



# Diagnostic safety of a machine learning-based automatic patient selection algorithm for stress-only myocardial perfusion SPECT

Evann Eisenberg, MD,<sup>a</sup> Robert J. H. Miller, MD,<sup>a,b</sup> Lien-Hsin Hu, MD,<sup>a,c</sup> Richard Rios, PhD,<sup>a</sup> Julian Betancur, PhD,<sup>a</sup> Peyman Azadani, MD,<sup>a</sup> Donghee Han, MD,<sup>a</sup> Tali Sharir, MD,<sup>d</sup> Andrew J. Einstein, MD, PhD,<sup>e</sup> Sabahat Bokhari, MD,<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>i</sup> Timothy M. Bateman, MD,<sup>j</sup> Sharmila Dorbala, MD, MPH,<sup>k</sup> Marcelo Di Carli, MD,<sup>k</sup> Joanna X. Liang, BA,<sup>a</sup> Yuka Otaki, MD, PhD,<sup>a</sup> Balaji K. Tamarappoo, MD, PhD,<sup>a</sup> Damini Dey, PhD,<sup>a</sup> Daniel S. Berman, MD,<sup>a</sup> and Piotr J. Slomka, PhD<sup>a</sup>

<sup>a</sup> Cedars-Sinai Medical Center, Los Angeles, CA

<sup>b</sup> University of Calgary, Calgary, AB, Canada

<sup>c</sup> Taipei Veterans General Hospital, Taipei, Taiwan

<sup>d</sup> Assuta Medical Centers, Tel Aviv, Israel

<sup>e</sup> Columbia University Irving Medical Center and New York-Presbyterian Hospital, New York, NY

<sup>f</sup> Sacred Heart Medical Center, Springfield, OR

<sup>g</sup> University of Ottawa Heart Institute, Ottawa, ON, Canada

<sup>h</sup> University Hospital Zurich, Zurich, Switzerland

<sup>i</sup> Yale University School of Medicine, New Haven, CT

<sup>j</sup> Cardiovascular Imaging Technologies LLC, Kansas City, MO

<sup>k</sup> Brigham and Women's Hospital, Boston, MA

Received May 17, 2021; accepted May 18, 2021

doi:10.1007/s12350-021-02698-4

**Background.** Stress-only myocardial perfusion imaging (MPI) markedly reduces radiation dose, scanning time, and cost. We developed an automated clinical algorithm to safely cancel unnecessary rest imaging with high sensitivity for obstructive coronary artery disease (CAD).

**Methods and Results.** Patients without known CAD undergoing both MPI and invasive coronary angiography from REFINE SPECT were studied. A machine learning score (MLS) for prediction of obstructive CAD was generated using stress-only MPI and pre-test clinical variables. An MLS threshold with a pre-defined sensitivity of 95% was applied to the automated patient selection algorithm. Obstructive CAD was present in 1309/2079 (63%) patients.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s12350-021-02698-4>.

This article includes a PowerPoint file that should be made available as ESM on SpringerLink.

The authors of this article have provided a PowerPoint file, available for download at SpringerLink, which summarizes the contents of the paper and is free for re-use at meetings and presentations. Search for the article DOI on SpringerLink.com.

Evann Eisenberg and Robert J. H. Miller have contributed equally as first author.

**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. The work was supported in part by the Dr. Miriam and Sheldon Adelson Medical Research Foundation.

Reprint requests: Piotr J. Slomka, PhD, Cedars-Sinai Medical Center, 8700 Beverly Boulevard, Ste. Metro 203, Los Angeles, CA 90048; [slomkap@cshs.org](mailto:slomkap@cshs.org)

1071-3581/\$34.00

Copyright © 2021 American Society of Nuclear Cardiology.

**MLS had higher area under the receiver operator characteristic curve (AUC) for prediction of CAD than reader diagnosis and TPD (0.84 vs 0.70 vs 0.78,  $P < .01$ ). An MLS threshold of 0.29 had superior sensitivity than reader diagnosis and TPD for obstructive CAD (95% vs 87% vs 87%,  $P < .01$ ) and high-risk CAD, defined as stenosis of the left main, proximal left anterior descending, or triple-vessel CAD (sensitivity 96% vs 89% vs 90%,  $P < .01$ ).**

**Conclusions.** The MLS is highly sensitive for prediction of both obstructive and high-risk CAD from stress-only MPI and can be applied to a stress-first protocol for automatic cancellation of unnecessary rest imaging. (J Nucl Cardiol 2022;29:2295–307.)

**Key Words:** SPECT-MPI • Machine learning • Stress-only • Coronary artery disease • Nuclear stress testing

#### Abbreviations

CAD	Coronary artery disease
ICA	Invasive coronary angiography
LAD	Left anterior descending
LCx	Left circumflex artery
LMCA	Left main coronary artery
MLS	Machine learning score
MPI	Myocardial perfusion imaging
SPECT	Single-photon emission computed tomography
SSS	Summed stress score
TPD	Total perfusion deficit

**See related editorial, pp. 2308–2310**

## INTRODUCTION

Single-photon emission computed tomography (SPECT) myocardial perfusion imaging (MPI) is the most utilized non-invasive imaging study for patients with known or suspected coronary artery disease (CAD),<sup>1</sup> with an estimated 15 million tests performed annually worldwide.<sup>2,3</sup> However, widespread use has led to concerns regarding the relatively high effective radiation dose associated with SPECT-MPI. In 2015, it was reported that a standard rest-stress MPI study carries an effective radiation dose of more than 10 mSv,<sup>2,4,5</sup> and contributed to over 10% of cumulative radiation from medical procedures in the United States.<sup>2,3</sup> Additional drawbacks include long procedure times and the subjectivity of visual interpretation for perfusion abnormalities.

Currently, most SPECT-MPI protocols utilize two separate image acquisitions, one with patient at stress (exercise or pharmacological) and another after rest. However, when stress images are normal, rest imaging provides no additional clinical value.<sup>5–8</sup> There are data from clinical trials, including over 20,000 patients, confirming that the prognosis of a normal stress-only MPI is no different than a normal rest-stress MPI.<sup>8</sup> Critically, eliminating unnecessary rest imaging reduces radiation exposure by up to 75% and laboratory time by up to 70%.<sup>5,6,8,9</sup> In a stress-first MPI protocol, stress

images are acquired and, when normal, rest scanning can be canceled, resulting in a “stress-only” procedure. However, successful implementation of stress-only MPI requires timely review of stress imaging which poses a significant problem for busy physicians. For this reason, stress-only MPI is severely underutilized in United States.<sup>2–4</sup> There is an unmet need for a simple and efficient method to cancel unnecessary rest imaging without increased work for the interpreting physician.

A machine learning approach that combines clinical and image data has been shown to improve both the diagnostic and prognostic value of SPECT-MPI results.<sup>10,11</sup> In this multicenter study we developed a high-sensitivity machine learning score (MLS) for prediction of obstructive CAD using stress-only MPI data and pre-test clinical variables for use in a stress-first MPI protocol. We recently demonstrated that the MLS could identify patients for stress-only imaging at low risk of major adverse cardiovascular events (MACE).<sup>12</sup> This novel application of machine learning can improve the utilization of stress-only MPI by automatically determining the need for the additional rest MPI imaging.

## METHODS

### Study Population

The REgistry of Fast Myocardial Perfusion Imaging with NExt generation SPECT (REFINE SPECT) is a multicenter, international collaborative database of both imaging and clinical variables.<sup>13</sup> The REFINE SPECT diagnostic registry includes 2079 patients from 9 centers without known CAD who underwent both SPECT-MPI and invasive coronary angiography (ICA) within 180 days between 2009 and 2014. Patients with a history of CAD, determined by a physician at each clinical site, were excluded. History of CAD was defined as either previous myocardial infarction or previous revascularization.<sup>14</sup> Pre-defined clinical variables as well as deidentified MPI data and ICA correlations were transferred to a single core laboratory (Cedars-Sinai Medical Center). The study was approved by the institutional review boards at each

participating institution and the overall study was approved by the institutional review board at Cedars-Sinai Medical Center.

## Clinical Data

Clinical information was obtained at each site from patients' medical records. Prior to collection, a homogenous classification of each "pre-test" clinical variable was defined to assure standardization among sites. Stress-test results, including patient symptoms and ECG data, were interpreted by the performing physician at each clinical site. Clinical and imaging parameters included in REFINE SPECT are shown in Supplemental Table 1.<sup>13</sup> Only stress imaging variables and pre-test clinical information collected as part of a standard SPECT-MPI protocol were considered for automated MLS development.

## Image Acquisition

Details of image acquisition and processing are available in the supplement.

## Invasive Coronary Angiography

ICA was performed according to standard clinical protocols within 180 days of MPI. The indication for ICA was determined by on-site cardiologists at each participating center. Follow-up for ICA results and revascularization details were collected from medical records (including all clinic visits, cardiology groups, insurance registries, and hospital visits) at each enrolling center. All coronary angiograms were visually interpreted by an on-site cardiologist. A luminal narrowing of  $\geq 70\%$  in the left anterior descending artery (LAD), left circumflex artery (LCx), or right coronary artery (RCA) or  $\geq 50\%$  stenosis of the left main coronary artery (LMCA) was considered significant. Dates of revascularization (if performed), revascularization type (percutaneous coronary intervention [PCI] or coronary artery bypass grafting [CABG]) and location (which artery) are also available.

## Visual Perfusion Analysis

Visual perfusion analysis was performed at each participating center at the time of clinical interpretation, with full details in the supplement.

## Machine Learning

The MLS was generated using automated variable selection by information gain ratio ranking and model building with a boosted ensemble algorithm which is then tested using a 10-fold cross-validation procedure. Machine Learning (ML) techniques were implemented in the open-source Waikato Environment for Knowledge Analysis (WEKA) platform 3.8.0

(University of Waikato, Hamilton, New Zealand).<sup>12</sup> A LogitBoost method was implemented in the WEKA platform for training and validation.

## Variable Selection

A combination of imaging variables, stress-test variables, and pre-test clinical variables were available for selection by the feature selection module (information gain ratio). The information gain ratio measures the effectiveness of each variable for building training models. Only variables that resulted in an information gain ratio  $> 0$  were included for model building.

## Model Building

Predictive classifiers for obstructive CAD were developed by an ensemble ("boosting") LogitBoost algorithm. The principle behind ML ensemble boosting is to combine the prediction of simple classifiers with weak performances to create a single strong classifier. These weak predictions are then combined in an ensemble (weighted majority voting) to derive an overall classifier, the MLS. More details of the model building process are available in the supplement.

## Cross-Validation

The performance and error estimation of the MLS (variable selection and LogitBoost) were assessed using a stratified 10-fold cross-validation procedure. With this approach, the total patient population is split into 10 equal groups, with similar prevalence of obstructive CAD, which act as a split-sample (train-test) method. An MLS is derived from each of the 90% (training cohort) and applied to the remaining 10% (testing cohort). This process is repeated 10 times, with different and non-overlapping data used for testing in each iteration, to ensure that the MLS is consistently evaluated on previously unseen data and not performed on the same population used for model building. These testing results are then concatenated.<sup>15,16</sup> This repeated procedure is done to obtain a more accurate and unbiased estimate of the diagnostic performance of the model. The repetition of testing removes the uncertainty associated with which cases are used for training and which are used for testing. The diagnostic accuracy the MLS in each individual testing cohort in Supplemental Figure 1.

## External Validation of the MLS

To assess the external validity of the proposed MLS, we performed an additional analysis in which 8 of the 9 sites were used to generate a new MLS which was then tested in the external held-out population.

## Missing Variables

For all variables with missing values, the default imputation was performed by the WEKA platform 3.8.1. These data were imputed with the population's mean value for continuous variables, or with a distinct 'missing' label for categorical variables. We also generated an MLS without variables with > 40% missing values to evaluate the impact of missing values, with results available in the supplement.

## Clinical Validation of Stress-Only MPI

The MLS was generated using stress-only MPI data but not all studies in the REFINE SPECT diagnostic registry were performed using a stress-first MPI protocol. Thus, an additional analysis was performed to determine its diagnostic performance for stress-first MPI using a clinical validation cohort. Details of the clinical validation procedure are available in the supplement.

## Statistical Analysis

Continuous variables are presented as mean  $\pm$  standard deviation (SD) or median ( $\pm$  interquartile range) as appropriate. Categorical variables are presented as numbers and proportions. The paired Student's *t* test (for normally distributed variables) or the Wilcoxon rank-sum test (for non-parametrically distributed variables) were used to compare differences between groups. Proportions were compared using Chi-Squared, McNemar's, or Friedman's Test, when appropriate. The predictive accuracy of the MLS, TPD, reader diagnosis and SSS for obstructive CAD were evaluated with a receiver operating characteristic (ROC) analysis and pairwise comparisons of AUC according to DeLong et al on a continuous scale.<sup>17</sup> For the clinical diagnostic algorithm, an ideal MLS would be highly sensitive for identifying obstructive and high-risk CAD from stress-only MPI data without increasing the frequency of unnecessary rest scans. Therefore, three thresholds for MLS were established which corresponded to sensitivities of 90%, 95%, and 99% for obstructive CAD based on ROC curve analysis. The calibration of the MLS was assessed with a calibration plot and the Brier score. A two-tailed *P* value < 0.05 was considered statistically significant. Statistical calculations were performed using R software version 3.4 and SPSS (IBM SPSS Software for Windows, Version 24.0).

## RESULTS

### Myocardial Perfusion Imaging

Baseline characteristics are presented in Table 1. MPI was performed using a D-SPECT scanner at four centers (*n* = 1170[56%]), GE Discovery NM 530c at three centers (*n* = 808[39%]), and GE Discovery NM/CT570c with attenuation correction at two centers (*n* = 101[5%]). A combination of stress and rest imaging

were performed in 2005 patients (96.4%), with a summary of image acquisition parameters in Table 2. The frequency of each missing variable is available in Supplemental Table 2.

Prior to stress imaging, patients underwent either symptom limited treadmill exercise testing without adjuvant pharmacologic stress (*n* = 805[39%]) or pharmacologic stress testing (1274[61%]) with radiotracer injection at peak exercise or during maximum hyperemia, respectively. Pharmacologic stress agents included adenosine (*n* = 185), dobutamine (*n* = 15), dipyridamole (*n* = 494), or regadenoson (*n* = 577). The administered stress activity was lower with stress-first imaging [*n* = 793(267  $\pm$  95 MBq or 7.24  $\pm$  2.6 mCi)] than when rest imaging was performed first [*n* = 1215(880  $\pm$  398 MBq or 23.8  $\pm$  10.8 mCi)] (*P* < .01). All other stress MPI results are shown in Table 3. Automatic contours were evaluated by an experienced technologist in all cases and corrected in 291 (14%) cases. Most changes were small and involved changes in the valve plane position. Two or more corrections were performed in a small proportion of the population (2.1%).

### Invasive Coronary Angiography

The average time interval between MPI and ICA was 28 days ( $\pm$  39). On ICA, obstructive CAD was identified in 1309 (63%) patients and involved 2423 coronary artery segments. Obstructive CAD involved the LAD in 73% of patients (*n* = 958), the LCx in 48% (*n* = 633), and the RCA in 53% (*n* = 689). High-risk CAD was present in 539 (26%) patients and was defined as LMCA stenosis (present in 143 patients), proximal LAD stenosis (present in 286 patients), or 3-vessel CAD (present in 223 patients).

### Analysis and Machine Learning Prediction

A summary of information gain ratio and variable selection is shown in Figure 1. AUC was calculated for prediction of obstructive CAD using the entire REFINE SPECT diagnostic cohort (*n* = 2079, prevalence = 63%). For prediction of any obstructive CAD, MLS had a superior AUC than both reader diagnosis and TPD (0.84 vs 0.70 vs 0.78, *P* < .01) (Figure 2). ROC results for 10 individual hold-out datasets for each of the 10 models are shown in Supplemental Figure 1. From the ROCs, three MLS thresholds were identified corresponding to pre-defined sensitivities of 90%, 95%, and 99%. The diagnostic properties of each MLS threshold were then compared to reader diagnosis, SSS, and TPD (Table 4).

All MLS thresholds had significantly higher sensitivity than TPD  $\geq$  1% or SSS > 0 (*P* < .01 for all). An MLS of 0.38 had a significantly greater specificity than

**Table 1.** Baseline characteristics

	<b>All patients (n = 2079)</b>	<b>With CAD (n = 1309)</b>	<b>Without CAD (n = 770)</b>	<b>P value</b>
Age	64.7 ± 11.2	65.7 ± 10.9	63 ± 11.6	< .01
BMI	28.6 ± 5.7	28.3 ± 5.1	29.1 ± 6.7	< .01
Male sex	1385 (66.5)	973 (74.3)	412 (53.4)	< .01
Hypertension	1422 (68.3)	918 (70.1)	504 (65.5)	.03
Hyperlipidemia	1297 (62.3)	849 (64.8)	448 (58.3)	< .01
Diabetes	610 (29.3)	412 (31.5)	198 (25.8)	< .01
Family history	694 (33.3)	447 (34.1)	247 (32.1)	.34
Peripheral vascular disease	123 (5.9)	92 (7.0)	31 (4.0)	< .01
Smoking	521 (25.0)	345 (26.3)	176 (22.9)	.08

Continuous variables reported as mean ± SD; categorical variables reported as n (%)  
BMI body mass index (kg/m<sup>2</sup>)

**Table 2.** SPECT-MPI acquisition

	<b>n (%)</b>
<i>Imaging acquisition</i>	
Stress-rest on same day	717 (34.5)
Rest-stress on same day	1221 (58.8)
Stress and rest on separate days	66 (3.2)
Stress-only	74 (3.6)
<i>Imaging protocol (stress)</i>	
2-position	1759 (84.6)
1-position NC	219 (10.5)
Supine + AC	101 (4.9)
<i>Imaging protocol (rest)</i>	
2-position	422 (20.2)
1-position NC	1483 (71.3)
Supine + AC	95 (4.6)
Stress gated	2077 (99.9)
Rest gated	1996 (96)

2-position (supine/upright or supine/prone)  
AC, attenuation corrected; NC, non-attenuation corrected

reader diagnosis or SSS but not TPD. An MLS of 0.29, which corresponded to sensitivity of 95% for obstructive CAD, had a sensitivity of > 95% for high-risk CAD. This threshold (0.29) had superior sensitivity for high-risk and obstructive CAD compared to an MLS of 0.38 ( $P < .01$ ). Although highly sensitive, an MLS threshold of 0.16 (corresponding to 99% sensitivity) did not identify significantly more cases with obstructive CAD than an MLS of 0.29 and had a significantly lower specificity. There was no difference in the negative

predictive value (NPV) for the MLS of 0.16 and 0.29 (0.80 vs 0.78, respectively) but both were significantly higher than the NPV for an MLS of 0.38 (0.73,  $P < .01$ ). An MLS of 0.29 was applied to the clinical algorithm. The MLS showed good calibration (Figure 3) with a Brier score of 0.16.

SSS were available in 1139 patients, 729 (64%) of whom had obstructive CAD. The ROC for prediction of obstructive CAD between SSS, TPD, and MLS are shown in Supplemental Figure 2.

Although developed for prediction of obstructive CAD, an MLS threshold of 0.29 also had superior sensitivity for detection of high-risk CAD than reader diagnosis or TPD (96% vs. 89% vs. 90%,  $P < .01$ ). The frequency of false-negative test results for obstructive and high-risk CAD by reader diagnosis, TPD, and MLS are shown in the Figure 4. The frequency of false-negative studies between SSS, TPD, and MLS are shown in Figure 5, in patients with available segmental scores. Of the 539 patients with high-risk CAD, the MLS was negative in 22 (4%). Of these, TPD was negative in 15 (68%) and by reader diagnosis in 14 (64%). When the stress images were reviewed as part of the internal validation cohort, 14 (64%) were considered normal.

The diagnostic properties of the MLS to identify stenosis  $\geq 50\%$  (Supplemental Figure 3), for prediction of future revascularization (Supplemental Results), for each center separately (Supplemental Figure 4), per vessel (Supplemental Figure 5), and when applied to different cameras and stress protocols (Supplemental Table 3) are available in the Supplemental Data.

**Table 3.** Stress and imaging results

Results	All patients (n = 2079)	With CAD (n = 1309)	Without CAD (n = 770)	P value
Resting HR (BPM)	72 ± 14	73 ± 14	71 ± 14	< .01
Stress HR (BPM)*	141 ± 20	139 ± 19	144 ± 22	< .01
Stress BP (systolic)*	166 ± 25	166 ± 24	167 ± 27	0.48
Stress BP (diastolic)*	79 ± 11	79 ± 11	79 ± 11	.61
% MPHR*	91 ± 11	90.5 ± 10.3	91.3 ± 11.0	.24
Exercise duration (min)*	7.6 ± 2.9	7.5 ± 2.8	8.0 ± 3.2	.02
SSS (N = 1139)	7.4 ± 6.7	9.5 ± 6.9	3.7 ± 4.3	< .01
TPD (N = 2079)	7.2 ± 8.8	9.8 ± 9.7	2.63 ± 4.1	< .01
MLS (N = 2079)	0.63 ± 0.27	0.75 ± 0.22	0.43 ± 0.21	< .01

Values presented in mean ± SD

BPM, beats per minute; BP, blood pressure (mmHg); CAD, coronary artery disease; HR, heart rate; MPHR, maximum predicted heart rate (%); SSS, summed stress score; TPD, total perfusion deficit; MLS, machine learning score

\*Stress HR, BP % MPHR and exercise duration shown for patients that underwent exercise stress testing (n = 805)

### External Validation Results

A new MLS was generated using the data from 8 sites (n = 1723) and then tested on the held-out site (n = 356). The new MLS was then compared to readers and TPD for prediction of obstructive CAD. This MLS had superior diagnostic accuracy compared to both readers and TPD for prediction of obstructive CAD (Supplemental Figure 6). Comparison of the MLS generated with and without variables with > 40% missing values are shown in Supplemental Figure 7.

### Automated Stress-First Clinical Diagnostic Algorithm

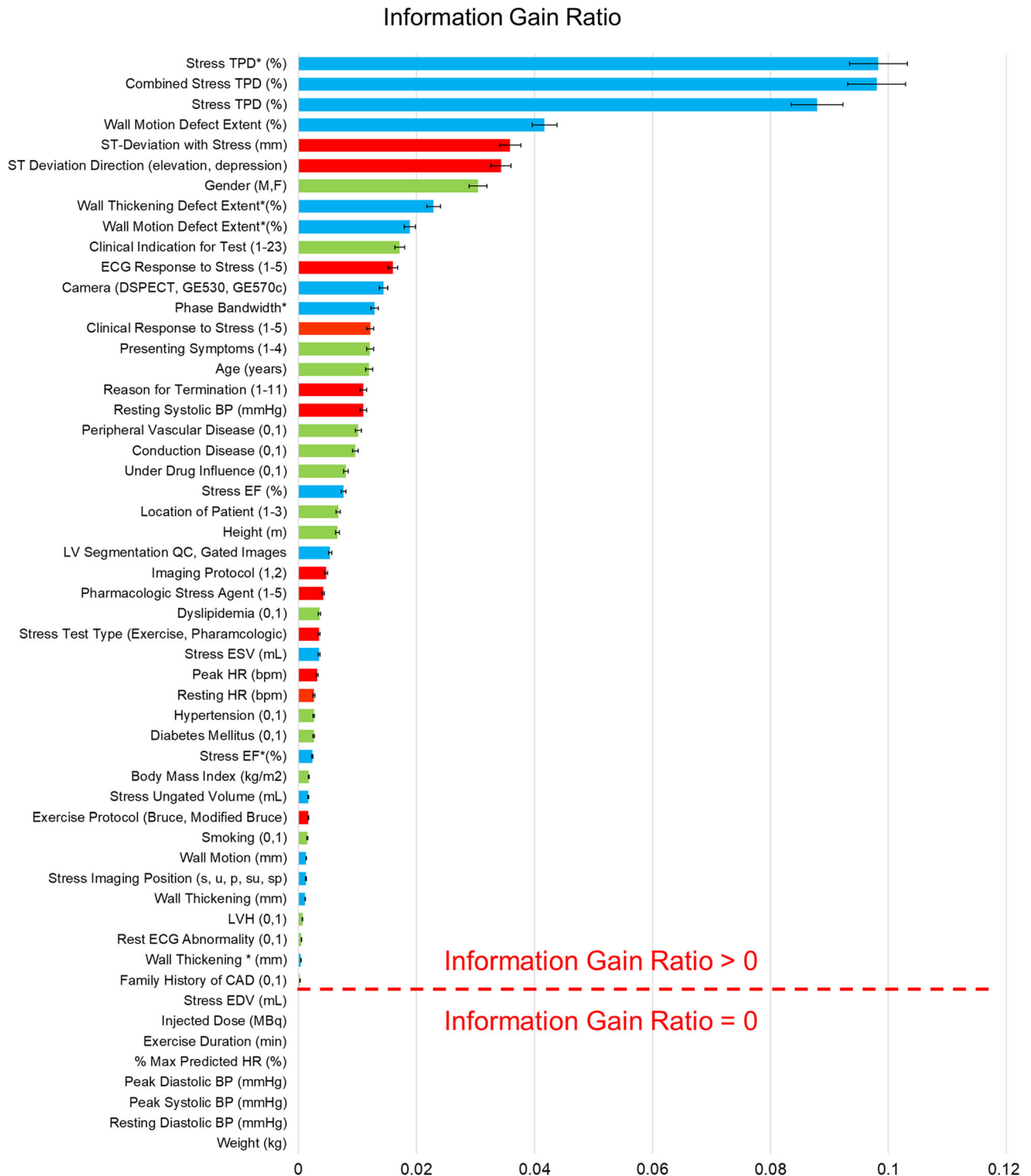
A clinical diagnostic algorithm for automated stress-first SPECT-MPI in Figure 6. After completion of stress testing, an MLS is automatically calculated from the stress MPI data and pre-test clinical variables. If the score is less than the decision threshold of 0.29, rest imaging can be omitted. If the MLS is greater than or equal to the applied threshold of 0.29, the patient should proceed with rest imaging.

### DISCUSSION

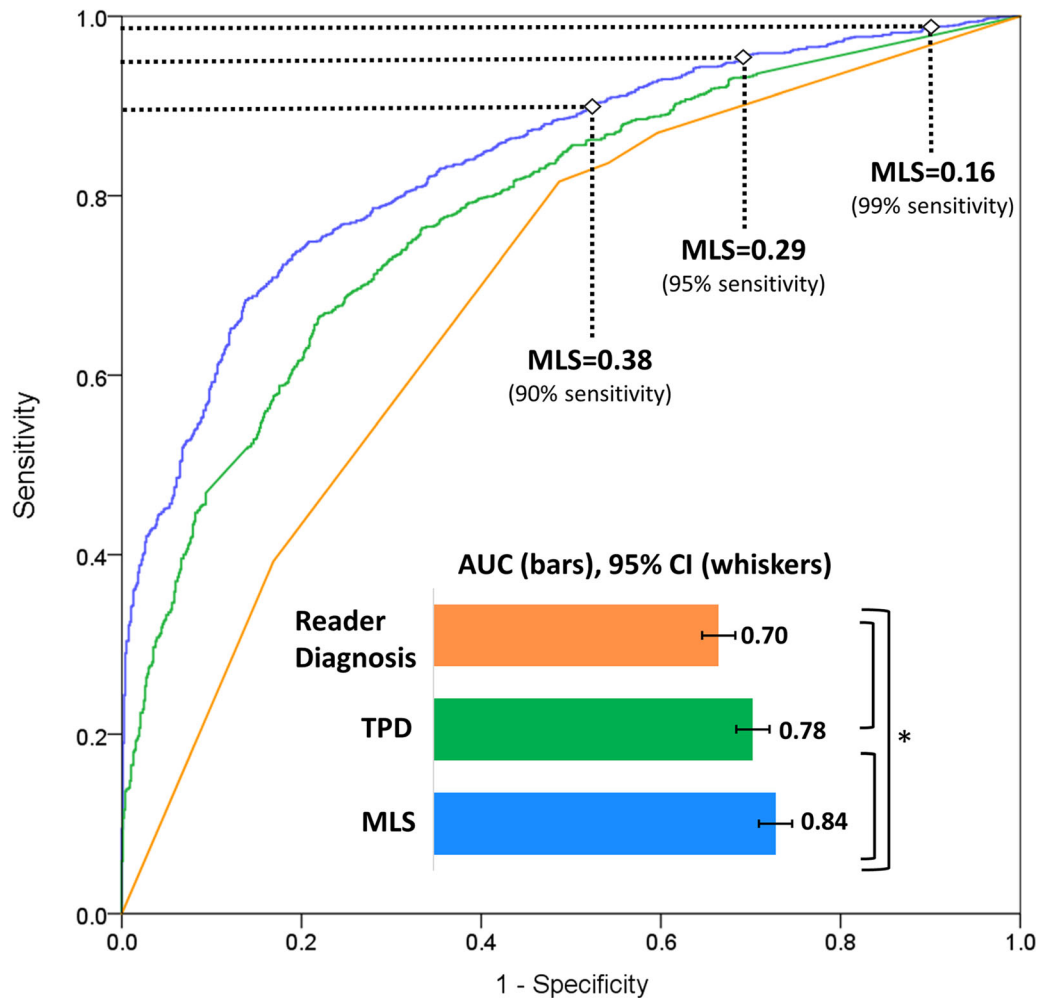
In this study, we applied machine learning to create a clinical diagnostic algorithm for stress-first MPI that automatically determines which patients are appropriate candidates for stress-only MPI, prior to physician review. An MLS threshold of 0.29, which is generated from stress-only MPI data and pre-test clinical variables demonstrated, superior diagnostic sensitivity for both

obstructive and high-risk CAD than reader diagnosis, and TPD. The MLS computation does not require additional clinical or imaging data collection beyond that of a standard SPECT-MPI protocol and can be readily incorporated into clinical laboratories. Trained on a large population, the MLS provides an adaptive model which can be applied to any patient population, omitting the need for pre-test patient selection. As shown in Figure 6, all clinical and imaging data, including the MLS are reviewed by the interpreting physician prior to finalizing a clinical report. This application of machine learning allows the physician to review all test results and, if necessary, the patient may return for additional rest imaging on a separate day.

The safety of stress-only MPI is well established,<sup>7,8,18</sup> but appropriate patient selection remains unclear. Current guidelines recommend performing stress-only MPI in patients with a “low pre-test probability of CAD”,<sup>1</sup> but do not provide a clinical algorithm for patient selection. Prior studies have proposed pre-test risk calculators that require additional clinical data collection prior to selecting the appropriate MPI protocol.<sup>9,19</sup> Duvall et al previously published pre-test clinical algorithms to predict which patients will have normal stress MPI results with over 80% accuracy.<sup>19</sup> These algorithms differ from the proposed MLS algorithm in that they use pre-test clinical variables to predict a successful stress-first MPI protocol, whereas the MLS uses pre-test clinical data, stress-test results and stress imaging data to produce a post-test clinical risk score. The MLS algorithm for automatic selection of patients for stress-only MPI who were at low risk of



**Figure 1.** Variable Selection. A combination of imaging (blue bar: 17 selected), stress-test (red bar: 12 selected) and clinical (green bar: 17 selected) variables ranked by their mean [95% CI] information gain ratio within 10-fold cross-validation for prediction of obstructive CAD. \*The alternative view (supine for D-SPECT or prone for Discovery).



**Figure 2.** Receiver operator characteristics for prediction of obstructive CAD in all patients (prevalence = 1309/2079 [63%]). \* $P < 0.01$  for AUC comparison by Delong Test. AUC, area under the receiver operating characteristic curve, CAD, coronary artery disease; CI, confidence interval; MLS, machine learning score; TPD, total perfusion deficit.

MACE was recently published.<sup>12</sup> In that study, a MLS was superior for prediction of all-cause mortality (N = 20,414, AUC: MLS 0.8 vs TPD 0.7 vs readers 0.68;  $P < 0.01$ ) and selected a similar proportion of patients for rest scan cancellation but with significantly lower all-cause mortality compared to clinical selection rules (1.3%, 1.2% and 1.0% vs 0.6%, 0.6% and 0.2%).<sup>12</sup> The combination of such robust prognostic prediction and the highly sensitive rule out of obstructive CAD could be a safe and efficient method for reducing unnecessary rest imaging in clinical laboratories

Stress-only imaging can greatly reduce patient radiation exposure while maintaining the robust prognostic value of a normal SPECT-MPI.<sup>2,3</sup> However, in 2015, it was estimated that stress-only MPI account for only 3% of studies in North America.<sup>3,4</sup> This can be

explained in part by the need for timely review of stress imaging which may be problematic for physicians with other clinical responsibilities. The ML algorithm provides a highly sensitive diagnostic tool that can automatically cancel unnecessary rest imaging and ultimately improve the utilization of stress-first MPI in busy cardiology practices. The primary endpoint of obstructive CAD on ICA was selected as most SPECT-MPI exams are performed clinically to rule out significant coronary artery stenosis. In the REFINE SPECT prognostic database (n = 20,414), the indication for exam was suspicion of obstructive CAD in more than 90% of cases.<sup>12</sup>

Although trained from the entire population, the diagnostic properties of the MLS were not different when generated from low-dose stress-first images



**Table 4.** Diagnostic testing characteristics for prediction of obstructive CAD from SPECT-MPI (n = 1309/2079[63%])

	<b>Reader diagnosis (%)</b>	<b>SSS &gt; 0 (%)</b>	<b>TPD ≥ 1% (%)</b>	<b>MLS ≥ 0.38 (%)</b>	<b>MLS ≥ 0.29 (%)</b>	<b>MLS ≥ 0.16 (%)</b>	<b>P value*</b>
Sensitivity† ‡§	87	88	87	90	95	99	< .01
Specificity† ‡§	40	38	45	48	31	7	< .01

SSS was available in 1139 patients (prevalence 729/1139 [64%]). For reader diagnosis, a score of 0 (normal) was considered negative

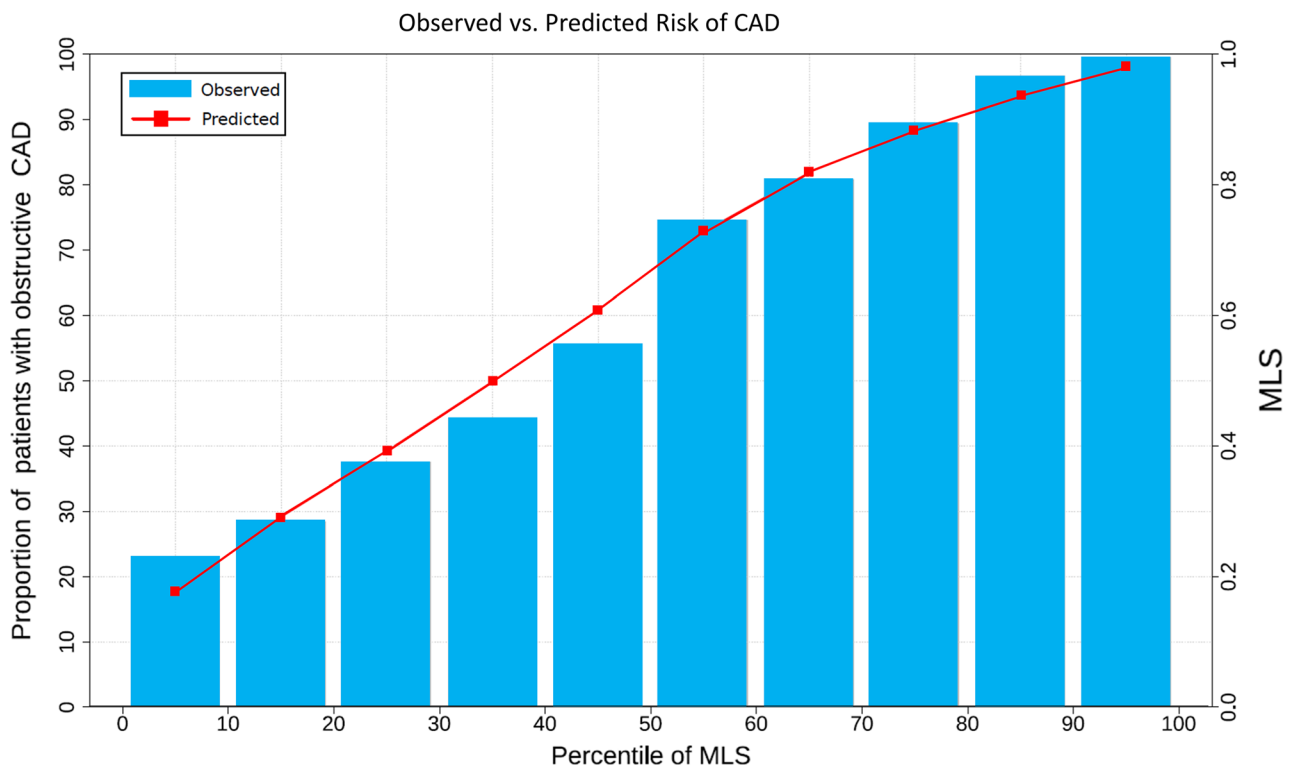
\*P value was calculated using Cochran's Q Test among all variables

†P < .01 by Cochran's Q Test for MLS 0.29, Reader, and TPD

‡P < .01 by McNemar's Test for MLS 0.38 and MLS 0.29

§P < .01 by McNemar's Test for MLS 0.29 vs MLS 0.16

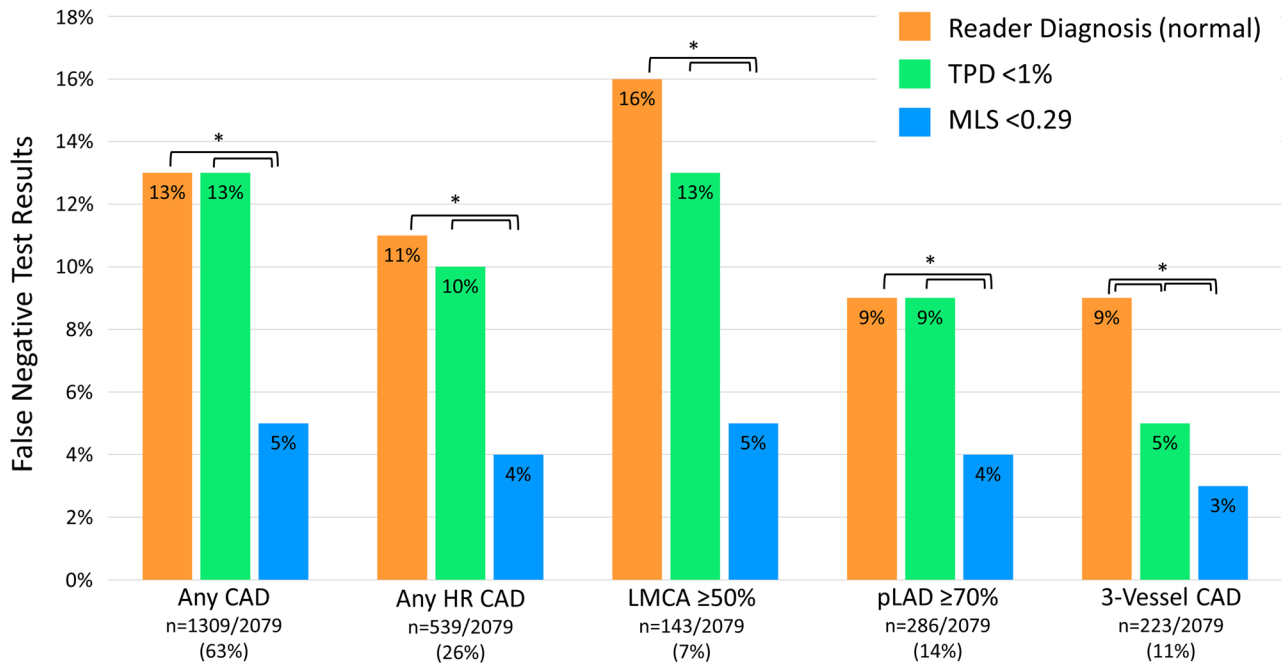
CAD, coronary artery disease; MLS, machine learning score; SSS, summed stress score; TPD, total perfusion deficit



**Figure 3.** Calibration plot showing the observed vs. predicted risk of obstructive CAD. The machine learning score (MLS) showed good calibration with a Brier score of 0.16. CAD, coronary artery disease.

compared to stress images acquired after rest imaging using higher isotope dosages. Thus, when applied to low-dose stress-first images, the MLS recognizes these differences and adjusts the feature selection algorithm to maintain its high diagnostic sensitivity for abnormal test results. Additionally, an MLS threshold of 0.29 was

associated with a diagnostic sensitivity of 95% for obstructive CAD from stress-only images, without the use of computed tomographic AC. The high sensitivity of the MLS offers a novel approach to increase the feasibility of stress-only MPI in laboratories without



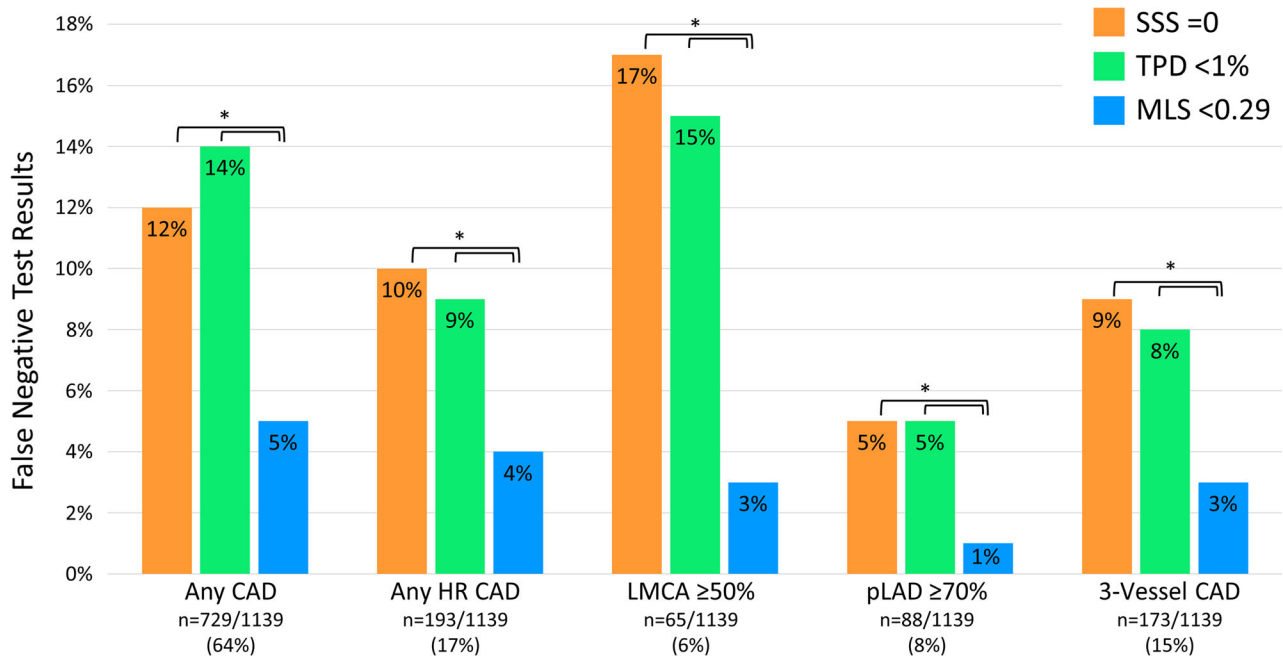
**Figure 4.** Diagnostic safety of automated cancellation of rest MPI in all patients. Frequency of false-negative test results among each method of interpretation for SPECT-MPI in total population (N = 2079). High-risk(HR) coronary artery disease (CAD) refers to 3-vessel CAD, LMCA stenosis  $\geq 50\%$ , or proximal LAD stenosis  $\geq 70\%$ . Three-vessel CAD refers to stenosis  $\geq 70\%$  of the LAD, LCx, and RCA. \* $P < .01$ . LCx, left circumflex artery; LMCA, left main coronary artery; pLAD, proximal left anterior descending artery; RCA, right coronary artery; TPD, total perfusion deficit; MLS, machine learning score.

access to attenuation correction and when applied to different stress imaging protocols.

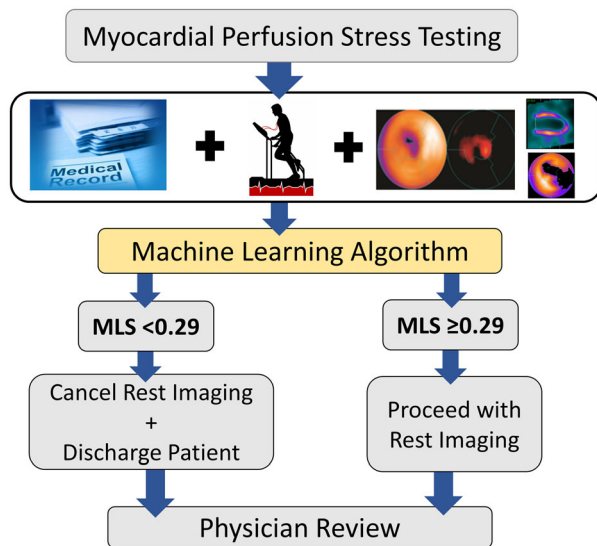
Another limitation of SPECT-MPI is the relatively low sensitivity for detection of stenosis in the LMCA or proximal LAD.<sup>20,21</sup> The CORE320 study previously reported a sensitivity of 75% for detection of LMCA stenosis.<sup>21</sup> Other studies have estimated the sensitivity of SPECT-MPI was just 54.5% and 55.0% for detection of stenosis in the LMCA and LAD, respectively.<sup>20</sup> In our study, the MLS threshold of 0.29 was associated with a diagnostic sensitivity of 95% for all obstructive CAD, including high-risk lesions in both the LMCA and proximal LAD. The MLS was trained using the REFINE SPECT diagnostic cohort, which has an overall prevalence of 63% for obstructive CAD, which is higher than the general population undergoing stress testing.<sup>22</sup> We included only patients who had gone for coronary angiography to ensure that coronary anatomy was accurately defined since SPECT-MPI may be interpreted as normal in patients with high-risk CAD, including those with low pre-test likelihood of disease.<sup>23</sup> This high-risk population is ideal for training the computer to learn features of high-risk CAD from stress-first MPI data and clinical variables, allowing the computer to

recognize abnormal stress MPI data that would have otherwise been considered normal by both TPD and readers. The incorporation of clinical variables to the MLS further improves its sensitivity of detecting obstructive CAD in images that would have been considered normal by readers. As shown in Figure 1, ischemic EKG changes were high predictors of an elevated MLS which is consistent with prior studies, especially given the high frequency of pharmacologic stress testing in the study population.<sup>24</sup> Similarly, post-stress wall motion abnormalities are associated with the presence of obstructive CAD,<sup>25</sup> and wall motion extent had the fourth highest information gain ratio.

A highly sensitive MLS is critical to safely cancel unnecessary rest imaging prior to physician review, but this was achieved at the cost of specificity. The specificity of the MLS reflects the frequency of patients who would have additional rest imaging without obstructive CAD. Importantly, in North America, rest imaging is performed prior to stress testing in over 92% of nuclear cardiology laboratories.<sup>2,10</sup> Thus, despite a lower specificity, the MLS threshold of 0.29 would lead to an overall reduction in unnecessary rest imaging and ultimately radiation exposure from SPECT-MPI. It is



**Figure 5.** Frequency of false-negative results in population with available segmental scores (N = 1139). High-risk (HR) coronary artery disease (CAD) refers to 3-vessel CAD, LMCA  $\geq 50\%$  or pLAD  $\geq 70\%$ . Three-vessel CAD refers to stenosis  $\geq 70\%$  of the LAD, LCx, and RCA. \* $P < .01$ . LCx, left circumflex artery; LMCA, left main coronary artery; pLAD, proximal left anterior descending artery; RCA, right coronary artery; SSS, summed stress score; TPD, total perfusion deficit; MLS, machine learning score.



**Figure 6.** Proposed automated stress-first SPECT-MPI algorithm. MLS, machine learning score.

also notable that the NPV of the proposed MLS was 78% despite a sensitivity of 95%. This is likely due to the high prevalence of obstructive CAD in the study population who were all referred for ICA. The

prevalence of abnormal stress perfusion, defined as stress TPD  $\geq 5\%$ , in this population was 64.7% compared to only 25.5% in the overall REFINE registry.<sup>26</sup> Assuming the prevalence of obstructive CAD is proportional to the prevalence of abnormal SPECT perfusion, we would estimate 25% of patients having obstructive CAD in the overall REFINE registry. Thus when applied to the general patient population referred for SPECT-MPI the NPV would be estimated at 95%.

There are limitations to the current study. Automatically generated myocardial contours were evaluated by an experienced technologist and corrected in a small subset of cases. Most adjustments were small, but the MLS algorithm would require technologist verification prior to generating an MLS. The information gain ratios provide a group importance for a given feature in the whole population but not an explanation for a given patient. Although not reported in this study, the individual importance of features can be determined to provide patient-specific explanations as recently demonstrated.<sup>11</sup> Another limitation is that the time interval between SPECT-MPI and ICA was relatively long. This interval is similar to prior studies with similar inclusion criteria,<sup>27</sup> and significant changes in coronary artery

stenosis are unlikely to occur over this timeframe. When generating the MLS, missing data was imputed using either population means or without the specific value, as described above. Despite this, the MLS algorithm was highly sensitive for obstructive and high-risk CAD and reflects clinical settings where not all data is available for MLS computation. Because all patients underwent ICA with 6 months of SPECT-MPI there is an overall selection bias of the study population. Although this high-risk population was ideal for training a computer to recognize abnormal SPECT-MPI, the true NPV is likely underestimated. Finally, although we studied a large, multicenter, international cohort, due to this selection bias, validation in a prospective, unselected population is needed

## CONCLUSION

In this study, machine learning was used to develop a diagnostic algorithm for stress-only MPI that automatically determines the need for additional rest images, prior to physician review. An MLS threshold of 0.29 was associated with superior diagnostic sensitivity for obstructive and high-risk CAD than SSS, TPD, or expert readers. The proposed MLS algorithm can improve implementation of stress-first MPI protocols, ultimately reducing cost, scanning times, and radiation dosages.

## NEW KNOWLEDGE GAINED

Machine learning can detect obstructive and high-risk CAD from stress MPI with high sensitivity and could be applied in an automated clinical algorithm to select patients for stress-only MPI without increased burden to the interpreting cardiologist.

## Disclosures

*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. The work was supported in part by the Dr. Miriam and Sheldon Adelson Medical Research Foundation. D.S.B., and P.J.S. participate in software royalties for QPS software at Cedars-Sinai Medical Center. P.J.S. has received research Grant support from Siemens Medical Systems. D.S.B., A.J.E., and E.J.M. have served as consultants for GE Healthcare. M.D.C. has received research Grant support from Spectrum-Dynamics and consulting honoraria from Sanofi and GE Healthcare. T.D.R. has received research Grant support from GE Healthcare and Advanced Accelerator Applications. A.J.E. and his institution have received research support from GE Healthcare, Philips Healthcare, and Toshiba*

*America Medical Systems. E.J.M. has served as a consultant for Bracco Inc, and he and his institution have received Grant support from Bracco Inc. All other authors have reported that they have no relationships relevant to the contents of this paper to disclose.*

## References

1. Wolk MJ, Bailey SR, Doherty JU, Douglas PS, Hendel RC, Kramer CM, et al. ACCF/AHA/ASE/ASNC/HFSA/HRS/SCAI/SCCT/SCMR/STS 2013 multimodality appropriate use criteria for the detection and risk assessment of stable ischemic heart disease. *J Card Fail* 2014;20:65-90.
2. Einstein AJ, Pascual TN, Mercuri M, Karthikeyan G, Vitola JV, Mahmarian JJ, et al. Current worldwide nuclear cardiology practices and radiation exposure: Results from the 65 country IAEA Nuclear Cardiology Protocols Cross-Sectional Study (INCAPS). *Eur Heart J* 2015;36:1689-96.
3. Mercuri M, Pascual TN, Mahmarian JJ, Shaw LJ, Dondi M, Paez D, et al. Estimating the reduction in the radiation burden from nuclear cardiology through use of stress-only imaging in the United States and worldwide. *JAMA Intern Med* 2016;176:269-73.
4. Jerome SD, Tilkemeier PL, Farrell MB, Shaw LJ. Nationwide laboratory adherence to myocardial perfusion imaging radiation dose reduction practices: A report from the intersocietal accreditation commission data repository. *JACC Cardiovasc Imaging* 2015;8:1170-6.
5. Gowd BM, Heller GV, Parker MW. Stress-only SPECT myocardial perfusion imaging: A review. *J Nucl Cardiol* 2014;21:1200-12.
6. Bhalodkar NC, Blum S. Normal stress-only versus standard stress/rest myocardial perfusion imaging: Similar patient mortality with reduced radiation exposure. *J Am Coll Cardiol* 2010;55:2611-2.
7. Ueyama T, Takehana K, Maeba H, Iwasaka T. Prognostic value of normal stress-only technetium-99m myocardial perfusion imaging protocol. Comparison with standard stress-rest protocol. *Circ J* 2012;76:2386-91.
8. Chang SM, Nabi F, Xu J, Raza U, Mahmarian JJ. Normal stress-only versus standard stress/rest myocardial perfusion imaging: Similar patient mortality with reduced radiation exposure. *J Am Coll Cardiol* 2010;55:221-30.
9. Gowdar S, Chaudhry W, Ahlberg AW, Henzlova MJ, Duvall WL. Triage of patients for attenuation-corrected stress-first Tc-99m SPECT MPI using a simplified clinical pre-test scoring model. *J Nucl Cardiol* 2018;25:1178-87.
10. 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. *JACC Cardiovasc Imaging* 2018;11:1000-9.
11. Hu LH, Betancur J, Sharir T, Einstein AJ, Bokhari S, Fish MB, et al. Machine learning predicts per-vessel early coronary revascularization after fast myocardial perfusion SPECT: results from multicentre REFINE SPECT registry. *Eur Heart J Cardiovasc Imaging* 2020;21:549-59.
12. Hu LH, Miller RJH, Sharir T, Commandeur F, Rios R, Einstein AJ et al. Prognostically safe stress-only SPECT myocardial perfusion imaging guided by machine learning: Report from REFINE SPECT. *Eur Heart J Cardiovasc Imaging* 2020.
13. 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). *J Nucl Cardiol* 2020;27:1010-21.
14. 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.
  15. Molinaro AM, Simon R, Pfeiffer RM. Prediction error estimation: A comparison of resampling methods. *Bioinformatics* 2005;21:3301-7.
  16. Motwani M, Dey D, Berman DS, Germano G, Achenbach S, Al-Mallah MH, et al. Machine learning for prediction of all-cause mortality in patients with suspected coronary artery disease: A 5-year multicentre prospective registry analysis. *Eur Heart J* 2017;38:500-7.
  17. 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.
  18. Duvall WL, Hiensch RJ, Levine EJ, Croft LB, Henzlova MJ. The prognosis of a normal TI-201 stress-only SPECT MPI study. *J Nucl Cardiol* 2012;19:914-21.
  19. Duvall WL, Baber U, Levine EJ, Croft LB, Henzlova MJ. A model for the prediction of a successful stress-first Tc-99m SPECT MPI. *J Nucl Cardiol* 2012;19:1124-34.
  20. Budoff MJ, Li D, Kazerooni EA, Thomas GS, Mieres JH, Shaw LJ. Diagnostic accuracy of noninvasive 64-row Computed Tomographic Coronary Angiography (CCTA) compared with Myocardial Perfusion Imaging (MPI): The PICTURE Study. A prospective multicenter trial. *Acad Radiol* 2017;24:22-9.
  21. George RT, Mehra VC, Chen MY, Kitagawa K, Arbab-Zadeh A, Miller JM, et al. Myocardial CT perfusion imaging and SPECT for the diagnosis of coronary artery disease: A head-to-head comparison from the CORE320 multicenter diagnostic performance study. *Radiology* 2015;274:626.
  22. Berman DS, Shaw LJ, Min JK, Hachamovitch R, Abidov A, Germano G, et al. SPECT/PET myocardial perfusion imaging versus coronary CT angiography in patients with known or suspected CAD. *Q J Nucl Med Mol Imaging* 2010;54:177-200.
  23. Nakanishi R, Gransar H, Slomka P, Arsanjani R, Shalev A, Otaki Y, et al. Predictors of high-risk coronary artery disease in subjects with normal SPECT myocardial perfusion imaging. *J Nucl Cardiol* 2016;23:530-41.
  24. Cosmai EM, Heller GV. The clinical importance of electrocardiographic changes during pharmacologic stress testing with radionuclide myocardial perfusion imaging. *J Nucl Cardiol* 2005;12:466-72.
  25. Emmett L, Iwanochko RM, Freeman MR, Barolet A, Lee DS, Husain M. Reversible regional wall motion abnormalities on exercise technetium-99m-gated cardiac single photon emission computed tomography predict high-grade angiographic stenoses. *J Am Coll Cardiol* 2002;39:991-8.
  26. Otaki Y, Betancur J, Sharir T, Hu L-H, Gransar H, Liang JX, et al. 5-year prognostic value of quantitative versus visual mpi in subtle perfusion defects. *JACC Cardiovasc Imaging* 2020;13:774-85.
  27. Fiechter M, Ghadri JR, Kuest SM, Pazhenkottil AP, Wolfrum M, Nkoulou RN, et al. Nuclear myocardial perfusion imaging with a novel cadmium-zinc-telluride detector SPECT/CT device: First validation versus invasive coronary angiography. *Eur J Nucl Med Mol Imaging* 2011;38:2025-30.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.