



Published in final edited form as:

*J Nucl Cardiol.* 2022 October ; 29(5): 2393–2403. doi:10.1007/s12350-022-03012-6.

## Machine Learning to Predict Abnormal Myocardial Perfusion from Pre-test Features

Robert J.H. Miller, MD<sup>a,b</sup>, M. Timothy Hauser, MD<sup>c</sup>, Tali Sharir, MD<sup>d</sup>, Andrew J. Einstein, MD, PhD<sup>e</sup>, Mathews B. Fish, MD<sup>f</sup>, Terrence D. Ruddy, MD<sup>g</sup>, Philipp A. Kaufmann, MD<sup>h</sup>, Albert J. Sinusas, MD<sup>i</sup>, Edward J. Miller, MD, PhD<sup>j</sup>, Timothy M. Bateman, MD<sup>j</sup>, Sharmila Dorbala, MD, MPH<sup>k</sup>, Marcelo Di Carli, MD<sup>k</sup>, Cathleen Huang, BA<sup>a</sup>, Joanna X. Liang, BA<sup>a</sup>, Donghee Han, MD<sup>a</sup>, Damini Dey, PhD<sup>a</sup>, Daniel S. Berman, MD<sup>a</sup>, Piotr J. Slomka, PhD<sup>a,\*</sup>

<sup>a</sup>Departments of Medicine (Division of Artificial Intelligence in Medicine), Imaging and Biomedical Sciences, Cedars-Sinai Medical Center, Los Angeles, California, United States

<sup>b</sup>Department of Cardiac Sciences, University of Calgary and Libin Cardiovascular Institute, Calgary, Alberta, Canada

<sup>c</sup>Section of Nuclear Cardiology, Department of Clinical Imaging, Oklahoma Heart Hospital, Oklahoma City, Oklahoma, United States

<sup>d</sup>Department of Nuclear Cardiology, Assuta Medical Centers, Tel Aviv, and Ben Gurion University of the Negev, Beer Sheba, Israel

<sup>e</sup>Division of Cardiology, Department of Medicine and Department of Radiology, Columbia University Irving Medical Center and New York-Presbyterian Hospital, New York, New York, United States

<sup>f</sup>Oregon Heart and Vascular Institute, Sacred Heart Medical Center, Springfield, Oregon, United States

<sup>g</sup>Division of Cardiology, University of Ottawa Heart Institute, Ottawa, Ontario, Canada

<sup>h</sup>Department of Nuclear Medicine, Cardiac Imaging, University Hospital Zurich, Zurich, Switzerland

<sup>i</sup>Section of Cardiovascular Medicine, Department of Internal Medicine, Yale University School of Medicine, New Haven, Connecticut, United States

<sup>j</sup>Cardiovascular Imaging Technologies LLC, Kansas City, Missouri, United States

<sup>k</sup>Department of Radiology, Division of Nuclear Medicine and Molecular Imaging, Brigham and Women's Hospital, Boston, Massachusetts, United States

### Abstract

**Background:** Accurately predicting which patients will have abnormal perfusion on MPI based on pre-test clinical information may help physicians make test selection decisions. We developed

\*Corresponding author: **Piotr Slomka, PhD, FACC**, Cedars-Sinai Medical Center, Department of Medicine (Division of Artificial Intelligence in Medicine), 8700 Beverly Blvd, Suite Metro 203, Los Angeles, CA 90048, Phone: 310-423-4348 Fax: 310-423-0173, Piotr.Slomka@cshs.org.

and validated a machine learning (ML) model for predicting abnormal perfusion using pre-test features.

**Methods:** We included consecutive patients who underwent SPECT MPI, with 20,418 patients from a multi-center (5 sites) international registry in the training population and 9,019 patients (from 2 separate sites) in the external testing population. The ML (extreme gradient boosting) model utilized 30 pre-test features to predict the presence of abnormal myocardial perfusion by expert visual interpretation.

**Results:** In external testing, the ML model had higher prediction performance for abnormal perfusion (area under receiver-operating characteristic curve [AUC] 0.762, 95% CI 0.750 – 0.774) compared to the clinical CAD consortium (AUC 0.689) basic CAD consortium (AUC 0.657), and updated Diamond-Forrester models (AUC 0.658,  $p < 0.001$  for all). Calibration (validation of the continuous risk prediction) was superior for the ML model (Brier score 0.149) compared to the other models (Brier score 0.165 to 0.198, all  $p < 0.001$ ).

**Conclusions:** ML can predict abnormal myocardial perfusion using readily available pre-test information. This model could be used to help guide physician decisions regarding non-invasive test selection.

### Keywords

Myocardial perfusion imaging; artificial intelligence; machine learning; CAD; PET; SPECT; Image Analysis

---

## INTRODUCTION

Myocardial perfusion imaging (MPI) is frequently used to diagnose or risk stratify patients with known or suspected coronary artery disease (CAD) (1–3). Abnormal regional myocardial perfusion can be used to detect obstructive CAD with high diagnostic accuracy (4). Additionally, the presence of abnormal regional myocardial perfusion can be used to identify a group of patients with a higher risk of major adverse cardiovascular events (5). As a result of these findings, the volume of SPECT MPI has grown to 15–20 million scans performed annually worldwide (6). However, the prevalence of abnormal perfusion has been decreasing since SPECT MPI was first implemented clinically, dropping from 40.9% in 1991 to 8.7% in 2009 (7). As a result, it is becoming increasingly important to ensure appropriate patient selection. This is particularly relevant since patients with normal regional myocardial perfusion on SPECT MPI may be more effectively risk stratified by using measures such as absolute myocardial blood flow (8, 9) or coronary artery calcification (10, 11). Additionally, patients with a low likelihood of abnormal perfusion may be candidates for stress-first imaging (and evaluated for stress only imaging) to reduce radiation exposure (12). However, SPECT (or PET) MPI could be particularly useful in patients with a high risk of ischemia since this may help target more aggressive therapies. For example, ischemia can be used to predict symptom benefit from revascularization (13) (14). However, to date, most pre-test risk models have focused on a patient's likelihood of having anatomically defined obstructive CAD rather than abnormal regional myocardial perfusion.

Most pre-test risk prediction models have also been developed using statistical regression methods (15). Machine learning (ML) has the potential to surpass these traditional methods by efficiently integrating a large number of variables. ML may also identify non-linear relationships and higher-order interactions between variables (16), which is not possible with traditional statistical methods. This approach has been used to predict risk of major adverse cardiovascular events and to identify patients for stress-only imaging who have a low probability of obstructive CAD (17) or adverse events (18). In this work, we propose a ML score to predict likelihood of abnormal perfusion solely from pre-test patient factors (without imaging information). The model was trained using patients from the multi-center, international REgistry of Fast Myocardial Perfusion Imaging with NExt generation SPECT (REFINE SPECT)(19), then tested using data from two external sites.

## MATERIALS AND METHODS

### Study Populations

The internal testing population from included consecutive patients from 5 sites undergoing SPECT MPI between 2009 and 2014 (n=20,418) as previously described (19). The external testing population included 9,019 consecutive patients from the University of Calgary (n=2,985) and Oklahoma Heart Hospital (n=6,034). The study protocol complied with the Declaration of Helsinki and was approved by the institutional review boards at each participating institution. The overall study was approved by the institutional review board at Cedars-Sinai Medical Center.

### Clinical Data

Clinical data was collected using site-specific mechanisms as previously described. Clinical indications for testing were recorded at the time of imaging. Clinical data included: age, sex, body mass index, past medical history, symptoms, and family history of CAD. Chest pain symptoms were classified as typical chest pain or atypical chest pain using standard definitions. Past medical history included: diabetes mellitus, hypertension, dyslipidemia, peripheral vascular disease, myocardial infarction (MI), percutaneous coronary intervention (PCI), coronary artery bypass grafting, and history of smoking. History of CAD was defined as previous myocardial infarction or previous revascularization(20).

### Imaging protocols and interpretation

Imaging protocols for the REFINE SPECT population have been described previously (19). Experienced cardiologists interpreted perfusion with access to clinical history, rest and stress hemodynamic and ECG findings, and all available imaging data at the time of clinical imaging. Visual subjective interpretation of stress perfusion was assessed with summed stress scores (SSS) using the 17-segment American Heart Association model(21), or overall reader interpretation.

De-identified image datasets were transferred to the core laboratory (Cedars-Sinai Medical Center) where automated quantitation of SPECT MPI studies was performed by experienced technologists (19). Myocardial contours were generated automatically with Quantitative Perfusion SPECT (QPS) /Quantitative Gated SPECT (QGS) software (Cedars-Sinai Medical

Center, Los Angeles, CA). Myocardial perfusion was quantified by total perfusion deficit (TPD) which incorporates severity and extent of perfusion abnormalities and is more reproducible compared to visual ischemia scoring (22, 23).

## Outcomes

Abnormal perfusion was defined as  $SSS > 3$  if segmental scoring was available and as reader interpretation of abnormal or definitely abnormal if not (24). All studies in the external population were assessed using SSS, with  $SSS > 3$  defined as abnormal (24). We also performed a sensitivity analysis, where abnormal perfusion was defined as stress TPD  $> 5\%$ .

## Machine Learning

ML considered 30 pre-test variables as outlined in Supplemental Table 1. The variables were categorized as demographic, past medical history, indication, ECG, and vital signs. Resting ECG findings were included, since those could be determined prior to testing, but stress ECG findings were not considered. The overall proportion of missing values within the dataset was 0.35%. Missing values were imputed with the population's median value for numerical variables, or with a distinct 'missing' label for categorical variables. For the model development, we utilized the currently leading ML method, extreme gradient boosting approach—XGBoost (25) implemented as *xgboost* package (version 0.82.1). XGBoost is a leading ML method which combines multiple decision trees as weak predictors in an ensemble learning method and has been shown to provide improved prediction performance compared to alternative methods (26). Variable importance was assessed using the built-in XGBoost gain metric, which is a measure of how important a particular feature is for the overall prediction performance.

## Internal Training and Testing

ML models were trained with stratified 10-fold cross-validation to separate data into training and testing samples (27, 28). The whole dataset was randomly split into 10 folds with a similar proportion of patients with abnormal perfusion in each fold. During internal testing, 8 folds (80% of the population) were used for testing, one-fold (10% of the population) for validation, and one-fold (10%) used for testing. The process was repeated 10 times, with a different fold used for testing and validation in each iteration. Thus, only unseen data was used for testing each model and a total of 10 models were assessed. The advantages of the 10-fold cross-validation over single split-sample approach are well documented in the ML literature and include reducing variance in prediction error, related to arbitrarily splitting data, leading to a more accurate estimate of model performance and maximizing the data for both training and validation, without overfitting or overlap between test and validation data (27).

## Comparison Risk Models

We compared prediction performance for the ML model to the CAD consortium basic (basic CAD) and clinical (clinical CAD) models (15) as well as the updated Diamond-Forrester (DF) model (29). The basic CAD model considers age, sex, and presenting symptoms (15).

The clinical CAD model considers age, sex, presenting symptoms, diabetes, hypertension, dyslipidemia, and smoking (15). Both of the CAD consortium models were developed in a population of 5,677 patients who underwent coronary computed tomography angiography or invasive angiography and were designed to predict obstructive CAD. The updated DF model, a contemporary modification of the original DF model(30), considers age, sex, and type of chest pain and was modified using a cohort of 2260 patients with chest pain referred for invasive coronary angiography to ensure better calibration in current populations (29). These models were all developed for use in patients with suspected CAD.

### External testing:

During external testing, the ML model was trained using all patients from the internal population. The model was then tested using patients from the two external sites.

### Statistical Analyses

Categorical variables were summarized as number (proportion) and compared with a Chi-square or Fisher Exact test as appropriate. Continuous variables were summarized as mean (standard deviation [SD]) if normally distributed and median (interquartile range [IQR]) otherwise. Receiver operating characteristic (ROC) curves for identifying patients with abnormal perfusion. Area under the ROC curve (AUC) was used to compare models using the method described by DeLong et al. (31). Model calibration was assessed with calibration graphs and Brier scores (32, 33). Cost estimates for cost per patient identified with abnormal perfusion were based on CPT re-imburement code 78451 (34). All statistical tests were two-sided with a p-value <0.05 considered significant. Analyses were performed with Stata version 14.2 and R 4.1.2.

## RESULTS

Details of the internal and external populations are shown in Table 1. As expected, there were significant differences between the two populations. In particular, patients in the external population were older and less likely to have previous CAD. An outline of the internal and external testing procedures is shown in Figure 1.

### Internal Testing

Of the 20418 patients included in the internal population, 4106 (20.1%) patients were classified as having abnormal perfusion. Prediction performance for abnormal perfusion was higher for ML (AUC 0.829, 95% CI 0.822 – 0.836) compared to the CAD clinical (AUC 0.716, 95% CI 0.708 – 0.725), CAD basic (AUC 0.699, 95% CI 0.691 – 0.708) and updated DF model (AUC 0.700, 95% CI 0.691 – 0.708, all p<0.001) (results in Figure 2). We performed a separate analysis where existing clinical scores were included as ML features, with no significant difference in ML prediction performance (AUC 0.829, 95% CI 0.822 – 0.836).

The calibration graph for the ML model is shown in Figure 3. The Brier score for ML (0.119) was superior compared to the other models (0.149 to 0.184, all p<0.001). The optimal threshold, determined by Youden index, was 0.196 (74% sensitive, 77% specific).

However, we also identified a highly sensitive (95% sensitive) threshold ( $>0.0731$ ,  $n = 6221$  [30.5%] below threshold, 37% specific; 97% negative predictive value [NPV]) and an ultra-sensitive (99% sensitive) threshold ( $>0.0452$ ,  $n = 2643$  [12.9%] below threshold, 16% specific; 98% NPV) for abnormal perfusion. The top 25 variables by feature importance (a measure of how important a variable is for the model predictions) from the internal cross-validation are shown in Figure 4.

### External Testing

Of the 9,019 patients included in the external testing population, 2,003 patients (22.2%) were classified as having abnormal perfusion. Prediction performance for abnormal perfusion was higher for ML (AUC 0.762, 95% CI 0.750 – 0.774) compared to the CAD clinical (AUC 0.689, 95% CI 0.677 – 0.702), CAD basic (AUC 0.657, 95% CI 0.643 – 0.670) and updated DF model (AUC 0.658, 95% CI 0.645 – 0.671, all  $p < 0.001$ ) (Figure 2).

Model calibration in the external testing population is shown in Figure 3. All models demonstrated modest calibration. The Brier score for ML (0.149) was superior compared to other models (Brier score 0.165 to 0.198, all  $p < 0.001$ ). The optimal threshold from the internal testing population ( $>0.196$ ,  $n = 2,394$ ) was 56% sensitive, 82% specific. The highly sensitive threshold ( $>0.0731$ ,  $n = 3089$  [34.2%] below threshold) had sensitivity 89% and specificity 41% (NPV 93%). The cost per patient with abnormal perfusion identified for patients below the highly sensitive threshold would be \$16,300, compared to \$4,007 above the cut-off. The ultra-sensitive threshold ( $>0.0452$ ,  $n = 1399$  [15.5%] below threshold) had sensitivity 96% and specificity 19% (NPV 95%). The cost per patient with abnormal perfusion identified for patients below the ultra-sensitive threshold would be \$22,997, compared to \$4,738 above the cut-off.

### Sensitivity Analyses

The ML model had higher prediction performance compared to traditional models in patients with a history of CAD (Supplemental Figure 1) and in patients without a history of CAD (Supplemental Figure 2). Additionally, the ML model had higher performance for predicting abnormal quantitative perfusion (stress TPD  $> 5\%$ , Supplemental Figure 3).

## DISCUSSION

We developed a ML model to predict abnormal perfusion on MPI using traditional pre-test features and evaluated it with both internal cross-validation and external testing. We demonstrated that the ML model was able to incorporate pre-test clinical information to improve prediction performance for abnormal regional myocardial perfusion compared to existing risk models. All models demonstrated modest calibration, but this was superior for the ML method compared to traditional methods. These results suggest the ML could be used to predict likelihood of abnormal perfusion on MPI using pre-test features to may help guide physicians when choosing diagnostic tests. Physicians could potentially utilize these models to identify high-risk patients in whom extent of myocardial ischemia may help guide aggressive therapies(13, 14, 35, 36). Alternatively, patients with a low likelihood of



abnormal MPI who have already been sent for testing could be scheduled for stress-first imaging.

We found that discrimination for abnormal perfusion was higher with ML compared to existing risk scores. This was an expected finding, since ML can objectively integrate a larger number of features to classify patients compared to existing prediction models. Additionally, existing risk models have primarily been developed to predict presence of obstructive CAD(15), which may not necessarily be predictive of perfusion abnormalities. In another study, Juarez-Orozco et al. developed a ML model to predict the presence of abnormal myocardial perfusion (37), but also considered functional imaging variables which were not utilized in the present model (since they would only be available after MPI was performed). During internal cross-validation, the model had an AUC of 0.72 (37). Rouhani et al. developed a conventional statistical model for identifying patients with normal MPI from pre-test features, demonstrating an AUC of 0.68 in the derivation cohort (38). External validation is a critical step for evaluating the generalizability of a model and typically demonstrates lower prediction performance (39) as was seen in our study. Assessing model calibration (40), a measure of how well predicted probability matches actual probability, is also important in external populations since this can significantly influence the clinical impact of a risk model. For example, a model could have excellent discrimination (when assessed as a continuous measure) but identify all of the patients as high-risk and therefore have low clinical utility. In our study, the ML model had better calibration compared to existing models in both internal and external testing.

The results of our study may also help physicians better understand which clinical predictors of abnormal perfusion are most important. Features such as prior CAD and exercise stress had high feature importance (a measure of how important a variable was for the model predictions) but are not considered in any of the comparison risk scores (which were developed for patients with suspected CAD) (15). We also confirmed the significance of other features, like typical chest pain (41), which are considered in existing models. While we included features with partially redundant information (for example, known CAD, previous MI, and previous revascularization) ML algorithms are inherently capable of handling interactions between these variables to maximize prediction performance (16). Lastly, out of a possible 23 test indications we identified that suspected angina, troponin elevation, and cardiomyopathy had the highest feature importance. The clinical information required for the ML model is routinely available from electronic medical records, which could be utilized to enable automatic risk predictions when ordering tests. This task could potentially be further simplified by considering ML models with reduced features (42), and integrating multiple imputation methods for handling missing values (43).

The proposed ML model could potentially be used by physicians to help guide test selection. Current guidelines suggest that physicians decide between testing options based on patient likelihood of obstructive CAD (1–3), but suggest that anatomic or functional imaging are appropriate in intermediate risk patients. Our model can help guide physicians to choosing SPECT MPI for patients where it is most likely to change management. In patients with a low probability of abnormal perfusion, subsequent management is more likely to change in response to information regarding absolute myocardial blood flow (8, 9) or anatomic

testing (10, 11). From a health care system perspective, the cost required to identify a patient with abnormal perfusion is significantly higher in patients with a low probability of abnormal perfusion and some of those patients could potentially be managed without further imaging (1–3). Nuclear cardiology laboratories could potentially use this tool to aid in patient triaging or to select patients already referred for SPECT MPI for stress-first, and potentially stress-only, imaging to reduce health care costs and radiation exposure. Lastly, physicians could potentially use the ML predictions, which provides a robust estimate of pre-test probability, when adjudicating equivocal MPI findings.

Our study has a few important limitations. The score was derived and tested in patients undergoing SPECT MPI and separate studies would be required to determine if the scores can also predict probability of abnormal myocardial in patients undergoing PET MPI or stress echocardiography. Additionally, while ML demonstrated superior discrimination compared to other risk score the overall prediction performance was poor for all models in patients with known CAD. This suggests that alternative approaches to prediction may be helpful in this important sub-group. Lastly, future studies are needed to determine if using these risk scores to guide diagnostic test selection improved clinical outcomes or reduced costs.

## CONCLUSIONS

ML demonstrated high prediction performance for abnormal myocardial perfusion using readily available pre-test information. This model could be used to help guide physician decisions regarding testing strategies.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## FUNDING

This research was supported in part by grant R01HL089765 from the National Heart, Lung, and Blood Institute/ National Institutes of Health (NHLBI/NIH) (PI: Piotr Slomka). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

## DISCLOSURES

Dr. Miller has received research support from Pfizer. Drs. Berman and Slomka participate in software royalties for QPS software at Cedars-Sinai Medical Center. Dr. Slomka has received research grant support from Siemens Medical Systems. Drs. Berman, Dorbala, Einstein, and Edward Miller have served as consultants for GE Healthcare. Dr. Einstein has served as a consultant to W. L. Gore & Associates. Dr. Dorbala has served as a consultant to Bracco Diagnostics; her institution has received grant support from Astellas. Dr. Di Carli has received research grant support from Spectrum Dynamics and consulting honoraria from Sanofi and GE Healthcare. Dr. Ruddy has received research grant support from GE Healthcare and Advanced Accelerator Applications. Dr. Einstein's institution has received research support from Attralus, GE Healthcare, Eidos Therapeutics, Pfizer, Roche Medical Systems, and W. L. Gore & Associates. The remaining authors have nothing to disclose.

## ABBREVIATIONS

<b>AUC</b>	area under the receiver operating characteristic curve
<b>CAC</b>	coronary artery calcium



<b>CAD</b>	coronary artery disease
<b>CI</b>	confidence interval
<b>ECG</b>	electrocardiogram
<b>MI</b>	myocardial infarction
<b>ML</b>	machine learning
<b>MPI</b>	myocardial perfusion imaging
<b>PCI</b>	percutaneous coronary intervention
<b>PET</b>	positron emission tomography
<b>REFINE SPECT</b>	REgistry of Fast Myocardial Perfusion Imaging with NExt generation SPECT
<b>SPECT</b>	single photon emission computed tomography
<b>SSS</b>	summed stress score
<b>TPD</b>	total perfusion deficit

## REFERENCES

- (1). Gulati M, Levy PD, Mukherjee D, Amsterdam E, Bhatt DL, Birtcher KK et al. 2021 AHA/ACC/ASE/CHEST/SAEM/SCCT/SCMR Guideline for the Evaluation and Diagnosis of Chest Pain: Executive Summary. *J Am Coll Cardiol*;0.
- (2). Knuuti J, Wijns W, Saraste A, Capodanno D, Barbato E, Funck-Brentano C et al. 2019 ESC Guidelines for the diagnosis and management of chronic coronary syndromes. *Eur Heart J* 2020;41:407–77. [PubMed: 31504439]
- (3). Gulati M, Levy PD, Mukherjee D, Amsterdam E, Bhatt DL, Birtcher KK et al. 2021 AHA/ACC/ASE/CHEST/SAEM/SCCT/SCMR Guideline for the Evaluation and Diagnosis of Chest Pain. *J Am Coll Cardiol* 2021;78:e187–e285. [PubMed: 34756653]
- (4). Nudi F, Iskandrian AE, Schillaci O, Peruzzi M, Frati G, Biondi-Zoccai G. Diagnostic Accuracy of Myocardial Perfusion Imaging With CZT Technology: Systemic Review and Meta-Analysis of Comparison With Invasive Coronary Angiography. *JACC Cardiovasc Imaging* 2017;10:787–94. [PubMed: 28330657]
- (5). Otaki Y, Betancur J, Sharir T, Hu L-H, Gransar H, Liang Joanna X et al. 5-Year Prognostic Value of Quantitative Versus Visual MPI in Subtle Perfusion Defects. *JACC Cardiovasc Imaging* 2020;13:774–85. [PubMed: 31202740]
- (6). Einstein AJ. Multiple opportunities to reduce radiation dose from myocardial perfusion imaging; 2013. p. 649–51.
- (7). Rozanski A, Gransar H, Hayes SW, Min J, Friedman JD, Thomson LE et al. Temporal trends in the frequency of inducible myocardial ischemia during cardiac stress testing: 1991 to 2009. *J Am Coll Cardiol* 2013;61:1054–65. [PubMed: 23473411]
- (8). Yoshinaga K, Manabe O, Tamaki N. Absolute quantification of myocardial blood flow. *Journal of nuclear cardiology : official publication of the American Society of Nuclear Cardiology* 2018;25:635–51. [PubMed: 27444500]
- (9). Murthy VL, Bateman TM, Beanlands RS, Berman DS, Borges-Neto S, Chareonthaitawee P et al. Clinical Quantification of Myocardial Blood Flow Using PET: Joint Position Paper of the SNMMI Cardiovascular Council and the ASNC. *Journal of nuclear cardiology :*

- official publication of the American Society of Nuclear Cardiology 2018;25:269–97. [PubMed: 29243073]
- (10). Danad I, Szymonifka J, Twisk JWR, Norgaard BL, Zarins CK, Knaapen P et al. Diagnostic performance of cardiac imaging methods to diagnose ischaemia-causing coronary artery disease when directly compared with fractional flow reserve as a reference standard: a meta-analysis. *Eur Heart J* 2017;38:991–8. [PubMed: 27141095]
  - (11). Rozanski A, Gransar H, Shaw LJ, Kim J, Miranda-Peats L, Wong ND et al. Impact of coronary artery calcium scanning on coronary risk factors and downstream testing the EISNER (Early Identification of Subclinical Atherosclerosis by Noninvasive Imaging Research) prospective randomized trial. *J Am Coll Cardiol* 2011;57:1622–32. [PubMed: 21439754]
  - (12). Einstein AJ, Johnson LL, DeLuca AJ, Kontak AC, Groves DW, Stant J et al. Radiation dose and prognosis of ultra-low-dose stress-first myocardial perfusion SPECT in patients with chest pain using a high-efficiency camera. *Journal of nuclear medicine : official publication, Society of Nuclear Medicine* 2015;56:545–51. [PubMed: 25745089]
  - (13). Spertus JA, Jones PG, Maron DJ, O'Brien SM, Reynolds HR, Rosenberg Y et al. Health-Status Outcomes with Invasive or Conservative Care in Coronary Disease. *N Engl J Med* 2020;382:1408–19. [PubMed: 32227753]
  - (14). Al-Lamee RK, Shun-Shin MJ, Howard JP, Nowbar AN, Rajkumar C, Thompson D et al. Dobutamine Stress Echocardiography Ischemia as a Predictor of the Placebo-Controlled Efficacy of Percutaneous Coronary Intervention in Stable Coronary Artery Disease. *Circulation* 2019;140:1971–80. [PubMed: 31707827]
  - (15). Genders TS, Steyerberg EW, Hunink MG, Nieman K, Galema TW, Mollet NR et al. Prediction model to estimate presence of coronary artery disease: retrospective pooled analysis of existing cohorts. *BMJ* 2012;344:e3485. [PubMed: 22692650]
  - (16). Ryo M, Rillig MC. Statistically reinforced machine learning for nonlinear patterns and variable interactions. *Ecosphere* 2017;8:e01976.
  - (17). Hu LH, Miller RJH, Sharir T, Commandeur F, Rios R, Einstein AJ et al. Prognostically safe stress-only single-photon emission computed tomography myocardial perfusion imaging guided by machine learning: report from REFINE SPECT. *European heart journal Cardiovascular Imaging* 2021;22:705–14. [PubMed: 32533137]
  - (18). Eisenberg E, Miller RJH, Hu LH, Rios R, Betancur J, Azadani P et al. Diagnostic safety of a machine learning-based automatic patient selection algorithm for stress-only myocardial perfusion SPECT. *J Nucl Cardiol* 2021.
  - (19). Slomka PJ, Betancur J, Liang JX, Otaki Y, Hu LH, Sharir T et al. Rationale and design of the REgistry of Fast Myocardial Perfusion Imaging with NExt generation SPECT (REFINE SPECT). *Journal of nuclear cardiology : official publication of the American Society of Nuclear Cardiology* 2020;27:1010–21. [PubMed: 29923104]
  - (20). Miller RJH, Klein E, Gransar H, Slomka PJ, Friedman JD, Hayes S et al. Prognostic significance of previous myocardial infarction and previous revascularization in patients undergoing SPECT MPI. *Int J Cardiol* 2020;313:9–15. [PubMed: 32349938]
  - (21). Cerqueira MD, Weissman NJ, Dilsizian V, Jacobs AK, Kaul S, Laskey WK et al. Standardized myocardial segmentation and nomenclature for tomographic imaging of the heart. A statement for healthcare professionals from the Cardiac Imaging Committee of the Council on Clinical Cardiology of the American Heart Association. *Circulation* 2002;105:539–42. [PubMed: 11815441]
  - (22). Slomka PJ, Nishina H, Berman DS, Akincioglu C, Abidov A, Friedman JD et al. Automated quantification of myocardial perfusion SPECT using simplified normal limits. *Journal of Nuclear Cardiology* 2005;12:66–77. [PubMed: 15682367]
  - (23). Xu Y, Hayes S, Ali I, Ruddy TD, Wells RG, Berman DS et al. Automatic and visual reproducibility of perfusion and function measures for myocardial perfusion SPECT. *J Nucl Cardiol* 2010;17:1050–7. [PubMed: 20963537]
  - (24). Otaki Y, Betancur J, Sharir T, Hu LH, Gransar H, Liang JX et al. 5-Year Prognostic Value of Quantitative Versus Visual MPI in Subtle Perfusion Defects: Results From REFINE SPECT. *JACC Cardiovascular imaging* 2020;13:774–85. [PubMed: 31202740]

- (25). Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '16; 2016. p. 785–94.
- (26). Pieszko K, Slomka PJ. Assessing Performance of Machine Learning. *JAMA Cardiology* 2021;6:1465-. [PubMed: 34586335]
- (27). Molinaro AM, Simon R, Pfeiffer RM. Prediction error estimation: a comparison of resampling methods. *Bioinformatics* 2005;21:3301–7. [PubMed: 15905277]
- (28). Betancur J, Otaki Y, Motwani M, Fish MB, Lemley M, Dey D et al. Prognostic Value of Combined Clinical and Myocardial Perfusion Imaging Data Using Machine Learning. *J Am Coll Cardiol Img* 2018;11:1000–9.
- (29). 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. *Eur Heart J* 2011;32:1316–30. [PubMed: 21367834]
- (30). Diamond GA, Forrester JS. Analysis of probability as an aid in the clinical diagnosis of coronary-artery disease. *N Engl J Med* 1979;300:1350–8. [PubMed: 440357]
- (31). DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;44:837–45. [PubMed: 3203132]
- (32). Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology* 2010;21:128–38. [PubMed: 20010215]
- (33). Brier GW. Verification of forecasts expressed in terms of probability. *Monthly weather review* 1950;78:1–3.
- (34). Services. HOPPSfMM. Medicare-Fee-for-Service-Paymen; 2022.
- (35). Azadani PN, Miller RJH, Sharir T, Diniz MA, Hu LH, Otaki Y et al. Impact of Early Revascularization on Major Adverse Cardiovascular Events in Relation to Automatically Quantified Ischemia. *JACC Cardiovascular imaging* 2021;14:644–53. [PubMed: 32828784]
- (36). Sharir T, Hollander I, Hemo B, Tsamir J, Yefremov N, Bojko A et al. Survival benefit of coronary revascularization after myocardial perfusion SPECT: The role of ischemia. *Journal of nuclear cardiology : official publication of the American Society of Nuclear Cardiology* 2021;28:1676–87. [PubMed: 31823328]
- (37). Juarez-Orozco LE, Knol RJJ, Sanchez-Catasus CA, Martinez-Manzanera O, van der Zant FM, Knuuti J. Machine learning in the integration of simple variables for identifying patients with myocardial ischemia. *Journal of nuclear cardiology : official publication of the American Society of Nuclear Cardiology* 2020;27:147–55. [PubMed: 29790017]
- (38). Rouhani S, Shahrani AA, Hossain A, Yam Y, Wells RG, deKemp RA et al. A Clinical Tool to Identify Candidates for Stress-First Myocardial Perfusion Imaging. *JACC Cardiovasc Imaging* 2020;13:2193–202. [PubMed: 32563652]
- (39). Hijazi W, Leslie W, Filipchuk N, Choo R, Wilton SB, James M et al. External Validation of the CRAX2MACE score. *J Nucl Cardiol* 2022;Epub ahead of print.
- (40). Van Calster B, McLernon DJ, van Smeden M, Wynants L, Steyerberg EW, Topic Group 'Evaluating diagnostic t et al. Calibration: the Achilles heel of predictive analytics. *BMC Med* 2019;17:230. [PubMed: 31842878]
- (41). Rozanski A, Miller RJH, Han D, Gransar H, Slomka P, Dey D et al. The prevalence and predictors of inducible myocardial ischemia among patients referred for radionuclide stress testing. *Journal of nuclear cardiology : official publication of the American Society of Nuclear Cardiology* 2021.
- (42). Rios R, Miller RJH, Hu LH, Otaki Y, Singh A, Diniz M et al. Determining a minimum set of variables for machine learning cardiovascular event prediction: results from REFINE SPECT registry. *Cardiovasc Res* 2021.
- (43). Rios R, Miller RJH, Manral N, Sharir T, Einstein AJ, Fish MB et al. Handling missing values in machine learning to predict patient-specific risk of adverse cardiac events: Insights from REFINE SPECT registry. *Computers in Biology and Medicine* 2022:105449. [PubMed: 35381453]

- (44). Nakao YM, Miyamoto Y, Higashi M, Noguchi T, Ohishi M, Kubota I et al. Sex differences in impact of coronary artery calcification to predict coronary artery disease. *Heart* 2018;104:1118–24. [PubMed: 29331986]
- (45). Grandhi GR, Mszar R, Cainzos-Achirica M, Rajan T, Latif MA, Bittencourt MS et al. Coronary Calcium to Rule Out Obstructive Coronary Artery Disease in Patients With Acute Chest Pain. *JACC Cardiovasc Imaging* 2021.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**NEW KNOWLEDGE GAINED**

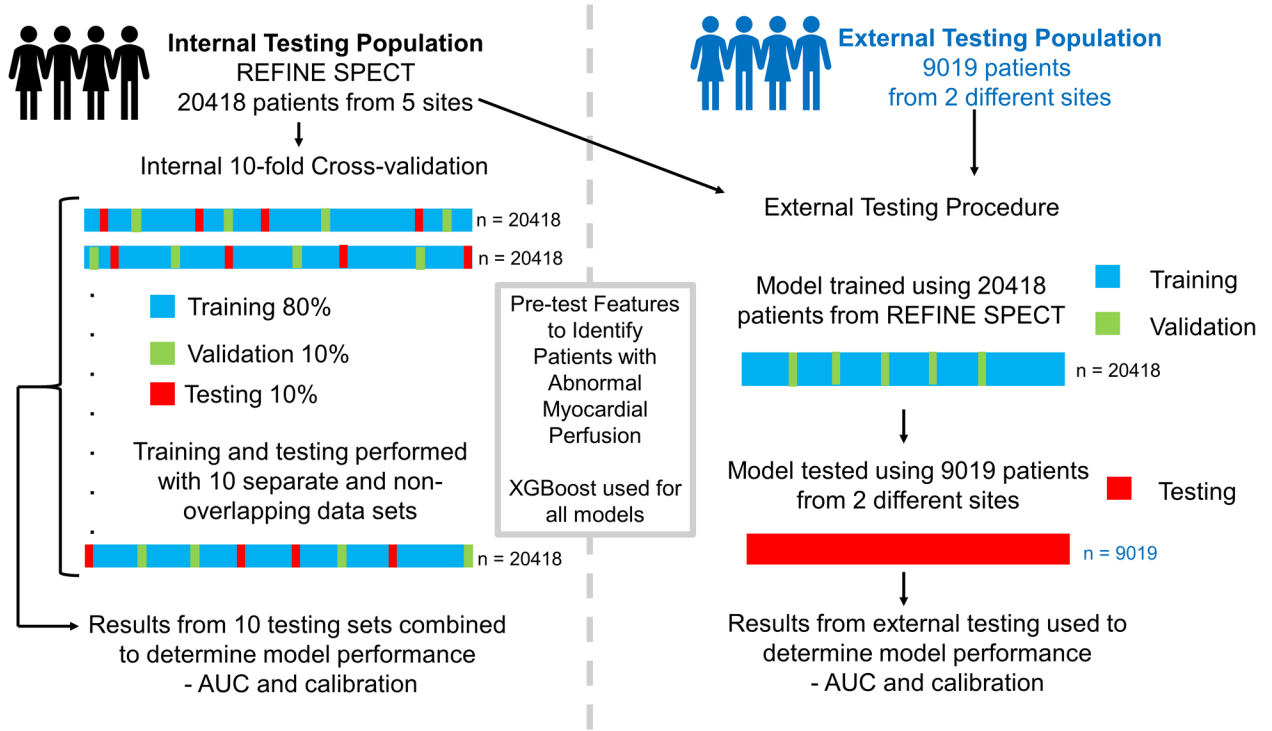
Using external testing, machine learning identified patients who are more likely to have abnormal myocardial perfusion using readily available clinical information. The model demonstrated higher prediction performance and better calibration compared to traditional risk models.

Author Manuscript

Author Manuscript

Author Manuscript

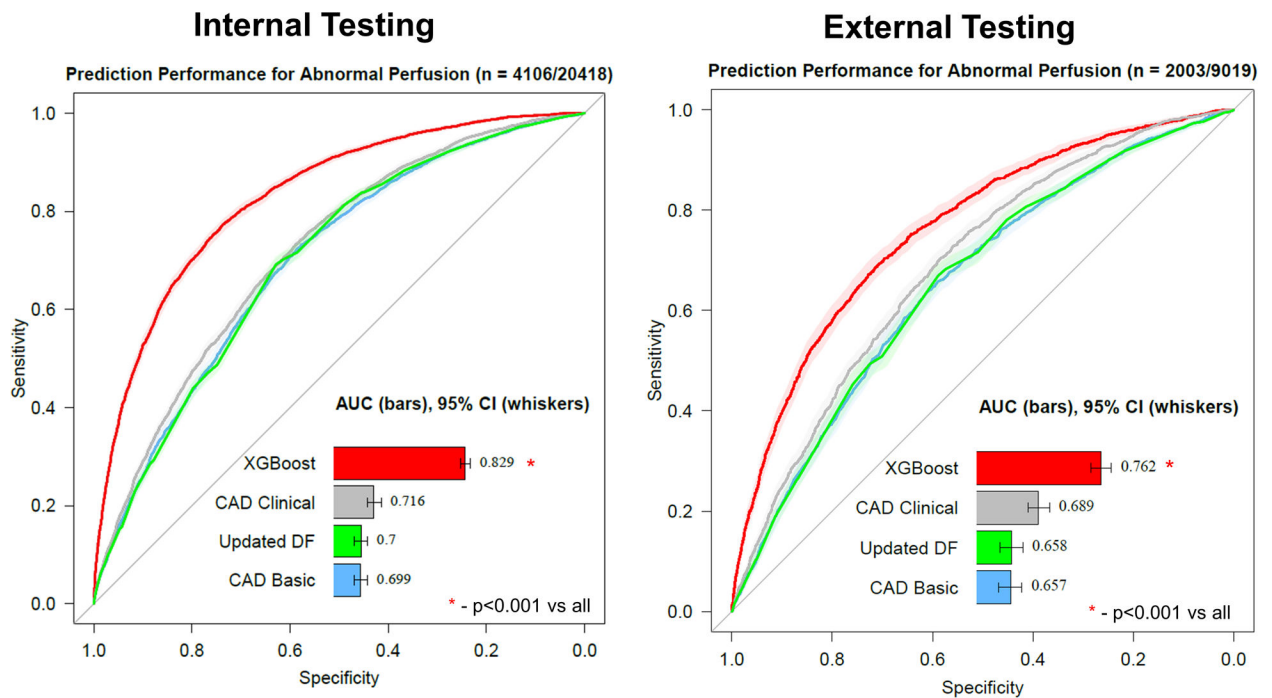
Author Manuscript



**Figure 1: Internal and External Testing Procedures**

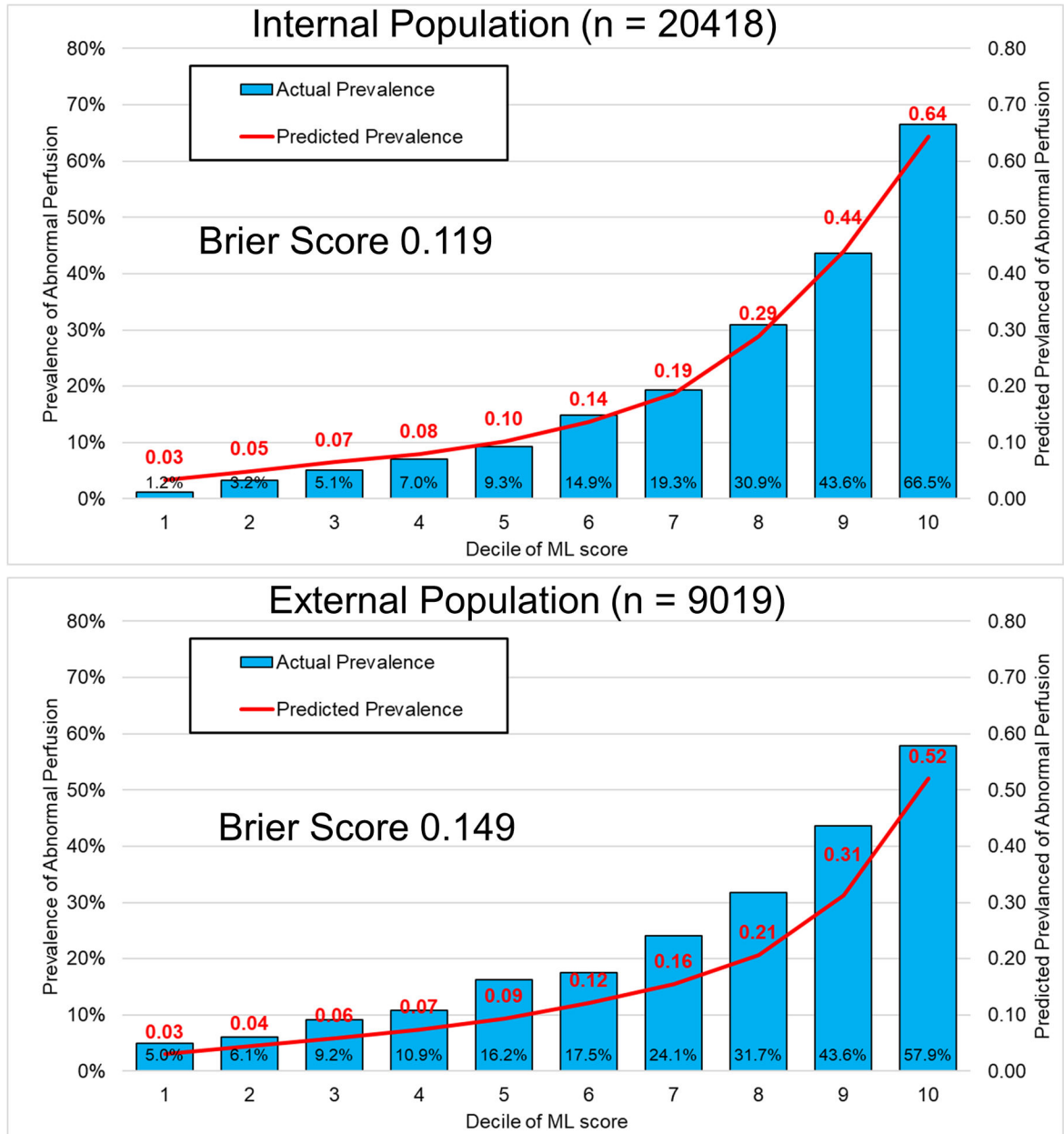
Summary of the internal and external testing procedures. Internal testing (left) was performed using 10-fold cross-validation where the internal testing population was randomly split into ten folds. In each of 10 iterations, 8 folds are used for training, one-fold for validation and one-fold for testing, with a different fold used for testing in each of 10 iterations. The results from the ten different testing sets are combined to provide an estimate of model performance. In external testing (right), model training and validation is performed with patients from REFINE SPECT and tested in two external sites.





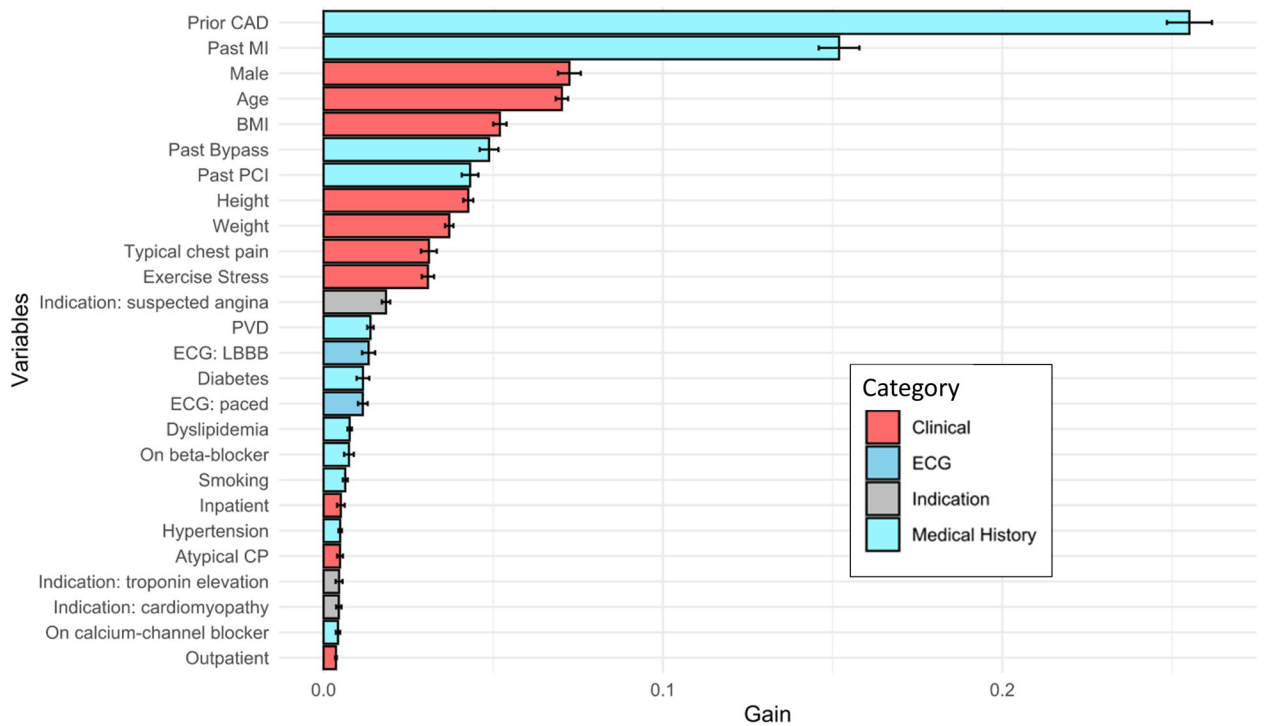
**Figure 2: Prediction performance**

Prediction performance for abnormal perfusion from pre-test information in internal testing (left) and external testing (right). The machine learning model (XGBoost) had higher prediction performance compared to the coronary artery disease (CAD) consortium basic or clinical models as well as the updated Diamond-Forrester (DF) model (all  $p < 0.001$ ). AUC – area under the receiver operating characteristic curve, CI – confidence interval.



**Figure 3: Model Calibration**

Calibration graph showing the predicted proportion of patients with abnormal perfusion and actual proportion of patients with abnormal perfusion by decile of machine learning (ML) score. The model calibration for the internal population (top) was higher compared to the external population (bottom), where there was a tendency to underestimate prevalence of abnormal perfusion.



**Figure 4: Feature Importance**

Feature importance ranked according to feature gain. Higher feature gain means that the variable was more useful for categorizing patients as having normal or abnormal perfusion. Prior coronary artery disease (CAD) included a history of previous myocardial infarction (MI), percutaneous coronary intervention (PCI), or coronary artery bypass grafting. BMI – body mass index, CP – chest pain, ECG – electrocardiogram, LBBB – left bundle branch block, PVD – peripheral vascular disease.

**Table 1:**

## Baseline population characteristics

	<b>Internal Population (N=20418)</b>	<b>External Population (N=9019)</b>	<b>p-value</b>
Age, median (IQR)	64 (56, 73)	68 (60, 75)	<0.001
Male, n (%)	11642 (57.0%)	4871 (54.0%)	<0.001
Typical chest pain, n (%)	1224 (6.0%)	133 (1.5%)	<0.001
Atypical chest pain, n (%)	4589 (22.5%)	3208 (35.6%)	<0.001
Outpatient, n (%)	18417 (90.2%)	8169 (90.6%)	0.326
Inpatient, n (%)	1535 (7.5%)	832 (9.2%)	<0.001
Emergency department, n (%)	466 (2.3%)	18 (0.2%)	<0.001
Body mass index, median (IQR)	27.3 (24.6, 30.9)	29.4 (25.8, 33.3)	<0.001
Past Medical History, n (%)			
Hypertension	12920 (63.3%)	5736 (63.6%)	0.600
Dyslipidemia	12903 (63.2%)	1979 (21.9%)	<0.001
Diabetes mellitus	5212 (25.5%)	2726 (30.2%)	<0.001
Peripheral vascular disease	2420 (11.9%)	2834 (31.4%)	<0.001
Prior CAD	5796 (28.4%)	1606 (17.8%)	<0.001
Prior myocardial infarction	2768 (13.6%)	677 (7.5%)	<0.001
Prior PCI	3967 (19.4%)	839 (9.3%)	<0.001
Smoking, n (%)	3875 (19.0%)	2841 (31.5%)	<0.001
Family History, n (%)	5642 (27.6%)	4720 (52.3%)	<0.001
Exercise stress, n (%)	9457 (46.3%)	5318 (59.0%)	<0.001
Abnormal perfusion, n (%)	4106 (20.1%)	2003 (22.2%)	<0.001

Population characteristics for the internal and external populations. CAD – coronary artery disease, IQR – interquartile range, PCI – percutaneous coronary intervention.