ESC European Society of Cardiology

European Heart Journal - Cardiovascular Imaging (2020) **21**, 549–559
doi:10.1093/ehjci/jez177

Machine learning predicts per-vessel early coronary revascularization after fast myocardial perfusion SPECT: results from multicentre REFINE SPECT registry

Lien-Hsin Hu[1,2], Julian Betancur[1], Tali Sharir[3,4], Andrew J. Einstein[5,6], Sabahat Bokhari[5,6], Mathews B. Fish[7], Terrence D. Ruddy[8], Philipp A. Kaufmann[9], Albert J. Sinusas[10], Edward J. Miller[10], Timothy M. Bateman[11], Sharmila Dorbala[12], Marcelo Di Carli[12], Guido Germano[1], Frederic Commandeur[1], Joanna X. Liang[1], Yuka Otaki[1], Balaji K. Tamarappoo[1], Damini Dey[1], Daniel S. Berman[1], and Piotr J. Slomka[1]*

[1]Department of Imaging, Medicine, and Biomedical Sciences, Cedars-Sinai Medical Center, 8700 Beverly Blvd, Los Angeles, CA 90048, USA; [2]Department of Nuclear Medicine, Taipei Veterans General Hospital, No. 201, Section 2, Shipai Rd, Taipei, Taiwan; [3]Department of Nuclear Cardiology, Assuta Medical Center, HaBarzel St 20, Tel Aviv, Israel; [4]Faculty of Health Sciences, Ben Gurion University of the Negev, Rager Blvd, 84105 Be'er Sheva, Israel; [5]Division of Cardiology, Department of Medicine, Columbia University Irving Medical Center, 622 W 168th St, New York, NY 10032, USA; [6]Department of Radiology and Herbert Irving Comprehensive Cancer Center, Columbia University Irving Medical Center, 622 W 168th St, New York, NY 10032, USA; [7]Department of Nuclear Medicine, Oregon Heart and Vascular Institute, Sacred Heart Medical Center, 3333 Riverbend Dr, Springfield, OR 97477, USA; [8]Division of Cardiology, University of Ottawa Heart Institute, 40 Ruskin St, Ottawa, ON K1Y 4W7, Canada; [9]Department of Nuclear Medicine, Cardiac Imaging, University Hospital Zurich, Rämistrasse 100, 8091 Zurich, Switzerland; [10]Department of Internal Medicine, Section of Cardiovascular Medicine, Yale University, 333 Cedar St, New Haven, CT 06510, USA; [11]Cardiovascular Imaging Technologies LLC, 4320 Wormhall Rd, Kansas City, 64111 MO, USA; and [12]Division of Nuclear Medicine and Molecular Imaging, Department of Radiology, Brigham and Women's Hospital, 75 Francis St, Boston, MA 02115, USA

Received 31 October 2018; editorial decision 22 May 2019; accepted 0 Month 0000; online publish-ahead-of-print 16 July 2019

Aims

To optimize per-vessel prediction of early coronary revascularization (ECR) within 90 days after fast single-photon emission computed tomography (SPECT) myocardial perfusion imaging (MPI) using machine learning (ML) and introduce a method for a patient-specific explanation of ML results in a clinical setting.

Methods and results

A total of 1980 patients with suspected coronary artery disease (CAD) underwent stress/rest ^{99m}Tc-sestamibi/tetrofosmin MPI with new-generation SPECT scanners were included. All patients had invasive coronary angiography within 6 months after SPECT MPI. ML utilized 18 clinical, 9 stress test, and 28 imaging variables to predict per-vessel and per-patient ECR with 10-fold cross-validation. Area under the receiver operator characteristics curve (AUC) of ML was compared with standard quantitative analysis [total perfusion deficit (TPD)] and expert interpretation. ECR was performed in 958 patients (48%). Per-vessel, the AUC of ECR prediction by ML (AUC 0.79, 95% confidence interval (CI) [0.77, 0.80]) was higher than by regional stress TPD (0.71, [0.70, 0.73]), combined-view stress TPD (AUC 0.71, 95% CI [0.69, 0.72]), or ischaemic TPD (AUC 0.72, 95% CI [0.71, 0.74]), all $P < 0.001$. Per-patient, the AUC of ECR prediction by ML (AUC 0.81, 95% CI [0.79, 0.83]) was higher than that of stress TPD, combined-view TPD, and ischaemic TPD, all $P < 0.001$. ML also outperformed nuclear cardiologists' expert interpretation of MPI for the prediction of early revascularization performance. A method to explain ML prediction for an individual patient was also developed.

Conclusion

In patients with suspected CAD, the prediction of ECR by ML outperformed automatic MPI quantitation by TPDs (per-vessel and per-patient) or nuclear cardiologists' expert interpretation (per-patient).

Keywords

coronary artery disease • early coronary revascularization • explainable machine learning • SPECT myocardial perfusion imaging • new-generation cardiac camera

* Corresponding author. Tel: 310-423-4348; Fax: 310-423-0173. E-mail: piotr.slomka@cshs.org

Introduction

In the era of great diversity and excellent performance of the diagnostic tools to evaluate stable ischaemic heart disease, non-invasive tests have been able to provide a comprehensive assessment of myocardial perfusion, heart function, and anatomical obstructive disease.[1] It is likely that in the future invasive coronary angiography (ICA) will be reserved only for patients whom an invasive procedure and/or coronary revascularization is needed, given the risk and the cost of ICA relative to non-invasive tests. Thus, an accurate prediction of revascularization procedure after a non-invasive test can potentially aid physicians in an appropriate selection of patients for ICA.

Prediction of early coronary revascularization (ECR) by myocardial perfusion imaging (MPI) has been investigated in a few studies.[2–5] Among those, a single-centre study by Arsanjani *et al.*[5] studied dual-tracer MPI on conventional gamma camera for the prediction of ECR. In that study, machine learning (ML) approach gave a comparable prediction to human experts on a per-patient basis. In this study, we aimed to develop per-vessel ML prediction of ECR after single-photon emission computed tomography (SPECT) MPI and compare such algorithms with standard perfusion measures. We utilized a diagnostic cohort from a recently established multicentre, international registry—*RE*gistry of *F*ast Myocardial Perfusion *I*maging with *NE*xt generation *SPECT* (REFINE SPECT)[6] with an expanded number of clinical and imaging variables. We also aimed to devise a novel method to explain the rationale of ML prediction of ECR in an individual patient.

Methods

Study population and data collection

This study included all currently available 2079 patients from 9 centres in the diagnostic cohort of the REFINE SPECT registry.[6] Patients were consecutively recruited at their respective centres with no previously known coronary artery disease (CAD), myocardial infarction, or previous coronary revascularization. All patients underwent MPI on a new-generation SPECT camera as well as ICA within 6 months, between 2008 and 2017. The MPI images, clinical information, and data on ICA and any revascularization procedure performed within 6 months of the MPI were collected (*Table 1*). Patients with cardiac transplant ($n = 4$), open-heart surgery ($n = 3$), ICA performed before MPI ($n = 63$), or incomplete data on revascularization or expert interpretation of MPI ($n = 29$) were excluded, leaving 1980 patients as the study population.

Imaging protocols

MPI studies were performed on new-generation cardiac scanners: D-SPECT (Spectrum-Dynamics, Israel), Discovery NM530c or NM/CT570c (GE Healthcare, Israel). Among all nine sites, the MPI protocols included ^{99m}Tc-sestamibi or ^{99m}Tc-tetrofosmin-based 1-day stress-and-rest, 2-day stress-and-rest, or stress-only. The stress protocols consisted of either symptom-limited exercise stress (of which 92.5% used the Bruce protocol) or pharmacologic stress (with regadenoson, adenosine, dipyridamole, or dobutamine) (*Table 1*). Images were obtained 10–60 min (for the stress imaging) and 5–60 min (for the rest imaging) after radiotracer injection. When applicable, images were acquired for patients in two positions (upright/supine for D-SPECT, or supine/prone for Discovery scanners,[6] which are later referred to as default/alternative views). Stress images acquisition time for each position was 4–6 min.

Image processing and automated quantitation

Reconstructed image datasets were automatically processed with the Quantitative Perfusion/Gated SPECT software (QPS/QGS) version 2015 (Cedars-Sinai Medical Center, CA).[7] Automatically generated myocardial contours were checked by experienced technologists who were blinded to the clinical data. Left ventricular contours (10.8% of all datasets) were adjusted manually if necessary.[8] Correlated imaging quantitation parameters regarding myocardial perfusion, function, and phase analysis were generated accordingly.[6] Total perfusion deficit (TPD) variables used in this study included stress TPD (quantified from the stress image), combined-view stress TPD (computed according to the defect concomitant in both positions, if both available),[9–11] and ischaemic TPD (stress TPD minus rest TPD) with the values of rest images assumed normal for cases without rest data.[12] Quantitation of the severity and extent of perfusion deficits were computed accordingly.

Visual perfusion assessment

Currently, the probability of revascularization is not formally evaluated in the clinical setting, therefore as a proxy, we utilized the standard clinical interpretation of perfusion (final expert interpretation with or without segmental scores) performed by experienced nuclear cardiologists with access to all imaging and clinical information during clinical reporting. In four sites, visual segmental scores including summed stress score (SSS), summed rest score (SRS), and summed difference score (SDS) as well as a four-point scoring clinical diagnosis (0-normal, 1-probably normal, 2-equivocal, and 3-abnormal perfusion) were reported during clinical reading. In the other five sites, only the four-point scoring system was used. The final expert interpretations were homogenized to a four-point scoring system: 0-normal (SSS = 0), 1-probably normal (SSS = 1), 2-equivocal (SSS = 2 or 3), and 3-abnormal (SSS $\geq$ 4). We also evaluated the results in a subpopulation with both SSS and SDS available.

ICA and revascularization

All patients underwent clinically indicated ICA with standard protocols within 6 months after MPI. Obstructive CAD was visually assessed and defined as stenosis $\geq$50% for the left main coronary artery, or stenosis $\geq$70% in the left anterior descending artery (LAD), left circumflex (LCx), and right coronary artery (RCA). Revascularization with either percutaneous coronary intervention or coronary artery bypass grafting was performed according to clinical indications. The endpoint, ECR, was defined as revascularization of any coronary arteries within 90 days after MPI on a per-vessel and per-patient basis.

Data pre-processing and ML

Variables for ML

Patient data were arranged in a stacked format for the subsequent per-vessel ML modelling. Variables of each patient were stacked on three territories with different regional values but sharing the same clinical and global imaging variables, resulting in a total of 5940 vessel observations. To avoid redundancy, ML variables were selected as follows: (i) for perfusion variables, only the extent and severity of perfusion deficit were used since they were highly correlated with TPD,[13] (ii) for stress/rest/ischaemic variables, only stress and ischaemic variables were used, (iii) for dimensional measurements, only LV volume was used, (iv) for function variables, only wall thickening was used, and (v) redundant clinical variables were also removed, e.g. body weight and height (body mass index kept). Overall, 18 clinical variables, 9 stress-test variables, and 28 imaging variables (17 global and 11 regional) were used for ML (Supplementary data online, *Table S1*).

Table 1 Patient baseline characteristics in the non-early revascularized and early revascularized groups

	No early revascularization (*n* = 1022)	%	Early revascularization (*n* = 958)	%	*P*-value
Obstructive disease	302	29.5	947	98.9	<0.005
ECR in LAD			675	70.5	N/A
ECR in LCx			425	44.4	N/A
ECR in RCA			400	41.8	N/A
Age (year, mean ± SD)	64 ± 11		65 ± 11		0.01
Male	594	58.1	718	74.9	<0.005
Diabetes mellitus	294	28.8	285	29.7	0.63
Hypertension	692	67.7	663	69.2	0.47
Dyslipidaemia	615	60.2	620	64.7	0.04
Smoking	268	26.2	234	24.4	0.36
Family history	324	31.7	338	35.3	0.09
Symptoms					
Symptoms other than chest pain	356	34.8	235	24.5	<0.005
Non-anginal chest pain	127	12.4	106	11.1	0.35
Atypical chest pain	303	29.6	324	33.8	0.05
Typical chest pain	153	15.0	257	26.8	<0.005
MPI protocol					
Stress dose (MBq)	674 ± 436		609 ± 446		<0.005
Rest dose (MBq)	425 ± 313		480 ± 309		<0.005
Exercise stress	367	35.9	382	39.9	0.07
Pharmacologic stress	654	64.0	575	60.0	0.07

ECR, early coronary revascularization; LAD, left anterior descending artery; LCx, left circumflex artery; MBq, megabecquerels; MPI, myocardial perfusion imaging; RCA, right coronary artery; SD, standard deviation.

Missing variables

For all variables with missing values, the default imputation was performed by the Waikato Environment for Knowledge Analysis platform 3.8.1 (WEKA) (University of Waikato, Hamilton, New Zealand). They were imputed with the population's mean value for numerical variables, or with a distinct 'missing' label for categorical variables.

Machine learning

The LogitBoost method implemented in the WEKA platform was used for training and validation of the ML model.[14] Four ML modelling steps were applied in sequence: (i) generation of a randomized, non-contaminated (with all three vessel observations from the same patient being sorted into the same fold) 10-fold dataset, (ii) automatic variable selection using only variables with an information gain ratio (IGR) >0, (iii) model building employing an ensemble LogitBoost algorithm,[15] and (iv) 10-fold cross-validation, currently the preferred method to assess ML performance.[16,17] The main advantages of 10-fold cross-validation compared with the conventional split-sample (test and validation) approach are: (i) reduces the variance in prediction error leading to a more accurate estimate of model performance; (ii) maximizes the use of data for both training and validation, without overfitting or overlap between test and validation data; and (iii) guards against testing hypotheses suggested by arbitrarily split data (Type III errors).[17]

In brief, validation tests are performed 10 times with each fold (10% of the data) being used in turn as the test set while the other 9-folds (remaining 90% of the data) as the training set. This results in 10 experimental models, and each is trained on 90% of the samples. Thus, only unseen data were used for testing of each model. The validation results from 10 experimental models are then concatenated to provide a measure of the overall performance.[17]

Generalization error: assessment of overfitting

To evaluate the overfitting in this work, we compared the performance of ML models with different thresholds of IGR by: (i) using entire population as the training set and (ii) modelling with the 10-fold cross-validation (hence, assessing the performance of 10 models on 10 unseen test datasets by these models). By comparing the difference between area under the ROC curve (AUC) derived between the first and second model, the generalization error was derived to evaluate the overfitting issue.

Explanation of rationale behind ML prediction

We devised a method for explaining the ML predictions on an individual patient and territory basis. Further details of this approach are provided in the Supplementary data online.

Statistical analysis

Paired receiver-operating characteristic (ROC) curves for the ECR prediction were computed for the ML score and for TPD (per-vessel and per-patient). In addition, the AUC with the non-parametric approach of DeLong *et al.*[18,19] was used to assess ML predictive gains. The performance of the expert interpretation was also plotted in the per-patient paired-ROC graph to indicate sensitivity and specificity at different thresholds. McNemar's chi-square test was used when comparing

Table 2 Stress-test results: clinical responses, quantitative results, and expert interpretation

	No early revascularization (n = 1022)	Early revascularization (n = 958)	P-value
Clinical and physiological responses			
Pharmacologic MPI (n = 1231)	655 (64)	576 (60)	0.07
Rest HR, beats/min	73 ± 15	72 ± 14	0.11
Stress peak HR, beats/min	99 ± 23	104 ± 24	<0.05
Rest SBP, mmHg	136 ± 22	138 ± 21	0.18
Stress peak SBP, mmHg	137 ± 25	141 ± 27	<0.05
ST deviation, mm	0.1 ± 0.4	0.3 ± 0.8	<0.05
Ischemic symptoms	70 (11)	76 (13)	0.08
Exercise MPI (n = 749)	367 (36)	382 (40)	–
Rest HR, beats/min	72 ± 13	70 ± 14	0.06
Stress peak HR, beats/min	143 ± 21	142 ± 16	0.22
Rest SBP, mmHg	133 ± 21	136 ± 18	<0.05
Stress peak SBP, mmHg	167 ± 25	168 ± 22	0.86
ST deviation, mm	0.8 ± 1.0	1.3 ± 1.2	<0.05
Ischemic symptoms	120 (33)	187 (49)	<0.05
MPI quantitative results			
Global			
Stress TPD, %	7.5 ± 7.5	14.8 ± 10.9	<0.05
Combined stress TPD, %	4.3 ± 6.5	11.0 ± 10.5	<0.05
Ischemic TPD, %	4.9 ± 4.8	10.7 ± 7.9	<0.05
Stress EXT, %	7.7 ± 8.3	20.2 ± 13.2	<0.05
Stress SEV, %	0.9 ± 0.5	1.6 ± 1.0	<0.05
Rest EF, %	58 ± 16	60 ± 13	<0.05
Transient ischemic dilation	1.01 ± 0.15	1.06 ± 0.16	<0.05
Regional			
Stress TPD, %	3.0 ± 4.3	7.2 ± 9.0	<0.05
Stress EXT, %	7.6 ± 11.7	17.4 ± 19.9	<0.05
Stress SEV, %	0.8 ± 0.7	1.5 ± 1.4	<0.05
Stress EF, %	58 ± 15	57 ± 13	0.69
Expert interpretation			
Normal (0)	332 (32)	102 (11)	<0.05
Probably normal (1)	64 (6)	32 (3)	<0.05
Equivocal (2)	56 (5)	32 (3)	<0.05
Abnormal (3)	570 (56)	792 (83)	<0.05

Values are n (%), mean ± SD, or %.
EF, ejection fraction; EXT, perfusion deficit extent; HR, heart rate; MPI, myocardial perfusion imaging; SBP, systolic blood pressure; SEV, perfusion deficit severity; ST, ST-segment; TPD, total perfusion deficit.

sensitivities. A P-value of < 0.05 was considered significant. R (v3.3.3, R Foundation, Austria) was used for merging the database and statistical analyses.

Results

Study population and stress-test results

Of the 1980 patients, 1249 (63%) patients were diagnosed with CAD by visually assessed ICA. A total of 958 (48%) patients underwent ECR (*Table 1*). In *Table 2*, clinical and physiological responses to the stress test, quantitative SPECT results, and the expert interpretation for MPI are shown. Clinical interpretation of MPI was provided by all nine sites (n = 1980) and four sites reported segmental scores as well (n = 1072), among which both SSS and SRS were available in 1014 cases.

Missing rates of variables

The missing rates for all 55 variables were provided in Supplementary data online, *Table S1*.

Variable selection

Among all 55 used variables, 49 variables were selected and are shown in *Figure 1* according to the ranking of IGR. A list of all 55 variables is shown in the Supplementary data online, Table S1.

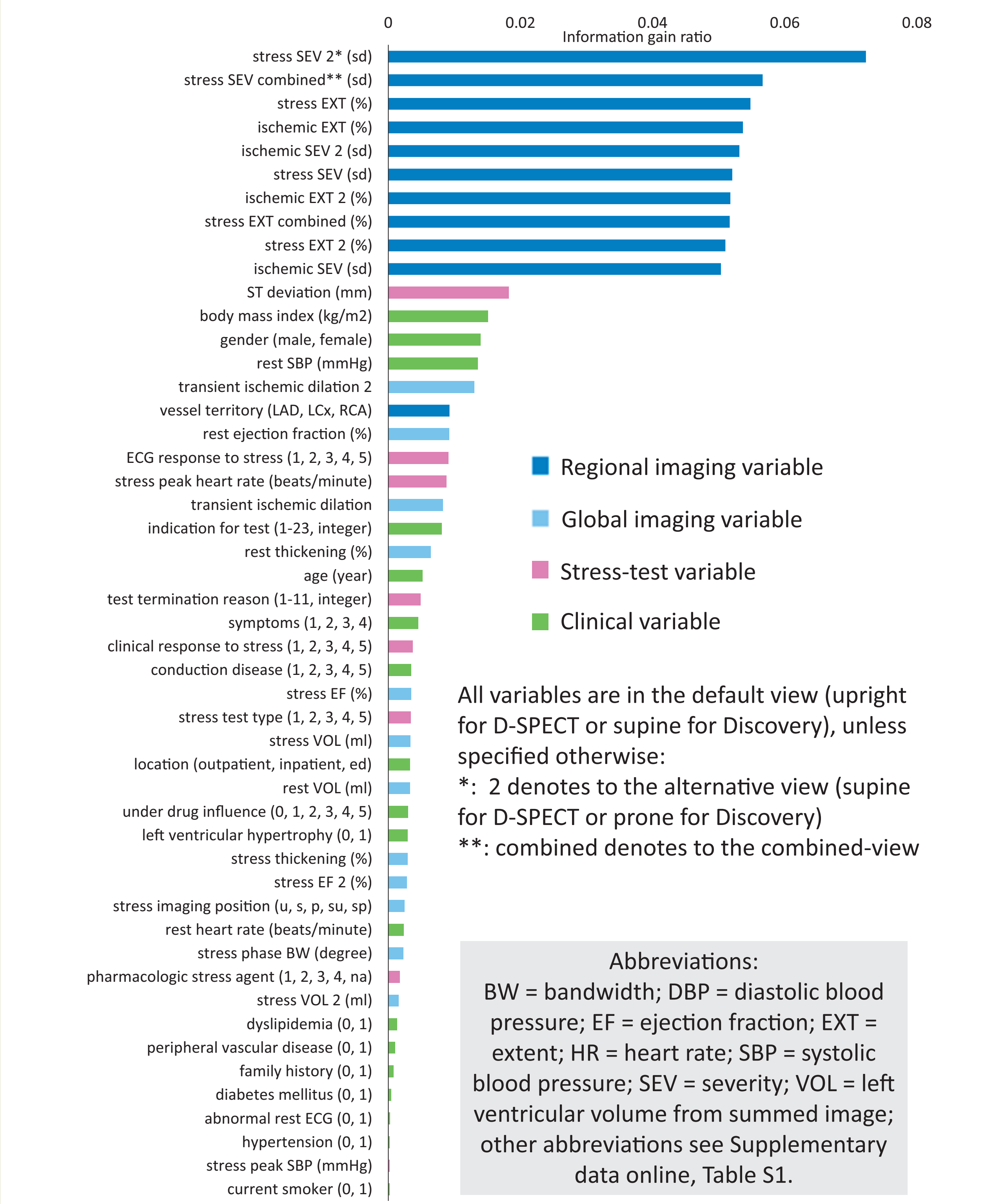

Figure 1 Selected variables ranking by IGR. The ML algorithm evaluated all 55 used variables independently to determine the IGR for each variable in each fold. Forty-nine out of 55 variables had IGR >0 and were selected. ML models were built with these selected variables. Most variables in the ranking are imaging variables (blue and light blue bars) with regional imaging variables (blue bars) leading while clinical and stress-test variables also play roles in the prediction. IGR, information gain ratio; ML, machine learning.

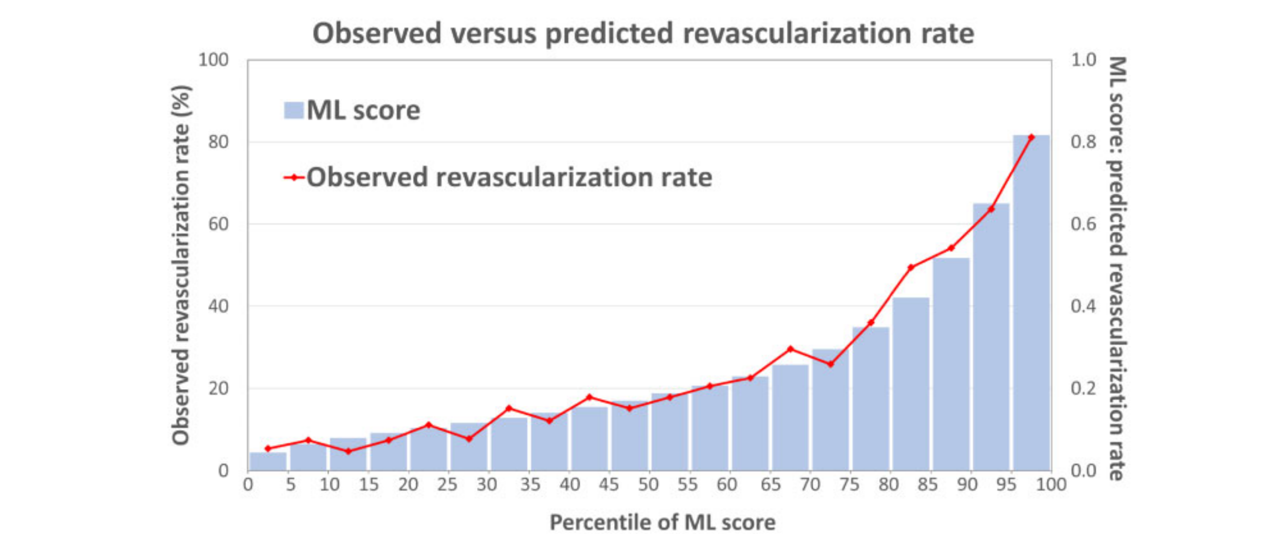

Figure 2 Observed and predicted revascularization rate. The observed revascularization rate (the red line) matched the predicted risk of revascularization (the blue bars) against every five percentile of ML score. ML, machine learning.

Table 3 Revascularization rate and ML scores

Site	Number of patients (*n*)	Obstructive disease (CAD [*n*])	No early revascularization (*n*)	Early revascularization (*n*, [%, in CAD])	Total revascularization (*n*)
Brigham and Women's	302	184	176	126 (67)	142
Cedars-Sinai	190	119	96	94 (77)	102
Aspire Foundation	278	179	131	147 (82)	156
Oregon Heart and Vascular Institute	334	199	169	165 (82)	178
Assuta Medical Center	313	221	108	205 (92)	205
Columbia University	113	47	79	34 (72)	37
Ottawa Heart Institute	360	246	220	140 (54)	223
Yale University	56	29	32	24 (79)	24
University of Zurich	34	25	11	23 (92)	24
Total	1980	1249	1022	958 (75)	1091
Per-vessel ML score, mean ± SD	0.25 ± 0.20	0.29 ± 0.22	0.18 ± 0.14	0.36 ± 0.24	0.32 ± 0.23
Per-patient ML score, mean ± SD	0.40 ± 0.24	0.47 ± 0.24	0.27 ± 0.17	0.53 ± 0.23	0.47 ± 0.24

CAD, coronary artery disease; ML, machine learning; SD, standard deviation.

Per-vessel and per-patient prediction

Figure 2 shows the ECR rate increases with per-patient ML score. *Table 3* shows revascularization results among the sites and mean ML scores in groups with or without ECR.

On a per-vessel basis, *Figure 3* shows the ROC curves of regional stress TPD, combined-view TPD, ischaemic TPD, and ML scores. ML outperformed TPD in all territories with the largest gain in LAD. In order to understand the contribution of non-imaging variables, we compared ML training utilizing all (55), imaging-only (28), and imaging-plus-limited-clinical (age, gender, and body mass index) (31) variables. The AUCs of ML prediction based on all variables vs. imaging-only variables vs. imaging-plus-limited-clinical variables were 0.785 (95% confidence interval (CI) [0.771, 0.798]) vs. 0.754 (95% CI [0.739, 0.769]) vs. 0.768 (95% CI [0.754, 0.783]) on a per-vessel basis (all $P < 0.01$).

On a per-patient basis, ML prediction had the highest AUC, as shown in *Figure 4*. If compared with the four-point-categorical expert interpretation and taking a diagnostic scoring >2 as the positive prediction of ECR, an improvement of sensitivity from 82.7% to 90.7% ($P < 0.001$) was noted at a specificity of 44.2%. The AUCs of ML prediction based on all variables vs. imaging-only variables vs. imaging-plus-limited-clinical variables were 0.812 (95% CI [0.793, 0.831]) vs. 0.786 (95% CI [0.766, 0.806]) vs. 0.799 (95% CI [0.779, 0.813]) on a per-vessel basis (all $P < 0.01$). A subpopulation analysis of the

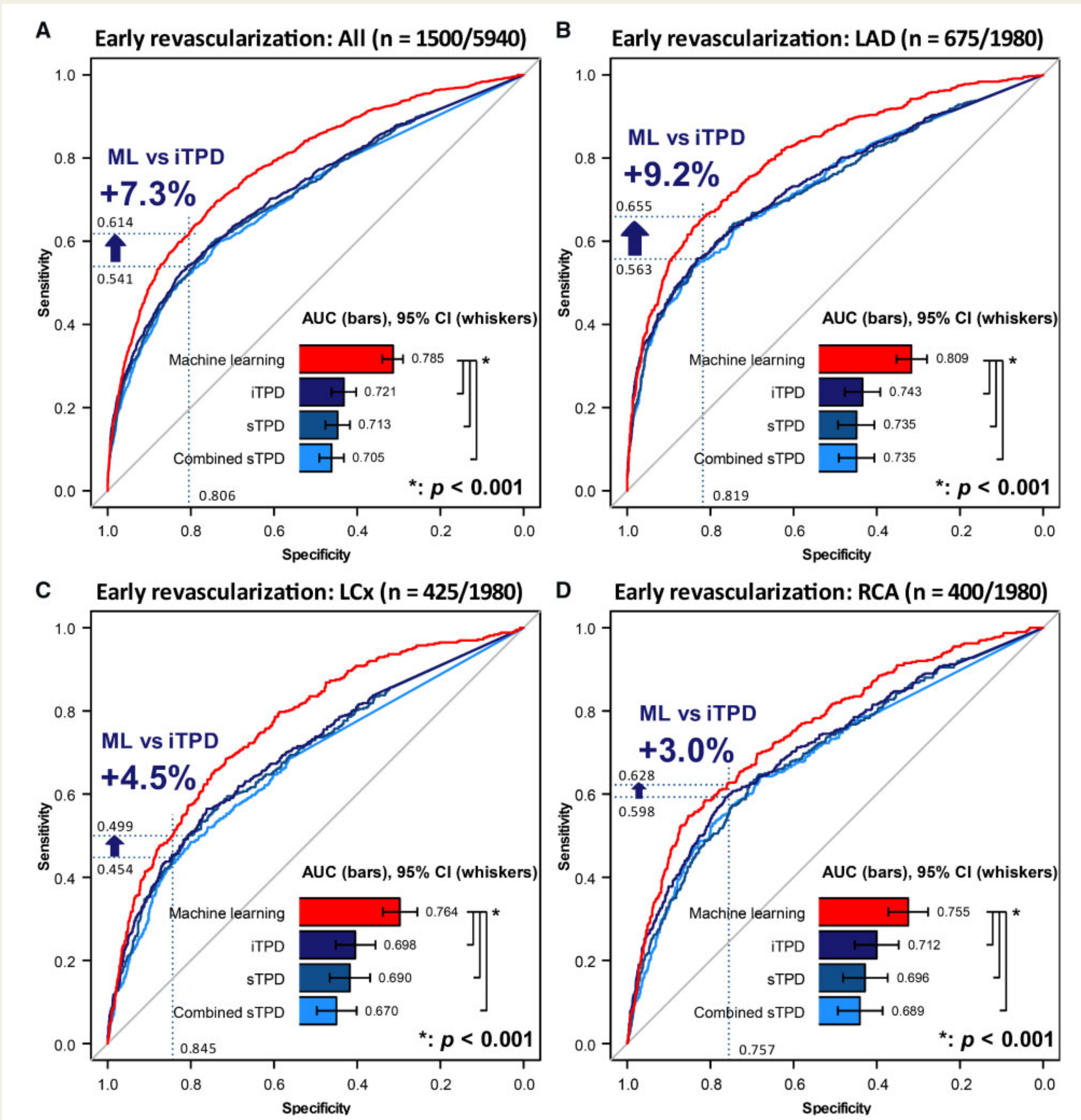

Figure 3 Paired ROC curves for the prediction of early revascularization: per-vessel result. Per-vessel results of predictions in all three coronaries (*A*), LAD (*B*), LCx (*C*), and RCA (*D*) territories. When operating with the same specificity of 80.6% using a cut-off of iTPD of 3%, ML shows a 7.3% improvement of sensitivity in the result of all three territories (A), and the improvement in each territory is also shown in (B–D) (blue arrows). The somewhat lower performance for RCA revascularization prediction might have been due to more common attenuation artefacts in that territory.[29] AUC, area under the ROC curve; CI, confidence interval; combined sTPD, regional combined-view stress TPD; iTPD, regional ischaemic TPD; LAD, left anterior descending artery; LCx, left circumflex artery; ML, machine learning; RCA, right coronary artery; ROC, receiver-operating characteristic curve.

performance of SSS and SDS in the Supplementary data online, *Figure S1* demonstrated that AUCs by TPD and visual summed scoring were not significantly different. ML had a better prediction than any of the other methods.

The generalization error was 3.0% as shown in the Supplementary data online, *Figure S2*.

Examples

Figure 5 shows examples of ML prediction and the rationale behind the prediction. ML prediction of ECR could have helped the decisions of referral to ICA, while the quantitative parameters were misleading. The patient-specific rationale for the prediction is shown at the bottom of the figure.

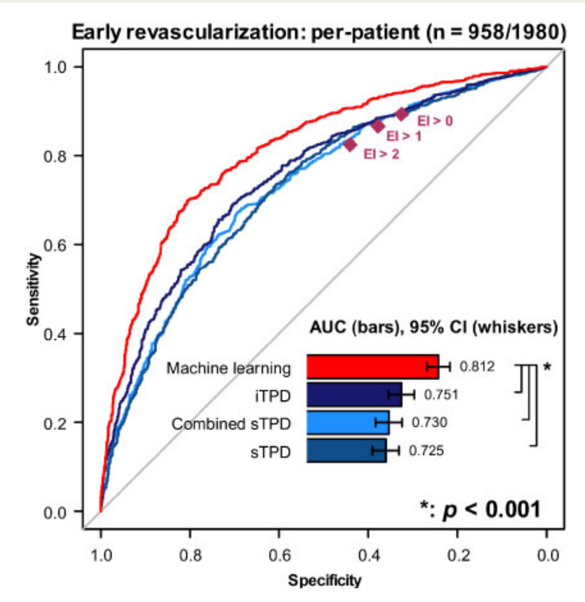

Figure 4 Paired ROC curves for the prediction of early revascularization: per-patient result. The per-patient result of predictions by retaining the maximal ML score per-territory was compared with TPD. ML has a higher area under the ROC curve (AUC) (all $P <$ 0.001). Higher sensitivity is also shown in comparison with expert interpretation (EI) at any diagnostic scoring from 0 to 3 (0-normal, 1-probably normal, 2-equivocal, and 3-abnormal). CI, confidence interval; combined sTPD, global combined-view stress TPD; EI, expert interpretation; iTPD, global ischaemic TPD; ML, machine learning; ROC, receiver-operating characteristic curve; sTPD, global stress TPD.

Discussion

This is the first attempt to use ML in predicting ECR after MPI using per-vessel modelling in a large multicentre study[5,20–22] and the first report of a novel method for the explanation of the ML result for a given patient. In our study, we utilized comprehensive imaging quantitative variables including perfusion, function, and phase analysis for prediction. As a marker of severe and extensive ischaemia,[23] myocardial stunning, was indirectly considered with both stress and rest perfusion and functional variables used for ML modelling. Our method based on the overall logit of each variable normalized to the ML score for the explanation of ML prediction may help clinicians understand the rationale behind the ML score for a given patient.

This study is different from the study by Arsanjani *et al.*[5] in several aspects. These include ML modelling (per-vessel vs. per-patient), the multicentre setting and number of cases (1980 vs. 713), the isotopes and scanners used (Tc-99m-based stress test on new-generation cameras in our study vs. dual isotope Tl-201/Tc-99m sestamibi rest/stress test on conventional Anger camera). Further, our study included patients between 2008 and 2017 after the publication of the Clinical Outcomes Utilizing Revascularization and Aggressive Drug Evaluation (COURAGE) trial[24] and might have reflected more contemporary revascularization guidelines.[25]

Utilizing patient clinical and MPI quantitative information, this model provides per-vessel probabilities of ECR performance for a given patient, before being possibly sent for the invasive angiography. Such a tool could serve as a handy 'expert consult system' to both nuclear cardiologists and interventional cardiologists, indicating the personalized likelihood that if invasive angiography is performed, revascularization will also be performed. This tool could be helpful to decide if the patient should undergo invasive angiography procedure. For example, a low chance of revascularization for the specific patient could perhaps save the patient from the subsequent angiographic procedure. Additionally, the predicted ECR outcome can be regarded as more prognostically informative than the prediction of obstructive CAD alone (used in many studies to date[7]), since interventional cardiologists decided to perform ECR.

While fractional flow reserve (FFR) standard could be potentially more definitive,[26] the decisions of ECR in this study were mostly based on the visual results of coronary angiography, clinical information, and MPI findings. The data in the registry did not include FFR results and such large multicentre data have not yet been available. An advantage of the per-vessel-based model is that it provides advice on which vascular distribution would be revascularized—this can be helpful in a cath lab to dilemmas with comparable/equivocal stenotic lesions in different coronary vessels.

Compared with TPD, the improvement in AUC by ML approach that integrates a large scale of variables seems modest. These improvements are, however, valuable since they may be close to the best possible prediction performance of revascularization after MPI. Further, AUC is also improved in comparison to human readings where multiple clinical and imaging variables are integrated mentally. Moreover, the demonstration of the rationale for the ML prediction may give more confidence to clinicians and allow easier introduction into clinical routine and acceptance by physicians. Nevertheless, physicians will have to decide ultimately how to take the probability score into account.

To assess overfitting in our ML process, we compared the performance of ML trained on entire dataset and on unseen data (via 10-fold cross-validation) as a function of the number of features and deriving the generalization error accordingly (Supplementary data online, *Figure S2*). The results show that the generalization error stabilizes at 3.0% and is not affected by the number of features, until the model lowers performance above a threshold of IGR >0.002.

Our study has some limitations. While ultimately we would like to predict who could benefit from the revascularization with longer-term outcomes,[27] we have not done it in this study, which merely predicts the fact of performing revascularization within 90 days of MPI. Subjective estimation for the probability of ECR was not performed clinically and we were therefore not able to compare between the best performance of human expert and ML. There has not been clinical standard established for this purpose. As a proxy, we utilized the clinical expert interpretation and segmental scores when available. Our subpopulation analysis revealed that the AUCs of the summed scores were not different from the AUCs of automatic quantitation. Hence, the automatic quantitation can be viewed as the currently best-available surrogate of human expert reading. The variables of SPECT myocardial blood flow were not used because such data were not available in the registry. Potentially adding absolute flow measurements could further improve the prediction of ECR.[28] While we performed a feature selection step, this was performed with a fixed threshold which rejected only truly redundant features,

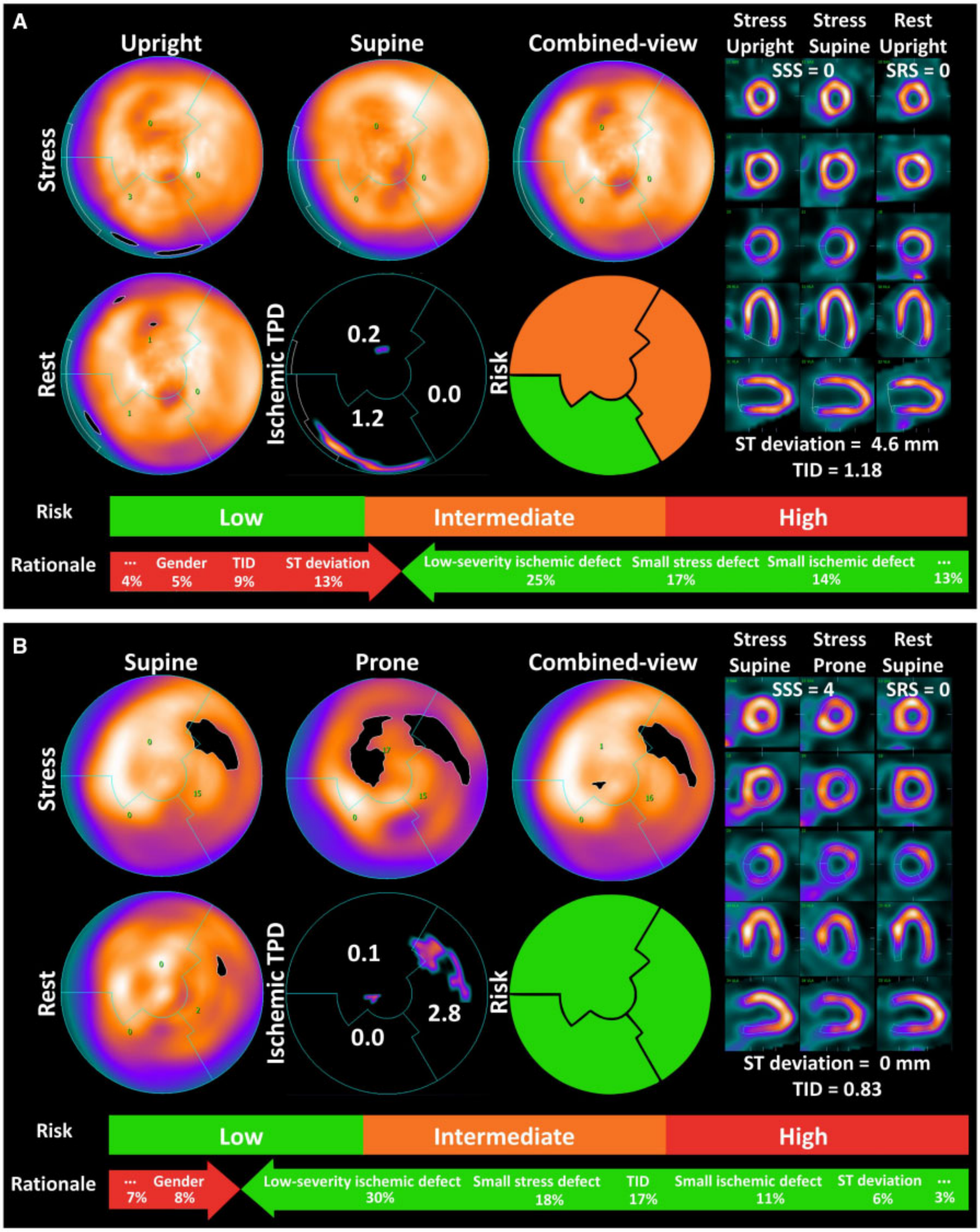

Figure 5 Prediction by ML: example and the rationale. ECR was correctly predicted in the LAD and LCx in an 82-year-old man with 80% stenoses in the LM and LAD and a 70% stenosis in proximal LCx (*A*). The findings on the perfusion scan were nearly normal—the typical presentation of balanced ischaemia. Despite the expert interpretation of the imaging as normal, the patient was sent for invasive angiography for a marked ST-segment deviation during the stress test. In case (*B*), ML correctly predicted no ECR in a 62-year-old man with a reversible perfusion defect in the LCx territory. The expert interpretation was abnormal. Invasive angiography revealed non-significant stenosis and no ECR was performed. The low, intermediate, and high risks indicate a 25%, 55%, and 75% risk of ECR, respectively. The rationale for ML prediction is shown at the bottom of the graph. This can be assessed on a per-vessel basis; to simplify, scales of an averaged contribution of leading variables are shown on a per-patient basis. Presenting ML rationale might help physicians in having confidence in suggestions from ML models. ECR, early coronary revascularization; LAD, left anterior descending artery; LCx, left circumflex artery; LM, left main coronary artery; ML, machine learning; RCA, right coronary artery; SRS, summed rest score; SSS, summed stress score; TID, transient ischemic dilation; TPD, total perfusion deficit.

we did not perform the optimization for the minimal number of features required to accurately predict the outcome. We could not distinguish elective or emergent revascularization in this study. However, the incidence of emergent revascularization should have accounted for a small population in our registry. Finally, all results reported in this work all came from the 10 models in the 10-fold cross-validation, each model tested in unseen data by this specific model. However, another cohort from other hospitals beyond the cases used from nine centres to provide external validation of the model would be required to demonstrate how the model generalizes to data from other centres, beyond the multicentre data used here.

Conclusions

In patients with suspected CAD, ML integrating clinical, stress testing, and MPI imaging data predicts early revascularization better than individual quantitative measures—stress TPD, combined-view stress TPD, ischaemic TPD—on a per-vessel and per-patient basis. ML also outperforms clinical interpretation by a human expert on a per-patient basis.

Supplementary data

Supplementary data are available at *European Heart Journal - Cardiovascular Imaging* online.

Acknowledgements

The authors want to thank all the people whose efforts allowed us to collect, to process, and to analyse the data in the National Institutes of Health-sponsored REFINE SPECT registry.

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: P.J.S.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Dr Hu also received the funding from Taipei Veterans General Hospital-National Yang-Ming University Excellent Physician Scientists Cultivation Programme, No. 106-V-A-007.

Conflict of interest: G.G., 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., S.D., A.J.E., and E.J.M. have served as consultants for GE Healthcare. S.D. has served as a consultant to Bracco Diagnostics; her institution has received grant support from Astellas. 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. D.S.B.'s institution has received grant support from HeartFlow. All other authors have reported that they have no relationships relevant to the contents of this paper to disclose.

References

1. Fihn SD, Gardin JM, Abrams J, Berra K, Blankenship JC, Dallas AP. 2012 ACCF/AHA/ACP/AATS/PCNA/SCAI/STS Guideline for the diagnosis and management of patients with stable ischemic heart disease: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines, and the American College of Physicians, American Association for Thoracic Surgery, Preventive Cardiovascular Nurses Association, Society for Cardiovascular Angiography and Interventions, and Society of Thoracic Surgeons. *J Am Coll Cardiol* 2012;**60**:e44–164.
2. Romero-Farina G, Candell-Riera J, Aguadé-Bruix S, Ferreira-Gonzalez I, Igual A, García-Dorado D. Relationship between myocardial perfusion-gated SPECT and the performance of coronary revascularization in patients with ischemic cardiomyopathy. *Clin Nucl Med* 2012;**37**:965–70.
3. Dong W, Wang Q, Gu S, Su H, Jiao J, Fu Y. Cardiac hybrid SPECT/CTA imaging to detect 'functionally relevant coronary artery lesion': a potential gatekeeper for coronary revascularization? *Ann Nucl Med* 2014;**28**:88–93.
4. Hachamovitch R, Hayes SW, Friedman JD, Cohen I, Berman DS. Comparison of the short-term survival benefit associated with revascularization compared with medical therapy in patients with no prior coronary artery disease undergoing stress myocardial perfusion single photon emission computed tomography. *Circulation* 2003;**107**:2900–7.
5. Arsanjani R, Dey D, Khachatryan T, Shalev A, Hayes SW, Fish M *et al*. Prediction of revascularization after myocardial perfusion SPECT by machine learning in a large population. *J Nucl Cardiol* 2015;**22**:877–84.
6. Slomka PJ, Betancur J, Liang JX, Otaki Y, Hu LH, Sharir T. Rationale and design of the REgistry of Fast Myocardial Perfusion Imaging with NExt generation SPECT (REFINE SPECT). *J Nucl Cardiol* 2018;doi:10.1007/s12350-018-1326-4.
7. Betancur J, Commandeur F, Motlagh M, Sharir T, Einstein AJ, Bokhari S *et al*. Deep learning for prediction of obstructive disease from fast myocardial perfusion SPECT: a multicenter study. *JACC Cardiovasc Imaging* 2018;**11**:1654–63.
8. Betancur J, Rubeaux M, Fuchs TA, Otaki Y, Arnson Y, Slipczuk L *et al*. Automatic valve plane localization in myocardial perfusion SPECT/CT by machine learning: anatomic and clinical validation. *J Nucl Med* 2017;**58**:961–7.
9. Nakazato R, Tamarappoo BK, Kang X, Wolak A, Kite F, Hayes SW *et al*. Quantitative upright-supine high-speed SPECT myocardial perfusion imaging for detection of coronary artery disease: correlation with invasive coronary angiography. *J Nucl Med* 2010;**51**:1724–31.
10. Ben-Haim S, Almukhailed O, Neill J, Slomka P, Allie R, Shiti D *et al*. Clinical value of supine and upright myocardial perfusion imaging in obese patients using the D-SPECT camera. *J Nucl Cardiol* 2014;**21**:478–85.
11. Nishina H, Slomka PJ, Abidov A, Yoda S, Akincioglu C, Kang X *et al*. Combined supine and prone quantitative myocardial perfusion SPECT: method development and clinical validation in patients with no known coronary artery disease. *J Nucl Med* 2006;**47**:51–8.
12. Henzlova MJ, Duvall WL, Einstein AJ, Travin MI, Verberne HJ. ASNC imaging guidelines for SPECT nuclear cardiology procedures: stress, protocols, and tracers. *J Nucl Cardiol* 2016;**23**:606–39.
13. Germano G, Berman D. *Clinical Gated Cardiac SPECT*. 2nd ed. Malden, MA, USA: Wiley-Blackwell; 2008. p384.
14. Witten I, Frank E, Hall M, Pal C. *Data Mining: Practical Machine Learning Tools and Techniques*. 4th ed. Cambridge, MA, USA: Morgan Kaufmann; 2016.
15. Friedman J, Hastie T, Tibshirani R. Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors). *Ann Statist* 2000;**28**:337–407.
16. Molinaro AM, Simon R, Pfeiffer RM. Prediction error estimation: a comparison of resampling methods. *Bioinformatics* 2005;**21**:3301–7.
17. 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.
18. 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.
19. Molodianovitch K, Faraggi D, Reiser B. Comparing the areas under two correlated ROC curves: parametric and non-parametric approaches. *Biom J* 2006;**48**:745–57.
20. Arsanjani R, Xu Y, Dey D, Fish M, Dorbala S, Hayes S *et al*. Improved accuracy of myocardial perfusion SPECT for the detection of coronary artery disease using a support vector machine algorithm. *J Nucl Med* 2013;**54**:549–55.
21. Han D, Lee JH, Rizvi A, Gransar H, Baskaran L, Schulman-Marcus J *et al*. Incremental role of resting myocardial computed tomography perfusion for predicting physiologically significant coronary artery disease: a machine learning approach. *J Nucl Cardiol* 2018;**25**:223–33.
22. Arsanjani R, Xu Y, Dey D, Vahistha V, Shalev A, Nakanishi R *et al*. Improved accuracy of myocardial perfusion SPECT for detection of coronary artery disease by machine learning in a large population. *J Nucl Cardiol* 2013;**20**:553–62.

23. Ben-Haim S, Gips S, Merdler A, Front A, Tamir A. Myocardial stunning demonstrated with rest and post-stress measurements of left ventricular function using dual-isotope gated myocardial perfusion SPECT. *Nucl Med Commun* 2004;**25**: 657–63.
24. Boden WE, O'Rourke RA, Teo KK, Hartigan PM, Maron DJ, Kostuk WJ *et al*. Optimal medical therapy with or without PCI for stable coronary disease. *N Engl J Med* 2007;**356**:1503–16.
25. Patel MR, Dehmer GJ, Hirshfeld JW, Smith PK, Spertus JA. ACCF/SCAI/STS/AATS/AHA/ASNC/HFSA/SCCT 2012 appropriate use criteria for coronary revascularization focused update: a report of the American College of Cardiology Foundation Appropriate Use Criteria Task Force, Society for Cardiovascular Angiography and Interventions, Society of Thoracic Surgeons, American Association for Thoracic Surgery, American Heart Association, American Society of Nuclear Cardiology, and the Society of Cardiovascular Computed Tomography. *J Am Coll Cardiol* 2012;**59**:857–81.
26. 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.
27. Hachamovitch R, Rozanski A, Shaw LJ, Stone GW, Thomson LE, Friedman JD *et al*. Impact of ischaemia and scar on the therapeutic benefit derived from myocardial revascularization vs. medical therapy among patients undergoing stress-rest myocardial perfusion scintigraphy. *Eur Heart J* 2011;**32**:1012–24.
28. Agostini D, Roule V, Nganoa C, Roth N, Baavour R, Parienti JJ *et al*. First validation of myocardial flow reserve assessed by dynamic (99m)Tc-sestamibi CZT-SPECT camera: head to head comparison with (15)O-water PET and fractional flow reserve in patients with suspected coronary artery disease. The WATERDAY study. *Eur J Nucl Med Mol Imaging* 2018;**45**:1079–90.
29. Singh B, Bateman TM, Case JA, Heller G. Attenuation artifact, attenuation correction, and the future of myocardial perfusion SPECT. *J Nucl Cardiol* 2007;**14**: 153–64.