Development of a model for angiography-based fractional flow reserve using machine learning in coronary artery disease

Geunhee Park

The Graduate School
Yonsei University
Department of Medicine

Development of a model for angiography-based fractional flow reserve using machine learning in coronary artery disease

A Master's Thesis Submitted
to the Department of Medicine
and the Graduate School of Yonsei University
in partial fulfillment of the
requirements for the degree of
Master of Medical Science

Geunhee Park

December 2024

This certifies that the Master's Thesis of Geunhee Park is approved

Thesis Supervisor Jung-Sun Kim

Thesis Committee Member Seng Chan You

Thesis Committee Member Jinyong Ha

The Graduate School
Yonsei University
December 2024

ACKNOWLEDGEMENTS

Firstly, I would like to express my deepest appreciation to my supervisor, Professor Jung-Sun Kim. His unwavering support and insightful guidance have been instrumental in shaping the direction of my research. I am profoundly grateful for the way he has mentored me, not only as an academic advisor but also as a role model in the field of cardiology. His exemplary career as a senior physician has inspired me, and his wealth of knowledge and experience has enriched my understanding of the subject. Thank you to him for sharing his wisdom and for encouraging me to pursue excellence in my work.

I am also thankful to the chairperson of the thesis committee, Professor Seng Chan You, for his meticulous guidance on the areas where my research was lacking and for collaboratively exploring the objectives, limitations, and potential solutions to the challenges I faced. I would like to extend my appreciation to a member of the thesis committee, Professor Jinyong Ha, for providing substantial technical support in building the machine learning model and for helping me achieve the goals of my research.

I am grateful to Professor Myeong-Ki Hong, Professor Donghoon Choi, Professor Young-Guk Ko, Professor Byeong-Keuk Kim, Professor Chul-Min Ahn, Professor Sung-Jin Hong, Professor Seung-

Jun Lee, Professor Yong-Joon Lee, Professor Sang-Hyup Lee, and Professor Joongmin Kim, who provided me with substantial support and encouragement, even during their busy schedules, helping me secure data and serving as excellent examples of dedicated researchers and great educators. Also, I would like to express my gratitude to my good friend, Professor Jae Young Kim, for his assistance with the statistical analysis.

I deeply appreciate my family, especially my parents, Sehyung Park and Hyunsuk Choi, and my sister, Giyoung Park, for taking good care of the children during my research process and for their encouragement. I would also like to thank my daughter, Susie Park, and my son, Jiwoo Park, for their understanding and support of their busy dad. Finally, I extend my gratitude and love to my wife, Dawoon Jeong, a good friend and senior researcher, for her unfailing support and encouragement throughout my master's thesis. I look forward to continuing our journey together, embracing the joys and challenges that life brings.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Development of a model for angiography-based fractional flow reserve using machine learning in coronary artery disease

Objectives: Guidelines recommend using pressure wire-based fractional flow reserve (FFR) when deciding to perform coronary intervention in patients with angina and intermediate lesions. However, several hurdles, including cost and safety concerns, have reduced the rate of FFR usage in clinical practice. Meanwhile, noninvasive FFR methods have been widely investigated for their safety and convenience. This study aimed to develop machine learning (ML) model, referred as ML-Angio-FFR model, using quantitative coronary angiography and clinical data that can predict FFR more safely as well as cost and time effectively.

Methods: The ML-Angio-FFR model was trained using the Random Forest algorithm with 42 features based on 1,459 patients with 2,439 intermediate lesions. All the features were ranked by feature importance score. The training and testing sets were divided in a 4:1 ratio. FFR ≤0.80 was considered ischemic driven lesion, while FFR >0.80 was considered lesion can be deferred intervention and the performance test was conducted using this threshold.

Results: The mean age of the enrolled patients was 67.1 years, and 1,004 (68.8%) were male. Of the included lesions, 1,481 (60.7%), 533 (21.9%), and 425 (17.4%) were located in left anterior descending artery, left circumflex artery, and right coronary artery, respectively. Using the 8 key features, the ML-Angio-FFR model was developed. There was a strong correlation between ML-Angio-FFR and wire-based FFR (r=0.7828; p<0.001), along with high diagnostic accuracy in the testing set (87%; area under the curve, 0.89). External validation revealed modest correlation (r=0.4940; p<0.001) and accuracy (62%; area under the curve, 0.73).

Conclusions: The ML-Angio-FFR model demonstrates a high correlation with pressure wire-based FFR and presents good diagnostic performance to predict FFR >0.80.

Key words : fractional flow reserve, machine learning, quantitative coronary angiography

1. INTRODUCTION

Fractional flow reserve (FFR) is one of the most widely investigated coronary physiology assessment. Moreover, FFR-guided percutaneous coronary intervention (PCI) is currently recommended by several guidelines for patients with angina who have intermediate coronary artery stenosis.[1,2] This approach can be applied not only to single coronary lesions but also in various clinical scenarios, such as multi-vessel coronary artery disease or acute coronary syndrome.[3-8] However, despite guidelines and numerous studies, FFR is still underutilized due to the time and cost associated for additional using pressure wire, as well as the use of medications to induce maximal hyperemia during the procedure.[9-11] In these regards, there have been increasing efforts to predict FFR without pressure wire, defined as noninvasive coronary physiology assessments.[12-14] Some noninvasive methods, such as coronary computed tomography angiography (CTA) based FFR prediction or angiography-based FFR prediction (without using pressure wire), have already been developed and validated for use in clinical practice.[15-17] These methods are considered safe and showed not only good correlation with pressure wire-based FFR, but also higher accuracy compared to traditional quantitative coronary angiography (QCA) based FFR prediction using diameter stenosis.[18-22] However, each method still requires dedicated software and formulas, along with additional cost and time.[23,24] Meanwhile, advances in computing power have led to the emergence of machine learning (ML) technology, there are several studies for predicting FFR in ML model with various features.[25,26] However, it is not yet feasible to use these methods for on-site analysis due to difficulties in applying them simultaneously with coronary angiography. In this era, this study aimed to develop a ML model, referred to as ML-Angio-FFR, to predict FFR using both lesion characteristics obtained from traditional 2D QCA and patient's baseline characteristics, which do not require additional software or preprocessing steps.

2. METHODS

2.1. Study design and population

The ML-Angio-FFR study is a retrospective single-center study aimed at validating the accuracy of ML model to predict FFR based on angiography, compared with pressure wire-based FFR assessment. This study was permitted by the Institutional Review Board of Severance Hospital, Seoul, Korea (4-2024-0245), and followed the ethical principles of the Declaration of Helsinki (2013). The requirement of informed consent was dispensed with the current study was retrospective and study subjects were de-identified according to confidentiality guidelines. A total of 1,955 patients who underwent invasive coronary angiography and wire-based FFR measurement for assessment of chest pain were screened from April 2018 to February 2024 in the single-center. These patients had a total of 3,885 lesions with intermediate coronary stenosis (40-70% stenosis) by visual estimation. The ML-Angio-FFR model was planned to develop based on per-vessel analysis, data preprocessing was conducted by lesion level. Of these lesions, 1,446 were excluded due to the following criteria: (1) poor quality of FFR data or lack of clinical data (n=20); (2) overlapped or severe tortuosity of coronary artery which was difficult to measure QCA (n=68); (3) tandem or mixed lesions (n=1,121); (4) small side branch evaluation (n=20); (5) proximal left main coronary artery evaluation (< 3 mm from the aorta) (n=33); (6) culprit lesion in acute myocardial infarction (n=43); (7) diagnosed with variant angina (n=6); and (8) outliers, defined as FFR <0.70 (n=135). Finally, 1,459 patients with 2,439 lesions were included in the study (**Fig. 1**).

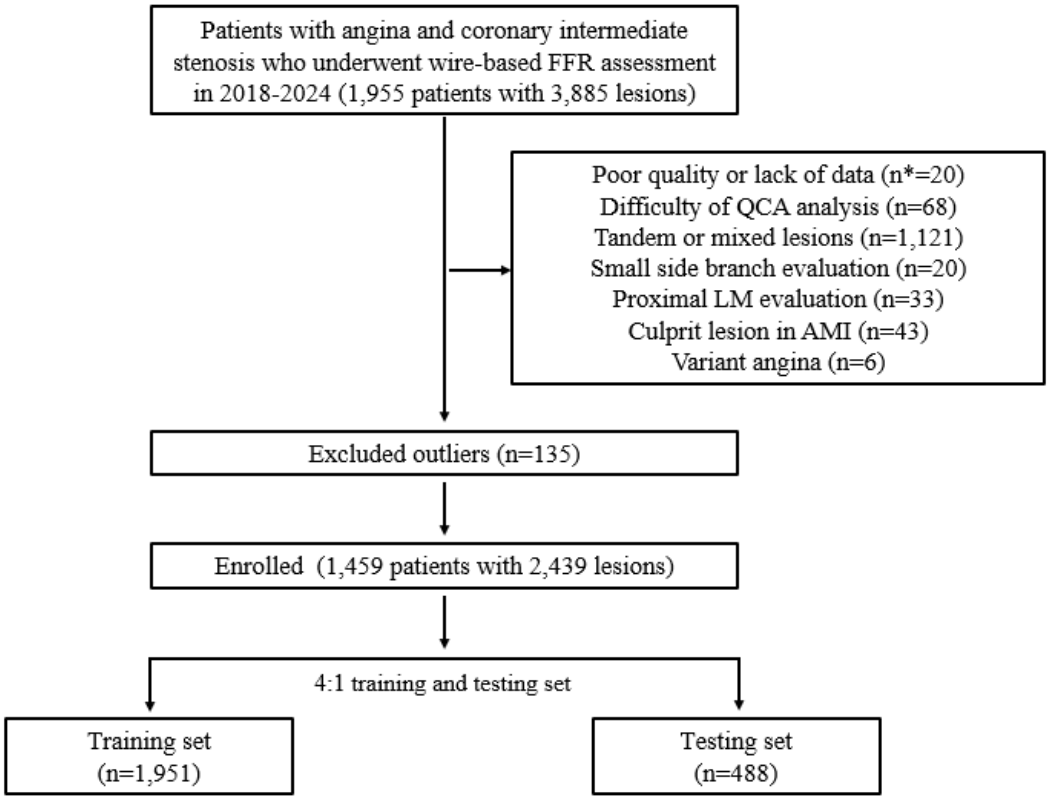


Figure 1. Flow chart of the study enrollment. Patients were screened from April 2018 to February 2024 and were excluded based on the criteria described above. Lesions with FFR <0.70 were considered outliers and were excluded. Finally, the dataset was randomly divided into training and testing sets in a 4:1 ratio.

Note: *, n=number of lesions.

Abbreviations: FFR, fractional flow reserve; QCA, quantitative coronary angiography; LM, left main coronary artery; AMI, acute myocardial infarction.

2.2. Image analysis

QCA analysis was conducted by trained specialists using an offline analytic system of quantitative coronary angiography (CAAS, Pie Medical Imaging, Maastricht, the Netherlands). The minimal lumen diameter was assessed with diastolic frames from a single-matched view that was neither foreshortened nor distorted, using a guiding catheter for magnification calibration to ensure accurate measurement of the smallest lumen diameter. Each lesion was classified into four categories, defined as lesion profile.[27] Lesion with a length of less than 20 mm was defined as “focal lesion”, while those of 20 mm or longer was defined as “diffuse lesion”. If multiple focal lesions were present in one major epicardial coronary artery, they were defined as "tandem lesions", and QCA was conducted for each included focal lesion. When both focal and diffuse lesions co-existed, or multiple diffuse lesions were present in one major epicardial coronary artery, they were defined as “mixed lesions". Target vessel was defined as left anterior descending artery (LAD), left circumflex artery (LCX), right coronary artery (RCA), and left main coronary artery. The lesion location was classified according to each vessel. The LAD included proximal, middle, distal, and diagonal branch. The LCX included proximal, distal, obtuse marginal branch, and ramus intermedius branch. The RCA included proximal, middle, distal, posterolateral branch, and posterior descending artery. If the lesion extends from proximal to middle, from middle to distal, or from proximal to distal portions, it was noted separately as LAD extended, LCX extended, or RCA extended.

2.3. Pressure wire-based FFR measurement

The pressure wire-based FFR measurements were conducted using a 0.014-inch pressure wire (Abbott Vascular, Santa Clara, CA, USA). After equalization process at aorta, the pressure wire was positioned most distal to the target lesions. Hyperemia was induced by intravenous infusion of 140 μg/kg/min of adenosine via an antecubital vein or intracoronary bolus injection of 2 mg nicorandil. FFR was determined by calculating the ratio of the mean pressure of distal coronary lesion to the mean pressure of aorta during maximal hyperemia. The lesion with FFR $\leq$0.80 was considered significant ischemic driven lesion, while FFR >0.80 was considered that the lesion can be deferred for PCI.

2.4. ML-FFR assessment based on angiography

In ML regression task, the model attempts to identify and quantify the relationships between variables by estimating how changes in one variable impact another. The process involved selecting relevant features, choosing an appropriate ML algorithm, training the model, assessing its performance, and tuning the hyperparameters. For this study, a Random Forest (RF) algorithm was applied to predict FFR values. As an ensemble technique, RF constructs multiple decision trees from random subsets of the data, and the final prediction is the combined output of all the trees. This technique helps prevent overfitting, reduces variance, and enhances accuracy even with imbalanced datasets.

A total of 42 features were selected for model development, based on patients' clinical characteristics and QCA data, to support the ML-Angio-FFR model in assessing coronary intermediate lesions. During data preprocessing, cases with FFR <0.70 were considered as outliers and removed to prevent the model from being distorted. Initially, ML model was trained by using comprehensive set of 42 features to establish baseline performance (**Table 1**), and compute feature importance scores using Gini importance derived from the RF model. After evaluating the importance ranking, top 8 features were selected for model development, ensuring that only the most relevant predictors were retained. The dataset consisted of 2,439 samples, which were split into training and testing sets using a stratified sampling technique to maintain the distribution across the subgroups: LAD, LCX, and RCA. The training set comprised 1,951 samples, and the remaining 488 samples were used for independent testing, with a training-to-testing ratio of 4:1. To ensure model stability and optimize performance, cross-validation (CV) alongside hyperparameter tuning was performed using GridSearchCV. The training set was partitioned into K folds, with K = 5 used in this study. The RF model was trained through multiple iterations of the 5-fold CV process, each time using different combinations of hyperparameters to fine-tune the model and ensure robust results. The final optimized hyperparameters, as summarized in **Table 2**, include *n_estimators* = 20, *max_depth* = 8, *min_samples_leaf* = 4, and *min_samples_split* = 3, while default values were retained for the remaining parameters. Using these optimized settings, the RF model was established. During the testing phase, the 488 lesions characterized by the top 8 features were evaluated using the trained models to predict FFR values. The performance of the RF model was assessed by applying Pearson correlation coefficient (*r*) analysis and calculating the mean absolute error (MAE) to quantify the agreement between predicted and wire-based FFR values.

Table 1. List of 42 features and their weight

Rank	Feature	Weight
1	Minimal lumen diameter	0.520
2	Target vessel	0.491
3	Diameter stenosis	0.472
4	Distal reference vessel diameter	0.292
5	Mean reference vessel diameter	0.251
6	Lesion profile	0.242
7	Lesion length	0.236
8	Proximal reference vessel diameter	0.182
9	Age	0.090
10	Lesion location	0.081
11	Platelet count	0.076
12	Hemoglobin	0.073
13	Height	0.068
14	Previous percutaneous coronary intervention	0.063
15	Male sex	0.062
16	Vessel territory of regional wall motion abnormality	0.059
17	Regional wall motion abnormality	0.053
18	White blood cell count	0.052
19	Low-density lipoprotein-cholesterol level	0.051
20	Aspartate aminotransferase level	0.051
21	Body surface area	0.048
22	Weight	0.038
23	Chronic kidney disease	0.035
24	Diabetes mellitus	0.033
25	Previous myocardial infarction	0.031
26	Previous stroke	0.031
27	Clinical presentation	0.030
28	Left ventricular ejection fraction	0.028
29	Creatinine level	0.027

30	Heart failure	0.027
31	End stage renal disease	0.025
32	Triglyceride level	0.022
33	Acute coronary syndrome	0.021
34	Previous coronary artery bypass surgery	0.020
35	Blood urea nitrogen level	0.018
36	Current smoker	0.018
37	Dyslipidemia	0.016
38	Hypertension	0.015
39	Alanine aminotransferase level	0.012
40	High-density lipoprotein-cholesterol level	0.008
41	Persistent atrial fibrillation	0.008
42	Body mass index	0.007

Note: Target vessel included LAD, LCX, and RCA. All of laboratory and echocardiography features were obtained peri-procedural periods. Total 42 features which were consist with 9 of angiography features, 20 of clinical features, 10 of laboratory features, and 3 of echocardiography features were included.

Abbreviations: LAD, left anterior descending artery; LCX, left circumflex artery; RCA, right coronary artery.

Table 2. Optimized hyperparameters for Random Forest model

Hyperparameter	Description	Value
n_estimators	Number of trees in Random Forest	20
max_depth	Maximum number of levels in tree	8
min_samples_split	Minimum number of samples required to split an internal node	3
min_samples_leaf	Minimum number of samples required at each leaf node	4
max_features	Maximum number of features to consider for splitting a node	log2
random_state	Controls randomness for result reproducibility	8

2.5. External validation

To evaluate the consistent performance of ML-Angio-FFR model, external validation was conducted using another single-center's data which includes 131 patients with 149 intermediate coronary lesions who underwent coronary angiography and invasive wire-based FFR. This data was permitted for use by the Institutional Review Board of National Health Insurance Service Ilsan Hospital, Goyang-si, Korea (2024-09-025).

2.6. Statistical analysis

Baseline clinical characteristics were analyzed on per-patient level. The other evaluations and ML modeling were analyzed on per-vessel level. Continuous variables were presented as mean ± standard deviation or median (interquartile range). These were analyzed using the Shapiro–Wilk test for normality, and based on the results, analyzed using Student's t-test or the Mann-Whitney U test. Categorical variables were presented as numbers (percentages) and analyzed using chi-square test or Fisher's exact test. Pearson correlation coefficient (*r*) analysis and the Bland-Altman analysis were performed to evaluate correlation between ML-Angio-FFR and wire-based FFR measurement. Receiver operating characteristic (ROC) curve analyses were conducted, and sensitivity, specificity, positive predictive value, negative predictive value, and diagnostic accuracy of the ML model were calculated to define diagnostic performance of ML-Angio-FFR model. R version 4.2.2 (R Foundation for Statistical Computing, Vienna, Austria) was used for the statistical analyses.

3. RESULTS

3.1. Baseline characteristics of patients and lesions

The baseline characteristics of the study population and enrolled lesions to develop ML model are presented **Table 3** and **Table 4**, respectively. The mean age of enrolled patients was 67.1 ± 9.5 years, 1,004 patients (68.8%) were male, and the mean body mass index was 24.5 ± 3.1 kg/m^2. Nine hundred sixty-six (66.2%), 576 (39.5%), and 102 (7.0%) patients had a medical history of hypertension, diabetes mellitus, and previous myocardial infarction, respectively, and 116 (8.0%) patients were current smokers. Three hundred twenty-two (22.1%) patients had heart failure, and 113 (7.7%) patients had chronic kidney disease (defined as estimated glomerular filtration ratio below 60). One thousand one hundred eighty-six (81.3%) patients presented with stable coronary artery disease, and 273 (18.7%) patients presented with acute coronary syndrome. Among the 2,439 target lesions, 1,481 (60.7%), 533 (21.9%), and 425 (17.4%) were placed in the LAD, LCX, and RCA, respectively. In echocardiographic findings, mean left ventricular ejection fraction was 64.0 ± 9.9 %, and 277 (19%) patients had regional wall motion abnormalities.

The median FFR was 0.87 (interquartile range [IQR], 0.82–0.93), with the median FFR for LAD at 0.84 (IQR, 0.80–0.88), LCX at 0.94 (IQR, 0.88–0.97), and RCA at 0.93 (IQR, 0.89–0.97). Ischemia driven lesions (FFR ≤0.80) were observed in 468 (19.2%) lesions. The median values for FFR in the ischemia driven and not significantly driven ischemia lesions were 0.76 (IQR, 0.74–0.78) and 0.89 (IQR, 0.85–0.94), respectively. The mean minimal lumen diameter was 1.6 ± 0.6 mm, mean reference vessel diameter was 3.0 ± 0.6 mm, and mean diameter stenosis was 49.0 ± 14.5 %. The mean lesion length was 21.0 ± 12.6 mm, including 1,436 (58.9%) of focal lesions, and 1,003 (41.1%) of diffuse lesions.

Table 3. Baseline characteristics of the study population

	Total (N=1,459)
Age, years	67.1 ± 9.5
Male	1,004 (68.8)
Height, cm	164.5 ± 8.4
Weight, kg	66.6 ± 11.2
Body mass index, kg/m^2	24.5 ± 3.1
Body surface area, m^2	1.7 ± 0.2
Hypertension	966 (66.2)
Diabetes mellitus	576 (39.5)
Dyslipidemia	1096 (75.1)
Heart failure	322 (22.1)
Persistent atrial fibrillation	21 (1.4)
Chronic kidney disease	113 (7.7)
End stage renal disease	45 (3.1)
Current smoker	116 (8.0)
Previous myocardial infarction	102 (7.0)
Previous percutaneous coronary intervention	391 (26.8)
Previous stroke	106 (7.3)
Previous coronary bypass surgery	17 (1.2)
Clinical presentation	
Stable coronary artery disease	1186 (81.3)
Acute coronary syndrome	273 (18.7)
Unstable angina	225 (15.4)
Non-ST-elevation myocardial infarction	47 (3.2)
ST-elevation myocardial infarction	1 (0.1)
Echocardiography finding	
Left ventricular ejection fraction, %	64 ± 10
Regional wall motion abnormality	277 (19.0)
Left anterior descending artery territory	81 (5.6)
Left circumflex artery territory	35 (2.4)

Right coronary artery territory	89 (6.1)
Multi-vessel territory	72 (4.9)
Laboratory finding	
Hemoglobin, g/dL	13.5 ± 3.8
White blood cell count, $10^3/\mu L$	6.7 ± 1.9
Platelet count, $10^3/\mu L$	217 ± 61
Blood urea nitrogen, mg/dL	18.1 ± 8.6
Creatinine, mg/dL	1.1 ± 1.2
AST, IU/L	28 ± 16
ALT, IU/L	27 ± 19
Triglyceride, mg/dL	128 ± 90
HDL-Cholesterol, mg/dL	46 ± 12
LDL-Cholesterol, mg/dL	79 ± 34

Note: Data are presented as mean ± standard deviation, or number (percentage).

Abbreviations: AST, Aspartate aminotransferase; ALT, Alanine aminotransferase; HDL, High-density lipoprotein; LDL, Low-density lipoprotein.

Table 4. Fractional flow reserve and angiographic findings of the enrolled lesions

	Total (N=2,439)
Lesion location	
LAD	1,481 (60.7)
Proximal LAD	424 (17.4)
Middle LAD	513 (21.0)
Distal LAD	46 (1.9)
LAD extended	382 (15.7)
Diagonal branch	116 (4.8)
LCX	533 (21.9)
Proximal LCX	257 (10.5)
Distal LCX	59 (2.4)
LCX extended	111 (4.6)
Obtuse marginal branch	106 (4.3)
RCA	425 (17.4)
Proximal RCA	129 (5.3)
Middle RCA	140 (5.7)
Distal RCA	75 (3.1)
RCA extended	79 (3.2)
Posterolateral branch	1 (0.0)
Posterior descending artery	1 (0.0)
Wire-based FFR data	
Median FFR	0.87 (0.82-0.93)
FFR at LAD	0.84 (0.80-0.88)
FFR at LCX	0.94 (0.88-0.97)
FFR at RCA	0.93 (0.89-0.97)
FFR ≤0.80	468 (19.2)
Median FFR in FFR ≤0.80	0.76 (0.74-0.78)
Median FFR in FFR >0.80	0.89 (0.85-0.94)
Hyperemia agent	
Intravenous adenosine	1,856 (76.1)

Intracoronary nicorandil	583 (23.9)
Quantitative coronary angiography data	
Reference vessel diameter, mm	3.0 ± 0.6
Proximal reference vessel diameter, mm	3.2 ± 0.6
Distal reference vessel diameter, mm	2.8 ± 0.6
Minimal lumen diameter, mm	1.6 ± 0.6
Diameter stenosis, %	49.0 ± 14.5
Lesion length, mm	21.0 ± 12.6
Lesion profile	
Focal	1,436 (58.9)
Diffuse	1,003 (41.1)

Note: Data are presented as mean ± standard deviation, median (interquartile range), or number (percentage).

Abbreviations: LAD, left anterior descending artery; LCX, left circumflex artery; RCA, right coronary artery; FFR, fractional flow reserve.

3.2. Major features and diagnostic performance of the ML-Angio-FFR model

The eight major features were selected from a total of 42 features based on RF feature importance. These selected 8 features as follows: target vessel, minimal lumen diameter, diameter stenosis, lesion location, lesion length, mean reference vessel diameter, distal reference vessel diameter, and proximal reference vessel diameter (**Fig. 2**). There was a strong correlation between ML-Angio-FFR and wire-based FFR (r=0.7828; p<0.001), and good agreement (MAE=0.0385) in the testing sets (**Fig. 3A**). The Bland-Altman plot also demonstrated good agreement in the testing sets, resulting in a mean difference of 0.002, and 95 % confidence limits ranging from -0.093 to 0.097 (**Fig. 3B**). In the external validation sets, a modest correlation between ML-Angio-FFR and wire based FFR (r=0.4940; p<0.001) and good agreement (MAE=0.0563) were observed (**Fig. 3C)**. The Bland-Altman plot demonstrated modest agreement in the external validation sets, resulting in a mean difference of 0.031, and 95 % confidence limits ranging from -0.094 to 0.155 (**Fig. 3D**).

Using a threshold of FFR >0.80, the performance tests were conducted. In the testing sets, the area under the curve (AUC) of the ROC curve was 0.89 (95% confidence interval [CI], 0.85-0.92). The sensitivity, specificity, positive predictive value, negative predictive value, and diagnostic accuracy in the testing sets were 97 % (95% CI, 96-99), 47 % (95% CI, 37-57), 88 % (95% CI, 85-91), 80 % (95% CI, 70-91), and 87 % (95% CI, 84-90), respectively (**Fig. 4A**). In the external validation sets, AUC of the ROC curve was 0.73 (95% CI, 0.64-0.81). The sensitivity, specificity, positive predictive value, negative predictive value, and diagnostic accuracy in the external validation sets were 97 % (95% CI, 92-100), 11 % (95% CI, 3-19), 62 % (95% CI, 53-70), 67 % (95% CI, 33-100), and 62 % (95% CI, 54-70), respectively (**Fig. 4B**).

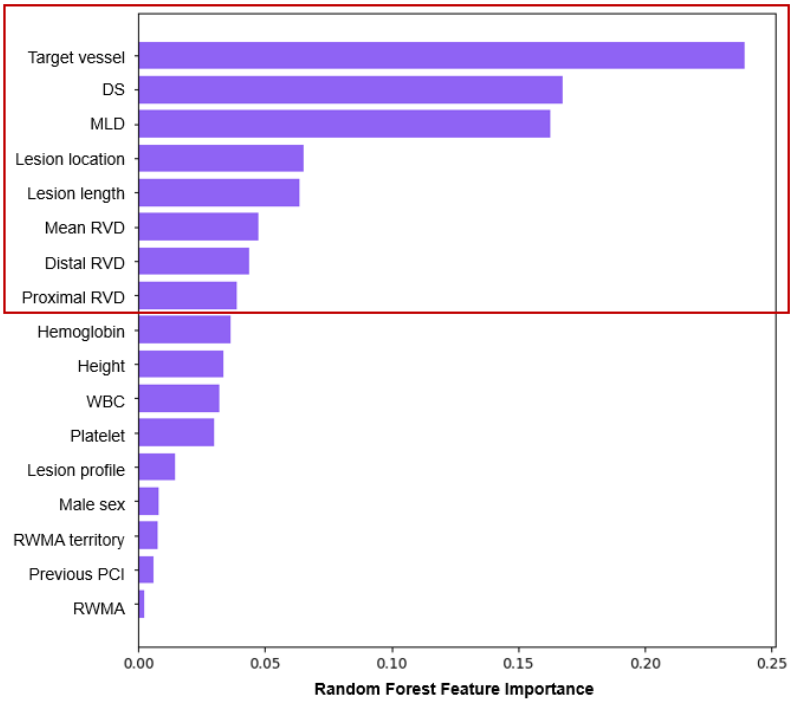


Figure 2. Random Forest feature importance for the selected 8 features. Feature importance was computed using the Gini importance, the top 8 features were selected for model development.
Abbreviations: DS, diameter stenosis; MLD, minimal lumen diameter; RVD, reference vessel diameter; WBC, white blood cell; RWMA, regional wall motion abnormality; PCI, percutaneous coronary intervention.

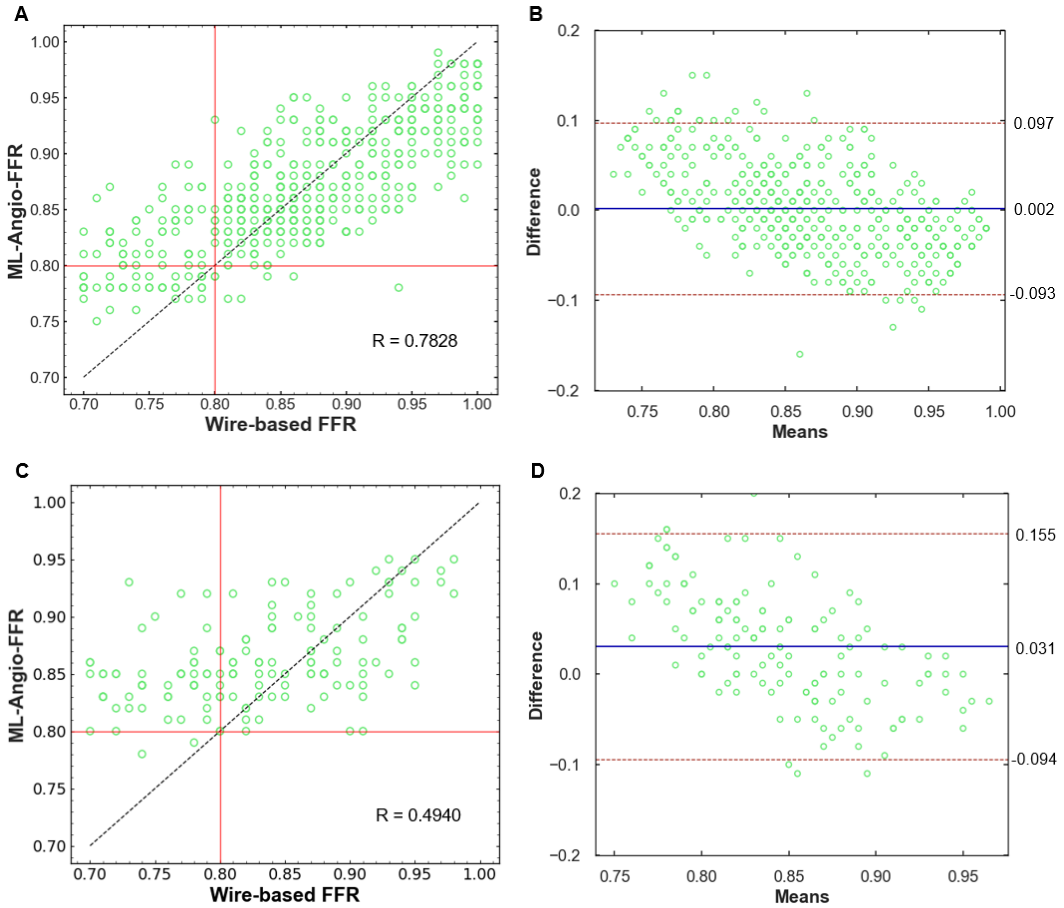
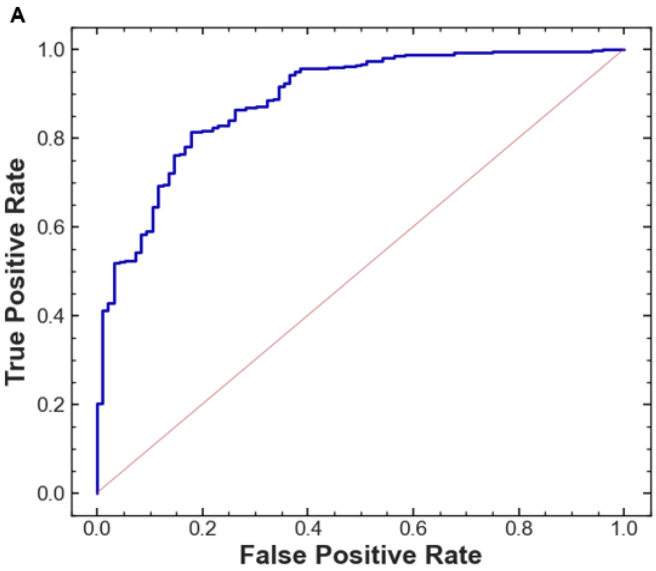
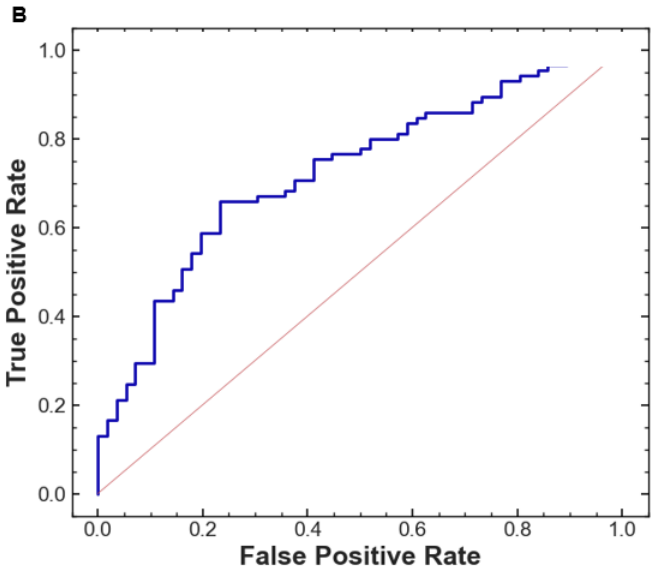


Figure 3. Correlation between ML-Angio-FFR and wire-based FFR. (A) The scatter plot demonstrates a strong correlation between ML-Angio-FFR and wire-based FFR in the testing sets, with a Pearson correlation coefficient of r=0.7828. (B) The Bland-Altman plot demonstrates a mean difference of 0.002, and 95% confidence limits between -0.093 and 0.097 in the testing sets. (C) The scatter plot demonstrates a modest correlation between ML-Angio-FFR and wire-based FFR in the external validation sets, with a Pearson correlation coefficient of r=0.4940. (D) The Bland-Altman plot demonstrates a mean difference of 0.031, and 95% confidence limits between -0.094 and 0.155. The mean (x-axis) was calculated as follows: (ML-Angio-FFR + wire-based FFR)/2, and the difference (y-axis) was calculated as follows: ML-Angio-FFR - wire-based FFR.

Abbreviation: FFR, fractional flow reserve.



Diagnostic characteristics	Value (95% CI)
AUC	0.89 (0.85-0.92)
Sensitivity	97 % (96-99)
Specificity	47 % (37-57)
PPV	88 % (85-91)
NPV	80 % (70-91)
Diagnostic accuracy	87 % (84-90)



Diagnostic characteristics	Value (95% CI)
AUC	0.73 (0.64-0.81)
Sensitivity	97 % (92-100)
Specificity	11 % (3-19)
PPV	62 % (53-70)
NPV	67 % (33-100)
Diagnostic accuracy	62 % (54-70)

Figure 4. Receiver operating characteristic curves of ML-Angio-FFR model and its diagnostic performance. (A) The ROC curve in the testing sets was constructed with FFR >0.80 as the threshold. (B) The ROC curve in the external validation sets was also constructed with FFR >0.80 as the threshold.

Abbreviations: AUC, area under the curve; PPV, positive predictive value; NPV, negative predictive value; ROC, receiver operating characteristic; FFR, fractional flow reserve.

4. DISCUSSION

The ML-Angio-FFR study was a retrospective single-center study for the development of an angiography-based FFR model using ML. This study demonstrated good diagnostic performance of ML-Angio-FFR model, showing high correlation and agreement with pressure wire-based FFR.

The pressure wire-based invasive FFR is a crucial tool in treatment strategies of coronary artery disease, supported by numerous previous research.[3-8] Consequently, there has been growing interest in developing noninvasive FFR methods that can more safely and cost-effectively predict wire-based FFR. FFT_{CT} (Heartflow, Inc., Redwood City, CA, USA) is one of the most widely used noninvasive FFR methods, based on computed tomography (CT)-derived FFR.[14] FFT_{CT} creates 3D vessel model from coronary CTA and applies computational fluid dynamics (CFD) to predict FFR.[14] It is beneficial as it enables selective conduct of invasive coronary angiography in patients with angina and intermediate stenosis.[14,15,22] Moreover, several studies have demonstrated that FFT_{CT} can reduce total healthcare costs,[15] additionally, it has shown equivalent or better results in clinical efficacy and safety compared to traditional diagnostic testing in stable patients suspected coronary artery disease.[28] However, the CFD algorithm relies on various assumptions about hemodynamics of coronary flow and the quality of coronary CTA.[29] One of the major limitations of CFD-based models is that their predictive accuracy may be reduced if there is a significant difference between these assumptions and actual blood flow conditions, or if the CT quality is poor.[29,30] Quantitative flow ratio (QFR) (Medis Medical Imaging System, Leiden, the Netherlands, and Pulse Medical Imaging Technology, Shanghai, China) is one of the effective angiography-based FFR methods which has shown strong correlation with wire-based FFR.[12,16,17,21] In addition, QFR took less time than wire-based FFR.[16] Moreover, in recent randomized controlled trial, QFR-guided PCI was improved 1-year clinical outcomes compared to standard angiography-guided PCI.[17] However, QFR requires specific angles during coronary angiography and coronary flow can influence the predicted FFR.[24] Vessel fractional flow reserve (vFFR) (CAAS, Pie Medical Imaging, Maastricht, the Netherlands) is another angiography-based FFR method which generates 3D QCA from two distinct views over 30° and can be conducted using the CAAS workstation.[13,18] It is beneficial that angiography images do not need to be transferred to other software for predicting FFR, and vFFR can also trace target vessel effectively, resulting in lower interobserver variability.[18,24] However, like other methods, it is

limited to use in heavily calcified or tortuous vessels and requires additional effort to obtain the appropriate angiographic views.[24] FFR_{angio} (CathWorks, Ltd, Kfar Saba, Israel) is also angiography-based FFR, which is needed two or more angiographic views with different angle for 3D reconstruction of entire coronary trees.[19] It is possible to conduct a comprehensive evaluation of all coronary arteries, which makes it beneficial for patients with multi-vessel disease, though it may be time-consuming in patients with single coronary artery disease.[19,24] Likewise, FFR_{angio} is based on CFD principles, it can be influenced by quality of image or contrast flow.[24]

Traditional 2D QCA is still widely used for stent sizing and is convenient for on-site measurement. However, if the decision of PCI is based solely on diameter stenosis (>50%) in QCA analysis, it leads to poor accuracy in predicting FFR ≤0.80 and a lower positive predictive value, often resulting in unnecessary PCI.[9,16,20,31] This study proposes a new ML model to predict FFR based on angiography, combining several features including traditional 2D QCA data with some additional information about lesions to identify ischemic driven lesions or lesions that can be deferred for PCI. With high sensitivity to predict FFR >0.80, this ML-Angio-FFR model can help identify patients who are not required to undergo invasive FFR, providing both cost and time efficiency.

This study, however, is not without limitations. The ML-Angio-FFR model was developed using single-center data, therefore, there were concerns about overfitting. The results of performance tests in the external validation sets slightly decreased compared to the testing sets. These findings suggest that the ML-Angio-FFR model may be overfitted. It is necessary to investigate additional study using multi-center data, which will help to overcome the possibility of overfitting. The ML-Angio-FFR model used QCA data with a single angiographic projection, which may lead to an over- or underestimation of the lesion profile. However, this is a pragmatic method used in real-world practice. Additionally, all angiographic views were selected by expert clinicians to define appropriate lesion characteristics and avoid foreshortening. Not all lesion profiles, such as tandem and mixed lesions, were included in ML-Angio-FFL model due to the lack of pull-back FFR data, even though they are commonly shown in clinical practice. Therefore, it can be used only in limited situations. Although the model demonstrated good diagnostic accuracy and sensitivity, the specificity value is lower than the other values; therefore, it may not be advisable to determine PCI solely based on the ML-Angio-FFR value at this time. Finally, since it has not been applied on-site clinical practice, additional examination will be needed for evaluation of interobserver variability.

5. CONCLUSION

The ML-Angio-FFR model demonstrates a high correlation with pressure wire-based FFR and presents good diagnostic performance to predict FFR >0.80 in off-site analysis. Additional training with data from other centers may be needed to enhance accuracy and avoid overfitting.

References

1. Lawton JS, Tamis-Holland JE, Bangalore S, Bates ER, Beckie TM, Bischoff JM, et al. 2021 ACC/AHA/SCAI Guideline for Coronary Artery Revascularization: A Report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. Circulation 2022;145:e18-e114.
2. Neumann FJ, Sousa-Uva M, Ahlsson A, Alfonso F, Banning AP, Benedetto U, et al. 2018 ESC/EACTS Guidelines on myocardial revascularization. Eur Heart J 2019;40:87-165.
3. Bech GJ, De Bruyne B, Pijls NH, de Muinck ED, Hoorntje JC, Escaned J, et al. Fractional flow reserve to determine the appropriateness of angioplasty in moderate coronary stenosis: a randomized trial. Circulation 2001;103:2928-34.
4. Pijls NH, van Schaardenburgh P, Manoharan G, Boersma E, Bech JW, van't Veer M, et al. Percutaneous coronary intervention of functionally nonsignificant stenosis: 5-year follow-up of the DEFER Study. J Am Coll Cardiol 2007;49:2105-11.
5. Tonino PA, De Bruyne B, Pijls NH, Siebert U, Ikeno F, van' t Veer M, et al. Fractional flow reserve versus angiography for guiding percutaneous coronary intervention. N Engl J Med 2009;360:213-24.
6. De Bruyne B, Pijls NH, Kalesan B, Barbato E, Tonino PA, Piroth Z, et al. Fractional flow reserve-guided PCI versus medical therapy in stable coronary disease. N Engl J Med 2012;367:991-1001.
7. Fearon WF, Zimmermann FM, De Bruyne B, Piroth Z, van Straten AHM, Szekely L, et al. Fractional Flow Reserve-Guided PCI as Compared with Coronary Bypass Surgery. N Engl J Med 2022;386:128-37.
8. Engstrøm T, Kelbæk H, Helqvist S, Høfsten DE, Kløvgaard L, Holmvang L, et al. Complete revascularisation versus treatment of the culprit lesion only in patients with ST-segment elevation myocardial infarction and multivessel disease (DANAMI-3—PRIMULTI): an open-label, randomised controlled trial. Lancet 2015;386:665-71.
9. Toth GG, Toth B, Johnson NP, De Vroey F, Di Serafino L, Pyxaras S, et al. Revascularization decisions in patients with stable angina and intermediate lesions: results of the international survey on interventional strategy. Circ Cardiovasc Interv 2014;7:751-9.
10. Shin DH, Kang HJ, Jang JS, Moon KW, Song YB, Park DW, et al. The Current Status of Percutaneous Coronary Intervention in Korea: Based on Year 2014 & 2016 Cohort of Korean Percutaneous Coronary Intervention (K-PCI) Registry. Korean Circ J 2019;49:1136-51.

11. G GT, Johnson NP, Wijns W, Toth B, Achim A, Fournier S, et al. Revascularization decisions in patients with chronic coronary syndromes: Results of the second International Survey on Interventional Strategy (ISIS-2). Int J Cardiol 2021;336:38-44.
12. Tu S, Westra J, Yang J, von Birgelen C, Ferrara A, Pellicano M, et al. Diagnostic Accuracy of Fast Computational Approaches to Derive Fractional Flow Reserve From Diagnostic Coronary Angiography: The International Multicenter FAVOR Pilot Study. J Am Coll Cardiol Intv 2016;9:2024-35.
13. Masdjedi K, van Zandvoort LJC, Balbi MM, Gijsen FJH, Ligthart JMR, Rutten MCM, et al. Validation of a three-dimensional quantitative coronary angiography-based software to calculate fractional flow reserve: the FAST study. EuroIntervention 2020;16:591-9.
14. Koo BK, Erglis A, Doh JH, Daniels DV, Jegere S, Kim HS, et al. Diagnosis of ischemia-causing coronary stenoses by noninvasive fractional flow reserve computed from coronary computed tomographic angiograms. Results from the prospective multicenter DISCOVER-FLOW (Diagnosis of Ischemia-Causing Stenoses Obtained Via Noninvasive Fractional Flow Reserve) study. J Am Coll Cardiol 2011;58:1989-97.
15. Douglas PS, De Bruyne B, Pontone G, Patel MR, Norgaard BL, Byrne RA, et al. 1-Year Outcomes of FFRCT-Guided Care in Patients With Suspected Coronary Disease: The PLATFORM Study. J Am Coll Cardiol 2016;68:435-45.
16. Xu B, Tu S, Qiao S, Qu X, Chen Y, Yang J, et al. Diagnostic Accuracy of Angiography-Based Quantitative Flow Ratio Measurements for Online Assessment of Coronary Stenosis. J Am Coll Cardiol 2017;70:3077-87.
17. Xu B, Tu S, Song L, Jin Z, Yu B, Fu G, et al. Angiographic quantitative flow ratio-guided coronary intervention (FAVOR III China): a multicentre, randomised, sham-controlled trial. Lancet 2021;398:2149-59.
18. Masdjedi K, Tanaka N, Van Belle E, Porouchani S, Linke A, Woitek FJ, et al. Vessel fractional flow reserve (vFFR) for the assessment of stenosis severity: the FAST II study. EuroIntervention 2022;17:1498-505.
19. Fearon WF, Achenbach S, Engstrom T, Assali A, Shlofmitz R, Jeremias A, et al. Accuracy of Fractional Flow Reserve Derived From Coronary Angiography. Circulation 2019;139:477-84.
20. Westra J, Tu S, Winther S, Nissen L, Vestergaard MB, Andersen BK, et al. Evaluation of Coronary Artery Stenosis by Quantitative Flow Ratio During Invasive Coronary Angiography:

The WIFI II Study (Wire-Free Functional Imaging II). Circ Cardiovasc Imaging 2018;11:e007107.

21. Westra J, Andersen BK, Campo G, Matsuo H, Koltowski L, Eftekhari A, et al. Diagnostic Performance of In-Procedure Angiography-Derived Quantitative Flow Reserve Compared to Pressure-Derived Fractional Flow Reserve: The FAVOR II Europe-Japan Study. J Am Heart Assoc 2018;7.
22. Nørgaard BL, Leipsic J, Gaur S, Seneviratne S, Ko BS, Ito H, et al. Diagnostic performance of noninvasive fractional flow reserve derived from coronary computed tomography angiography in suspected coronary artery disease: the NXT trial (Analysis of Coronary Blood Flow Using CT Angiography: Next Steps). J Am Coll Cardiol 2014;63:1145-55.
23. Curzen N, Nicholas Z, Stuart B, Wilding S, Hill K, Shambrook J, et al. Fractional flow reserve derived from computed tomography coronary angiography in the assessment and management of stable chest pain: the FORECAST randomized trial. Eur Heart J 2021;42:3844-52.
24. Scoccia A, Tomaniak M, Neleman T, Groenland FTW, Plantes A, Daemen J. Angiography-Based Fractional Flow Reserve: State of the Art. Curr Cardiol Rep 2022;24:667-78.
25. Cho H, Lee JG, Kang SJ, Kim WJ, Choi SY, Ko J, et al. Angiography-Based Machine Learning for Predicting Fractional Flow Reserve in Intermediate Coronary Artery Lesions. J Am Heart Assoc 2019;8:e011685.
26. Cha JJ, Nguyen NL, Tran C, Shin WY, Lee SG, Lee YJ, et al. Assessment of fractional flow reserve in intermediate coronary stenosis using optical coherence tomography-based machine learning. Front Cardiovasc Med 2023;10:1082214.
27. Scarsini R, Fezzi S, Leone AM, De Maria GL, Pighi M, Marcoli M, et al. Functional Patterns of Coronary Disease: Diffuse, Focal, and Serial Lesions. J Am Coll Cardiol Intv 2022;15:2174-91.
28. Douglas PS, Nanna MG, Kelsey MD, Yow E, Mark DB, Patel MR, et al. Comparison of an Initial Risk-Based Testing Strategy vs Usual Testing in Stable Symptomatic Patients With Suspected Coronary Artery Disease: The PRECISE Randomized Clinical Trial. JAMA Cardiol 2023;8:904-14.
29. Taylor CA, Fonte TA, Min JK. Computational fluid dynamics applied to cardiac computed tomography for noninvasive quantification of fractional flow reserve: scientific basis. J Am Coll Cardiol 2013;61:2233-41.
30. Tesche C, De Cecco CN, Albrecht MH, Duguay TM, Bayer RR, 2nd, Litwin SE, et al. Coronary CT Angiography-derived Fractional Flow Reserve. Radiology 2017;285:17-33.

31. Park SJ, Kang SJ, Ahn JM, Shim EB, Kim YT, Yun SC, et al. Visual-functional mismatch between coronary angiography and fractional flow reserve. J Am Coll Cardiol Intv 2012;5:1029-36.

Abstract in Korean

관상동맥질환 환자에서 머신 러닝을 통한 관상동맥 조영술 기반 분획 혈류 예비력 측정 모델 수립

배경: 협심증을 호소하는 환자에게서 관상동맥의 중등도 협착이 관찰된다면, 압력 철선 기반의 분획 혈류 예비력을 평가하여 관상동맥 중재술 등의 치료 방침을 결정하도록 최신의 가이드라인에서 권고하고 있다. 그러나 실제 임상에서는 시간, 비용 및 안전상의 문제로 분획 혈류 예비력 검사 시행이 저조한 실정이다. 이러한 배경에서, 비침습적으로 이런 분획 혈류 예비력을 예측하기 위한 연구들도 많이 이루어지고 있다. 본 연구는 정량적 관상동맥 조영술과 환자의 임상 정보를 이용하여 비침습적으로 분획 혈류 예비력을 예측할 수 있는 머신 러닝 모델 (ML-Angio-FFR)을 수립하는 것을 목표로 하고 있다.

방법: ML-Angio-FFR 모델은 1,459명의 협심증 환자의 2,439개의 중등도 관상동맥 병변을 가지고 머신 러닝 기반의 랜덤 포레스트 알고리즘으로 훈련되었다. 첫 분석시에는 42개의 특성들이 포함되었다. 훈련 세트와 테스트 세트는 4:1 비율로 나누어졌다. 분획 혈류 예비력 0.80 이하의 병변은 허혈을 유발하는 병변으로, 0.80 초과는 중재술을 연기하고 약물 치료를 먼저 시행하는 병변으로 간주하고, 0.80 초과를 기준으로 모델 성능 테스트를 수행하였다.

결과: 환자의 평균 나이는 67.1세였고, 1,004명은 (68.8%) 남성이었다. 좌전하행지 1,481개 (60.7%), 좌회선지 533개 (21.9%), 우관상동맥 425개 (17.4%)의 병변이 포함되었다. 8개의 주요 특징을 사용하여 최종 ML-Angio-FFR 모델이 개발되었고, 테스트 세트에서 압력 철선 기반 분획 혈류 예비력과의 강한 상관 관계를 보였다 (r=0.7828, p<0.001). 또한 모델 성능 평가에서 진단 정확도는 87%로 확인되었고, 수신자 조작 특성 곡선의 곡선 아래 면적은 0.89로 나타났다. 외부 자료 검증에서도 준수한 상관 관계 및 진단 정확도를 보였으며 (r=0.4940, p<0.001; 62%), 수신자 조작 특성 곡선의 곡선 아래 면적은 0.73으로 나타났다.

결론: ML-Angio-FFR 모델은 기존의 압력 철선 기반 분획 혈류 예비력과 높은 상관관계를 보이며 허혈을 유발하지 않는 병변을 예측하는 것에 준수한 진단 성능을 나타낸다.

핵심되는 말 : 분획 혈류 예비력, 머신 러닝, 정량적 관상동맥 조영술