HHS Public Access

Author manuscript

Circ Cardiovasc Imaging. Author manuscript; available in PMC 2023 March 23.

Published in final edited form as:

Circ Cardiovasc Imaging. 2022 September ; 15(9): e014526. doi:10.1161/CIRCIMAGING.122.014526.

Deep learning for explainable estimation of mortality risk from myocardial positron emission tomography images

Ananya Singh, MS[a], **Jacek Kwiecinski, MD, PhD**[a,b], **Robert JH Miller, MD**[a,c], **Yuka Otaki, MD, PhD**[a], **Paul B. Kavanagh, MS**[a], **Serge D. Van Kriekinge, PhD**[a], **Tejas Parekh, BS**[a], **Heidi Gransar, MS**[a], **Konrad Pieszko, MD, PhD**[a,d], **Aditya Killekar, MS**[a], **Ramyashree Tummala, MD**[a], **Joanna X. Liang, BA**[a], **Marcelo Di Carli, MD**[e], **Daniel S. Berman, MD**[a], **Damini Dey, PhD**[a], **Piotr J. Slomka, PhD**[a]

[a.]Departments of Medicine (Division of Artificial Intelligence in Medicine), Imaging and Biomedical Sciences, Cedars-Sinai Medical Center, Los Angeles, CA, USA

[b.]Department of Interventional Cardiology and Angiology, Institute of Cardiology, Warsaw, Poland

[c.]Department of Cardiac Sciences, University of Calgary, Calgary AB, Canada

[d.]Department of Interventional Cardiology and Cardiac Surgery, Collegium Medicum, University of Zielona Góra, Zielona Góra, Poland

[e.]Division of Nuclear Medicine and Molecular Imaging, Department of Radiology, Brigham and Women's Hospital, Boston, MA, USA

Abstract

Background: We aim to develop an explainable deep learning (DL) network for the prediction of all-cause mortality (ACM) directly from positron emission tomography (PET) myocardial perfusion imaging (MPI) flow and perfusion polar map data and evaluate it using prospective testing.

Methods: A total of 4,735 consecutive patients referred for stress and rest ^{82}Rb PET between 2010–2018 were followed-up for ACM for 4.15 [2.24–6.3] years. DL network utilized polar maps of stress and rest perfusion, myocardial blood flow, myocardial flow reserve (MFR), and spill-over fraction combined with cardiac volumes, singular indices, and sex. Patients scanned from 2010–2016 were used for training and validation. The network was tested in a set of 1,135 patients scanned from 2017–2018 to simulate prospective clinical implementation.

Results: In prospective testing, the area under the receiver operating characteristic curve for ACM prediction by DL (0.82[95% CI: 0.77–0.86]) was higher than ischemia (0.60[95% CI: 0.54–0.66], p<0.001), MFR (0.70[95% CI: 0.64–0.76], p<0.001) or a comprehensive logistic regression model (0.75[95% CI: 0.69–0.80], p < 0.05). The highest quartile of patients by DL had an annual ACM rate of 11.87% and had a 16.8[95% CI: 6.12–46.3%, p<0.001]-fold increase in the risk of death compared to the lowest quartile patients. DL showed a 21.6% overall reclassification improvement as compared to established measures of ischemia.

Address for correspondence: Piotr J. 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, Piotr.Slomka@cshs.org, Phone: 310-423-4348 Fax: 310-423-0173, Twitter: @Piotr_JSlomka.

Conclusion: The DL model trained directly on polar maps allows improved patient risk stratification in comparison to established methods for PET flow or perfusion assessments.

Keywords

Artificial intelligence; all-cause mortality; deep learning; myocardial flow reserve; positron emission tomography; risk stratification

INTRODUCTION

Positron emission tomography (PET) myocardial perfusion imaging (MPI) provides a non-invasive assessment of myocardial blood flow (MBF) in addition to regional myocardial perfusion and left ventricular (LV) function[1, 2]. PET quantification of MBF and myocardial flow reserve (MFR) allows robust diagnosis of obstructive coronary artery disease (CAD)[3] and serves as a powerful risk stratification tool to guide management of patients with suspected or known atherosclerotic coronary artery disease[4, 5, 6, 7, 8, 9]. This noninvasive quantitative assessment of coronary vasodilator function provides incremental risk stratification over the combination of comprehensive clinical assessment, LV systolic function, and semiquantitative measures of myocardial ischemia and scar for identification of patients at risk of cardiac mortality[4]. However, it is challenging to optimally integrate the multiparametric imaging information derived from cardiac PET images (i.e., regional perfusion, MBF, and function data) for risk assessment at the point of care.

The recent introduction of highly efficient, image-based artificial intelligence (AI) methods is revolutionizing medical image processing[10]. AI has been used to improve risk stratification[11, 12] and deep learning has been used to provide automated predictions for presence of obstructive CAD[13]. These advances enable automatic image analysis and integration of rich multidimensional PET datasets for enhanced risk stratification. In this study, we developed a novel explainable deep-learning (DL) network for the prediction of all-cause mortality (ACM) directly from quantitative MBF, myocardial perfusion, and function PET imaging data and tested it in a prospective regimen.

METHODS

Study Population

To the extent allowed by data sharing agreements and institutional review boards (IRB) protocols, the data and code from this manuscript will be shared upon written request. We included 4,761 consecutive patients referred for rest/pharmacologic stress MPI with ^{82}Rubidium (^{82}Rb) PET on clinical grounds at Cedars-Sinai Medical Center from 2010–2018. 26 patients were excluded due to errors in acquisition or corruption of data. For patients with multiple studies during the study period, only the first exam was considered. Studies collected between 2010–2016 were used for training and model optimization. Studies between 2017–2018 were held out for prospective testing. In this set, we excluded patients who underwent early revascularization (percutaneous coronary intervention or coronary artery bypass surgery) within 90-days from the index PET scan. Clinical and demographic data were prospectively collected on the day of the PET scan. The first author

had full access to all the data in the study and takes responsibility for its integrity and data analysis. IRB approval was obtained and the study complies with the Declaration of Helsinki and was approved by the institutional review board at Cedars-Sinai Medical Center.

Image Acquisition

All patients underwent same-day rest-stress gated ^{82}Rb PET MPI on a hybrid Biograph 64 (n=4,729, Siemens Healthcare, Erlangen, Germany) PET/CT scanner or GE Discovery 710 (n=6, GE Healthcare, Wakesha, Wisconsin) scanner. A 6-minute rest list-mode acquisition was started immediately before the injection of weight-based 925–1,850 MBq (25–50 mCi) of ^{82}Rb for each rest and stress study. The estimated total radiation dose to the patient of ^{82}Rb for the entire rest/stress study was 1.5–3 mSv. Pharmacologic stress was then performed using intravenous regadenoson (n = 4,298), adenosine (n = 423), dobutamine (n = 9) or dipyridamole (n = 5). A 6-minute stress list-mode acquisition was started concurrently with the ^{82}Rb injection. Adenosine infusion lasted 7-minutes with ^{82}Rb injected at 1.5-minutes into the adenosine administration with concurrent initiation of the dynamic acquisition. A low-dose helical CT was acquired prior to the PET acquisition for attenuation correction.

Semiquantitative Assessment of Regional Myocardial Perfusion

Semiquantitative expert interpretation was performed during clinical reporting, in accordance with societal guidelines, by experienced board-certified nuclear cardiologists with knowledge of all available data including stress and rest perfusion, gated functional data, and clinical information[14]. The regional myocardium perfusion assessment was performed according to a five-point scale (0=normal to 4=absence of detectable tracer uptake) and 17-segment model[14]. Segmental scoring was guided by quantitative analysis (QPET, Cedars-Sinai, Los Angeles, CA) and with knowledge of clinical information. Summed stress, rest and difference myocardial perfusion scores (SSS, SRS and SDS) were generated. Perfusion stress and rest defect were defined as SSS and SRS respectively. SDS reflecting ischemic burden was calculated as the difference between the summed stress and rest scores. SDS was converted to percent myocardium ischemic by dividing summed scores by 68 for studies and multiplying by 100[15]. Moderate-to-severe ischemia was defined as ≥10% of the left ventricular myocardium[16].

Automated Image Segmentation

LV contours were positioned automatically as previously described[17]. For MBF and perfusion analyses, LV contours were determined from the attenuation-corrected high-resolution summed dynamic image data, skipping first 2-minutes for improved valve-plane definition.

Myocardial blood flow quantification

Quality control was performed by a trained technician at the time of image acquisition. Rest/stress MBF and MFR calculated with a 1-tissue compartment kinetic model were obtained with clinical software (QPET, Cedars-Sinai, Los Angeles, California)[17, 18]. MBF and spillover fraction from blood to myocardium were computed by numeric optimization.

Stress and rest flow values in mL/g/min were computed for each sample on the polar map. MFR was computed as the ratio of stress over MBF.

Deep learning input data and architecture

The overview of the DL approach is shown in the Figure 1. DL was trained using the stress and rest polar maps of myocardial perfusion, MBF, MFR and spill-over fraction(SOF) polar maps obtained from QPET, combined with the following imaging numerical variables: sex, shape indexes[19], and LV end-systolic and diastolic volumes. Sex was extracted automatically from digital imaging and communications in medicine (DICOM) image headers. Cardiac volumes and shape indexes were quantified automatically from stress end-systolic and diastolic images. Shape index was defined as the ratio of the maximum 3D short- and long- axis dimensions of the left ventricle[19]. The DL network consisted of convolutional layers followed by a flattening and fully-connected layers. Polar images were input to the convolutional layers which extracts latent spatial information from neighboring pixels in adjacent layers. The normalized numerical imaging variables were concatenated to the first fully connected layer. Extraction of information from the image header and image quantification can be incorporated into the DL pipeline to ensure fully-automated predictions as demonstrated in our previous work[13]. The output of the DL network was the prediction of ACM following sigmoid activation expressed as annualized mortality risk. Additional details of the model architecture are shown in Supplemental Figure 1. To elucidate the potential added value of flow for ACM prediction, another DL model including only perfusion images was trained and evaluated. DL was developed using Python 3.7.3 and PyTorch 1.4.0. Multivariate logistic regression analysis was performed for comparison to the DL model, which included age, sex, MFR, stress and rest TPD, MBF, SOF, cardiac volumes and shape indexes.

Individual attention maps and polar map rankings

We incorporated attention maps in the architecture using gradient-weighted class activation mapping (Grad-CAM)[20]. The algorithm uses the gradients of a prediction flowing into the final convolutional layer to produce an attention map highlighting regions of importance to the network for a prediction. To determine the importance of input polar maps, SHapley Additive exPlanations[21] values per image input were obtained, which assigns each feature an importance value for a particular prediction. These values were converted to percentage contribution per image input using Grad-CAM attention. The output attenion maps and polar map rankings facilitate the explainability of DL predictions for an individual patient.

Clinical follow-up

The cohort was followed for early revascularization and ACM. Follow-up for ACM was obtained using internal hospital records, Social Security Death Index (SSDI), National Death Index (NDI), and California Non-comprehensive Death information until December, 2020. Information regarding early revascularization were collected from hospital records and adjudicated by site physicians, according to standard criteria[22].

Internal model testing

Image datasets of 3,521 patients scanned between 2010–2016, were randomly divided into 10 distinct folds, using a 90% -training and 10% -testing stratified split to perform 10-fold internal repeated testing (Supplemental Figure 2), similar to our previous work[13]. This method of testing conservatively maximizes data available for training while ensuring evaluation of the entire dataset and maintaining an equal proportion of ACM cases in each set. Further details are provided in the supplement. To elucidate the potential added value of age for ACM prediction, another DL model including age was trained and evaluated. The logistic regression model was developed using the same set of patients used for internal repeated testing.

Prospective Testing

To evaluate the performance of DL applied to new real world data, we used a cohort of patients scanned between 2017–2018 as a prospective testing set. The logistic regression model was tested in the same dataset.

Thresholds for comparisons of DL

Patients were divided into corresponding low-/high-risk DL subgroups by idenitifying thresholds, set to match the same proportion of patients as the established respective clinical thresholds for ischemia (≥10%)[23, 24] and MFR (≥1.8)[25].

Statistical Analysis

Categorical variables are presented as frequencies (percentages) and continuous variables as medians (interquartile range). Variables were compared using a Pearson χ^2 statistic for categorical variables and a Wilcoxon rank-sum or Kruskal-Wallis test for continuous variables. We assessed the distribution of data with the Shapiro-Wilk test. Prediction of ACM by ischemia, MFR, and DL were evaluated using the analysis and pairwise comparisons of the areas under the receiver operating characteristic curve (AUC) according to DeLong et al.[26]. The results of ACM prediction by DL were visualized with Kaplan-Meier survival curves. Risk discrimination was assessed by time-dependent net reclassification improvement (NRI) at 2-years as described by Pencina et al.[27]. Confidence intervals were calculated using bootstrapping. Additionally, comparisons of mortality rate between groups of low- and high-risk DL, MFR and ischemia were performed using the “partiallyoverlapping” package of R (version 2.0).

A two-tailed p-value <0.05 was considered statistically significant. All analyses including survival were performed using “nricens” package (version 1.6) and “survival” package (version 3.2–13) in R (version 4.0.3, R Foundation, Vienna, Austria) within R studio version 1.4.1717 (RStudio, Boston, MA).

RESULTS

Patient Characteristics

Patient characteristics are shown in Table 1. Women represented about 40% of the cohort. Cardiovascular risk factors were prevalent in both groups including diabetes (35% and 32%),

hypertension (79% and 77%) and dyslipidemia (66% and 72%). About a quarter of the patients had a history of prior CAD. Total of 4,723 consecutive patients referred for rest and stress ^{82}Rb PET between 2010–2018 were followed for ACM for a median of 4.1 [IQR: 2.2–6.3] years. Of these, 3,521 patients (792 deaths, years 2010–2016) were used for training and optimization of the algorithm. After excluding early revascularization, the prospective testing set consisted of 1,135 patients (85 deaths, years 2017–2018).

Predictive Performance

In the prospective testing set, we compared the performance of DL with traditional clinical measures and a comprehensive logistic regression model. ACM prediction by DL (0.82[95% CI: 0.77–0.86]) was higher than the ischemia, derived from SDS (0.60[95% CI: 0.54–0.66], p<0.001), MFR (0.70[95% CI: 0.64–0.76], p<0.001) or logistic regression (0.75[95% CI: 0.69–0.80], p<0.05) (Figure 2). Adding age to the pure image driven DL model did not significantly improve the prediction of ACM (0.83[95% CI: 0.79–0.87]). We have also performed separate analyses within predefined subpopulations. These include patients with and without perfusion abnormalities (SSS=0 versus SSS>1), as well as a gender specific analysis (Figure 3). Across these populations the DL model maintained the high diagnostic performance (all AUC >0.77). A separate DL model trained with only perfusion images had an inferior predictive performance (0.76[95% CI: 0.71–0.81], p<0.001) compared to the comprehensive DL model which included flow images.

Survival Analysis

Kaplan-Meier curves for patients with low and high DL scores matched with ischemia thresholds (<10% and ≥10%) (corresponding DL threshold:0.645), are shown in Figure 4. Deep learning led to improved risk reclassification of patients who experienced ACM (20%[95% CI: 34.7–6.7%, p<0.01]) and patients who did not experience ACM (1.6%[95% CI: 3.6 – −0.3%, p=0.06]), with an overall net reclassification improvement of 21.6%[95% CI: 36.7–8.2%, p<0.01]. The risk associated with high DL risk (14.3 events/ 100 patient years) which was present in 102 patients was similar to the risk associated with having >23% myocardial ischemia (13.5 events/ 100 patient years) which was present in 9 patients. Kaplan-Meier curves for patients with low and high DL scores matched with low and high MFR (≥1.8 and <1.8) (corresponding DL threshold was 0.531), are shown in Figure 5. Here again, DL led to improved risk reclassification of patients who experienced ACM (22.4%[95% CI: 38.9–7.2%, p<0.01]) and patients who did not experience ACM (1.8%[95% CI: 4.2 – −0.6%, p = 0.07]), with overall net reclassification improvement of 24.2%[95% CI: 38.9–8.2%, p=0.036]. Additionally, annualized mortality for the low- and high-score groups based on DL and ischemia and DL and MFR are shown in Supplemental Figure 3. The highest quartile of patients by DL had an annual ACM rate of 11.87% and had a 16.8-fold [95% CI: 6.12–46.3%, p<0.001] increase in the risk of death, compared to the lowest quartile patients which had an annual ACM rate of 0.7% (Supplemental Figure 4). The highest quartile of patients by MFR had a 4.5-fold [95% CI: 2.35–8.7%, p<0.001] and by Ischemia had a 2.77-fold [95% CI: 1.5–5.11%, p<0.001] compared to the respective lowest quartile patients.

Individualized risk inference by the algorithm

Results were generated in approximately 1-second per-patient during testing. All results were generated in batch mode. For each patient, DL highlighted regions associated with ACM risk directly on PET polar maps using Grad-CAM, along with a continuous probability of ACM risk and patient specific contribution of each polar map to the outcome prediction. In the overall prospective testing population, MFR map was most commonly the highest contributing polar map, towards risk prediction by DL (Table 2). Case example is shown in Supplemental Figure 5 with individual risk prediction, attention map and polar maps ranked by importance in contribution towards risk.

DISCUSSION

We demonstrated that an explainable DL model, combining multiparametric PET imaging data, has superior accuracy for mortality prediction compared to established measures of myocardial perfusion or flow. The model attained high accuracy in prospective testing on a separate dataset by objectively integrating absolute flow data and relative perfusion maps, along with other imaging parameters, a task which is challenging to accomplish at the point of care. Importantly, our DL model was derived directly from polar maps, without the need for derivation and selection of multiple quantitative measurements and without the need for using normal limits. DL also incorporates methods to introduce “explainability” in prognostic predictions by highlighting regions contributing to DL score on polar maps and ranking the relative contribution from different inputs for a specific patient. This would allow physicians to understand and improve confidence in the DL results, which may potentially help to overcome the perception of AI as a “black box”.

PET MPI provides robust risk stratification by assessing both relative and absolute myocardial perfusion. In our study, DL demonstrated improved risk stratification compared to either imaging marker alone or a comprehensive logistic regression model. Our approach showed a 16.8-fold increased rate of death in patients in the highest quartile compared to the lowest quartile of DL risk. Compared to standard measurements of ischemia burden, the model tested in a separate sample improved overall patient risk reclassification by over 21.6%. Likewise, the DL model led to a 24.2% risk re-reclassification over MFR measurements. While risk estimation by DL is beneficial to patients in both low- and high-risk groups of ischemia and MFR, it is better suited for identification of high-risk patients. Importantly, as seen in Figure 3, our approach maintained high predictive performance for PET MPI scans with visually normal perfusion, and performed equally well for men and women. Overall, these results suggest that a DL model incorporating regional perfusion, flow and other imaging characteristics has the potential to significantly improve risk stratification and enable dissemination of advanced PET MPI analysis outside experienced centers.

The complex interplay of various perfusion, flow and functional imaging estimates from PET has been extensively studied, but the optimal method for integrating multiple measures remains elusive. Ziadi et al. demonstrated that MFR provides independent and incremental prediction for major adverse cardiac events when combined with visually assessed regional perfusion[6]. Similarly, Murthy et al. found improved classification of the risk of cardiac death

by incorporating MFR and ischemia[2]. Gould et al. established that coronary flow capacity, which integrates regional stress flow and MFR into one variable, can identify patients at high risk of cardiovascular events[28] and also proposed the use of longitudinal flow and the base to apex perfusion gradient[29] in addition to global flow measures. Gupta et al. demonstrated that integrating MFR and maximal MBF with customized thresholds could aid in stratifying cardiovascular mortality[30]. Importantly, the above studies did not separate data for model development and testing; therefore, the benefit of these parameters may be overestimated. In contrast, we employed rigorous temporal data separation for development and evaluation. In our approach, by supplying perfusion, flow, and flow reserve polar images directly to the DL model, we leverage the totality of imaging data Incorporating numerical image derived data such as sex in the DL model, adds information missing from polar map data, improving the reliability of the predictions. Similar approaches have been applied to integrate body mass index and sex into DL models[13, 31].This approach allows the DL model to implicitly account for complex interactions between flow, perfusion, and function, without the need to derive or establish custom image parameters. Additionally, DL is inherently capable of handling non-linear associations between variables and outcomes which may be particularly relevant to ^{82}Rb MBF measurements given its low extraction fraction. In our study, the MFR polar maps and stress perfusion were the most important contributors to predictions on a population level. However, on a per-patient basis, the importance of DL inputs can vary, suggesting that objective methods to integrate all parameters are critical when predicting patient-specific risk.

This DL model was developed to predict risk of ACM and indicate which imaging feature is driving the prediction, which we speculate might be important to guide management decisions. While several observational studies have suggested that ischemia, assessed by regional myocardial perfusion, may identify patients who benefit from revascularization[16, 25, 32], this was not the case in the ISCHEMIA trial[33]. However, other parameters provided by PET may help with prediction of revascularization benefit. For instance, Taqueti et al. have previously shown that global MFR modified the effect of revascularization, such that only patients with low MFR appeared to benefit from revascularization[7]. In a large cohort, Patel et al. reported potential benefit from revascularization in patients with MFR<1.8[5]. Likewise, observational studies have shown that measures of coronary flow capacity identified patients with reduced mortality after revascularization[28]. A DL approach, similar to the one presented here, could potentially be used to integrate all available PET imaging data to optimally select patients who might benefit from revascularization.

AI has the potential to address the lack of integrative tools for flow, perfusion and functional PET imaging data, but for integration into clinical practice, predictions need to be explained to physicians to enable clinical oversight, ensure the validity of the inference, and provide a framework for management decisions. By identifying areas of polar maps contributing to risk predictions and ranking the polar maps according to their contribution to prediction of outcomes, our DL model may allow physicians to better understand the key drivers of the individual risk. This may also allow physicians to identify potential errors in risk predictions by DL or areas that require additional scrutiny on the clinical interpretation. By informing the clinician which unique prognostic imaging markers are driving the overall

risk score for a given patient (flow reserve, flow, or perfusion), the proposed tool could potentially aid clinical decision making including individualized allocation to therapy. Previously such considerations were addressed by establishing high-risk thresholds and considering a limited number of variables. This multiparametric independent data then had to be weighed during interpretation by the reading physician for establishing the diagnosis and recommendation[28, 30]. The proposed explainable AI approach has the potential to inform and aid the physician in a quantitative and objective fashion.

STUDY LIMITATIONS

Our study has limitations. It is a post-hoc analysis of single-center data. Thus, we were not able to perform external testing of the developed model, but instead conducted prospective testing. Our finding requires confirmation in larger, external datasets. The model was developed and tested for ^{82}Rb perfusion data only studies; the generalizability of our findings to other PET MPI tracers remains to be evaluated. Additionally, we included patients who received a variety of pharmacologic stress agents without correction for differences in change in peak blood flow. While a more homogenous population may improve prediction performance, this suggests the model is broadly generalizable. We used all-cause death and not cardiac mortality as the end point; however, identifying underlying cause of death is often of limited reliability[34] and would result in fewer events. We excluded patients who underwent early revascularization, since revascularization can alter myocardial perfusion patterns and the subsequent association with cardiac outcomes. Since myocardial perfusion is a major driver of decision to pursue revascularization, including these patients would likely bias the results towards improved risk stratification by DL. Future studies should also aim to evaluate whether inclusion of clinical, biomarker and multimodality imaging data could further enhance risk stratification.

CONCLUSIONS

The study demonstrated that a deep learning model trained directly with quantitative myocardial blood flow and conventional perfusion polar map image, derived from PET images, improved risk stratification and reclassification compared to conventional prognostic approaches. We also showed that it is possible to highlight specific polar map image regions associated with increased clinical risk and rank polar maps according to their prognostic importance for the specific patient. This provides an explanation of the patient-specific risk estimation at the point of care and may help with translation of AI tools into practice.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

We would like to thank Mark Lemley, Mark Hyun, Keiichiro Kuronuma and all the individuals involved in the collection, processing, and analysis of this data.

SOURCES OF 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

Drs. Berman, Slomka, Van Kriekinge and Mr. Kavanagh participate in software royalties for nuclear cardiology software at Cedars-Sinai Medical Center. Dr. Slomka has received research grant support from Siemens Medical Systems. Dr. Berman has served as consultant for GE Healthcare. Dr. Di Carli has received research grant support from Spectrum Dynamics and consulting honoraria from Sanofi and GE Healthcare. The remaining authors have no relevant disclosures.

Non-standard Abbreviations and Acronyms

ACM	all-cause mortality
AI	artificial intelligence
DL	deep learning
MBF	myocardial blood flow
MFR	myocardial flow reserve
MPI	myocardial perfusion imaging
PET	positron emission tomography

REFERENCES:

1. Gould KL, et al. Anatomic versus physiologic assessment of coronary artery disease. Role of coronary flow reserve, fractional flow reserve, and positron emission tomography imaging in revascularization decision-making. J Am Coll Cardiol 62, 1639–1653 (2013). [PubMed: 23954338]
2. Murthy VL, et al. Clinical quantification of myocardial blood flow using PET: Joint position paper of the SNMMI cardiovascular council and the ASNC.). Springer (2018).
3. Driessen RS, et al. Comparison of Coronary Computed Tomography Angiography, Fractional Flow Reserve, and Perfusion Imaging for Ischemia Diagnosis. J Am Coll Cardiol 73, 161–173 (2019). [PubMed: 30654888]
4. Murthy VL, et al. Improved cardiac risk assessment with noninvasive measures of coronary flow reserve. Circulation 124, 2215–2224 (2011). [PubMed: 22007073]
5. Patel KK, et al. Myocardial blood flow reserve assessed by positron emission tomography myocardial perfusion imaging identifies patients with a survival benefit from early revascularization. Eur Heart J 41, 759–768 (2020). [PubMed: 31228200]
6. Ziadi MC, et al. Impaired myocardial flow reserve on rubidium-82 positron emission tomography imaging predicts adverse outcomes in patients assessed for myocardial ischemia. J Am Coll Cardiol 58, 740–748 (2011). [PubMed: 21816311]
7. Taqueti VR, et al. Global Coronary Flow Reserve Is Associated With Adverse Cardiovascular Events Independently of Luminal Angiographic Severity and Modifies the Effect of Early Revascularization. Circulation 131, 19–+ (2015). [PubMed: 25400060]
8. Taqueti VR, et al. Excess cardiovascular risk in women relative to men referred for coronary angiography is associated with severely impaired coronary flow reserve, not obstructive disease. Circulation 135, 566–577 (2017). [PubMed: 27881570]
9. Herzog BA, et al. Long-term prognostic value of 13N-ammonia myocardial perfusion positron emission tomography added value of coronary flow reserve. J Am Coll Cardiol 54, 150–156 (2009). [PubMed: 19573732]

10. Dey D, et al. Artificial Intelligence in Cardiovascular Imaging: JACC State-of-the-Art Review. J Am Coll Cardiol 73, 1317–1335 (2019). [PubMed: 30898208]
11. Rios R, 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 145, 105449 (2022).
12. Rim TH, et al. Deep-learning-based cardiovascular risk stratification using coronary artery calcium scores predicted from retinal photographs. The Lancet Digital Health 3, e306–e316 (2021). [PubMed: 33890578]
13. Otaki Y, et al. Clinical Deployment of Explainable Artificial Intelligence of SPECT for Diagnosis of Coronary Artery Disease. JACC Cardiovasc Imaging, (2021).
14. Cerqueira MD, 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 105, 539–542 (2002). [PubMed: 11815441]
15. Rozanski A, et al. Temporal trends in the frequency of inducible myocardial ischemia during cardiac stress testing: 1991 to 2009. J Am Coll Cardiol 61, 1054–1065 (2013). [PubMed: 23473411]
16. 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 107, 2900–2907 (2003). [PubMed: 12771008]
17. Nakazato R, et al. Automated quantitative Rb-82 3D PET/CT myocardial perfusion imaging: normal limits and correlation with invasive coronary angiography. J Nucl Cardiol 19, 265–276 (2012). [PubMed: 22203445]
18. Dekemp RA, et al. Multisoftware reproducibility study of stress and rest myocardial blood flow assessed with 3D dynamic PET/CT and a 1-tissue-compartment model of 82Rb kinetics. J Nucl Med 54, 571–577 (2013). [PubMed: 23447656]
19. Abidov A, et al. Left ventricular shape index assessed by gated stress myocardial perfusion SPECT: initial description of a new variable. J Nucl Cardiol 13, 652–659 (2006). [PubMed: 16945745]
20. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-cam: Visual explanations from deep networks via gradient-based localization. In: Proceedings of the IEEE international conference on computer vision) (2017).
21. Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. In: Proceedings of the 31st international conference on neural information processing systems) (2017).
22. Hicks KA, et al. 2017 cardiovascular and stroke endpoint definitions for clinical trials. Circulation 137, 961–972 (2018). [PubMed: 29483172]
23. Hachamovitch R, 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. European Heart Journal 32, 1012–1024 (2011). [PubMed: 21258084]
24. Group ITR, et al. International Study of Comparative Health Effectiveness with Medical and Invasive Approaches (ISCHEMIA) trial: Rationale and design. Am Heart J 201, 124–135 (2018). [PubMed: 29778671]
25. Patel KK, et al. Extent of Myocardial Ischemia on Positron Emission Tomography and Survival Benefit With Early Revascularization. J Am Coll Cardiol 74, 1645–1654 (2019). [PubMed: 31558246]
26. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 44, 837–845 (1988). [PubMed: 3203132]
27. Pencina MJ, D'Agostino RB Sr., D'Agostino RB Jr., Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Stat Med 27, 157–172; discussion 207–112 (2008). [PubMed: 17569110]

28. Gould KL, et al. Mortality Prediction by Quantitative PET Perfusion Expressed as Coronary Flow Capacity With and Without Revascularization. Jacc-Cardiovascular Imaging 14, 1020–1034 (2021). [PubMed: 33221205]

29. Gould KL, Johnson NP. Coronary Physiology Beyond Coronary Flow Reserve in Microvascular Angina: JACC State-of-the-Art Review. J Am Coll Cardiol 72, 2642–2662 (2018). [PubMed: 30466522]

30. Gupta A, et al. Integrated Noninvasive Physiological Assessment of Coronary Circulatory Function and Impact on Cardiovascular Mortality in Patients With Stable Coronary Artery Disease. Circulation 136, 2325–2336 (2017). [PubMed: 28864442]

31. Chen X, et al. CT-free attenuation correction for dedicated cardiac SPECT using a 3D dual squeeze-and-excitation residual dense network. Journal of Nuclear Cardiology, 1–16 (2021).

32. Azadani PN, et al. Impact of Early Revascularization on Major Adverse Cardiovascular Events in Relation to Automatically Quantified Ischemia. JACC Cardiovasc Imaging 14, 644–653 (2021). [PubMed: 32828784]

33. Maron DJ, et al. Initial Invasive or Conservative Strategy for Stable Coronary Disease. N Engl J Med 382, 1395–1407 (2020). [PubMed: 32227755]

34. Mant J, et al. Clinicians didn't reliably distinguish between different causes of cardiac death using case histories. Journal of Clinical Epidemiology 59, 862–867 (2006). [PubMed: 16828682]

SUPPLEMENTAL REFERENCE:

1. Kingma DP, Ba J. Adam: A method for stochastic optimization. arXiv preprinarXiv:14126980 2014

CLINICAL PERSPECTIVE:

Explainable AI could be incorporated during routine clinical care, automatically providing mortality risk predictions and explaining them with attention maps and patient specific contribution of each input polar map. Future prospective studies may assess if explainable AI leads to improvement in clinical interpretation, and subsequently improves patient management and clinical outcomes.

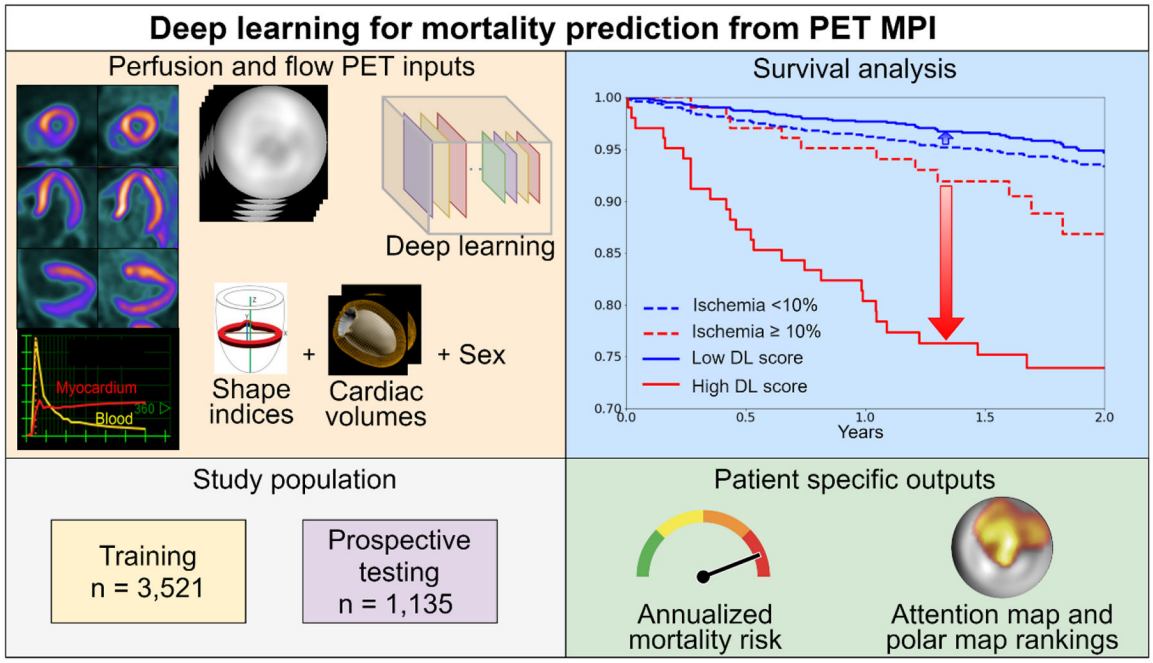

Figure 1: Deep learning for mortality prediction from PET MPI

The stress and rest raw perfusion and flow polar maps derived from the MPI slices and flow images, were input to the DL. Shape indices, cardiac volumes and sex were added as well. The population was divided into training and testing set based on year of scan. The patient specific outputs of the network are the annualized mortality risk and attention map with polar map rankings. Compared to the widely used 10% ischemia threshold, DL was matched by proportion and outperformed ischemia assessments shown by Kaplan-Meier survival graphs.

DL: deep learning, MPI: myocardial perfusion imaging, PET: positron emission tomography

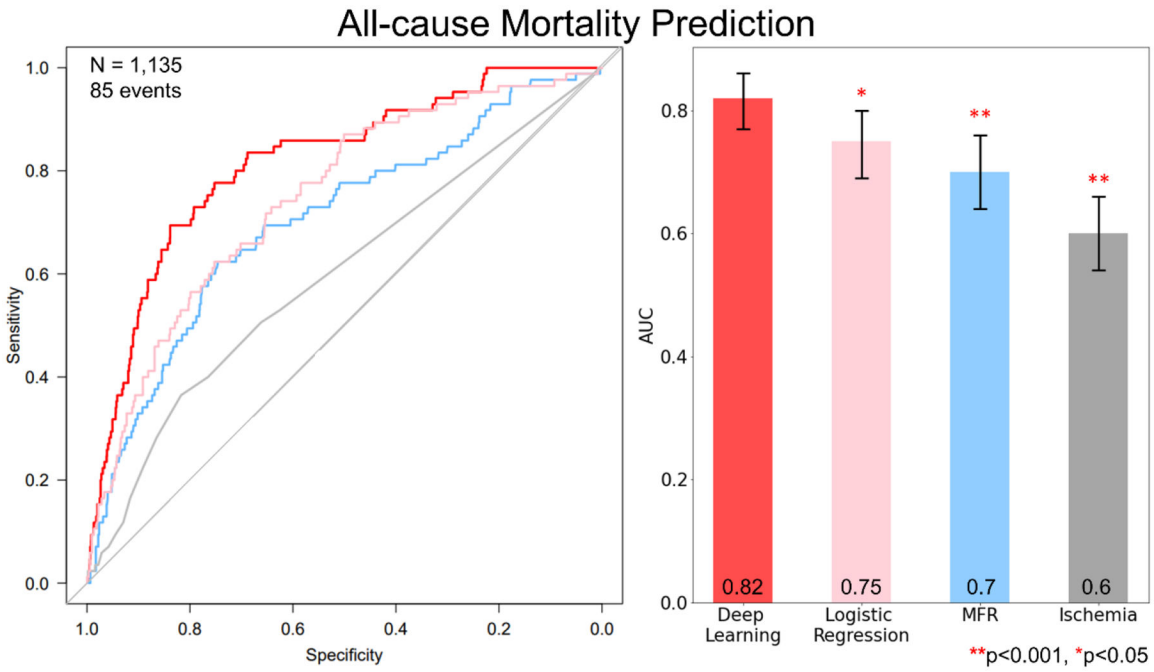

Figure 2: Per-patient AUC (left) and bar plot with AUC values (right) for the prediction of ACM in the prospective testing set by DL, MFR and ischemia

Prediction by DL had a significantly larger area under receiver-operating characteristic curve than measures of percent myocardial ischemia, MFR and logistic regression in the prospective testing set.

AUC: area under curve, DL: deep learning, MFR: myocardial flow reserve

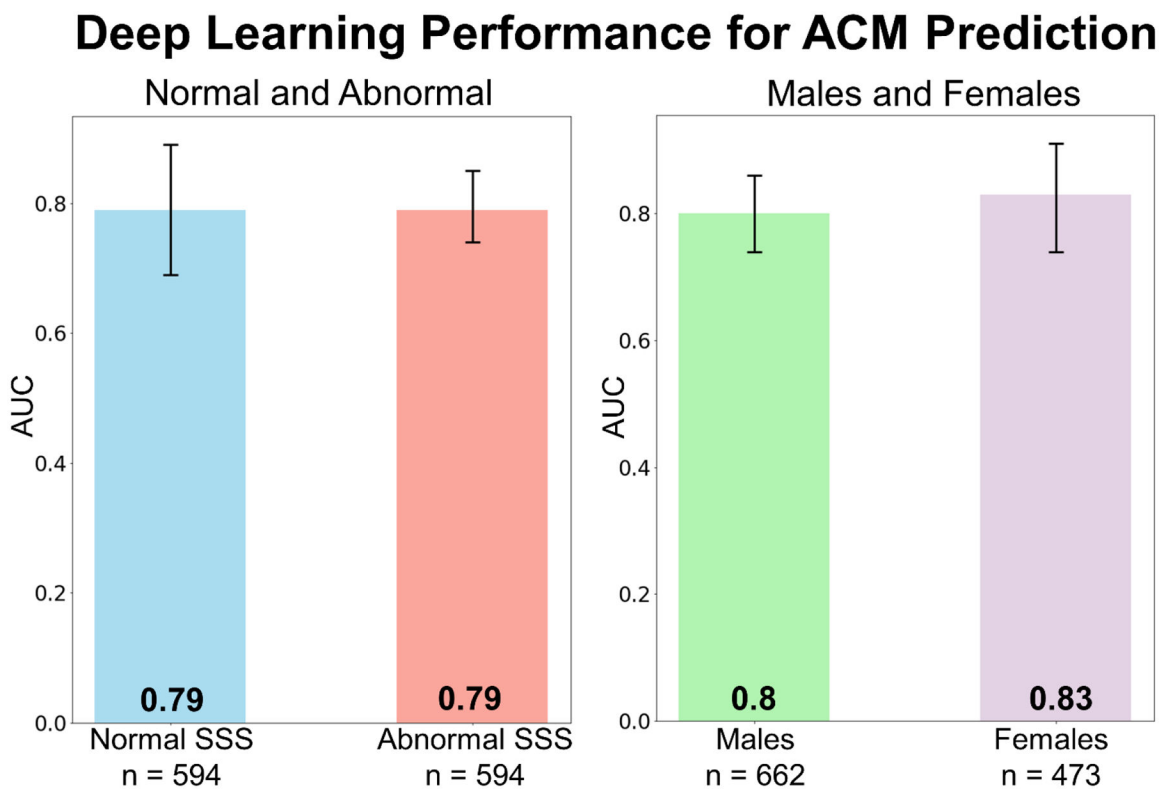

Figure 3: Predictive performance of DL for prediction of ACM in different subgroups

Prediction by DL was similar for patients with normal SSS (= 0) and abnormal SSS (left). The prediction of ACM for females and males was similar as well (right) ensuring generalizability.

DL: deep learning, SSS: summed stress score, ACM: all-cause mortality, AUC: area under curve

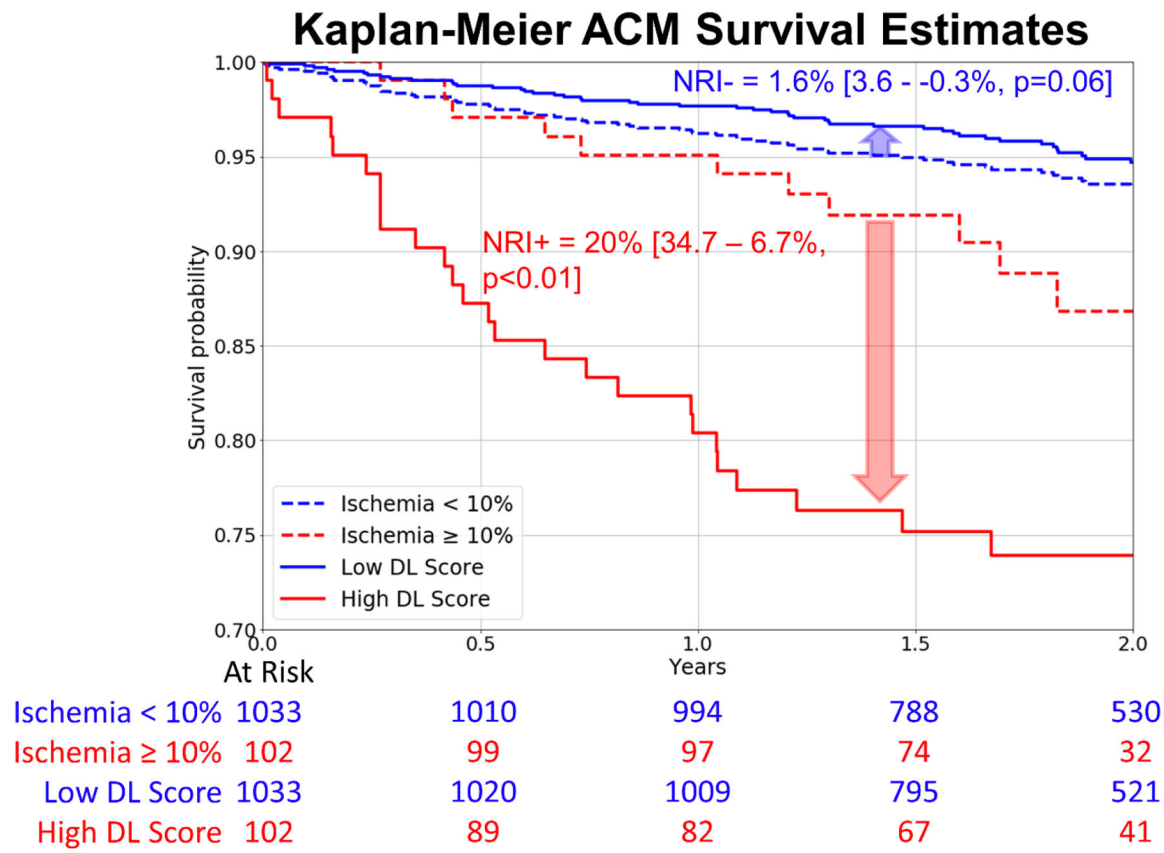

Figure 4: Kaplan-Meier Survival estimates in the prospective testing set for high (red) and low (blue) ischemia matched to DL

Compared to the widely used 10% ischemia threshold, DL was matched by proportion and outperformed ischemia assessments.

ACM: all-cause mortality, DL: deep learning, NRI: net reclassification improvement

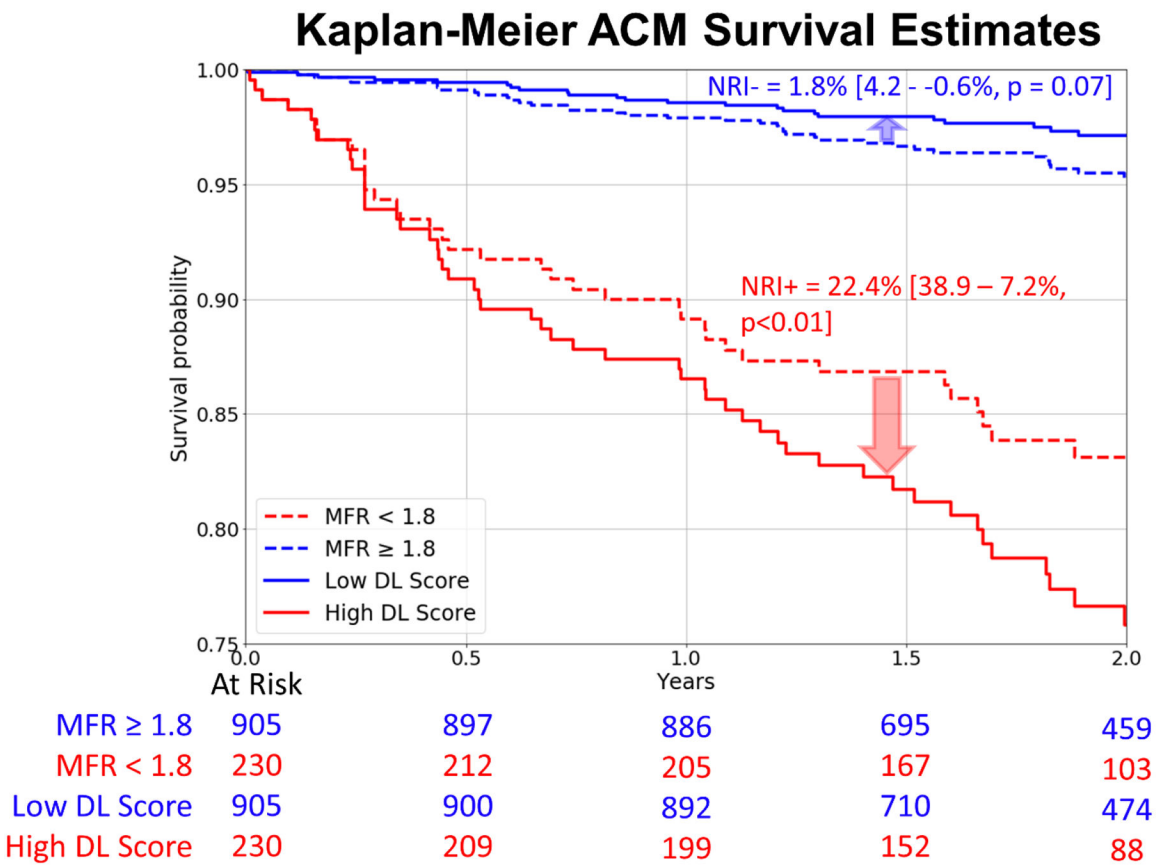

Figure 5: Kaplan-Meier Survival estimates in the prospective testing set for low (red) and high (blue) MFR matched to DL

Compared to the MFR threshold of 1.8, DL was matched by proportion and outperformed MFR assessments.

ACM: all-cause mortality, DL: deep learning, NRI: net reclassification improvement, MFR: myocardial flow reserve

Table 1:

Patient Characteristics of the training and prospective testing datasets

Characteristic	Training dataset, N = 3,521	Prospective testing dataset, n = 1,135	p-values
Age (years)	71 [64 – 79]	71 [65 – 78]	0.908
Females, n(%)	1,443 (41)	473 (42)	0.681
BMI (kg/m^2)	27 [24 – 31]	27 [24 – 32]	0.511
Hypertension, n(%)	2,765 (79)	869 (77)	0.164
Diabetes mellitus, n(%)	1,215 (35)	364 (32)	0.132
Smoking, n(%)	250 (7)	98 (9)	0.087
Family history of CAD, n(%)	493 (14)	254 (22)	<0.001
Prior PCI, n(%)	851 (24)	270 (24)	0.9
Dyslipidemia, n(%)	2,335 (66)	813 (72)	<0.001
Chest pain, n(%)	1,892 (54)	586 (52)	0.216
Dyspnea, n(%)	1,047 (30)	420 (37)	<0.001
Both chest pain and dyspnea, n(%)	521 (15)	231 (20)	<0.001
Statins, n(%)	2,140 (61)	712 (63)	0.24
Beta-blockers, n(%)	1,771 (50)	507 (45)	<0.001
Stress-induced equivocal ECG changes, n(%)	132 (4)	74 (7)	<0.001
Stress-induced ischemic ECG changes, n(%)	102 (3)	33 (3)	0.985
Stress MBF, ml/g/min	2.5 [1.9–3.2]	2.6 [1.9–3.2]	0.118
Rest MBF, ml/g/min	1.1 [0.9–1.4]	1.1 [0.8–1.3]	<0.001
MFR	2.3 [1.7–2.9]	2.5 [1.9–3.1]	<0.001
Ischemia >10%, n(%)	403 (11.4)	102 (9)	0.021
Stress EF (%)	68.5 [56.6–76.1]	69.2 [59.6–76.8]	0.009
Rest EF (%)	64.7 [53.5–72.2]	65.9 [56.1–72.9]	0.008
Stress ED shape index	0.7 [0.6–0.7]	0.7 [0.6–0.7]	0.099
Annualized event rate	6%	3.7%	<0.001

Categorical value is expressed as n(%). Continuous values are expressed as median[interquartile range]. Variables were compared using a Pearson χ^2 statistic for categorical variables and Kruskal-Wallis test for continuous variables.

BMI: body mass index, CAD: coronary artery disease, ECG: electrocardiogram, ED: end diastolic, EF: ejection fraction, IQR: interquartile range, MBF: myocardial blood flow, MFR: myocardial flow reserve, PCTA: Percutaneous transluminal coronary angioplasty

Table 2:

Percentage of highest contributing polar maps ranking for individual predictions among subsets of testing population

Population \ Polar maps	MFR	Stress Perfusion	Stress Flow	Rest Perfusion
Prospective testing, n=1,135	73%	18%	7%	4%
Prior CAD, n=360	82%	10%	6%	1%
No prior CAD, n=775	73%	18%	6%	3%
Men, n=662	75%	15%	5%	3%
Women, n=473	70%	23%	5%	3%

Indicates percentage of patients with specific polar maps ranked highest (top 4 rankings) for individual prediction for prospective testing and, population with and without prior CAD, men and women.

CAD: coronary artery disease, MFR: myocardial flow reserve