels, respectively. Thus, there was insufficient data for evaluating the LCx or RCA. Although we included significant CV risk factors for CAD, further specified and various CV risk factors should be included to enhance the performance of this model. However, in the ridge-regression analysis model using extensive complete datasets, we could get excellent performance for CAD prediction even based on clinical risk factors. Due to the lack of available CAC score data, there were insufficient datasets utilized in the training or validation process for developing Radiomics scoring-based CAC prediction or the CNN-based ECG analysis model. Among 5,643 patients’ datasets, only 559 patients’ data were used to create these models, and this was the leading cause of low prediction performance, particularly in the ECG model. In

Radiomics scoring-based CAC prediction model, we manually delineated the cardiac contour to extract radiomics features; thus, an automatic segmentation algorithm is required to build a fully automated system to report the predicted CAC score. The development of a CNNs based segmentation algorithm for this process is required for future applications. An external validation was not performed; the study population in the present study was from two well-known multicenter CCTA trials including patients with stable chest pain; thus, patients from a single center would not meet the criteria for the qualified external validation cohort. Therefore, the study cohort was divided into training and validation cohorts, and the baseline characteristics were meticulously matched between cohorts to substitute for the external validation cohort. In future work, it will be necessary to establish a qualified external validation cohort for the external validation process.

V. CONCLUSION

In conclusion, we developed the integrated various-type feature-based AI-gatekeeper solution for CAD prediction using “primary tests” composed of CV risk factors, CXR, and ECG. This novel and innovative AI solution could turn common clinical data from primary tests into vital information by revealing previously unrecognized patterns, enhancing the diagnostic accuracy of CAD, and thereby reducing unnecessary downstream tests. This novel method might be readily applicable in clinical practice and contribute to reducing medical costs and radiation hazards, playing a role as the gatekeeper for CCTA and improving the pre-test probability of CAD, particularly in patients with stable chest pain.

REFERENCES

1. Roth GA, Abate D, Abate KH, Abay SM, Abbafati C, Abbasi N, et al. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: a systematic analysis for the Global Burden of Disease Study 2017. The Lancet 2018;392:1736-88.
2. Lloyd-Jones D, Adams RJ, Brown TM, Carnethon M, Dai S, De Simone G, et al. Heart disease and stroke statistics—2010 update: a report from the American Heart Association. Circulation 2010;121:e46-e215.
3. Min JK, Sharma A, Nicolo D. Economic Considerations for Coronary CT Angiography. Current Cardiovascular Imaging Reports 2010;3:390-5.
4. Patel MR, Peterson ED, Dai D, Brennan JM, Redberg RF, Anderson HV, et al. Low diagnostic yield of elective coronary angiography. New England Journal of Medicine 2010;362:886-95.
5. Patel MR, Dai D, Hernandez AF, Douglas PS, Messenger J, Garratt KN, et al. Prevalence and predictors of nonobstructive coronary artery disease identified with coronary angiography in contemporary clinical practice. American heart journal 2014;167:846-52. e2.
6. Arbab-Zadeh A, Hoe J. Quantification of coronary arterial stenoses by multidetector CT angiography in comparison with conventional angiography: methods, caveats, and implications. JACC: Cardiovascular Imaging 2011;4:191-202.
7. Sabarudin A, Sun Z. Coronary CT angiography: diagnostic value and clinical challenges. World journal of cardiology 2013;5:473.
8. Chang H-J, Lin FY, Gebow D, An HY, Andreini D, Bathina R, et al. Selective referral using CCTA versus direct referral for individuals referred to invasive coronary angiography for suspected CAD: a randomized, controlled, open-label trial. JACC: Cardiovascular Imaging 2019;12:1303-12.

9. Rizvi A, Knaapen P, Leipsic J, Shaw LJ, Andreini D, Pontone G, et al. Rationale and Design of the CREDENCE Trial: computed TomogRaphic evaluation of atherosclerotic DEtermiNants of myocardial IsChEmia. BMC cardiovascular disorders 2016;16:1-10.
10. Lee JM, Koo BK, Shin ES, Nam CW, Doh JH, Hu X, et al. Clinical outcomes of deferred lesions with angiographically insignificant stenosis but low fractional flow reserve. Journal of the American Heart Association 2017;6:e006071.
11. Lee S-E, Chang H-J, Rizvi A, Hadamitzky M, Kim Y-J, Conte E, et al. Rationale and design of the Progression of AtheRosclerotic PlAque DetermIned by Computed TomoGraphic Angiography IMaging (PARADIGM) registry: a comprehensive exploration of plaque progression and its impact on clinical outcomes from a multicenter serial coronary computed tomographic angiography study. American heart journal 2016;182:72-9.
12. Agatston AS, Janowitz WR, Hildner FJ, Zusmer NR, Viamonte M, Detrano R. Quantification of coronary artery calcium using ultrafast computed tomography. Journal of the American College of Cardiology 1990;15:827-32.
13. Blair KJ, Allison MA, Morgan C, Wassel CL, Rifkin DE, Wright CM, et al. Comparison of ordinal versus Agatston coronary calcification scoring for cardiovascular disease mortality in community-living individuals. The international journal of cardiovascular imaging 2014;30:813-8.
14. Azour L, Kadoch MA, Ward TJ, Eber CD, Jacobi AH. Estimation of cardiovascular risk on routine chest CT: Ordinal coronary artery calcium scoring as an accurate predictor of Agatston score ranges. Journal of cardiovascular computed tomography 2017;11:8-15.
15. Van Griethuysen JJ, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational radiomics system to decode the radiographic phenotype. Cancer research 2017;77:e104-e7.

16. Oikonomou EK, Siddique M, Antoniades C. Artificial intelligence in medical imaging: a radiomic guide to precision phenotyping of cardiovascular disease. Cardiovascular Research 2020;116:2040-54.
17. Sauerbrei W, Royston P, Binder H. Selection of important variables and determination of functional form for continuous predictors in multivariable model building. Statistics in medicine 2007;26:5512-28.
18. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. Journal of statistical software 2010;33:1.
19. Reeves TJ, Oberman A, Jones WB, Sheffield LT. Natural history of angina pectoris. The American journal of cardiology 1974;33:423-30.
20. Kumpuris AG, Quinones MA, Kanon D, Miller RR. Isolated stenosis of left anterior descending or right coronary artery: relation between site of stenosis and ventricular dysfunction and therapeutic implications. The American Journal of Cardiology 1980;46:13-20.
21. Lim HF, Dreifus LS, Kasparian H, Najmi M, Balis G. Chest pain, coronary artery disease and coronary cine-arteriography. Chest 1970;57:41-6.
22. Koenker R, Bassett Jr G. Regression quantiles. Econometrica: journal of the Econometric Society 1978:33-50.
23. Buchinsky M. Quantile regression, Box-Cox transformation model, and the US wage structure, 1963–1987. Journal of econometrics 1995;65:109-54.
24. Yu K, Lu Z, Stander J. Quantile regression: applications and current research areas. Journal of the Royal Statistical Society: Series D (The Statistician) 2003;52:331-50.
25. Kim MS, Lee BS, Lee HS, Lee SH, Lee J, Kim W. Robust estimation of outage costs in South Korea using a machine learning technique: Bayesian Tobit quantile regression. Applied Energy 2020;278:115702.
26. Wehby GL, Murray JC, Castilla EE, Lopez-Camelo JS, Ohsfeldt RL. Prenatal care effectiveness and utilization in Brazil. Health policy and planning 2009;24:175-88.

27. Beyerlein A. Quantile regression—opportunities and challenges from a user's perspective. American journal of epidemiology 2014;180:330-1.
28. Cleophas TJ, Zwinderman AH. Quantile Regression in Clinical Research: Springer Cham; 2022.
29. Diamond GA, Forrester JS. Analysis of probability as an aid in the clinical diagnosis of coronary-artery disease. New England Journal of Medicine 1979;300:1350-8.
30. Genders TS, Steyerberg EW, Alkadhi H, Leschka S, Desbiolles L, Nieman K, et al. A clinical prediction rule for the diagnosis of coronary artery disease: validation, updating, and extension. European heart journal 2011;32:1316-30.
31. Rovai D, Neglia D, Lorenzoni V, Caselli C, Knuuti J, Underwood SR, et al. Limitations of chest pain categorization models to predict coronary artery disease. The American Journal of Cardiology 2015;116:504-7.
32. Genders TS, Coles A, Hoffmann U, Patel MR, Mark DB, Lee KL, et al. The external validity of prediction models for the diagnosis of obstructive coronary artery disease in patients with stable chest pain: insights from the PROMISE trial. JACC: Cardiovascular Imaging 2018;11:437-46.
33. Hecht HS. Coronary artery calcium scanning: past, present, and future. JACC: Cardiovascular Imaging 2015;8:579-96.
34. Greenland P, LaBree L, Azen SP, Doherty TM, Detrano RC. Coronary artery calcium score combined with Framingham score for risk prediction in asymptomatic individuals. Jama 2004;291:210-5.
35. Budoff MJ, Shaw LJ, Liu ST, Weinstein SR, Tseng PH, Flores FR, et al. Long-term prognosis associated with coronary calcification: observations from a registry of 25,253 patients. Journal of the American College of Cardiology 2007;49:1860-70.
36. Budoff MJ, Nasir K, McClelland RL, Detrano R, Wong N, Blumenthal RS, et al. Coronary calcium predicts events better with absolute calcium scores than age-sex-

race/ethnicity percentiles: MESA (Multi-Ethnic Study of Atherosclerosis). Journal of the American College of Cardiology 2009;53:345-52.

37. Haberl R, Becker A, Leber A, Knez A, Becker C, Lang C, et al. Correlation of coronary calcification and angiographically documented stenoses in patients with suspected coronary artery disease: results of 1,764 patients. Journal of the American College of Cardiology 2001;37:451-7.
38. Budoff MJ, Diamond GA, Raggi P, Arad Y, Guerci AD, Callister TQ, et al. Continuous probabilistic prediction of angiographically significant coronary artery disease using electron beam tomography. Circulation 2002;105:1791-6.
39. Winther S, Schmidt SE, Mayrhofer T, Bøtker HE, Hoffmann U, Douglas PS, et al. Incorporating coronary calcification into pre-test assessment of the likelihood of coronary artery disease. Journal of the American College of Cardiology 2020;76:2421-32.
40. Schmermund A, Denktas AE, Rumberger JA, Christian TF, Sheedy PF, Bailey KR, et al. Independent and incremental value of coronary artery calcium for predicting the extent of angiographic coronary artery disease: comparison with cardiac risk factors and radionuclide perfusion imaging. Journal of the American College of Cardiology 1999;34:777-86.
41. Shavelle DM, Budoff MJ, LaMont DH, Shavelle RM, Kennedy JM, Brundage BH. Exercise testing and electron beam computed tomography in the evaluation of coronary artery disease. Journal of the American College of Cardiology 2000;36:32-8.
42. Dedic A, Rossi A, Ten Kate G, Neefjes L, Galema T, Moelker A, et al. First-line evaluation of coronary artery disease with coronary calcium scanning or exercise electrocardiography. International journal of cardiology 2013;163:190-5.
43. Al-Mallah MH, Aljizeeri A, Alharthi M, Alsaileek A. Routine low-radiation-dose coronary computed tomography angiography. European Heart Journal Supplements 2014;16:B12-B6.

44. Lee S-E, Sung JM, Andreini D, Budoff MJ, Cademartiri F, Chinnaiyan K, et al. Differential association between the progression of coronary artery calcium score and coronary plaque volume progression according to statins: the Progression of AtheRosclerotic PlAque DetermIned by Computed TomoGraphic Angiography Imaging (PARADIGM) study. European Heart Journal-Cardiovascular Imaging 2019;20:1307-14.

45. Doris M, Newby DE. Coronary CT angiography as a diagnostic and prognostic tool: perspectives from the SCOT-HEART trial. Current cardiology reports 2016;18:1-8.

46. Nazir MS, Nicol E. The SCOT-HEART trial: cardiac CT to guide patient management and improve outcomes. Cardiovascular Research 2019;115:e88-e90.

47. Lee W, Yoon YE, Cho S-Y, Hwang I-C, Kim S-H, Lee H, et al. Sex differences in coronary artery calcium progression: The Korea Initiatives on Coronary Artery Calcification (KOICA) registry. PloS one 2021;16:e0248884.

48. Lahham A, Issa A, ALMasri H. Patient radiation dose from chest X-ray examinations in the west bank—Palestine. Radiation protection dosimetry 2018;178:298-303.

49. Hubbell FA, Greenfield S, Tyler JL, Chetty K, Wyle FA. The impact of routine admission chest x-ray films on patient care. New England Journal of Medicine 1985;312:209-13.

50. Davies HD, WANG EE-L, Manson D, Babyn P, Shuckett B. Reliability of the chest radiograph in the diagnosis of lower respiratory infections in young children. The Pediatric infectious disease journal 1996;15:600-4.

51. Hopstaken R, Witbraad T, Van Engelshoven J, Dinant G. Inter-observer variation in the interpretation of chest radiographs for pneumonia in community-acquired lower respiratory tract infections. Clinical radiology 2004;59:743-52.

52. Lambin P, Leijenaar RT, Deist TM, Peerlings J, De Jong EE, Van Timmeren J, et al. Radiomics: the bridge between medical imaging and personalized medicine. Nature reviews Clinical oncology 2017;14:749-62.
53. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. Radiology 2016;278:563.
54. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, Van Stiphout RG, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. European journal of cancer 2012;48:441-6.
55. Au-Yong I, Higashi Y, Giannotti E, Fogarty A, Morling JR, Grainge M, et al. Chest radiograph scoring alone or combined with other risk scores for predicting outcomes in COVID-19. Radiology 2022;302:460.
56. Vial A, Stirling D, Field M, Ros M, Ritz C, Carolan M, et al. The role of deep learning and radiomic feature extraction in cancer-specific predictive modelling: a review. Transl Cancer Res 2018;7:803-16.
57. Sakib S, Tazrin T, Fouda MM, Fadlullah ZM, Guizani M. DL-CRC: deep learning-based chest radiograph classification for COVID-19 detection: a novel approach. Ieee Access 2020;8:171575-89.
58. Tamal M, Alshammari M, Alabdullah M, Hourani R, Alola HA, Hegazi TM. An integrated framework with machine learning and radiomics for accurate and rapid early diagnosis of COVID-19 from Chest X-ray. Expert systems with applications 2021;180:115152.
59. Fusco R, Grassi R, Granata V, Setola SV, Grassi F, Cozzi D, et al. Artificial intelligence and COVID-19 using chest CT scan and chest X-ray images: machine learning and deep learning approaches for diagnosis and treatment. Journal of Personalized Medicine 2021;11:993.
60. Kamel PI, Yi PH, Sair HI, Lin CT. Prediction of coronary artery calcium and cardiovascular risk on chest radiographs using deep learning. Radiology: Cardiothoracic Imaging 2021;3:e200486.

ABSTRACT (IN KOREAN)

관상동맥질환 선별을 위한 이종 특징 기반 'AI Gatekeeper' 솔루션 개발

< 지도교수 장 혁 재 >

연세대학교 대학원 의학과

박 형 복

서론: 현 임상현장에서는 관상동맥 질환 선별을 위해 일차 검사로 심혈관 위험요소 문진, 흉부방사선 촬영, 심전도가 일반적으로 시행되며 이후 추가적으로 고가의 정밀검사를 시행하고 있다. 그러나 추가 검사들의 경우 일차 진료 현장에서는 즉시 시행하기 어렵고 비용 효과성이 떨어지는 문제점이 있다. 이에 본 연구는 인공지능 기반의 이종 특징 기반 융합 관상동맥 선별 솔루션을 개발하여 일차 진단 검사들의 정확도를 높이고자 한다.

방법: 본 연구는 CONSERVE (NCT01810198), CREDENCE (NCT02173275), 3V FFR-FRIENDS (NCT01621438), PARADIGM (NCT02803411)의 4가지 임상 연구 데이터를 분석하였다. 총 5,643명을 기계학습 기반 능형 회귀분석 기법을 사용하여 위험인자 기반의 관상동맥질환 예측 모델을 계발하였고 베이지안 분위회귀분석을 사용하여 연속적인 PARADIGM 데이터의 분석을 통해 심혈관 위험인자와 관상동맥 협착 및 진행과의 포괄적인 상관성 분석을 시도하였다. 또한 관상동맥 석회화 지수 정보를 가진 559명의 환자들을 대상으로 라디오믹스 스코어 기반의 중등도 이상의 석회화 지수 (≥100) 예측 모델 및 딥러닝 기반의 ECG에서 관상동맥질환 예측 모델을 계발하였다.

결과: (1) 능형 회귀모델에서 위험인자 기반 관상동맥질환 예측은 AUC 0.75 (95% CI 0.69-0.81)의 양호한 성능을 보여 주었고 베이지안 분위

회귀모델에서 세 혈관 DS의 90% 백분위는 41%–50% 이었고 DS 변화의 90% 백분위는 5.6%–7.3%이었다. 전형적인 협심증 증상은 LAD의 90% 백분위와 연관성을 보였고 당뇨는 LCx DS의 75% 혹은 90% 백분위와 연관성을 보였으며 이상지질혈증은 RCA DS의 90% 백분위와 연관성이 관찰되었다. 기타 증상은 LCx 및 RCA와 연관성이 관찰되었다. 고밀도지단백 콜레스테롤의 경우 환자별 분석에서 역동적인 연관성을 보였다. (2) 라디오믹스 스코어는 100 이상의 관상동맥 석회화 지수를 예측하는 가장 중요한 인자였다 (Odds ratio = 2.33; 95% Confidence interval [CI] = 1.62-3.44; $p < 0.001$). 또한 라디오믹스 기반 기계학습 모델은 AUC 0.84 (95% CI = 0.79-0.87)의 우수한 성능을 석회화 지수 100 이상을 예측하였다. (3) 딥러닝 기반의 허혈성 변화 예측 심전도 모델은 트레이닝 코호트 (training cohort)와 검증 코호트 (validation cohort) 각각 AUC 0.60 (95% CI 0.54-0.66) 및 AUC 0.60 (95% CI 0.50-0.69)의 중등도의 성능을 보여주었으나 임상인자를 추가했을 때에는 트레이닝 코호트와 검증 코호트 각각 AUC of 0.71 (95% CI 0.66-0.76) 및 0.67 (95% CI 0.59-0.69)로 성능 향상을 보였다. (4) 위험인자, 흉부방사성 영상, 심전도를 통합한 최종 모델의 성능은 외부검증 (external validation)에서 AUC 0.77 (민감도 73%, 특이도 69%)로 측정되었다.

결론: 본 연구를 통해 일차 검사인 위험인자, 흉부방사성 영상, 심전도를 이용하여 관상동맥질환 선별을 위한 이종 특징 기반 AI Gatekeeper 솔루션을 계발하였다. 이는 특히 일차 진료 현장에서 널리 활용되어 관상동맥질환의 검사 전 확률 (Pre-test probability)을 향상시킬 수 있을 것이다.

핵심되는 말 : 인공지능, 관상동맥질환, 심혈관위험요소, 흉부방사선촬영, 심전도