Self-supervised deep representation learning of a foundation transformer model enabling efficient ECG-based assessment of cardiac and coronary function with limited labels

Jonathan B. Moody[1], Alexis Poitrasson-Rivière[1], Jennifer M. Renaud[1], Tomoe Hagio[1], Fares Alahdab[2], Mouaz H. Al-Mallah[2], Michael D. Vanderver[1], Edward P. Ficaro[1,3], and Venkatesh L. Murthy[3]

2024-05-21

[1] INVIA, LLC, Ann Arbor, MI, USA

[2] Houston Methodist DeBakey Heart and Vascular Center, Houston, TX USA

[3] University of Michigan, Division of Cardiovascular Medicine, Department of Internal Medicine, Ann Arbor, MI, USA

Background: Although deep learning methods have shown great promise for identification of structural and functional cardiac abnormalities using electrocardiographic data, these methods are data hungry, posing a challenge for critically important tasks where ground truth labels are relatively scarce. Impaired coronary microvascular and vasomotor function is difficult to identify with standard clinical methods of cardiovascular testing such as coronary angiography and noninvasive single photon emission tomography (SPECT) myocardial perfusion imaging (MPI). Gold standard data from positron emission tomography (PET) are gaining emphasis in clinical guidelines but are expensive and only available in relatively limited centers. We hypothesized that signals embedded within resting and stress electrocardiograms (ECGs) identify individuals with microvascular and vasomotor dysfunction. **Methods**: We developed and pretrained a self-supervised foundation vision transformer model using a large database of unlabeled ECG waveforms (N=800,035). We then fine-tuned the foundation model for two clinical tasks: the difficult problem of identifying patients with impaired myocardial flow reserve (AI-MFR), and the relatively easier problem of detecting impaired LVEF (AI-LVEF). A second ECG database was labeled with task-specific annotations derived from quantitative PET MPI (N=4167). Diagnostic accuracy of AI predictions was tested in a holdout set of patients undergoing PET MPI (N=1031). Prognostic evaluation was performed in the PET holdout cohort, as well as independent cohorts of patients undergoing pharmacologic or exercise stress SPECT MPI (N=6635). **Results**: The diagnostic accuracy of AI-MFR with SSL pretraining increased significantly compared to *de novo* supervised training (AUROC, sensitivity, specificity: 0.758, 70.1%, 69.4% vs. 0.632, 66.1%, 57.3%, $p < 0.0001$). SSL pretraining also produced a smaller increase in AI-LVEF accuracy (AUROC, sensitivity, specificity: 0.946, 89.4%, 85.9% vs. 0.918, 87.6%, 82.5%, $p < 0.02$). Abnormal AI-MFR was found to be significantly associated with mortality risk in all three test cohorts (Hazard Ratio (HR) 2.61 [95% CI 1.83, 3.71], $p < 0.0001$, PET cohort; HR 2.30 [2.03, 2.61], $p < 0.0001$, pharmacologic stress SPECT cohort; HR 3.76 [2.36, 5.99], $p < 0.0001$, exercise stress SPECT cohort). **Conclusion**: SSL pretraining of a vision transformer foundation model enabled identification of signals predictive of impaired MFR, a hallmark of microvascular and vasomotor dysfunction, and impaired LV function in resting and stress ECG waveforms. These signals are powerful predictors of prognosis in patients undergoing routine

noninvasive stress testing and could enable more efficient diagnosis and management of these common conditions.

1 Introduction

Recent AI models of electrocardiogram (ECG) tracings have shown promise as low-cost screening tools for the detection of left ventricular systolic dysfunction and a variety of cardiac structural abnormalities (*1*,*2*). These models have been derived with ECG annotations based on structural labels related to the size and shape of the heart chambers (*3*). ECG signals have also traditionally been used as noninvasive tests of myocardial ischemia and infarction (*4*), and thus, may also have relevance for identifying tissue and cellular properties such as perfusion and potassium balance between intra- and extracellular compartments (*5*,*6*).

Quantitative positron emission tomography (PET) is the gold standard for noninvasive characterization of myocardial ischemia and coronary microvascular dysfunction (*7*). Inspired by recent AI models of left ventricular systolic dysfunction, we sought to use ECG waveforms for AI prediction of tissue- and molecular-level phenomena related to myocardial blood flow and the uptake and retention of rubidium-82 (a potassium analog). Such models have not, to our knowledge, been previously reported due in part to the complexity and limited availability of appropriate ECG annotations based on quantitative PET measurements. Indeed, in preliminary AI studies using simple convolutional neural networks our results were promising but limited by the relatively small labeled dataset that was available for supervised training.

Self-supervised learning (SSL) is a recent paradigm shift in AI which seeks to learn powerful task-agnostic representations from unlabeled data (*8*). Such foundation models can then undergo task-specific fine-tuning for multiple downstream tasks with much smaller amounts of annotated data (*9*). This approach underlies many contemporary state-of-the-art large language and computer vision models (*10*,*11*). In this work we present an example of an SSL foundation model based on the vision transformer architecture and designed for ECG-based assessment of cardiac and coronary function. After pretraining the SSL model on a large database of unlabeled ECG waveforms, we demonstrate state-of-the-art performance for two task-specific applications: the detection of impaired myocardial flow reserve (MFR) and left ventricular ejection fraction (LVEF).

2 Methods

2.1 Data sources

Two ECG databases were used for AI model development and validation. The publicly available MIMIC-IV-ECG repository v1.0 (*12*,*13*) provides a large database of standard 12-lead resting ECGs of 10 sec duration. These waveforms were used as unlabeled data for pretraining the SSL ECG foundation model.

A second database, an MPI registry at the University of Michigan Frankel Cardiovascular Center, consists of all consecutive patients who underwent clinically indicated stress testing with cardiac PET-CT or SPECT-CT myocardial perfusion imaging (MPI). This database was used to generate annotated data for task-specific fine-tuning of the foundation model. For the present study, patients were included with at least one rest-stress MPI exam. For patients with more than one exam, the earliest evaluable exam was included. Patients were excluded if they had a history of heart transplantation, missing or uninterpretable image data, or missing demographic or hemodynamic data.

Patients that had undergone both PET-CT and SPECT-CT exams were included in the PET cohort only and were excluded from the SPECT cohort to avoid label leakage. The SPECT cohort was used as an independent test population for prognostic evaluation. All patient data was de-identified and informed consent was waived under an exemption from the University of Michigan Institutional Review Board.

2.2 Noninvasive stress testing

Rest and stress cardiac ^{82}Rb PET exams were performed according to clinical guidelines for MPI testing (*14*) and measurement of MBF (*15*,*16*) as previously described (*17*).

Cardiac ^{99m}Tc-sestamibi SPECT exams were performed according to guidelines (*18*) as previously described (*19*). Left ventricular ejection fraction (LVEF) (*20*) and stress total perfusion deficit (TPD) (*21*) were routinely estimated during all imaging exams. Stress was induced either pharmacologically with intravenous bolus administration of regadenoson (0.4 mg), by treadmill exercise using a Bruce, modified Bruce, or Cornell protocol, or by a combination of intravenous regadenoson and low-level treadmill exercise. All PET exams

were conducted with regadenoson vasodilator stress while SPECT exams employed a wider variety of stress protocols. Heart rate, systolic, and diastolic blood pressure were monitored continuously during imaging.

Twelve-lead ECG waveforms of 10 sec duration and 500 Hz sampling rate were recorded at baseline immediately before stress testing and at 1 min intervals during stress (in supine position for pharmacologic or combination pharmacologic/low-level exercise stress) or once during each exercise stage (for exercise stress). Rest and stress ECG data were used for AI model development.

2.3 Data annotations

ECG annotations were derived from concurrently acquired PET image data. Myocardial blood flow (MBF) was measured from 30-frame dynamic PET images at rest and stress with a clinically validated, commercially available 1-tissue compartmental model, and MFR was computed as the ratio of stress over rest MBF and averaged globally over the entire left ventricular (*22*,*23*). Left ventricular volumes were measured from 16-frame ECG-gated PET images at rest using a clinically validated, commercially available algorithm, and LVEF was computed as one minus the volume ratio at systole over diastole (*20*). Each ECG waveform was labeled with two task-specific annotations: presence or absence of impaired MFR, defined as global MFR less than 2.0, which is a widely used threshold for prognostically significant MFR impairment (*24*,*25*); and presence or absence of impaired LV function defined as LVEF less than 35% (*26*).

2.4 Prognostic outcome

The primary patient outcome for prognostic evaluation of the AI models was mortality from all causes. The vital status of each patient was determined by integrating data from death certificates and hospital records.

2.5 Deep learning model

Model development

ECG waveform data was used directly as input to a Vision Transformer neural network (*27*) adapted for multichannel, 1-dimensional input (1dViT) (Figure S1 (a)). Waveform patches

of fixed 100 msec width were flattened, linearly projected into an embedding space of dimension 800, combined with a trainable class token, and summed with trainable positional encoding (*27*). A series of twelve Transformer Encoder blocks followed, each incorporating pre-layer normalization (*28*), 8 self-attention heads, and a multi-layer perceptron with GELU activation (*29*) and hidden dimension of 3200 (Figure S1 (a)). The final classifier head consisted of global average pooling, layer normalization, and a linear classifier. The number of input ECG leads was reduced from twelve to eight since the three augmented unipolar leads are linear combinations of the bipolar leads, and by Einthoven's triangle equation only two of the three bipolar leads are linearly independent (*30*). ECG preprocessing consisted of missing value imputation and waveform normalization to zero mean and unit standard deviation. For LVEF prediction, a single 1dViT backbone was used with resting ECG waveform input. For MFR prediction, a dual-1dViT fusion model (Figure S1 (b)) was used with resting and stress ECG waveform input. The 1dViT backbone contained 92.7 million trainable parameters and was developed using the pytorch framework (v2.1.1) (*31*).

SSL model pretraining

For SSL pretraining, the 1dViT backbone was paired with a small Transformer Decoder to perform a pretext task of masked signal modeling (*32*,*33*). The decoder had eight Transformer blocks similar to the 1dViT encoder blocks but with embedding dimension of 256 and 16 self-attention heads. Embedded input patches were randomly masked with a masking ratio of 60%, and only non-masked patches entered the encoder (*33*). Before input to the decoder, encoded patches were combined with a trainable mask token for each masked patch and summed with positional encoding. The input waveform was reconstructed at the level of raw time steps with an L_2 loss (*33*). The AdamW optimizer was used (*34*) with an initial learning rate of 1e-5, linear warmup for 5 epochs to 1e-3 followed by a cosine decay schedule (*35*). Pretraining was performed for 300 epochs with a batch size of 384.

Model fine-tuning

The SSL pretrained 1dViT encoder was subsequently fine-tuned on labeled data for each downstream task. Only the top 4 Transformer blocks were trained during fine-tuning, while freezing the remaining 8 blocks. A binary cross-entropy loss function was utilized with the AdamW optimizer and linear warmup from 1e-5 to 5e-5 followed by cosine decay. Fine-tuning was performed for up to 100 epochs with batch sizes of 256 and 128 for LVEF and MFR prediction tasks, respectively. Fine-tuning iterations were stopped early when validation loss failed to decrease for ten consecutive epochs. For comparison, the same two models without SSL-pretraining were tuned and validated *de novo* with the same labeled data and training setup, except the peak learning rate during linear warmup was set to 1e-3 as used for SSL pretraining.

Model evaluation

Evaluations were performed with the PET holdout test cohort after completion of AI model fine-tuning. Model performance for each downstream task was computed as the diagnostic accuracy of the binary outcomes and area under the ROC curve (AUROC) relative to the true task-specific data labels. Prognostic evaluation of the AI models was also performed in the PET holdout and SPECT MPI test cohorts (Figure 1). The two SPECT cohorts (pharmacologic and exercise stress) were tested separately as the appropriateness of the model for exercise stress ECG data was uncertain and considered exploratory.

2.6 Statistical methods

For diagnostic evaluation, the DeLong test (*36*) was used to compare AUROCs, and a generalized McNemar test (*37*) for joint comparison of sensitivity and specificity. To facilitate comparison between ROC curves, a threshold was selected as the point on the curve closest to the point of perfect discrimination (sensitivity=1, specificity=1). For each prognostic cohort, Cox proportional hazards models and Kaplan-Meier curves were created to test the association of AI predicted impairment with all-cause mortality risk. Adjustment of Cox models was pre-specified for patient demographics, baseline risk factors, and standard MPI findings on the basis of judgment and prior work (*24*,*38*). Baseline covariates included patient age, sex, body mass index (BMI), diabetes, hypertension, hyperlipidemia,

known coronary artery disease (CAD, history of myocardial infarction (MI) or previous percutaneous coronary intervention (PCI) or coronary artery bypass graft), LVEF, and stress TPD as a combined relative measure of ischemia or scar. Subgroup analysis was performed stratifying by age (< 60y), BMI (< 30kg/m^2), sex, diabetes, dyslipidemia, hypertension, known CAD, LVEF ($< 50\%$), and stress TPD ($< 5\%$).

Model discrimination was assessed using likelihood ratio χ^2 and c-index. The change in discrimination after adding AI predicted impairments to the baseline Cox models was assessed with continuous net reclassification improvement (NRI) (*39*). Non-nested Cox models were compared using Fine's test (*40*).

Continuous variables are summarized as mean±SD or median [1st – 3rd quartiles]. Welch's unequal variances t-test or Wilcoxon rank sum tests were used as appropriate for comparisons of continuous parameters, and χ-squared tests were used to compare categorical variables. Two-sided p-values less than 0.05 were considered statistically significant. Analyses were performed using R Statistical Software (v4.1.0) (*41*) with packages `survival` (v3.2-11) (*42*), `rms` (v6.2-9) (*43*), `pROC` (v1.18.0) (*44*), `DTComPair` (v1.2.4) (*45*), `survminer` (v0.4.9) (*46*), `nricens` (v1.6) (*47*), `forestploter` (v1.1.0) (*48*), `gtsummary` (v1.5.2) (*49*), and `Publish` (v2021.05.25) (*50*).

3 Results

3.1 Patient population

The MIMIC-IV-ECG repository provided 800,035 diagnostic resting ECGs acquired in 161,352 patients. For SSL pretraining, ECG data were split into training and validation subsets (97.5%/2.5%) and no other patient data were used. Of 12,764 patients in the MPI registry, 5202 patients underwent stress testing with PET-CT, and 7303 underwent SPECT-CT (Figure 1). The PET cohort was randomly split into training, validation, and holdout test subsets with ratios 60:20:20%. AI model diagnostic and prognostic evaluation was performed using the PET holdout test cohort (N=1031), as well as 5102 patients who underwent pharmacologic stress SPECT MPI and 1533 patients who underwent exercise stress SPECT MPI.

3.2 Baseline characteristics

Baseline characteristics are shown in Table 1 stratified by patient cohort. As expected, the PET MPI cohorts for model derivation and holdout test had nearly identical clinical characteristics. However, the pharmacologic stress SPECT MPI cohort differed meaningfully from the PET MPI cohort, reflecting heterogeneity in patient referral and practice patterns. SPECT patients were older, with lower rates of obesity, diabetes, and history of MI, and higher rates of hypertension, hyperlipidemia, and mortality than the PET cohorts. Baseline characteristics of the exercise stress SPECT MPI cohort is shown in Table S1.

3.3 AI model diagnostic evaluation

In the PET holdout test cohort, the incidence of impaired MFR (<2) was 48% (N=461) and impaired LVEF (<35%) was 11% (N=113). Within this cohort, the AUROC for both tasks was significantly increased for models based on the SSL-pretrained foundation model compared to *de novo* training without SSL (AI-MFR, p<0.0001; AI-LVEF, p=0.0176), and diagnostic accuracy was similary increased (AI-MFR, p<0.0001; AI-LVEF, p=0.0078) (Figure 2, Table S2, Table S3). However, the absolute gain was much greater for AI-MFR.

3.4 Prognostic assessment

AI model predictions of abnormal AI-MFR with and without SSL-pretraining were both significantly associated with risk of death in adjusted Cox models (Hazard Ratio (HR) 2.61 and 1.66, Table 2). However, SSL-based predictions provided better model fit (p=0.0212), approaching that of PET-measured MFR impairment (HR 3.18, Table 2, Figure 3). Cox models of SSL-based AI-MFR and PET-MFR were not significantly different (p=0.779), and provided similar discrimination (c-index 0.717 vs 0.725) and overall NRI (0.538 vs 0.534) (Table 2, Table S6). In contrast, PET-measured LVEF impairment was not a predictor of mortality risk in this population (possibly due to collider bias from referral patterns), and only SSL-based AI-LVEF was significantly associated with mortality (HR 1.66, Table 2, Figure S2 (b)).

Similar prognostic results were observed in the pharamcologic stress SPECT cohort (Table S4, Table S6, Figure S3 (b)). A Cox model of SSL-based AI-MFR again provided better model

fit than that without SSL (p<0.0001). Intriguingly, AI-MFR remained a strong predictor of outcomes in the exploratory SPECT exercise stress cohort (HR 3.76, $p < 0.0001$) despite the dramatically different type of stress (treadmill exercise versus vasodilator medications), much lower event rate and broadly healthier population (Table S5, Figure S4 (b)). Overall NRI was also significantly positive (Table S6) .

3.5 Subgroup analysis

In subgroups stratified by patient characteristics, abnormal AI-MFR was consistently a strong predictor across subgroups and remained significantly associated with higher mortality risk in the PET holdout test cohort (Table S7 and Table S8) and both pharmacologic (Table S9) and exercise stress (Table S10) SPECT cohorts. AI-MFR was a significant independent predictor of mortality in women, patients with diabetes, hypertension, LV dysfunction, and those with minimal or no stress perfusion abnormality (TPD $< 5\%$) representing patients with likely diffuse coronary microvascular dysfunction.

4 Discussion

We have applied a data-efficient SSL approach to develop, to our knowledge, the first deep neural network with the ability to detect impaired MFR, demonstrating that complex tissue- and molecular-level pathophysiology results in characteristic electrophysiologic changes detectable in surface ECG tracings. These results demonstrate that SSL and vision transformer architectures can enable training of clinically relevant tools with far less gold standard data than previously believed. Importantly, this may facilitate the development of deep learning based low-cost testing strategies for advanced diagnostics previously thought to be inaccessible or extremely challenging due to the need for extremely large training datasets.

4.1 AI considerations

Our results demonstrate the value of the self-supervised learning paradigm for clinical AI applications. Using a large publicly available database of unlabeled ECG data, we pretrained a single SSL foundation model, which we then fine-tuned for two distinct clinical tasks. Our annotated dataset of 4167 PET MPI patients, though relatively large as a clinical PET population, is small by AI model training standards. However, for task-specific fine-tuning

this annotated data was sufficient to achieve a diagnostic accuracy for impaired LVEF (AUROC 0.946) (Figure 2 (b), Table S3) comparable to a previous AI study (AUROC 0.932) (*26*) which *required ten times more annotated data* (44,959 patients). Our SSL foundation model also significantly improved the diagnostic accuracy of detecting impaired MFR compared to *de novo* supervised training (Figure 2 (a), Table S2).

Although our AI models yielded highly accurate predictions of AI-LVEF (Table S3), and moderately accurate predictions of AI-MFR (Table S2), the clinical evaluation of AI-MFR showed consistently stronger prediction of mortality risk than AI-LVEF, approaching the prognostic value of gold standard PET measurements of MFR (Table 2, Figure 3 (b)). The lower prognostic value of AI-LVEF may be due in part to the low prevalence of impaired LVEF in the test cohorts (11%, in PET holdout; 3.5% in pharmacologic SPECT; and 0.5% in exercise SPECT). This underscores the importance of evaluating clinically meaningful endpoints in AI applications. Interestingly, abnormal AI-MFR and AI-LVEF remained significant predictors of mortality risk even in the subset of patients with preserved MFR>2 and LVEF>35% (Table 3), and this was true in both the PET holdout and pharmacologic stress SPECT cohorts. This suggests that ECG abnormalities were detected by the AI model before overt clinical impairment of MFR or LVEF. A similar result has been reported in two previous studies (*26*,*51*) with AI models predicting four-fold higher risk of future LV systolic dysfunction in subjects with normal LVEF.

4.2 Clinical implications

Coronary microvascular and vasomotor dysfunction (CMVD) develops as a result of aging, diabetes, and a range of other cardiometabolic diseases (*52*,*53*) and causes myocardial ischemia and symptomatic angina without obstructive coronary disease (*54*,*55*). Approximately half of patients referred for invasive cardiac evaluation are found to have no obstructive coronary disease (*56*), and of these, more than two-thirds have some form of CMVD (*54*). Further, CMVD is associated with markedly increased rates of adverse cardiac outcomes (*57*,*58*).

Cardiac stress testing with quantitative PET MPI is considered the noninvasive gold standard for assessing CMVD (*7*). Although guidelines broadly recommend PET testing for

a range of clinical conditions (*61*), its accessibility remains limited and many regions are underserved by PET (*62*). In resource constrained environments, AI-MFR could provide a cost-effective initial test to direct patients who would benefit most from PET testing. Further, AI-MFR could be useful as a low-cost adjunct to other more accessible stress testing modalities such as stress SPECT, stress echocardiography, and exercise ECG, providing valuable additional information on CMVD phenotyping.

In clinical evaluation of the AI-MFR metric, we demonstrated consistently strong prognostic performance in three distinct clinical populations: a holdout test cohort undergoing PET MPI for which MFR measurements were available (Table 2), and two independent SPECT MPI cohorts undergoing either pharmacologic or exercise stress, and for whom MFR measurements were not possible (Table S4, Table S5). Our results demonstrate that impaired AI-MFR is independently associated with higher risk of death after adjusting for clinical risk factors and standard MPI measurements. Impaired AI-MFR improved the prognostic discrimination of standard MPI in terms of increased c-index and continuous Net Reclassification Improvement in all three cohorts, indicating the strong generalizability of this AI model. In subgroup analysis of patients with minimal or no stress perfusion abnormalities (stress TPD<5, Table S8), AI-MFR remained a consistent predictor of adverse outcomes, consistent with the identification of impaired MFR with coronary microvascular dysfunction (*63*).

Ahmad, et al. (*64*) recently performed a ground-breaking study on a related question. Using logistic regression and rest ECG data, they built a prediction model to detect coronary microvascular dysfunction identified with invasive coronary vasomotor testing. Although such invasive testing is precise, it is rarely performed and is subject to marked referral and selection biases. Further, invasive testing is rarely performed in all three coronary arteries and thus measures may reflect regional rather than global cardiac abnormalities. Given the limited sample size, simpler model architecture, and lack of stress perturbation data, the performance of their prediction model was limited (*64*). We believe the higher performance of our approach may in part be due to integration of ECGs from both resting and stress conditions. Prior studies have demonstrated that broad metabolic and physiologic changes are detectible with a single episode of exercise (*65*).

4.3 Limitations

Although we evaluated our AI models in three non-overlapping clinical test populations, all were from a single, high-volume health system. Further external validation in additional populations is warranted to confirm robustness.

The present study did not evaluate the diagnostic performance of the AI models for detection of obstructive CAD. We speculate that detecting AI-MFR impairment could potentially improve the diagnostic accuracy of standard relative MPI interpreted visually by clinicians for the presence of high risk left main or multi-vessel obstructive CAD, as seen for PET-measured MFR in prior studies (*63*,*66*,*67*).

Global MFR does not distinguish CMVD from myocardial perfusion impairments due to focal epicardial or diffuse CAD. Our recent work has shown that regional quantitative PET measures combined with relative perfusion and global MFR can be highly effective at assessing the additive risk of diffuse or microvascular disease (*68*). Further work will be necessary to integrate these results with the present ECG AI model predictions to further characterize CMVD.

4.4 Conclusion

Using a data-efficient self-supervised foundation model with a vision transformer architecture, we have demonstrated that ECG waveforms at rest and during stress carry important clinical information related to tissue and cellular properties, namely, coronary microvascular and vasomotor function. The same foundation model identified structural abnormalities associated with left ventricular systolic dysfunction. These deep learning metrics have prognostic value similar to that of gold standard measurements.

5 Acknowledgement

The authors acknowledge the Regents of the University of Michigan for the use of de-identified clinical data for this study.

6 Funding

VLM is supported by grants R01AG059729 from the National Institute on Aging, U01DK123013 from the National Institute of Diabetes and Digestive and Kidney Disease,

and R01HL136685 from the National Heart, Lung, and Blood Institute as well as the Melvyn Rubenfire Professorship in Preventive Cardiology.

7 Author Declarations

JBM, APR, JMR, TH, and MDV are employees of INVIA. JMR is a consultant for Jubilant Radiopharma and receives royalties from the licensing of FlowQuant software. EPF is a stockholder in INVIA. VLM has received research grants and consulting fees from Siemens Healthineers and serves as a scientific advisor for Ionetix and owns stock options in the same. He owns stock in GE and Cardinal Health, and has received payments for consulting from INVIA.

8 Ethics Approval

This is an observational study and informed consent was waived under an exemption from the University of Michigan Institutional Review Board.

9 References

1. Bjerkén LV, Rønborg SN, Jensen MT, Ørting SN, Nielsen OW. Artificial intelligence enabled ECG screening for left ventricular systolic dysfunction: a systematic review. *Heart Fail Rev*. 2023;28:419-430.

2. Siontis KC, Noseworthy PA, Attia ZI, Friedman PA. Artificial intelligence-enhanced electrocardiography in cardiovascular disease management. *Nat Rev Cardiol*. 2021;18:465-478.

3. Vaid A, Johnson KW, Badgeley MA, et al. Using deep-learning algorithms to simultaneously identify right and left ventricular dysfunction from the electrocardiogram. *JACC: Cardiovascular Imaging*. 2022;15:395-410.

4. Mirvis DM, Goldberger AL. Electrocardiography. In: Braunwald E, Zipes DP, Libby P, eds. Heart Disease: A Textbook of Cardiovascular Medicine. Vol 1. 6th edition. Philadelphia: WB Saunders, Philadelphia; 2001:82-128.

5. Parker JO, Chiong MA, West RO, Case RB. The effect of ischemia and alterations of heart rate on myocardial potassium balance in man. *Circulation*. 1970;42:205-217.

6. Sejersted OM, Sjøgaard G. Dynamics and consequences of potassium shifts in skeletal muscle and heart during exercise. *Physiol Rev*. 2000;80:1411-1481.

7. Feher A, Sinusas AJ. Quantitative assessment of coronary microvascular function: dynamic single-photon emission computed tomography, positron emission tomography, ultrasound, computed tomography, and magnetic resonance imaging. *Circulation: Cardiovascular Imaging*. 2017;10:e006427.

8. Ericsson L, Gouk H, Loy CC, Hospedales TM. Self-supervised representation learning: introduction, advances, and challenges. *IEEE Signal Processing Magazine*. 2022;39:42-62.

9. Balestriero R, Ibrahim M, Sobal V, et al. A cookbook of self-supervised learning. June 2023.

10. Raffel C, Shazeer N, Roberts A, et al. Exploring the limits of transfer learning with a unified text-to-text transformer. *J Mach Learn Res*. 2020;21:140:5485-140:5551.

11. Bao H, Dong L, Piao S, Wei F. BEiT: BERT pre-training of image transformers. September 2022.

12. Gow B, Pollard T, Nathanson LA, et al. MIMIC-IV-ECG: diagnostic electrocardiogram matched subset. 2023.

13. Johnson AEW, Bulgarelli L, Shen L, et al. MIMIC-IV, a freely accessible electronic health record dataset. *Sci Data*. 2023;10:1.

14. Dilsizian V, Bacharach SL, Beanlands RS, et al. ASNC imaging guidelines/SNMMI procedure standard for positron emission tomography (PET) nuclear cardiology procedures. *J Nucl Cardiol*. July 2016:1-40.

15. Murthy VL, Bateman TM, Beanlands RS, et al. Clinical quantification of myocardial blood flow using PET: Joint position paper of the SNMMI Cardiovascular Council and the ASNC. *J Nucl Cardiol*. 2018;25:269-297.

16. Bateman TM, Heller GV, Beanlands R, et al. Practical guide for interpreting and reporting cardiac PET measurements of myocardial blood flow: an Information Statement from the American Society of Nuclear Cardiology, and the Society of Nuclear Medicine and Molecular Imaging. *J Nucl Cardiol*. 2021;28:768-787.

17. Arida-Moody L, Moody JB, Renaud JM, et al. Effects of two patient-specific dosing protocols on measurement of myocardial blood flow with 3D ^{82}Rb cardiac PET. *Eur J Nucl Med Mol Imaging*. 2021;48:3835-3846.

18. Dorbala S, Ananthasubramaniam K, Armstrong IS, et al. Single photon emission computed tomography (SPECT) myocardial perfusion imaging guidelines: Instrumentation, acquisition, processing, and interpretation. *Journal of Nuclear Cardiology*. 2018;25:1784-1846.

19. Hagio T, Poitrasson-Rivière A, Moody JB, et al. "Virtual" attenuation correction: improving stress myocardial perfusion SPECT imaging using deep learning. *European Journal of Nuclear Medicine and Molecular Imaging*. 2022;49:3140-3149.

20. Ficaro EP, Lee BC, Kritzman JN, Corbett JR. Corridor4DM: the Michigan method for quantitative nuclear cardiology. *J Nucl Cardiol*. 2007;14:455-65.

21. Slomka PJ, Nishina H, Berman DS, et al. Automated quantification of myocardial perfusion SPECT using simplified normal limits. *J Nucl Cardiol*. 2005;12:66-77.

22. Lortie M, Beanlands RSB, Yoshinaga K, Klein R, Dasilva JN, DeKemp RA. Quantification of myocardial blood flow with ^{82}Rb dynamic PET imaging. *Eur J Nucl Med Mol Imaging*. 2007;34:1765-74.

23. Moody JB, Murthy VL, Lee BC, Corbett JR, Ficaro EP. Variance estimation for myocardial blood flow by dynamic PET. *IEEE Transactions on Medical Imaging*. 2015;34:2343-2353.

24. Murthy VL, Naya M, Foster CR, et al. Improved cardiac risk assessment with noninvasive measures of coronary flow reserve. *Circulation*. 2011;124:2215-2224.

25. Ziadi MC, deKemp RA, Williams KA, 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*. 2011;58:740-748.

26. Attia ZI, Kapa S, Lopez-Jimenez F, et al. Screening for cardiac contractile dysfunction using an artificial intelligence–enabled electrocardiogram. *Nat Med*. 2019;25:70-74.

27. Dosovitskiy A, Beyer L, Kolesnikov A, et al. An image is worth 16x16 words: transformers for image recognition at scale. June 2021.

28. Ba JL, Kiros JR, Hinton GE. Layer normalization. July 2016.

29. Hendrycks D, Gimpel K. Gaussian error linear units (GELUs). June 2023.

30. Salinet JL, Luppi Silva O. Chapter 2 - ECG Signal Acquisition Systems. In: do Vale Madeiro JP, Cortez PC, da Silva Monteiro Filho JM, Brayner ARA, eds. Developments and Applications for ECG Signal Processing. Academic Press; 2019:29-51.

31. Paszke A, Gross S, Massa F, et al. Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems*. 2019;32.

32. Xie Z, Zhang Z, Cao Y, et al. SimMIM: a simple framework for masked image modeling. In:; 2022:9653-9663.

33. He K, Chen X, Xie S, Li Y, Dollár P, Girshick R. Masked autoencoders are scalable vision learners. In:; 2022:16000-16009.

34. Kingma DP, Ba J. Adam: a method for stochastic optimization. *arXiv:14126980 [cs]*. December 2014.

35. Loshchilov I, Hutter F. SGDR: stochastic gradient descent with warm restarts. May 2017.

36. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988;44:837-845.

37. Lachenbruch PA, Lynch CJ. Assessing screening tests: extensions of McNemar's test. *Statistics in Medicine*. 1998;17:2207-2217.

38. Gupta A, Taqueti VR, van de Hoef TP, et al. Integrated noninvasive physiological assessment of coronary circulatory function and impact on cardiovascular mortality in patients with stable coronary artery disease. *Circulation*. 2017;136:2325-2336.

39. Leening MJG, Vedder MM, Witteman JCM, Pencina MJ, Steyerberg EW. Net Reclassification Improvement: Computation, interpretation, and controversies: A literature review and clinician's guide. *Ann Intern Med*. 2014;160:122-131.

40. Fine JP. Comparing nonnested Cox models. *Biometrika*. 2002;89:635-648.

41. R Core Team. R: A language and environment for statistical computing. 2021.

42. Therneau TM. A package for survival analysis in R. R package version 3.2-11. August 2021.

43. Harrell Jr FE. rms: Regression modeling strategies. R package version 6.2-0. 2021.

44. Robin X, Turck N, Hainard A, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*. 2011;12:77.

45. Stock C, Hielscher T, Discacciati A. DTComPair: comparison of binary diagnostic tests in a paired study design. R package version 1.2.4. 2024.

46. Kassambara A, Kosinski M, Biecek P. survminer: Drawing survival curves using "ggplot2." R package version 0.4.9. 2021.

47. Inoue E. nricens: NRI for risk prediction models with time to event and binary response data. R package version 1.6. 2018.

48. Dayimu A. forestploter: Create flexible forest plot. R package version 1.1.0. 2023.

49. Sjoberg DD, Whiting K, Curry M, Lavery JA, Larmarange J. Reproducible summary tables with the gtsummary package. *The R Journal*. 2021;13:570-580.

50. Gerds TA, Ozenne B. Publish: format output of various routines in a suitable way for reports and publication. R package version 2021.05.25. 2021.

51. König S, Hohenstein S, Nitsche A, et al. Artificial intelligence-based identification of left ventricular systolic dysfunction from 12-lead electrocardiograms: external validation and advanced application of an existing model. *European Heart Journal - Digital Health*. 2024;5:144-151.

52. Mannarino T, D'Antonio A, Assante R, et al. Regional myocardial perfusion imaging in predicting vessel-related outcome: interplay between the perfusion results and angiographic findings. *Eur J Nucl Med Mol Imaging*. 2022;50:160-167.

53. Del Buono MG, Montone RA, Camilli M, et al. Coronary microvascular dysfunction across the spectrum of cardiovascular diseases. *Journal of the American College of Cardiology*. 2021;78:1352-1371.

54. Sara JD, Widmer RJ, Matsuzawa Y, Lennon RJ, Lerman LO, Lerman A. Prevalence of coronary microvascular dysfunction among patients with chest pain and nonobstructive coronary artery disease. *J Am Coll Cardiol Intv*. 2015;8:1445-1453.

55. Masi S, Rizzoni D, Taddei S, et al. Assessment and pathophysiology of microvascular disease: recent progress and clinical implications. *European Heart Journal*. 2021;42:2590-2604.

56. Ford TJ, Ong P, Sechtem U, et al. Assessment of vascular dysfunction in patients without obstructive coronary artery disease. Why, how, and when. *JACC: Cardiovascular Interventions*. 2020;13:1847-1864.

57. Jespersen L, Hvelplund A, Abildstrøm SZ, et al. Stable angina pectoris with no obstructive coronary artery disease is associated with increased risks of major adverse cardiovascular events. *European Heart Journal*. 2012;33:734-744.

58. Gdowski MA, Murthy VL, Doering M, Monroy-Gonzalez AG, Slart R, Brown DL. Association of isolated coronary microvascular dysfunction with mortality and major

adverse cardiac events: A systematic review and meta-analysis of aggregate data. *Journal of the American Heart Association*. 2020;9:e014954.

59. Arbelo E, Protonotarios A, Gimeno JR, et al. 2023 ESC Guidelines for the management of cardiomyopathies: Developed by the task force on the management of cardiomyopathies of the European Society of Cardiology (ESC). *European Heart Journal*. 2023;44:3503-3626.

60. Delgado V, Ajmone Marsan N, de Waha S, et al. 2023 ESC Guidelines for the management of endocarditis: Developed by the task force on the management of endocarditis of the European Society of Cardiology (ESC) Endorsed by the European Association for Cardio-Thoracic Surgery (EACTS) and the European Association of Nuclear Medicine (EANM). *European Heart Journal*. 2023;44:3948-4042.

61. Virani SS, Newby LK, Arnold SV, et al. 2023 AHA/ACC/ACCP/ASPC/NLA/PCNA Guideline for the Management of Patients With Chronic Coronary Disease. *Journal of the American College of Cardiology*. 2023;82:833-955.

62. Sanghani R, Al-Mallah MH, Thompson R. Challenges and strategies to enable access to cardiac positron emission tomography in different parts of the world: The North American perspective. *Journal of Nuclear Cardiology*. 2024;31:101790.

63. Poitrasson-Rivière A, Moody JB, Renaud JM, et al. Integrated myocardial flow reserve (iMFR) assessment: optimized PET blood flow quantification for diagnosis of coronary artery disease. *European Journal of Nuclear Medicine and Molecular Imaging*. 2023;51:136-146.

64. Ahmad A, Shelly-Cohen M, Corban MT, et al. Machine learning aids clinical decision-making in patients presenting with angina and non-obstructive coronary artery disease. *European Heart Journal - Digital Health*. 2021;2:597-605.

65. Nayor M, Shah RV, Miller PE, et al. Metabolic architecture of acute exercise response in middle-aged adults in the community. *Circulation*. 2020;142:1905-1924.

66. Kajander S, Joutsiniemi E, Saraste M, et al. Clinical value of absolute quantification of myocardial perfusion with ^{15}O-water in coronary artery disease. *Circulation: Cardiovascular Imaging*. 2011;4:678-684.

67. Ziadi MC, deKemp RA, Williams K, et al. Does quantification of myocardial flow reserve using rubidium-82 positron emission tomography facilitate detection of multivessel coronary artery disease? *J Nucl Cardiol*. 2012;19:670-680.

68. Moody JB, Poitrasson-Rivière A, Renaud JM, et al. Integrated myocardial flow reserve (iMFR) assessment: diffuse atherosclerosis and microvascular dysfunction are more strongly associated with mortality than focally impaired perfusion. *Eur J Nucl Med Mol Imaging*. 2023;51:123-135.

10 Tables

Table 1: Baseline characteristics of patients who underwent pharmacologic stress testing by myocardial perfusion imaging stratified by cohort. AI models were developed with the Model Derivation cohort and tested in the PET Holdout Test and SPECT Pharmacologic Stress cohorts.

	PET MPI Cohort		**SPECT MPI Cohort**	**p-value**[2]
	Model Derivation, N = 4,167[1]	**Holdout Test, N = 1,031**[1]	**Pharmacologic Stress, N = 5,102**[1]	
Demographics				
Age (y)	63 (12)	63 (12)	66 (12)	<0.001
Male sex	2,312 (55)	590 (57)	2,874 (56)	0.53
Race				<0.001
White	3,404 (82)	840 (81)	4,127 (81)	
Black	512 (12)	115 (11)	504 (9.9)	
Other	251 (6.0)	76 (7.4)	471 (9.2)	
Cardiovascular Risk Factors				
Body mass index (kg/m^2)	34 (10)	33 (9)	30 (8)	<0.001
Obesity	2,539 (61)	599 (58)	2,299 (45)	<0.001
Hypertension	3,179 (76)	776 (75)	4,017 (79)	0.005
Diabetes	1,652 (40)	422 (41)	1,676 (33)	<0.001
Hyperlipidemia	2,747 (66)	673 (65)	3,699 (73)	<0.001
Family history of coronary artery disease	2,807 (67)	681 (66)	1,715 (34)	<0.001
Current smoker	336 (8.1)	99 (9.6)	1,135 (22)	<0.001
History of myocardial infarction	799 (19)	200 (19)	714 (14)	<0.001
Prior percutaneous coronary intervention	816 (20)	201 (19)	952 (19)	0.50
Prior coronary artery bypass graft	418 (10)	102 (9.9)	491 (9.6)	0.80
Symptoms				
Angina	1,819 (44)	431 (42)	2,204 (43)	0.56
Dyspnea	665 (16)	172 (17)	1,077 (21)	<0.001
Hemodynamics				
Heart rate, rest (bpm)	73 (14)	73 (13)	69 (13)	<0.001
SBP, rest (mm Hg)	132 (23)	131 (24)	143 (25)	<0.001
DBP, rest (mm Hg)	71 (24)	71 (22)	80 (13)	<0.001

	PET MPI Cohort		SPECT MPI Cohort	p-value[2]
	Model Derivation, N = 4,167[1]	**Holdout Test, N = 1,031**[1]	**Pharmacologic Stress, N = 5,102**[1]	
Rate pressure product (mm Hg bpm)	9,658 (2,496)	9,524 (2,471)	9,853 (2,522)	<0.001
Heart rate, stress (bpm)	93 (20)	92 (28)	101 (21)	<0.001
SBP, stress (mm Hg)	120 (31)	119 (30)	161 (27)	<0.001
DBP, stress (mm Hg)	64 (39)	63 (36)	80 (14)	<0.001
PET Measurements				
LV ejection fraction, rest (%)	57 (15)	58 (15)	---	0.80
LV ejection fraction, stress (%)	61 (16)	61 (16)	65 (13)	<0.001
Total perfusion deficit, stress (%)	7 (10)	7 (11)	5 (9)	<0.001
Myocardial blood flow, rest (ml/min/g)	1.08 (0.39)	1.08 (0.41)	---	0.93
Myocardial blood flow, stress (ml/min/g)	2.17 (0.83)	2.16 (0.82)	---	0.79
Myocardial flow reserve	2.11 (0.72)	2.12 (0.74)	---	0.88
Outcomes				
All-cause mortality	696 (17)	172 (17)	1,118 (22)	<0.001

[1] Mean (SD); n (%)

[2] One-way ANOVA; Pearson's Chi-squared test

Table 2: Prognostic evaluation in the PET holdout cohort (N=1031 patients). Cox Proportional Hazards models of ECG AI model predictions (AI-MFR, AI-LVEF) and PET measurements and risk of all-cause mortality.

	Impaired MFR			Impaired LVEF		
	AI-MFR No SSL	**AI-MFR SSL**	**PET-MFR**	**AI-LVEF No SSL**	**AI-LVEF SSL**	**PET-LVEF**
Hazard Ratio (95% CI)	1.66 (1.18, 2.34)	2.61 (1.83, 3.71)	3.18 (2.19, 4.61)	1.31 (0.79, 2.17)	1.66 (1.04, 2.63)	1.35 (0.86, 2.13)
χ^2	75.4 p = 0.0029	97.6 p < 0.0001	108.3 p < 0.0001	65.2 p = 0.3168	68.3 p = 0.0417	65.8 p = 0.2019
c-index	0.692 p = 0.0103	0.717 p = 0.0010	0.725 p = 0.0007	0.673 p = 0.9491	0.674 p = 0.8205	0.673 p = 0.8974

Base model included patient age, sex, BMI, cardiovascular risk factors, history of myocardial infarction, percutaneous coronary intervention, and coronary artery bypass graft, LVEF (MFR base model only) and standard myocardial perfusion imaging findings. For each model, either ECG AI model predictions or PET-measured impairment were added to the base model.
The base model for Impaired MFR had χ^2 of 66.5 on 9 d.f. and c-index of 0.670.
The base model for Impaired LVEF had χ^2 of 64.1 on 8 d.f. and c-index of 0.673.
LVEF = left ventricular ejection fraction
MFR = myocardial flow reserve

Table 3: Prognostic evaluation in subsets of patients with preserved MFR>2 and LVEF>35%. Cox Proportional Hazards models of ECG AI model predictions (AI-MFR, AI-LVEF) and risk of all-cause mortality demonstrating that even when direct clinical measurements are available, these AI measures may have added prognostic value.

	Observations	Events	HR†	95% CI
PET holdout				
AI-MFR	536	42	2.01	(1.09, 3.70)
AI-LVEF	918	142	2.33	(1.40, 5.07)
Pharm. stress SPECT				
AI-LVEF	4923	1055	1.99	(1.49, 2.65)

† Adjusted for patient age and sex.
MFR and LVEF were determined by image-based measurements.

11 Figures

Figure 1: CONSORT diagram.

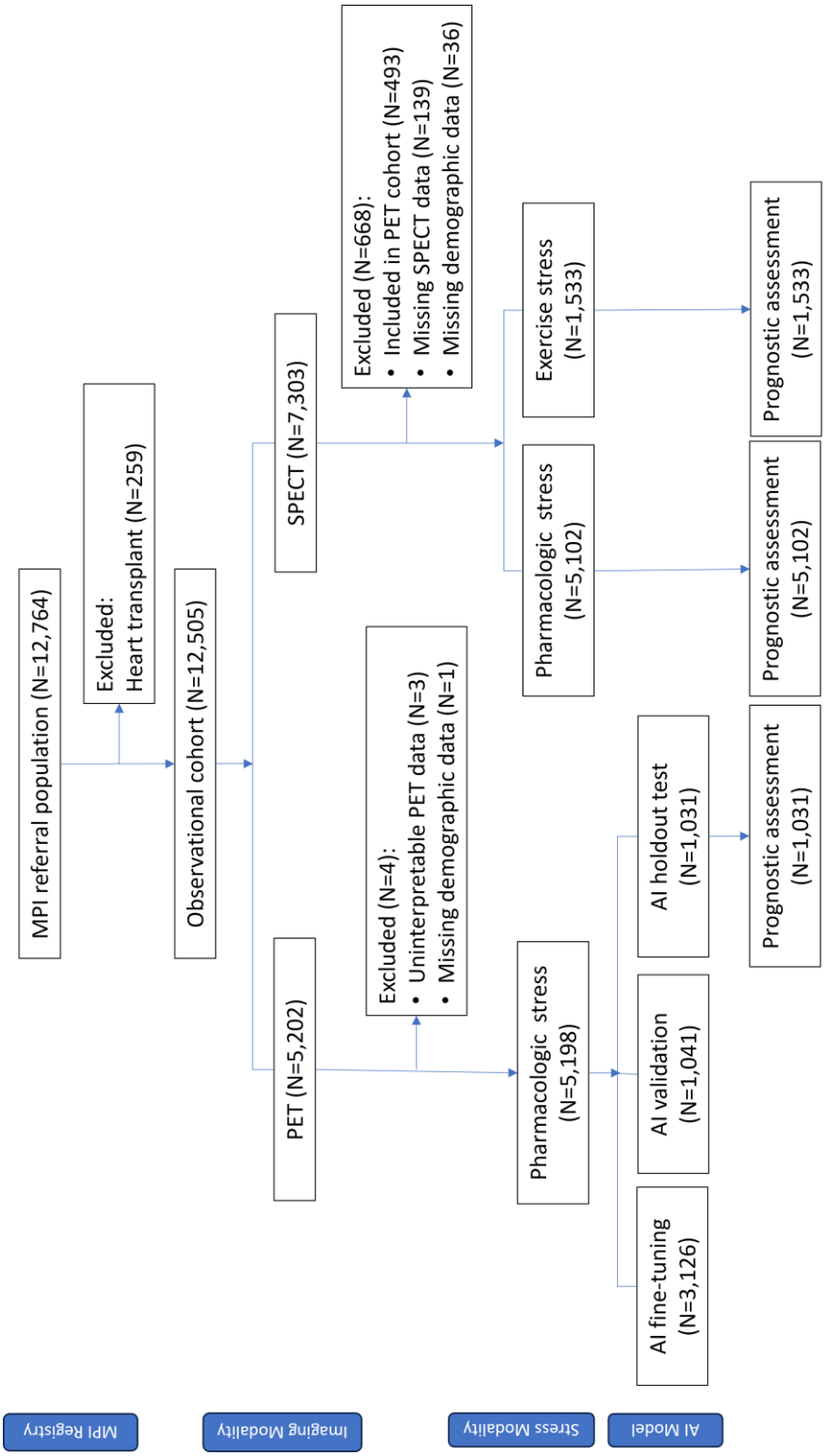
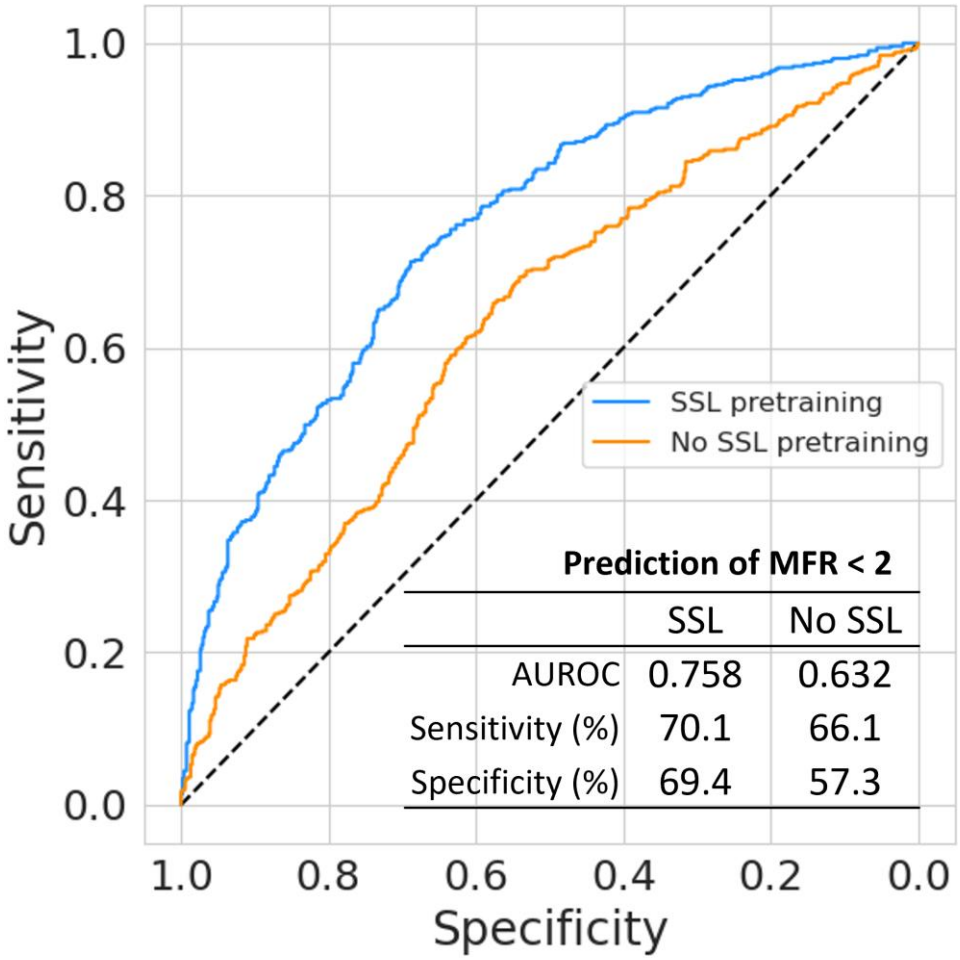


Figure 2: Diagnostic evaluation in the PET holdout test cohort (N=1031 patients). AI model predictions of abnormal MFR (a) and abnormal LVEF (b) with and without the SSL-pretrained foundation model. SSL-based models were fine-tuned and validated on the PET derivation cohort (N=4167 patients); models with no SSL pretraining were tuned and validated de novo in the same derivation cohort. See Table S2 and Table S3 for 95% confidence intervals and p-values.

(a)



(b)

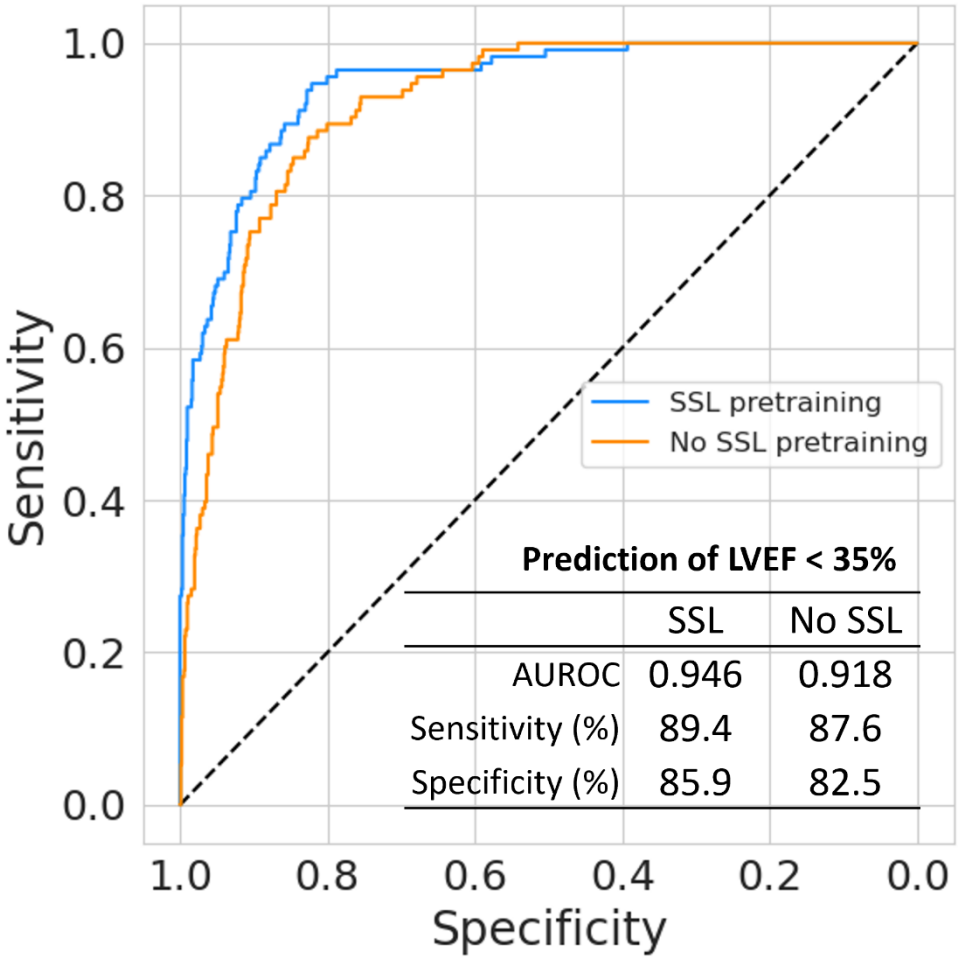
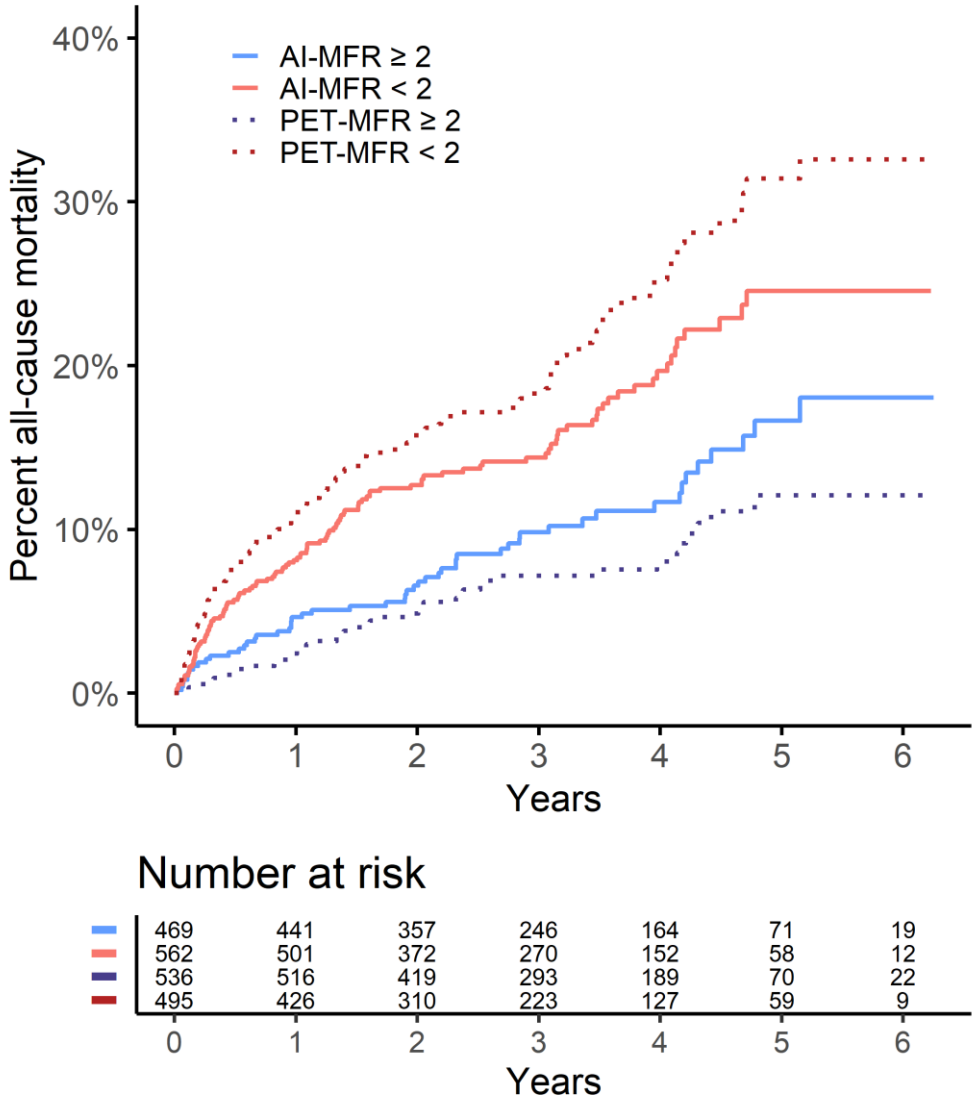
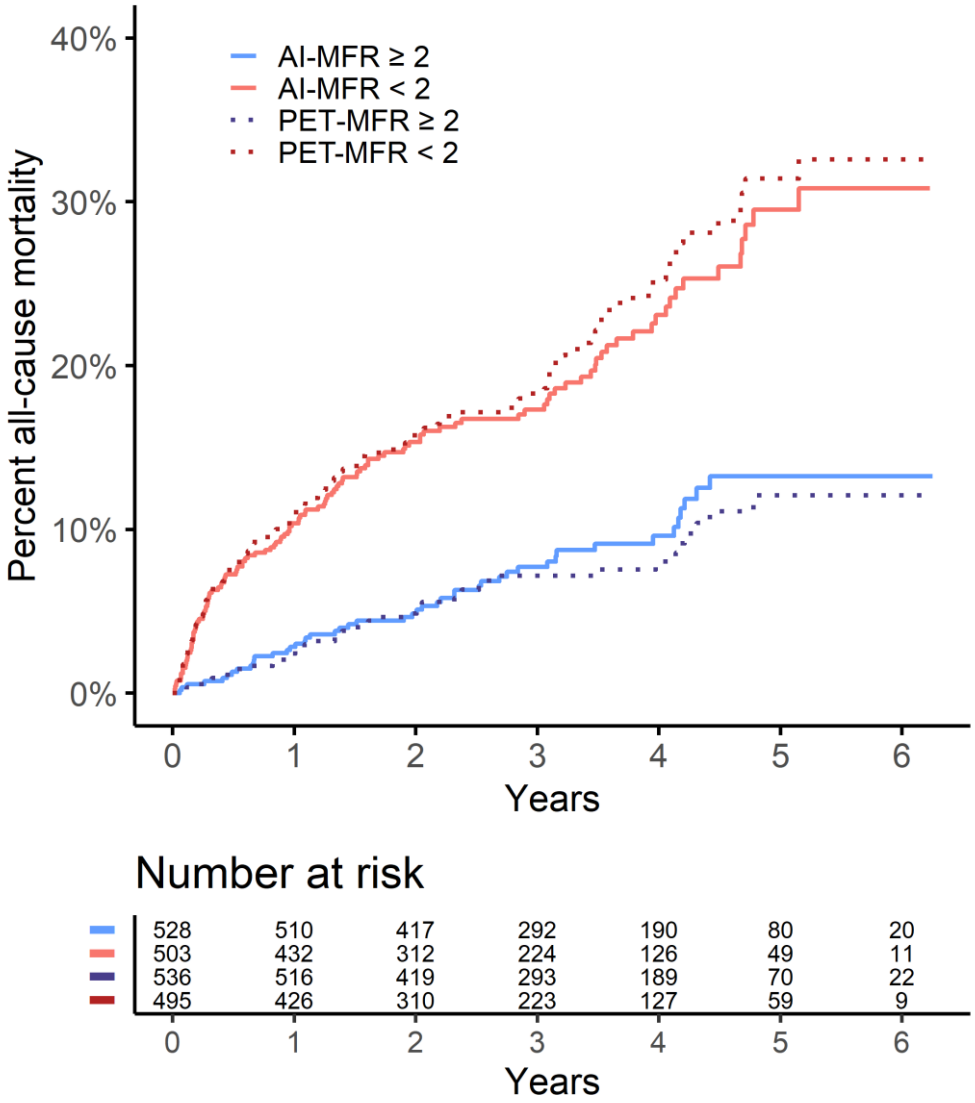


Figure 3: Adjusted incidence of all-cause mortality between ECG AI model predictions of abnormal MFR (AI-MFR, solid lines) without (a), and with (b) the SSL-pretrained foundation model. Dotted lines show a model of PET-measured MFR. Evaluation was performed on the PET holdout test cohort (N=1031 patients).

(a)



(b)



12 Supplemental Data

12.1 Supplemental Tables

Table S1: Baseline characteristics of patients who underwent SPECT myocardial perfusion imaging stratified by stress modality.

	Pharmacologic Stress, N = 5,102[1]	**Exercise Stress, N = 1,533**[1]	**p-value**[2]
Demographics			
Age (y)	66 (12)	61 (11)	<0.001
Male sex	2,874 (56)	974 (64)	<0.001
Race			<0.001
White	4,127 (81)	1,279 (83)	
Black	504 (9.9)	89 (5.8)	
Other	471 (9.2)	165 (11)	
Cardiovascular Risk Factors			
Body mass index (kg/m2)	29.6 (7.7)	28.8 (5.2)	<0.001
Obesity	2,299 (45)	593 (39)	<0.001
Hypertension	4,017 (79)	907 (59)	<0.001
Diabetes	1,676 (33)	317 (21)	<0.001
Hyperlipidemia	3,699 (73)	1,029 (67)	<0.001
Family history of coronary artery disease	1,715 (34)	576 (38)	0.004
Current smoker	1,135 (22)	207 (14)	<0.001
History of myocardial infarction	714 (14)	137 (8.9)	<0.001
Prior percutaneous coronary intervention	952 (19)	242 (16)	0.010
Prior coronary artery bypass graft	491 (9.6)	68 (4.4)	<0.001
Symptoms			
Angina	2,204 (43)	855 (56)	<0.001
Dyspnea	1,077 (21)	305 (20)	0.30
Hemodynamics			
Heart rate, rest (bpm)	69 (13)	66 (11)	<0.001
SBP, rest (mm Hg)	143 (25)	135 (18)	<0.001
DBP, rest (mm Hg)	80 (13)	80 (11)	0.39
Rate pressure product (mm Hg bpm)	9,853 (2,522)	8,920 (2,059)	<0.001
Heart rate, stress (bpm)	101 (21)	145 (21)	<0.001
SBP, stress (mm Hg)	161 (27)	190 (26)	<0.001
DBP, stress (mm Hg)	80 (14)	80 (14)	0.18

	Pharmacologic Stress, N = 5,102[1]	**Exercise Stress, N = 1,533**[1]	**p-value**[2]
SPECT Measurements			
LV ejection fraction, stress (%)	65 (13)	68 (10)	<0.001
Total perfusion deficit, stress (%)	5 (9)	3 (6)	<0.001
Outcomes			
All-cause mortality	1,118 (22)	91 (5.9)	<0.001

[1] Mean (SD); n (%)

[2] Welch Two Sample t-test; Pearson's Chi-squared test

Table S2: Diagnostic evaluation in the PET holdout test cohort (N=1031 patients). ECG AI model predictions of abnormal MFR with and without the SSL-pretrained foundation model.

	Estimate	SE	95% CI	p value†
No SSL				
AUROC	0.632	0.0173	(0.598, 0.666)	
Sensitivity	0.661	0.0213	(0.619, 0.702)	
Specificity	0.573	0.0214	(0.531, 0.615)	
PPV	0.588	0.0209	(0.547, 0.629)	
NPV	0.646	0.0219	(0.603, 0.689)	
SSL				
AUROC	0.758	0.0148	(0.729, 0.787)	< 0.0001
Sensitivity	0.701	0.0206	(0.661, 0.741)	< 0.0001‡
Specificity	0.694	0.0199	(0.655, 0.733)	
PPV	0.679	0.0207	(0.639, 0.720)	< 0.0001*
NPV	0.715	0.0198	(0.677, 0.754)	0.0002*

† SSL vs. No SSL
‡ Joint comparison of sensitivity and specificity by generalized McNemar test.
* Generalized score test.

Table S3: Diagnostic evaluation in the PET holdout test cohort (N=1031 patients). ECG AI model predictions of abnormal LVEF with and without the SSL-pretrained foundation model.

	Estimate	SE	95% CI	p value†
No SSL				
AUROC	0.918	0.0109	(0.897, 0.939)	
Sensitivity	0.876	0.0310	(0.815, 0.937)	
Specificity	0.825	0.0126	(0.800, 0.849)	
PPV	0.381	0.0301	(0.322, 0.440)	
NPV	0.982	0.0048	(0.972, 0.991)	
SSL				
AUROC	0.946	0.0097	(0.927, 0.965)	0.0176
Sensitivity	0.894	0.0290	(0.837, 0.951)	0.0078‡
Specificity	0.859	0.0115	(0.837, 0.882)	
PPV	0.439	0.0327	(0.375, 0.503)	0.0040*
NPV	0.985	0.0043	(0.977, 0.993)	0.5510*

† SSL vs. No SSL
‡ Joint comparison of sensitivity and specificity by generalized McNemar test.
* Generalized score test.

Table S4: Prognostic evaluation in the pharmacologic stress SPECT cohort (N=5102 patients). Cox Proportional Hazards models of ECG AI model predictions and risk of all-cause mortality.

	Impaired MFR		Impaired LVEF	
	AI-MFR No SSL	**AI-MFR SSL**	**AI-LVEF No SSL**	**AI-LVEF SSL**
Hazard Ratio (95% CI)	1.65 (1.46, 1.87)	2.30 (2.03, 2.61)	1.61 (1.21, 2.13)	1.47 (1.16, 1.87)
χ^2	473.3 $p < 0.0001$	577.0 $p < 0.0001$	364.2 $p = 0.0022$	363.9 $p = 0.0026$
c-index	0.684 $p = 0.0090$	0.705 $p < 0.0001$	0.665 $p = 0.1688$	0.663 $p = 0.9171$

Base model included patient age, sex, BMI, cardiovascular risk factors, history of myocardial infarction, percutaneous coronary intervention, and coronary artery bypass graft, LVEF (MFR base model only) and standard myocardial perfusion imaging findings. For each model, ECG AI model predictions were added to the base model.
The base model for Impaired MFR had χ^2 of 410.5 on 9 d.f. and c-index of 0.671.
The base model for Impaired LVEF had χ^2 of 364.2 on 8 d.f. and c-index of 0.663.
LVEF = left ventricular ejection fraction
MFR = myocardial flow reserve

Table S5: Prognostic evaluation in the exercise stress SPECT cohort (N=1533 patients). Cox Proportional Hazards models of ECG AI model predictions and risk of all-cause mortality.

	Impaired MFR		Impaired LVEF	
	AI-MFR No SSL	**AI-MFR SSL**	**AI-LVEF No SSL**	**AI-LVEF SSL**
Hazard Ratio (95% CI)	2.11 (1.39, 3.21)	3.76 (2.36, 5.99)	2.67 (0.61, 11.66)	2.43 (0.56, 10.66)
χ^2	56.8 p = 0.0005	71.6 p < 0.0001	39.6 p = 0.2282	39.4 p = 0.6543
c-index	0.736 p = 0.0050	0.723 p = 0.0901	0.691 p = 0.7413	0.691 p = 0.7250

Base model included patient age, sex, BMI, cardiovascular risk factors, history of myocardial infarction, percutaneous coronary intervention, and coronary artery bypass graft, LVEF (MFR base model only) and standard myocardial perfusion imaging findings. For each model, ECG AI model predictions were added to the base model.
The base model for Impaired MFR had χ^2 of 44.7 on 9 d.f. and c-index of 0.689.
The base model for Impaired LVEF had χ^2 of 38.1 on 8 d.f. and c-index of 0.689.
LVEF = left ventricular ejection fraction
MFR = myocardial flow reserve

Table S6: Continuous Net Reclassification Improvement (NRI) of ECG AI model MFR predictions with and without the SSL-pretrained foundation model.

	Overall NRI	Event NRI	Non-Event NRI
PET holdout			
No SSL AI-MFR	0.314 (0.133, 0.497)	0.354 (0.192, 0.517)	-0.040 (-0.107, 0.034)
SSL AI-MFR	0.538 (0.372, 0.702)	0.433 (0.274, 0.579)	0.105 (0.043, 0.168)
PET-MFR	0.534 (0.355, 0.706)	0.374 (0.206, 0.537)	0.161 (0.092, 0.224)
Pharm. stress SPECT			
No SSL AI-MFR	0.390 (0.319, 0.466)	0.130 (0.069, 0.196)	0.260 (0.231, 0.290)
SSL AI-MFR	0.594 (0.518, 0.670)	0.154 (0.088, 0.221)	0.441 (0.441, 0.470)
Exercise stress SPECT			
No SSL AI-MFR	0.427 (0.162, 0.699)	-0.010 (-0.266, 0.255)	0.437 (0.390, 0.486)
SSL AI-MFR	0.445 (0.179, 0.694)	-0.235 (-0.494, 0.0001)	0.681 (0.642, 0.718)

Table S7: Subgroup analysis in the PET holdout test cohort. ECG AI model predictions of MFR (AI-MFR) without SSL-pretrained foundation model and mortality risk.

Subgroup	AI-MFR<2	AI-MFR≥2	HR (95% CI)
All Patients	562	469	1.66 (1.18 - 2.34)
Age (p=0.340)			
<60	186	172	2.16 (1.08 - 4.31)
≥60	283	390	1.49 (1.02 - 2.18)
BMI (p=0.980)			
obese	186	246	1.63 (1.01 - 2.64)
non-obese	283	316	1.62 (1.02 - 2.58)
Sex (p=0.967)			
Female	222	219	1.61 (0.97 - 2.69)
Male	247	343	1.63 (1.05 - 2.56)
Diabetes (p=0.395)			
False	306	303	1.42 (0.90 - 2.25)
True	163	259	1.89 (1.15 - 3.13)
Dyslipidemia (p=0.766)			
False	193	165	1.73 (1.01 - 2.96)
True	276	397	1.56 (1.02 - 2.40)
Hypertension (p=0.830)			
False	135	120	1.54 (0.83 - 2.83)
True	334	442	1.66 (1.11 - 2.49)
Known CAD (p=0.719)			
False	357	340	1.56 (1.03 - 2.36)
True	112	222	1.77 (0.99 - 3.17)
LVEF (p=0.415)			
<50	36	188	1.23 (0.60 - 2.53)
≥50	433	374	1.74 (1.19 - 2.56)
Stress TPD (p=0.618)			
<5	354	326	1.75 (1.12 - 2.73)
≥5	115	236	1.47 (0.88 - 2.45)

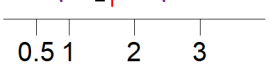


Table S8: Subgroup analysis in the PET holdout test cohort. ECG AI model predictions of MFR (AI-MFR) with SSL-pretrained foundation model and mortality risk.

Subgroup	AI-MFR<2	AI-MFR≥2	HR (95% CI)
All Patients	503	528	2.61 (1.83 - 3.71)
Age (p=0.752)			
<60	212	146	2.83 (1.45 - 5.52)
≥60	316	357	2.50 (1.67 - 3.75)
BMI (p=0.418)			
obese	216	216	2.99 (1.80 - 4.97)
non-obese	312	287	2.26 (1.41 - 3.62)
Sex (p=0.472)			
Female	237	204	2.24 (1.33 - 3.77)
Male	291	299	2.87 (1.81 - 4.56)
Diabetes (p=0.927)			
False	354	255	2.55 (1.61 - 4.04)
True	174	248	2.63 (1.55 - 4.45)
Dyslipidemia (p=0.425)			
False	204	154	3.08 (1.75 - 5.40)
True	324	349	2.31 (1.50 - 3.58)
Hypertension (p=0.816)			
False	145	110	2.75 (1.45 - 5.20)
True	383	393	2.52 (1.67 - 3.80)
Known CAD (p=0.272)			
False	391	306	2.98 (1.92 - 4.64)
True	137	197	2.02 (1.17 - 3.49)
LVEF (p=0.997)			
<50	66	158	2.58 (1.22 - 5.46)
≥50	462	345	2.58 (1.74 - 3.84)
Stress TPD (p=0.199)			
<5	399	281	3.12 (1.97 - 4.96)
≥5	129	222	1.99 (1.19 - 3.32)

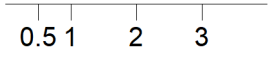


Table S9: Subgroup analysis in the pharmacologic stress SPECT cohort. ECG AI model predictions of MFR (AI-MFR) with SSL-pretrained foundation model and mortality risk.

Subgroup	AI-MFR<2	AI-MFR≥2	HR (95% CI)
All Patients	1728	3374	2.30 (2.03 - 2.61)
Age (p=0.858)			
<60	928	373	2.42 (1.84 - 3.19)
≥60	2446	1355	2.49 (2.17 - 2.86)
BMI (p=0.360)			
obese	1840	963	2.37 (2.03 - 2.77)
non-obese	1534	765	2.66 (2.18 - 3.23)
Sex (p=0.267)			
Female	1566	662	2.69 (2.22 - 3.27)
Male	1808	1066	2.34 (2.00 - 2.74)
Diabetes (p=0.446)			
False	2447	979	2.57 (2.20 - 3.01)
True	927	749	2.34 (1.92 - 2.84)
Dyslipidemia (p=0.976)			
False	996	407	2.47 (1.95 - 3.13)
True	2378	1321	2.48 (2.15 - 2.86)
Hypertension (p=0.757)			
False	838	247	2.59 (1.92 - 3.49)
True	2536	1481	2.46 (2.14 - 2.81)
Known CAD (p=0.412)			
False	2598	1089	2.39 (2.05 - 2.78)
True	776	639	2.65 (2.15 - 3.27)
LVEF (p=0.061)			
<50	218	399	1.88 (1.39 - 2.55)
≥50	3156	1329	2.60 (2.27 - 2.97)
Stress TPD (p=0.702)			
<5	2745	1103	2.51 (2.17 - 2.91)
≥5	629	625	2.39 (1.90 - 2.99)

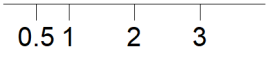
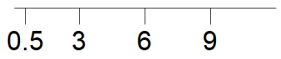


Table S10: Subgroup analysis in the exercise stress SPECT cohort. ECG AI model predictions of MFR (AI-MFR) with SSL-pretrained foundation model and mortality risk.

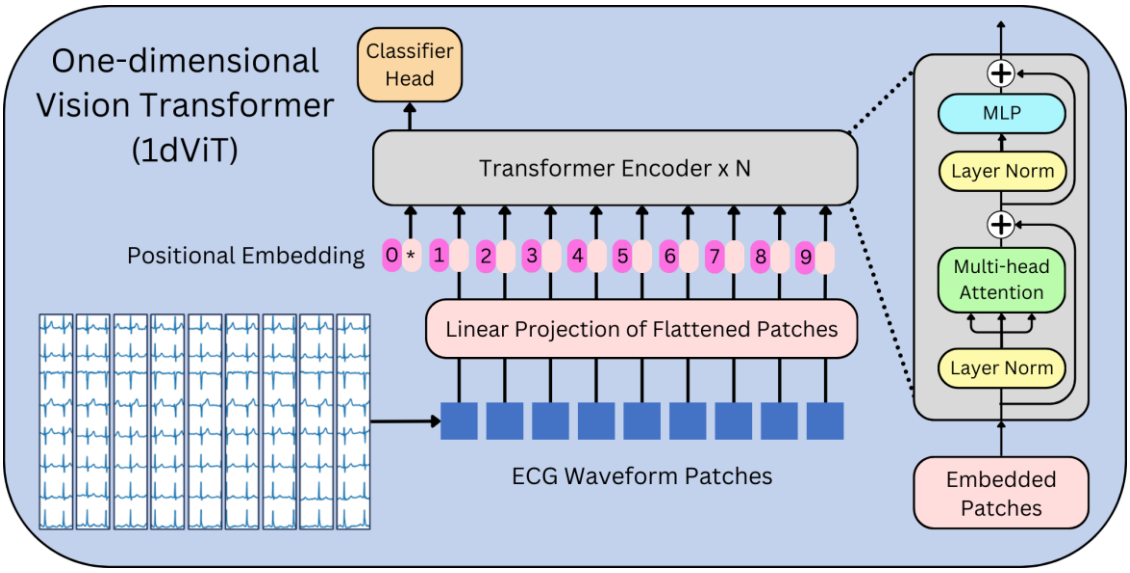
Subgroup	AI-MFR<2	AI-MFR≥2	HR (95% CI)
All Patients	161	1372	3.76 (2.36 - 5.99)
Age (p=0.846)			
<60	658	55	3.35 (1.39 - 8.09)
≥60	714	106	3.70 (2.17 - 6.31)
BMI (p=0.916)			
obese	846	94	3.54 (1.98 - 6.32)
non-obese	526	67	3.71 (1.83 - 7.52)
Sex (p=0.152)			
Female	511	48	5.59 (2.68 - 11.69)
Male	861	113	2.88 (1.64 - 5.06)
Diabetes (p=0.408)			
False	1116	100	4.08 (2.37 - 7.00)
True	256	61	2.74 (1.23 - 6.10)
Dyslipidemia (p=0.575)			
False	463	41	3.01 (1.35 - 6.72)
True	909	120	3.97 (2.24 - 7.04)
Hypertension (p=0.884)			
False	593	33	3.35 (1.11 - 10.12)
True	779	128	3.66 (2.20 - 6.08)
Known CAD (p=0.419)			
False	1108	113	3.22 (1.85 - 5.58)
True	264	48	4.81 (2.06 - 11.22)
LVEF (p=0.833)			
<50	45	26	3.24 (1.08 - 9.70)
≥50	1327	135	3.68 (2.23 - 6.08)
Stress TPD (p=0.341)			
<5	1218	112	4.12 (2.43 - 6.99)
≥5	154	49	2.57 (1.12 - 5.89)



12.2 Supplemental Figures

Figure S1: ECG AI model architecture. (a) The one-dimensional Vision Transformer (1dViT) model; (b) Fusion model combining two 1dViT channels for rest and stress ECG input.

(a)



(b)

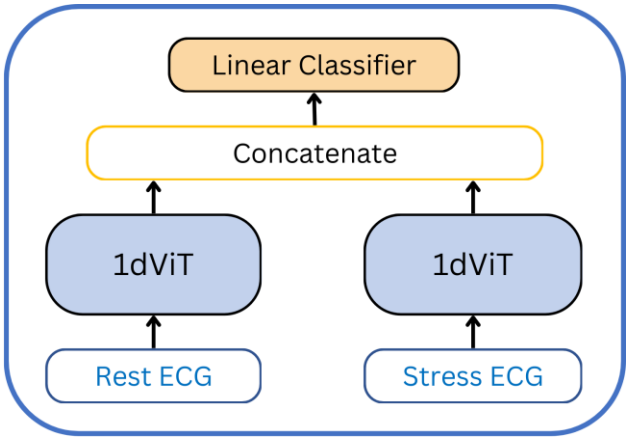
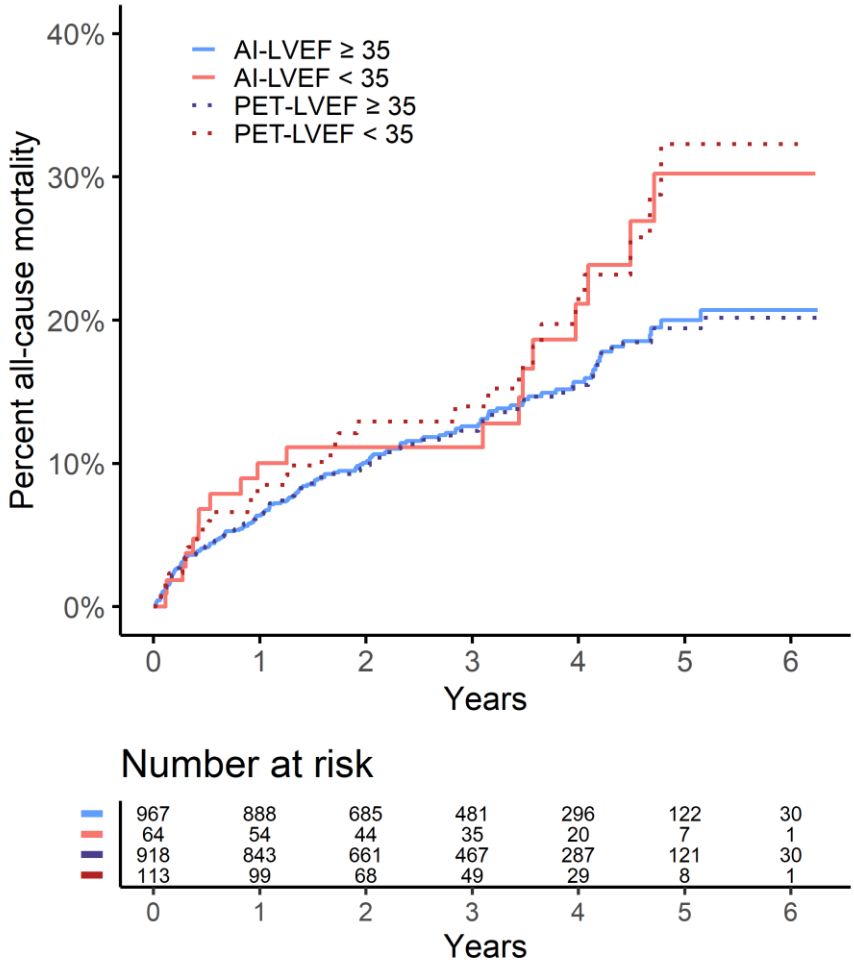


Figure S2: Adjusted incidence of all-cause mortality between ECG AI model predictions of abnormal LVEF (AI-LVEF, solid lines) without (a), and with (b) the SSL-pretrained foundation model. Dotted lines show a model of PET-measured LVEF. Evaluation was performed on the PET holdout test cohort (N=1031 patients).

(a)



(b)

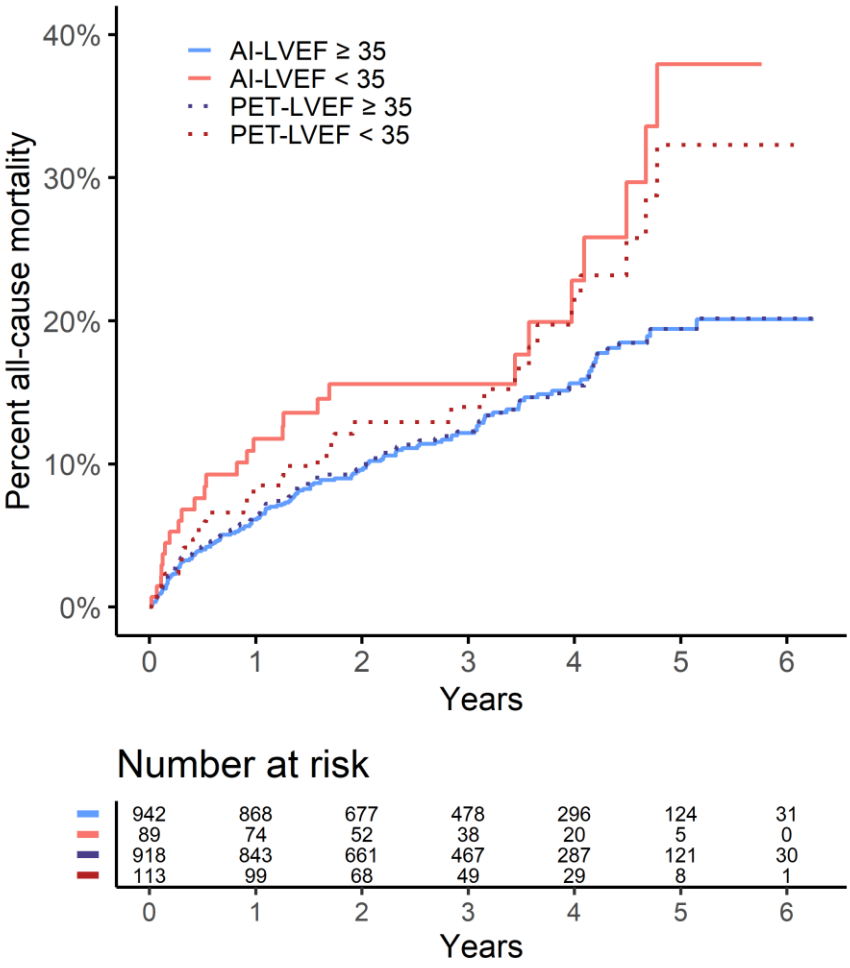
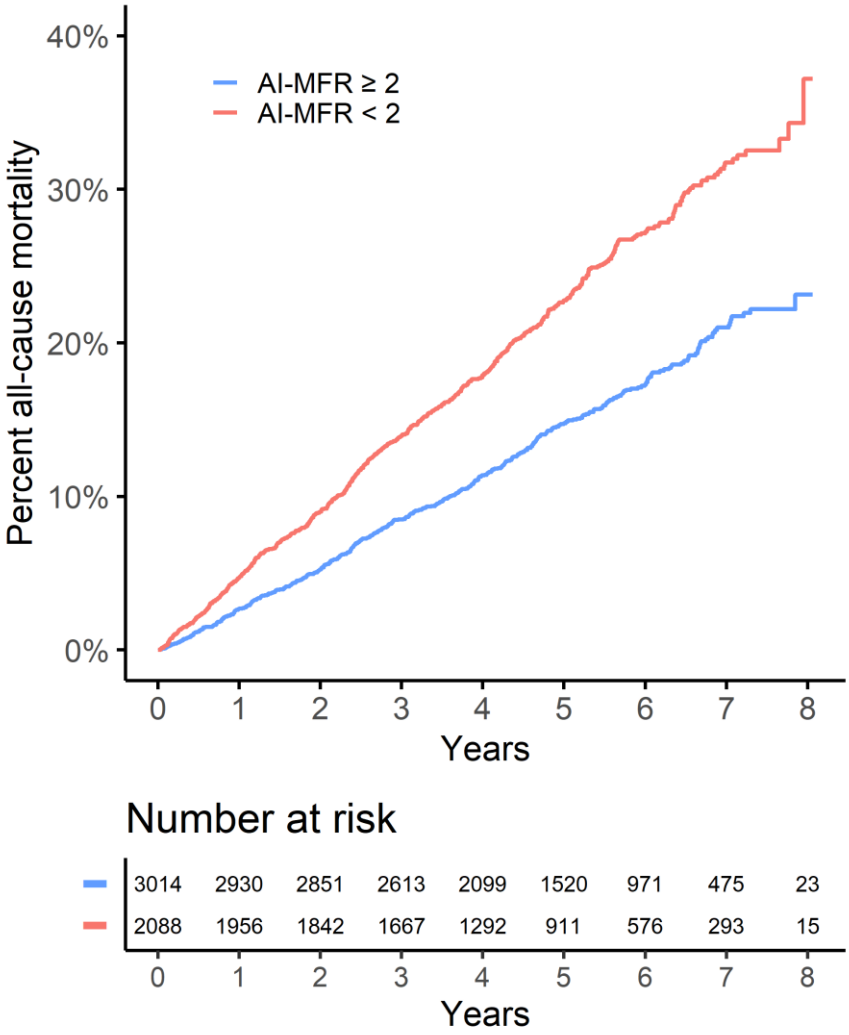


Figure S3: Adjusted incidence of all-cause mortality between ECG AI model predictions of abnormal MFR (AI-MFR) without (a), and with (b) the SSL-pretrained foundation model. Evaluation was performed on the pharmacologic stress SPECT cohort (N=5102 patients).

(a)



(b)

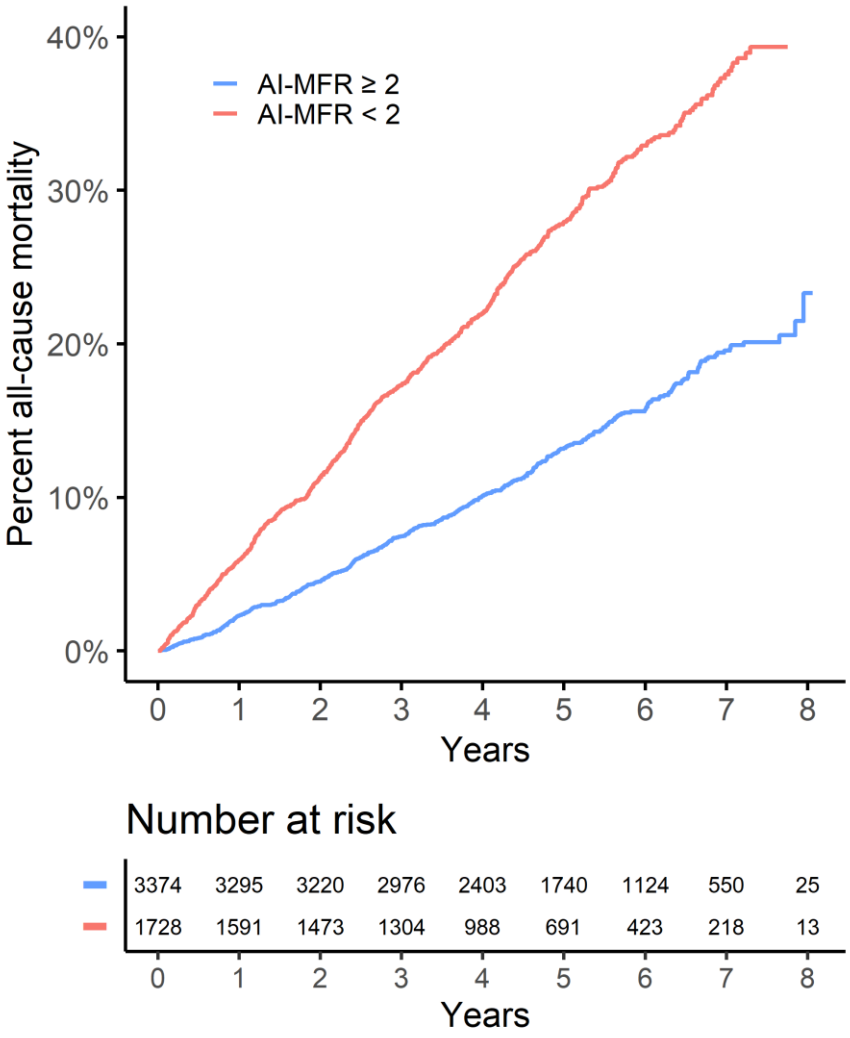
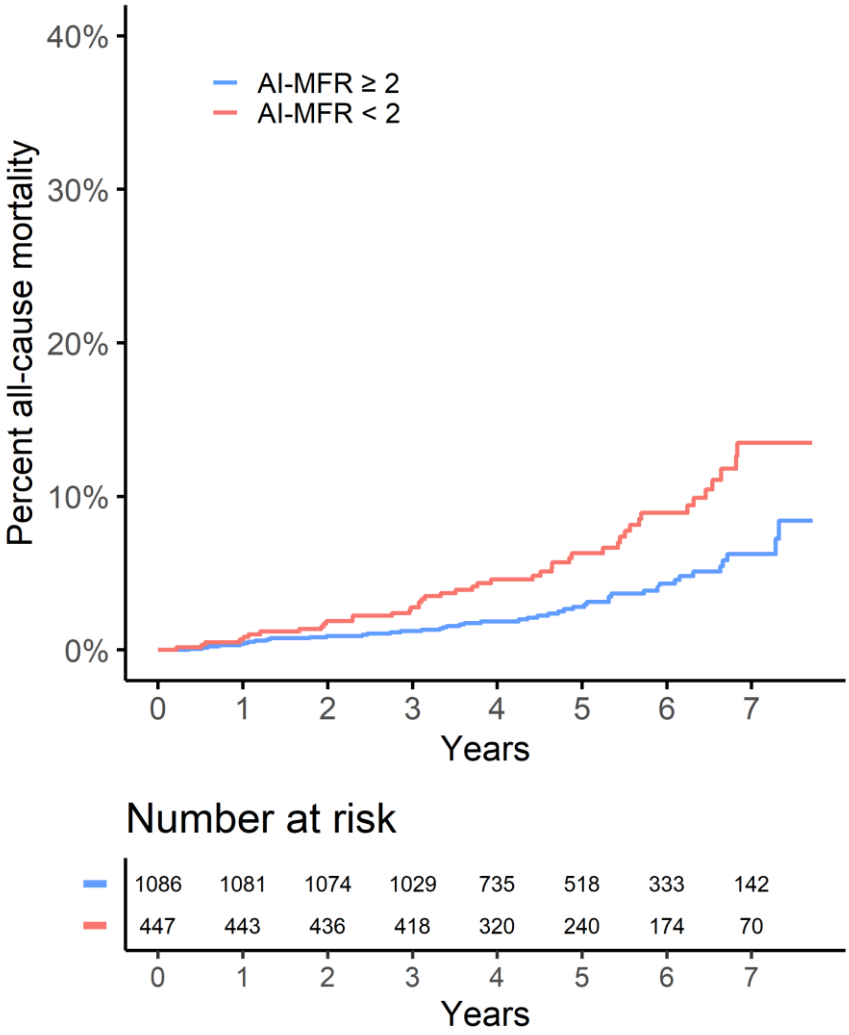


Figure S4: Adjusted incidence of all-cause mortality between ECG AI model predictions of abnormal MFR (AI-MFR) without (a), and with (b) the SSL-pretrained foundation model. Evaluation was performed on the exercise stress SPECT cohort (N=1533 patients).

(a)



(b)