Machine Learning for the ECG Diagnosis and Risk Stratification of Occlusion Myocardial Infarction at First Medical Contact

Results from ECG-SMART Observational Trial

By

Salah S. Al-Zaiti, PhD;[1]* Christian Martin-Gill, MD, MPH;[1-2] Jessica K. Zègre-Hemsey, PhD, RN;[3] Zeineb Bouzid, MS, PhD(c);[1] Ziad Faramand, MD, MS;[3] Mohammad O. Alrawashdeh, PhD;[5] Richard E. Gregg, MS;[6] Stephanie Helman, RN, PhD(c);[1] Nathan Riek, MS;[1] Karina Kraevsky-Phillips, RN, PhD(c),[1] Gilles Clermont, MD;[1,7] Murat Akcakaya, PhD;[1] Susan M. Sereika, PhD;[1] Peter Van Dam, PhD;[8] Stephen W. Smith, MD;[9] Yochai Birnbaum, MD;[10] Samir Saba, MD;[1-2] Ervin Sejdic, PhD;[11] and Clifton W. Callaway, MD, PhD[1-2]

From

(1) University of Pittsburgh, Pittsburgh, PA, USA; (2) University of Pittsburgh Medical Center (UPMC), Pittsburgh, PA, USA; (3) University of North Carolina – Chapel Hill, NC, USA; (4) Northeast Georgia Health System, GA, USA; (5) Harvard Medical School, Boston, MA, USA; (6) Philips Healthcare, Cambridge, MA, USA; (7) VA Healthcare System, Pittsburgh, PA, USA; (8) University Medical Center Utrecht, The Netherlands; (9) Hennepin Healthcare and University of Minnesota, Minneapolis, MN, USA; (10) Baylor College of Medicine, Houston, TX, USA; and (11) University of Toronto & North York General Hospital, Toronto, Canada

Trial Registration: ClinicalTrials.gov Identifier: NCT04237688

Funding: National Institute of Health grants # R01HL137761 (SSA), UL1TR001857 (SSA), K23NR017896 (JZH), and KL2TR002490 (JZH)

Conflicts of Interest: none

*** Corresponding Author:** ssa33@pitt.edu

Abstract: 149 / 150

Word count: ~~34~~,~~836~~ 097 / 4,000 (excluding abstract, methods, and references)

References: 60 / 60

Display items: 6 / 6

Extended data display items: ~~7~~ 8 / 10

ABSTRACT

Patients with occlusion myocardial infarction (OMI) and no ST-elevation on presenting ECG are increasing in numbers. These patients have a poor prognosis and would benefit from immediate reperfusion therapy, but we currently have no accurate tools to identify them during initial triage. Herein, we report the first observational cohort study to develop machine learning models for the ECG diagnosis of OMI. Using 7,313 consecutive patients from multiple clinical sites, we derived and externally validated an intelligent model that outperformed practicing clinicians and other widely used commercial interpretation systems, significantly boosting both precision and sensitivity. Our derived OMI risk score provided enhanced rule-in and rule-out accuracy relevant to routine care, and when combined with the clinical judgment of trained emergency personnel, this score helped correctly reclassify one in three patients with chest pain. ECG features driving our models were validated by clinical experts, providing plausible mechanistic links to myocardial injury.

The ECG diagnosis of acute coronary syndrome (ACS) in patients with acute chest pain is a longstanding challenge in clinical practice.[1-4] Guidelines primarily focus on ST-segment elevation (STE) for discerning patients with ST-elevation myocardial infarction (STEMI) vs. other forms of ACS.[5-8] A biomarker-driven approach is recommended in the absence of STE on the presenting ECG. This diagnostic paradigm has two important limitations. First, around 24%−35% of patients with non-STEMI have total coronary occlusion, referred to as occlusion myocardial infarction (OMI), and require emergent catheterization.[9-13] This vulnerable group, in contrast to ~~acute myocardial infarction~~ACS with an open artery ~~(non-OMI)~~ (**Extended Data Fig. 1**), suffers from unnecessary diagnostic and treatment delays that are associated with higher mortality.[14-17] This excess risk can be mitigated with enhanced diagnostic criteria. Although important ECG signatures of OMI are frequently described in the literature,[18-21] they are subtle, involve the entire QRST complex, and are spatial in nature (i.e., changes diluted across multiple leads).[22-24] Visual inspection of ECG images by clinical experts, thus, is suboptimal and leads to a high degree of variability in ECG interpretation.[25-27]

The second limitation is that cardiac biomarkers, including conventional or high sensitivity troponin (hs-cTn), cannot differentiate OMI until peak level is reached, which is too late to salvage myocardium. Positive troponin results (>99th percentile limit) come with a high false positive rate, and approximately one-third of patients remain in a biomarker-indeterminate “observation zone” even after serial sampling.[28,29] More importantly, ~25% of acute myocardial infarction cases have a negative initial hs-cTn, which is observed in both the STEMI and OMI subgroups.[30] Consequently, 25%-30% of patients with OMI are not treated in a timely fashion, and around 63% (IQR 38%-81%) of patients evaluated for chest pain at the emergency department are admitted to the hospital because of an inconclusive initial assessment.[31] These diagnostic limitations have created a costly, inefficient clinical practice paradigm where most patients with chest pain are over-monitored while some patients with OMI have delayed diagnosis and treatment, potentially contributing to the 14%−22% excess risk of mortality seen in the non-STE ACS group (NSTE-ACS).[15,32,33]

In our prior work, we designed prototype algorithms for AI-enabled ECG analysis and demonstrated the clinical feasibility of screening for ACS in the prehospital setting.[34,35] Herein, we describe the first multisite, prospective, observational cohort study to evaluate the diagnostic accuracy of machine learning for the ECG diagnosis and risk stratification of OMI at first medical contact in an observer-independent approach (**Extended Data Fig. 2**). Our intelligent models were derived and externally validated on 7,313 patients with chest pain from multiple clinical sites in the United States. The results demonstrate the superiority of machine learning in detecting subtle ischemic ECG changes indicative of OMI, outperforming practicing clinicians and other widely used commercial ECG interpretation software. Our derived OMI risk score provides ~~superior~~ enhanced rule-in and rule-out accuracy when compared to the HEART score, helping correctly reclassify one in three patients with chest pain. We identified the most important ECG features driving our model's classifications and identified plausible mechanistic links to myocardial injury.

RESULTS

Sample Characteristics

After excluding patients with cardiac arrest, ventricular tachyarrhythmias, confirmed prehospital STEMI, and duplicate ECGs, our derivation cohort included 4,026 consecutive patients with chest pain (age 59±16 years, 47% females, 5.2% OMI). The two external validation cohorts together included 3,287 patients (age 60±15 years, 45% females, 6.4% OMI) (**Fig. 1** and **Table 1**). Most patients in the derivation and validation cohorts were in normal sinus rhythm (>80%) and around 10% were in atrial fibrillation. Around 3% of patients had left bundle branch block (LBBB) and ~10% had ECG-evidence of left ventricular hypertrophy (LVH). The derivation and validation cohorts were comparable in terms of age, sex, baseline clinical characteristics, and 30-day cardiovascular mortality. The validation cohort, however, had more Black and Hispanic minorities and a slightly higher rate of ACS and OMI. ~~The~~

~~presence of OMI, defined as a culprit coronary artery with a TIMI flow grade of 0-1, was adjudicated from charts by independent reviewers blinded to all ECG analyses. A TIMI flow grade of 2 with significant coronary narrowing (>70%) and peak 4th generation (not high sensitivity) troponin of 5-10 ng/mL was also indicative of OMI.~~

Algorithm Derivation and Testing

The positive class for model training was the presence of OMI, defined as a culprit coronary artery with a TIMI flow grade of 0-1, as adjudicated from charts by independent reviewers blinded to all ECG analyses. A TIMI flow grade of 2 with significant coronary narrowing (>70%) and peak 4th generation (not high sensitivity) troponin of 5-10 ng/mL was also indicative of OMI. The negative class for model training was the absence of OMI, which included all other non-ACS etiologies or those with non-coronary occlusive ACS subtypes.

Input data for model training was based on prehospital 12-lead ECGs obtained at first medical contact. We selected 73 morphological ECG features out of 554 temporal-spatial metrics using a hybrid data-driven and domain expertise approach.[18] Using these features, ten classifiers were trained to learn ischemic patterns between ACS and non-ACS groups and to estimate the probability of OMI: regularized logistic regression, linear discriminant analysis, support vector machine, Gaussian Naïve Bayes, random forest, gradient boosting machine, extreme gradient boosting, stochastic gradient descent logistic regression, k-nearest neighbors, and artificial neural networks. We chose these classifiers ~~because they learn different mathematical representations in the data,~~ in order to maximiz~~e~~ing the chance of finding the best fitting ~~modeling~~approach ~~approach~~ for learning the mathematical representation relating complex ECG data to underlying physiology.

The random forest model achieved the best bias-variance tradeoff for training and internal testing. We compared the random forest against the ECG interpretation of practicing clinicians and against the performance of a commercial ECG interpretation system that is FDA-cleared for “Acute MI” diagnosis. On the hold-out test set, the random forest model (AUROC 0.91 [95% CI 0.87-0.96])

outperformed both practicing clinicians (AUROC 0.79 [95% CI 0.73-0.76], p<0.001) and the commercial ECG system (AUROC 0.78 [95% CI 0.70-0.85], p<0.001) (**Fig. 2A**).

Next, we used probability density plots for OMI(+) and OMI(-) classes to denote the optimal separation margins for risk prediction. As recommended by guidelines,[6] we defined a risk score to identify patients at low risk (OMI score <5), intermediate risk (OMI score 5-20), and high risk (OMI score >20), with these cutoffs yielding excellent separation between classes (Log-rank chi-square 133.04, df=2, p<0.001) (**Fig. 2B, left panel**). Our OMI score classified 74.4% of patients as low-risk and 4.6% as high-risk. Using the low-risk group in a rule-out strategy yielded a sensitivity of 0.91 and a negative predictive value (NPV) of 0.993, with an overall missed event rate of 0.5%. Using high-risk class for a rule-in strategy yielded a specificity of 0.976 and a positive predictive value (PPV) of 0.514, with an overall false discovery rate of 2%. Finally, we compared this OMI score to the HEART score, which uses patient history, ECG data, age, risk factors, and troponin values (**Fig. 2B, right panel**). Our OMI score, which is based on ECG data alone, classified 66% more patients as low risk than the HEART score with a comparable false negative rate <1%, and classified fewer patients as high-risk and with much higher precision (51% vs. 33%). The OMI score also triaged 50% fewer patients as intermediate risk and still got better discrimination for OMI detection (11.2% vs. 5.6%).

Model Explainability

We used Tree SHAP algorithms to generate an importance ranking that explains the output of the random forest model based on SHAP values estimated for the top 25 features (**Fig. 3A**). The features with the greatest impact on classification output included slight ST-depression in leads V1, V2, I, and aVL; slight ST-elevation in leads III and V4-V6; loss of concave pattern in anterior leads; T wave enlargement in II and aVF and T flattening or inversion in I and aVL; prolonged T_{peak}-T_{end} interval; T axis deviation; increased repolarization dispersion; and distorted directions of activation and recovery patterns. Most of these ECG patterns can be mechanistically linked to cardiac ischemia, suggesting their clinical value as plausible features for OMI detection.

To better visualize these global ECG patterns detected by our model, we created pooled population median beats for the OMI(+) class (n=414 ECGs), and superimposed these median beats on the pooled population median beats of patients with normal sinus rhythm and OMI(-) status (n=9,072 ECGs) (**Fig. 3B**). Findings from this figure agree with the patterns derived from the SHAP values described above. Specifically, this figure illustrates that OMI is associated with ST-depression and T flattening in V1-V2, I, and aVL; slight ST-elevation in the anterior leads with loss in concave pattern; peaked T wave in inferior leads; T_{peak}-T_{end} prolongation (seen in many leads); global repolarization dispersion (seen as peaked T in some leads and flattening in others); T axis deviation (away from the left ventricle), and distorted activation and recovery patterns (seen in the horizontal plane as loss of R wave progression in precordial leads with increased T wave discordance). Due to prevalent multivessel disease in this cohort, these OMI patterns remained relatively consistent regardless of culprit location.

Nevertheless, to examine local explainability of feature importance, we used force plots on individual cases to identify the features that met the contribution threshold of the random forest model on a given ECG. These force plots were also examined by study investigators to further corroborate on the clinical validity of model predictions. **Extended Data Fig. 3** shows a selected example of a 12-lead ECG with its corresponding force plot for the local features contribution.

Formatted: Font: Bold

External Validation

We tested the final lock-out model on 3,287 patients from two independent external clinical sites. Machine learning engineers were blinded to outcome data from other sites, and the pre-populated model predictions were independently evaluated by the clinical investigators. Our model generalized well and maintained high classification performance (AUROC 0.873 [95% CI 0.85-0.90]), outperforming the classification performance of the commercial ECG system (AUROC 0.75 [95% CI 0.71-0.79], p<0.001) and practicing clinicians (AUROC 0.80 [95% CI 0.77-0.83], p<0.001) (**Fig. 4A**). Our OMI risk score was a strong predictor of OMI, independent from, age, sex, and other coronary risk factors (OR

10.6 [95% CI 6.78-16.64] for high-risk class and OR 2.85 [95% CI 1.91-4.28] for intermediate-risk class) (**Fig. 4B**). This risk score triaged 69% of patients in the low-risk group at a false-negative rate of 1.3% and identified 5.1% of patients as high-risk at acceptable true positive rate >50%. The overall sensitivity, specificity, PPV, and NPV for the OMI rule-in and rule-out strategy were 0.86 (95% CI 0.81-0.91), 0.98 (95% CI 0.97-0.99), 0.54 (95% CI 0.46-0.62), and 0.99 (95% CI 0.98-0.99), respectively. This diagnostic accuracy remained relatively similar across subgroups based on age, sex, race, comorbidities, and baseline ECG findings, indicating the lack of aggregation bias (**Fig. 4C**). In comparison, the sensitivity, specificity, PPV, and NPV for ECG overread by practicing clinicians were 0.58, 0.93, 0.36, and 0.97, and for the commercial ECG system 0.79, 0.80, 0.22, and 0.98, respectively.

Next, we ~~used decision analysis to~~ evaluate the incremental gain of our derived risk score in re-classifying patients at first medical contact (**Fig. 5**). ~~To simulate i~~Initial assessment by emergency personnel ~~, we used~~ was based on the modified HEAR score (History, ECG, Age, and Risk factors) to triage patients into low, intermediate, and high-risk groups.[36] At baseline, emergency personnel triaged 48% of patients as low risk with a NPV of 99.0% and triaged 3% of patients as high risk with a PPV of 54.1%. Nearly 50% of patients remained in an indeterminate observation zone. Applying our OMI risk score would help triage 45% more patients as low risk while keeping the NPV at 98.8% and would help detect 85% more cases with OMI while keeping PPV at 50.0%. The OMI score would also help reduce the number of patients in the indeterminate observation zone by more than half. These numbers translate into a net reclassification improvement (NRI) index of 41% (95% CI 33%-50%). To validate this incremental clinical utility, we manually reviewed ECGs reclassified correctly as OMI(+) (**Extended data Fig. ~~3~~4**). Many of these ECGs showed subtle or nonspecific changes that were nondiagnostic as per guidelines,[5] suggesting potential value in boosting provider's confidence when interpreting "fuzzy" ECGs.

Finally, we investigated the potential sources of false negatives in the validation data. Among those with missed OMI events (n=28, 0.9%), many patients had high-frequency noise and baseline

wander on their initial ECG (n=13/28, 46%) or had low voltage ECG (n=14/28, 50%), and most patients (n=24/28, 86%) had benign ECGs without any diagnostic ST-T changes (**Extended Data Fig. ~~4~~5**). Moreover, we found no significant differences between false negatives and true positives in terms of demographics or clinical characteristics, with the exception that most false negatives had a history of a prior myocardial infarction (93% vs. 27%). The latter finding was intriguing given that our OMI model was slightly less specific in patients with known coronary artery disease (**Fig. 4C**). This finding aligns with recent evidence showing diminished NPV in patients with chest pain and known CAD.[37]

Screening for Any ACS Event

We further built a model to screen for any potential ACS event at first medical contact. Using the same set of ECG features, we trained and optimized a random forest classifier that denotes the likelihood of any ACS event. The model performed well during training (AUROC 0.88 [95% CI 0.87-0.90]) and generalized well during internal testing (AUROC 0.80 [95% CI 0.76-0.84]), outperforming both the commercial ECG interpretation system (AUROC 0.62 [95% CI 0.55-0.68], p<0.001) and practicing clinicians (AUROC 0.66 [95% CI 0.59-0.72], p<0.001) (**Extended Data Fig. ~~5~~6**). On external validation, the model continued to generalize well (AUROC 0.79 [95% CI 0.76-0.81]), outperforming the commercial system (AUROC 0.68 [95% CI 0.65-0.71], p < 0.001) and practicing clinicians (AUROC 0.72 [95% CI 0.69-0.74], p < 0.001). Our derived risk score provided a suboptimal rule-out classification for any ACS event (sensitivity 68.2% and NPV 92.5%) but provided superior rule-in accuracy (specificity 98.9% and PPV 82.5%).

DISCUSSION

In this study, we developed and validated a machine learning algorithm for the ECG diagnosis of OMI in consecutive patients with chest pain recruited from multiple clinical sites in the United States. This model outperformed practicing clinicians and other commercial interpretation systems. The derived risk score provided superior rule-in and rule-out accuracy for OMI, boosting the sensitivity by 7 to 28 percentage points and the precision by 18 to 32 percentage points compared to reference standards.

When combined with the judgment of experienced emergency personnel, our derived OMI risk score helped correctly reclassify one in three patients with chest pain. To our knowledge, this is the first study using machine learning methods and novel ECG features to optimize OMI detection in patients with acute chest pain and negative STEMI pattern on their baseline ECG at first medical contact.

Mapping myocardial ischemia, a problem of regional metabolic derangement, to coronary occlusion, a problem of diminished blood flow due to an atherosclerotic plaque rupture, is a complex process.[1] Essentially, ischemia disproportionately distorts action potentials in different myocardial segments, resulting in tissue-scale currents, often called 'injury' currents. Prior studies have mapped significant ST-elevation to transmural injury currents associated with total coronary occlusion. This has historically driven the current paradigm dichotomy of STEMI vs. *'others'* (any ACS other than STEMI) in determining who might benefit from emergent reperfusion therapy. However, nearly 65% of patients with ACS present with no ST-elevation on their baseline ECG,[35,38] and among the latter group, 24%−35% have total coronary occlusion requiring emergent catheterization.[9-13] Thus, determining who would benefit from reperfusion therapy remains an adjudicated diagnosis.

Conceptually, injury currents produced by ischemic cardiac cells are summative in nature, explaining how ST amplitude changes can get attenuated on the surface ECG (**Extended Data Fig. ~~6~~7**). These injury currents, however, distort the propagation of both excitation and recovery pathways, altering the configuration of the QRS complex and the STT waveform altogether.[39] Thus, a more comprehensive approach for the ECG detection of ischemia should focus on (1) evaluating temporal characteristics over entire waveform segments rather than the voltage at a given time point (e.g., J+80), and (2) evaluating lead-to-lead spatial characteristics in waveform morphology rather than absolute changes in isolated ECG leads.[1]

This study has identified several ECG patterns indicative of acute coronary occlusion beyond the criteria recommended by clinical guidelines.[5] Intriguingly, these ECG patterns overlap with those described in the literature. A consensus report in 2012 identified few ECG patterns that should be

241 treated as STEMI equivalent during acute pain episodes: ST-depression in V1 to V3; small inverted T
242 waves in V1 to V3; deep negative T waves in precordial leads; widespread ST-depression, and
243 prominent positive T waves.[20] Similar ECG patterns were also described more recently: ST-depression
244 in V1 to V4 (versus V5-V6); reciprocal ST-depression with maximal ST-depression vector towards the
245 apex (leads II and V5, with reciprocal STE in aVR); subtle ST-elevation; acute pathologic Q waves;
246 hyperacute T waves; and loss of terminal S wave.[21] Many of these expert-driven patterns rely on
247 assessing the proportion of repolarization amplitudes or area under the QRS amplitude. They also rely
248 heavily on the visual assessment of waveform morphology and can introduce a high degree of
249 subjectivity and variability among ECG interpreters. We demonstrated that the machine learning
250 models described herein not only outperform practicing clinicians in identifying OMI, but also provided
251 an objective, observer-independent approach to quantify subtle ECG patterns associated with OMI.

252 Many of the data-driven features identified by our machine learning model are subtle and cannot
253 be easily appreciated by clinical experts. T feature indices were among these most important features,
254 including T_{peak}-T_{end} interval prolongation, T wave flattening, and T wave characteristics at the inflection
255 point preceding T_{peak} (**Fig. 3A**). Mechanistically, ischemic injury currents interfere with signal
256 propagation leading to longer activation time.[40] These late activation potentials lead to a loss of terminal
257 S wave and longer recovery time, both manifesting as T wave flattening, shifted T peak, and loss of
258 concavity at the initial T wave (**Fig. 3B**). These STEMI-equivalent patterns were previously described in
259 the literature as small or negative T waves with widespread ST-depression or subtle ST- elevation.[20,21]
260 Another important subtle feature identified by our model was increased ventricular repolarization
261 dispersion, measured using the ratio between the principal components of the STT waveforms (i.e.,
262 PCA metrics), the direction of the T axis, and the angle between activation and recovery pathways
263 (e.g., total-cosine-R-to-T). Injury currents disproportionately affect the duration and velocity of
264 repolarization across different myocardial segments,[41] resulting in lead-to-lead variability in the
265 morphology of the STT waveform.[22-24,39,42] These high-risk ECG patterns were previously described as
266 a mixture of deep negative T waves and prominent / hyperacute T waves or reciprocal T wave

changes.[20,21] Despite their subtle nature, our machine learning model provided a more comprehensive, quantitative approach to evaluating this inter-lead variability in repolarization morphology.

Machine learning is well-suited to address many challenges in 12-lead ECG interpretation. Myocardial ischemia distorts the duration and amplitude of the Q wave, R peak, R`, QRS complex, ST segment, and T wave, as well as the morphology and configuration of these waveforms (e.g., upsloping, down-sloping, concavity, symmetry, notching, etc.). These distortions are lead-specific yet come with dynamic inter-lead correlations. Thus, ECG interpretation involves many complex aspects and parameters, making it a highly dimensional, decision space problem.[1] Few experienced clinicians excel in such pattern recognition,[21] which explains why so many OMI cases are not reperfused in a timely way; this is also why simple, rule-based commercial systems that use simple regression models are suboptimal for OMI detection. Machine learning algorithms can provide powerful tools to solve such highly dimensional, non-linear mathematical representations found in 12-lead ECG data.

Although the literature on machine learning for the ECG diagnosis of coronary disease is ubiquitous, it comes with many serious limitations. First, many studies focused on detecting the known STEMI group or other subtle ACS phenotypes[34,35,43,44] rather than the critical group without ST-elevation, which is not classified as STEMI and is therefore excluded from STEMI databases. Second, most prior work used open-source ECG datasets like PTB and PTB-XL,[45] which are highly selected datasets that focus on ECG-adjudicated diagnoses. Our unique cohorts included unselected, consecutive patients with clinical profiles and disease prevalence like that seen in real-world settings. Third, many studies used a full range of input features based on both ECG data and clinical data elements (e.g., patient history, physical exam abnormalities, laboratory values, diagnostic tests),[46-49] which limits the applicability to real-world settings. Fourth, to our knowledge, most studies used a single derivation cohort for training and testing,[50] without the use of an independent validation cohort. Finally, prior studies paid little attention to model explainability,[51] shedding little light on novel markers and

pathways of ischemia than what is already known. Without explanation aids of clinical meaningfulness, machine learning models for ECG interpretation would have limited clinical utility.[52]

This study has important clinical implications. Our machine learning model can be integrated into systems of care for real-time deployment where risk score assignments can be made readily available to clinicians right at time of ECG acquisition. Such enhanced decision support can help emergency personnel identify 85% more patients with critical coronary occlusion despite the absence of a STEMI pattern on the presenting ECG and without any loss in precision. Our models can also help inform care in more than 50% of patients in whom the initial assessment is indeterminate, placing 45% more patients in the low-risk group for OMI without any loss in NPV. This incremental gain in rule-in and rule-out accuracy can help re-allocate critical emergency resources to those in utmost need while optimizing the clinical workflow. This can impact numerous decisions at first medical contact, including targeted prehospital interventions, catheterization lab activation, administration of anti-ischemic therapies, hospital destination decisions, the need for medical consults, referrals for expedited diagnostic testing (e.g., echocardiogram, imaging scans), and early discharge decisions. Furthermore, until now, clinicians never had sensitive nor highly specific tools that would allow the ultra-early identification of OMI in the absence of a STEMI pattern. Such enhanced diagnostics can allow the design and implementation of prospective interventional trials to assess the therapeutic effectiveness of targeted interventions in this vulnerable group (e.g., early upstream $P2Y_{12}$ inhibitor administration,[53] emergent vs. delayed reperfusion therapy,[54] glucose-insulin-potassium infusion,[55] etc.).

Several limitations merit consideration. First, the engineered features we used for building our models are based on a manufacturer-specific software. There are known discrepancies between manufacturers in ECG preprocessing and metrics computation, which means that our models would need retraining and validation when using different software for ECG signal processing. Alternatively, deep neural networks can be used to directly analyze raw ECG signal without explicit feature engineering. However, these techniques require much larger sample size for model derivation (e.g.,

>10,000) and might not yield a meaningful improvement over feature engineering-based machine learning approaches for traditional 12-lead ECG based diagnosis.[56] Second, we found slight differences between the derivation and validation cohorts, specifically in terms of disease prevalence and practicing clinicians' accuracy in ECG interpretation. These cohorts came from two different regions in the U.S., and EMS systems follow state-specific protocols. It is possible that discrepancies in EMS protocols and in-hospital practices resulted in slight differences in the types and proportions of patients that receive prehospital 12-lead ECGs, as well as in their outcome adjudications. Yet, it is reassuring that our models continued to generalize well between the study sites. Third, it is worth noting that our model for screening for "any ACS event" only boosted the performance of the rule-in arm of the derived risk score. This means that a low-risk determination by our model suggests that a given patient would unlikely have OMI, but they might still have a less subtle phenotype of NSTE-ACS that does not require reperfusion therapy. It is likely that serial ECG testing might improve the detection of ~~this group~~missed events where a patients might switch to a higher risk category in the following hours,[34] but this remains to be confirmed. Coronary occlusion is a dynamic process that evolves over time, so an initial low risk class by our models should not lead to a lower level of active monitoring. Finally, although this study used prospective patients, all analyses were completed asynchronously with patient care. Prospective validation where OMI probabilities and decision support is provided in real time is warranted.

In conclusion, we developed and externally validated machine learning models for the ECG diagnosis of OMI in 7,313 patients with chest pain from multiple sites in the United States. The results demonstrated the superiority of machine learning in detecting subtle ischemic ECG changes indicative of OMI in an observer-independent approach. These models outperformed practicing clinicians and commercial ECG interpretation software, significantly boosting both precision and recall. Our derived OMI risk score provided ~~superior~~ enhanced rule-in and rule-out accuracy when compared to HEAR score, and when combined with the clinical judgment of trained emergency personnel, this score helped correctly reclassify one in three patients with chest pain. The ECG features driving our models were

evaluated, providing plausible mechanistic links to myocardial injury. Future work should focus on the prospective validation where OMI probabilities and decision support is provided in real time.

ACKNOWLEDGMENTS

This study was funded by grants from the National Institute of Health (NIH), National Heart, Lung, and Blood Institute (NHLBI), National Center for Advancing Translational Sciences (NCATS), and National Institute for Nursing Research (NINR) through grants # R01HL137761 (SSA), UL1TR001857 (SSA), K23NR017896 (JZH), and KL2TR002490 (JZH).

AUTHOR CONTRIBUTION

SSA, CMG, JZH, SS, ES, and CWC conceived the study, secured funding, and supervised the research. YB and SWS advised on the scientific direction of the study. SSA, JZH, ZF, MOA, KKP, and SH supervised dataset creation and annotation. SSA, CMG, JZH, ZF, MOA, and SH supervised clinical outcomes adjudication. REG, SSA, MA, ZB, PVD, and NR supervised ECG signal processing and feature extraction. SSA, ZB, NR, SMS, and ES performed feature engineering, machine learning modeling, statistical analysis, and results interpretation. SSA drafted the manuscript. All authors critically revised the manuscript for important intellectual content. All authors provided their final approval of the version to be published. All authors are accountable for the work.

DECLARATION OF INTERESTS

US Patent # 10820822, Owner: University of Pittsburgh, Inventors: SSA, ES, and CWC.

DATA AVAILABILITY

The ECG SMART trial makes use of extracted ECG features to train and evaluate a random forest classifier to denote the probability of OMI. The Python codes to evaluate these models along with the derivation and external validation datasets are available through GitHub (https://github.com/zeineb-bouzid/sharing-github-nature-medicine.git). Researchers wishing the source binary files to design alternative models should contact the corresponding author to arrange for proper approvals and institutional data use agreements.

CODE AVAILABILITY

The Python codes to evaluate these models along with the derivation and external validation datasets are available through GitHub (https://github.com/zeineb-bouzid/sharing-github-nature-medicine.git).

ONLINE METHODS

Ethics Statement

The derivation cohort included prehospital data from the City of Pittsburgh Bureau of Emergency Medical Services (EMS) and in-hospital data from three tertiary care hospitals from the University of Pittsburgh Medical Center (UPMC) healthcare system: UPMC Presbyterian Hospital, UPMC Shadyside Hospital, and UPMC Mercy Hospital (Pittsburgh, Pennsylvania, USA). All consecutive eligible patients were recruited under a waiver of informed consent. This observational trial was approved by the institutional review board of the University of Pittsburgh and was registered in www.ClinicalTrials.gov (identifier # NCT04237688). The analyses described in this paper were prespecified by the trial protocol that was funded by the National Institute of Health. The first external validation cohort included data from Orange County EMS (Chapel Hill, North Carolina, USA). This study actively consented eligible patients and was approved by the institutional review board of the University of North Carolina at Chapel Hill. The second external validation cohort included data from Mecklenburg County EMS and Atrium Health (Charlotte, North Carolina, USA). Data were collected through a healthcare registry and all consecutive eligible patients were enrolled under a waiver of informed consent. This study was also approved by the institutional review board of the University of North Carolina at Chapel Hill. These two external cohorts were very comparable and were, therefore, combined into one cohort.

Study Design & Data Collection

This was a prospective, observational cohort study. The methods for each study cohort were described in detail elsewhere.[57,58] All study cohorts enrolled adult patients with an emergency call for non-traumatic chest pain or anginal equivalent symptoms (arm, shoulder, jaw pain, shortness of breath, diaphoresis, syncope). Eligible patients were transported by an ambulance and had at least one recorded prehospital 12-lead ECG. There were no selective exclusion criteria based on sex, race,

comorbidities, or acuity of illness. For this prespecified analysis, we only included non-duplicate ECGs from unique patient encounters, and we removed patients with prehospital ECGs showing ventricular tachycardia or ventricular fibrillation (i.e., these patients are managed by ACLS algorithms). We also removed patients with confirmed prehospital STEMI, which included machine-generated ***ACUTE MI*** warning, EMS-documentation of STEMI, and medical consult for potential CATH lab activation.

Independent reviewers extracted data elements from hospital systems on all patients meeting eligibility criteria. If a prehospital ECG had no patient identifiers, we used a probabilistic matching approach to link each encounter with the correct hospital record. This previously validated data linkage protocol was based on the ECG-stamped birth date, sex, and date/time logs, as well as based on EMS dispatch logs and receiving hospital records. All probabilistic matches were manually reviewed by research specialists for accuracy. The match success rate ranged from 98.6% to 99.8%.

Clinical Outcomes

Adjudications were made by independent reviewers at each local site after reviewing all available medical records within 30 days of the indexed encounter. Reviewers were blinded from all ECG analyses and models' predictions. OMI was defined as coronary angiographic evidence of an acute culprit lesion in at least one of the three main coronary arteries (LAD, LCX, RCA) or their primary branches with TIMI flow grade of 0-1. TIMI flow grade of 2 with significant coronary narrowing > 70% and peak troponin of 5-10.0 ng/mL was also considered indicative of OMI.[17,21] These adjudications were made by two independent reviewers. The Kappa coefficient statistic between the two reviewers was 0.771 (i.e., substantial agreement). All disagreements were resolved by a third reviewer.

ACS was defined per the fourth universal definition of myocardial infarction as the presence of symptoms of ischemia (i.e. diffuse discomfort in the chest, upper extremity, jaw, or epigastric area for more than 20 minutes) and at least one of the following criteria: (1) subsequent development of labile, ischemic ECG changes (e.g., ST changes, T inversion) during hospitalization; (2) elevation of cardiac troponin (i.e., > 99th percentile) during the hospital stay with rise and/or drop on serial testing; (3)

coronary angiography demonstrating greater than 70% stenosis, with or without treatment; and/or (4) functional cardiac evaluation (stress testing) that demonstrates ECG, echocardiographic, or radionuclide evidence of focal cardiac ischemia.[5] Patients with type 2 MI and pre-existing subacute coronary occlusion were labeled as negative for ACS and OMI. This included around 10% of patients with positive troponin but with no rise and/or drop in concentration on serial testing (i.e., chronic leak) or with troponin leak attributed to non-coronary occlusive conditions such as pericarditis. On a randomly selected small subset of patients (n=1,209), the Kappa coefficient statistic for ACS adjudication ranged from 0.846 to 0.916 (i.e., substantial to perfect agreement).

ECG Methods

Prehospital ECGs were obtained in the field by paramedics as part of routine care. ECGs were acquired using either Heart Start MRX (Philips Healthcare) or LIFEPAK-15 (Physio-Control Inc.) monitor-defibrillator devices. All digital 12-lead ECGs were acquired at a sampling rate of 500 s/s (0.05-150 Hz) and transmitted to the respective EMS agency and receiving hospital. Digital ECG files were exported in XML format and stored in a secondary server at each local site. ECG images were de-identified and manually annotated by independent reviewers or research specialists; ECGs with poor quality or missing leads were removed from the study. Next, digital XML files were transmitted to the Philips Advanced Algorithm Research Center for offline analysis (Cambridge, Massachusetts, USA).

ECG featurization was described in detail elsewhere.[18] Briefly, ECG signal preprocessing and feature extraction were performed using a manufacturer-specific software (Philips DXL diagnostic 12/16 lead ECG analysis program). ECG signals were first preprocessed to remove noise, artifacts, and baseline wander. Ectopic beats were removed, and median beats were calculated for each lead. Next, we used the root mean square (RMS) signal to identify global waveform fiducials, including the onset, offset, and peak of the P wave, QRS complex, and T wave. Lead-specific fiducials were then identified to further segment individual waveforms into Q, R, R`, S, S`, and J point.

We then computed a total of 554 ECG features based on (1) the amplitude, duration, area, slope and/or concavity of global and lead-specific waveforms; (2) the QRS and T axes and angles in the frontal, horizontal, spatial, XY, XZ, and YZ planes, including directions at peak, inflection point, and initial / terminal loops; (3) eigenvalues of the principal components of orthogonal ECG leads (I, II, V1-V6), including PCA ratios for individual ECG waveform segments; and (4) T loop morphology descriptors. Features with zero distribution were removed to prevent representation bias.

Next, we previously identified an optimal parsimonious list of the most important ECG features that are mechanistically linked to cardiac ischemia as described in detail elsewhere.[18] Briefly, to prevent omitted-feature bias, we used a hybrid approach that combines domain knowledge with a data-driven strategy. First, clinical scientists identified 24 classical features that are known to correlate with cardiac ischemia (i.e., lead-specific ST-80 and T wave amplitudes). Next, starting with a comprehensive list of 554 candidate features, we used data-driven algorithms (e.g., recursive feature elimination and LASSO) to identify 198 supplemental features potentially related to ischemia. LASSO selects features with non-zero coefficients after L1 norm regularization, and recursive feature elimination uses repeated regression iterations to identify the features that have significant impact on model predictions. We then examined the feature pairs in this expanded list of 222 features and removed features with very high collinearity scores that contains redundant information (e.g., we kept QTc if both QT and QTc were selected by the model). Finally, we used feature importance ranking to identify the most parsimonious subset of features that are complementary and can boost the classification performance. This hybrid approach eventually yielded a subset of 73 features that can serve as plausible markers of ischemia~~Clinical scientists initially reviewed a list of 554 features and marked the ones that are known to correlate with cardiac ischemia. This list was then expanded by supplemental features identified by data-driven algorithms (e.g., recursive feature elimination and LASSO). The clinical scientists then reviewed the expanded list to examine feature pairs with high collinearity and retained the subset of~~

~~features that are complementary and can serve as plausible markers of ischemia. This approach eventually yielded a subset of 73 features that was shown to boost classification performance~~.[18]

Machine Learning Methods

We followed best practices recommended by "ROBUST-ML" and "ECG-AI stress test" checklists to design and benchmark our machine learning algorithms.[51,59] To prevent measurement bias, ECG features were manually reviewed to identify erroneous calculations. Physiologically plausible outliers were replaced with ±3 SD. On average, each feature had a 0.34% missingness rate (range 0.1% to 1.6%). Thus, we imputed missing values with the mean, median, or mode of that feature after consultation with clinical experts. ECG metrics were then z-score normalized and used as input features in machine learning models. The derivation and validation datasets were cleaned independently to prevent data leakage. Both cohorts were recruited over the same time window, suggesting the lack of temporal bias. To prevent potential mismatch with intended use, input features for model development included only ECG data plus the machine-stamped age. No other clinical data were used for model building.

We randomly split the derivation cohort into an 80% training set and a 20% internal testing set. On the training set, we fit 10 machine learning classifiers: regularized logistic regression, linear discriminant analysis, support vector machine, Gaussian Naïve Bayes, random forest, gradient boosting machine, extreme gradient boosting, stochastic gradient descent logistic regression, k-nearest neighbors, and artificial neural networks. Each classifier was optimized over 10-fold cross validation to finetune hyperparameters. After selecting optimal hyperparameters, models were re-trained on the entire training subset to derive final weights and create a lockout model to evaluate on the holdout test set. We calibrated our classifiers to produce a probabilistic output which can be interpreted as a confidence level (probability risk score). Trained models were compared using the AUROC curve with Wilcoxon signed-rank test for pairwise comparisons. ROC-optimized cutoffs were chosen using Youden index, and classifications on confusion matrix were compared using McNemar's test.

The random forest classifier (RF) achieved high accuracy on the training set (low bias) with a relatively small drop in performance on the test set (low variance), indicating an acceptable bias-variance tradeoff and low risk of overfitting (**Extended Data Fig. ~~7~~8**). Although the support vector machine (SVM) model had lower variance on the test set, when compared with the RF model, there were no significant differences in AUROC (Delong's test) or their binary classifications (McNemar's test). Moreover, there were no differences between the RF and SVM models in terms of Kolmogorov-Smirnov goodness-of-fit (0.716 vs. 0.715) or the Gini purity index (0.82 vs. 0.85). Due to its scalability and intuitive architecture, we chose the probability output of the RF model to build our derived OMI score. We generated density plots of these probability scores for positive and negative classes and selected classification thresholds for low, intermediate, and high-risk groups based on prespecified NPV > 0.99 and TPR > 0.50. Finally, we used the lock-out random forest classifier to generate probability scores and risk classes on the completely unseen external validation cohort. The code to generate probability scores is included with the supplemental materials of this manuscript.

Reference Standard

To reduce the risk of evaluation bias, we benchmarked our machine learning models against multiple reference standards used during routine care in clinical practice. First, we used a commercial, FDA-approved, ECG interpretation software (Philips DXL diagnostic algorithm) to denote the likelihood of ischemic myocardial injury. This likelihood (yes/no) was based on a composite of the followings: (1) diagnostic codes for ">>>Acute MI<<<", including descriptive statements that denote "acute", "recent", "age indeterminate", "possible" or "probable"; and (2) diagnostic codes for ">>>Acute Ischemia<<<", including descriptive statements that denote "possible", "probable", or "consider". Diagnostic statements that denoted "old" [infarct], "nonspecific" [ST depression], or "secondary to" [LVH or high heart rate] were excluded from this composite reference standard.

We also used practicing clinicians' overread of ECGs to denote the likelihood of ischemic myocardial injury on a given ECG (yes/no) when a STEMI pattern does not exist, which is congruent

with how ED physicians evaluate these patients in clinical practice. Independent physician reviewers annotated each 12-lead ECG image as per the fourth universal definition of MI criteria,[5] including two contiguous leads with ST-elevation (≥ 0.2 mV for V2-V3 in men ≥ 40 years and ≥ 2.5 mm in men < 40 years; ≥ 0.15 mV for V2-V3 in women; or ≥ 0.1 mV in other leads) or ST-depression (new horizontal or down-sloping depression ≥ 0.05 mV); with or without T wave inversion (> 0.1 mV in leads with prominent R wave or R/S ratio > 1). Reviewers were also prompted to use their clinical judgment to identify highly suspicious ischemic changes (e.g., reciprocal changes, hyperacute T waves), as well as to account for potential confounders (e.g., bundle branch blocks, early repolarization). On a randomly selected subset of patients in the derivation cohort (n=1,646), the Kappa coefficient statistic between two emergency physicians who interpreted the ECGs was 0.568 (i.e., moderate agreement). A third reviewer was used to adjudicate discrepancies on this randomly selected subset. Similarly, on a randomly selected subset of patients in the external validation cohort (n=375), the Kappa coefficient statistic between the two board-certified cardiologists who interpreted the ECGs was 0.690 (i.e., substantial agreement).

Finally, given that clinicians largely depend on risk scores to triage patients in the absence of STEMI, which would significantly affect how OMI patients are diagnosed and treated in clinical practice, we compared our derived OMI risk score against the HEART ~~risk~~ score. This score is commonly used in US hospitals and it has been well-validated for triaging patients in the emergency department.[60] The HEART score is based on the patient's **H**istory at presentation, **E**CG interpretation, **A**ge, **R**isk factors, and initial **T**roponin values (range 0-10). This score places patients in low (0-3), intermediate (4-6), and high-risk (7-10) groups. Given that troponin results are not usually available at first medical contact, we used a modified HEAR score after dropping the **T**roponin values, which has also been previously validated for use by paramedics prior to hospital arrival.[36] The comparison against the HEART score herein focused on establishing the incremental gain of using the derived OMI score over routine care at initial triage. We compared how the new risk classes assigned by our derived OMI score

agree with or differ from the risk classes assigned by the HEART score, which could inform potential incremental gain over routine care.

Statistical Analysis

Descriptive statistics were reported as mean ± standard deviation or n (%). Missing data was assessed for randomness and was handled during ECG feature selection (see Machine Learning Methods section above). Normality of distribution was assessed prior to hypothesis testing where deemed necessary. ECG features were z-score normalized as part of standard input architectures for machine learning models. Comparisons between cohorts were performed using chi-square (for discrete variables) and independent samples t-test or Mann-Whitney U test (for continuous variables). The level of significance was set at alpha 0.05 for two-tailed hypothesis testing where applicable.

All diagnostic accuracy values were reported as per STARD recommendations (Reporting Guidelines for Diagnostic Accuracy Studies). We reported classification performance using AUROC curve, sensitivity (recall), specificity, PPV (precision), and NPV, along with 95% confidence interval (CI) where applicable. For 10-fold cross validation, we compared the multiple classifiers using the Wilcoxon signed-rank test (for AUROC curves) and McNemar's test (for confusion matrices). We derived low-, intermediate-, and high-risk categories for the final classifier using Kernel density plot estimates between classes. The adequacy of these risk classes was evaluated using Log-rank chi-square of accumulative risk for clinically important outcomes over the length of stay during the indexed admission.

For assessing the incremental gain in classification performance, we compared the AUROC of the final model against reference standards using DeLong's test. For ease of comparison, the confidence bounds for AUROC of the reference standards (commercial system and practicing clinicians) were generated using 1000 bootstrap samples. To place the incremental gain value in a broader context of the clinical workflow, ~~We~~ we ~~then~~ also computed the Net Reclassification Improvement (NRI) index of our model against the HEAR score during the initial assessment at first

medical contact. Risk scores are an integral part of clinical workflow in patients with suspected ACS who do not meet STEMI criteria. As per STARD recommendations (Reporting Guidelines for Diagnostic Accuracy Studies), the NRI Index evaluates the net gain between up-triage and down-triage when correctly reclassifying risk class assignments of an “old” test (HEART score) using a “new” test (the derived OMI score).

We used logistic regression to identify the independent predictive value of OMI risk classes. We used variables significant in univariate analysis and then built multivariate models with stepwise backward selection method using Wald chi-square criteria. We reported odds ratios with 95% CI for all significant predictors. All analyses were completed using Python v3.8.5 and SPSS v24.

REFERENCES

1. Al-Zaiti S, Macleod MR, Van Dam PM, Smith SW, Birnbaum Y. Emerging ECG Methods for Acute Coronary Syndrome Detection: Recommendations & Future Opportunities. *Journal of Electrocardiology*. 2022;74:65-72.

2. Birnbaum Y, Wilson JM, Fiol M, de Luna AB, Eskola M, Nikus K. ECG diagnosis and classification of acute coronary syndromes. *Annals of Noninvasive Electrocardiology*. 2014;19(1):4-14.

3. Goodacre S, Pett P, Arnold J, et al. Clinical diagnosis of acute coronary syndrome in patients with chest pain and a normal or non-diagnostic electrocardiogram. *Emergency medicine journal.* 2009;26(12):866-870.

4. Ioannidis JP, Salem D, Chew PW, Lau J. Accuracy and clinical effect of out-of-hospital electrocardiography in the diagnosis of acute cardiac ischemia: a meta-analysis. *Annals of emergency medicine.* 2001;37(5):461-470.

5. Thygesen K, Alpert JS, Jaffe AS, et al. Fourth universal definition of myocardial infarction (2018). *European Heart Journal.* 2018:ehy462-ehy462.

6. Gulati M, Levy PD, Mukherjee D, et al. 2021 AHA/ACC/ASE/CHEST/SAEM/SCCT/SCMR Guideline for the Evaluation and Diagnosis of Chest Pain. *Journal of the American College of Cardiology.* 2021;78(22):e187-e285.

7. Levine GN, Bates ER, Blankenship JC, et al. 2015 ACC/AHA/SCAI focused update on primary percutaneous coronary intervention for patients with ST-elevation myocardial infarction: an update of the 2011 ACCF/AHA/SCAI guideline for percutaneous coronary intervention and the 2013 ACCF/AHA guideline for the management of ST-elevation myocardial infarction. *Journal of the American College of Cardiology.* 2016;67(10):1235-1250.

Formatted: Font: (Default) Arial

8. Amsterdam EA, Wenger NK, Brindis RG, et al. 2014 AHA/ACC Guideline for the Management of Patients With Non–ST-Elevation Acute Coronary Syndromes: Executive Summary. *Circulation.* 2014;130(25):2354-2394.

9. Dixon WC, Wang TY, Dai D, et al. Anatomic distribution of the culprit lesion in patients with non–ST-segment elevation myocardial infarction undergoing percutaneous coronary intervention: findings from the National Cardiovascular Data Registry. *Journal of the American College of Cardiology.* 2008;52(16):1347-1348.

10. Wang TY, McCoy LA, Bhatt DL, et al. Multivessel vs culprit-only percutaneous coronary intervention among patients 65 years or older with acute myocardial infarction. *American heart journal.* 2016;172:9-18.

11. Karwowski J, Gierlotka M, Gąsior M, et al. Relationship between infarct artery location, acute total coronary occlusion, and mortality in STEMI and NSTEMI patients. *Polish Archives of Internal Medicine.* 2017;127(6):401-411.

12. Figueras J, Otaegui I, Marti G, et al. Area at risk and collateral circulation in a first acute myocardial infarction with occluded culprit artery. STEMI vs non-STEMI patients. *International Journal of Cardiology.* 2018;259:14-19.

13. Tanaka T, Miki K, Akahori H, et al. Comparison of coronary atherosclerotic disease burden between ST-elevation myocardial infarction and non-ST-elevation myocardial infarction: Non-culprit Gensini score and non-culprit SYNTAX score. *Clinical Cardiology.* 2021;44(2):238-243.

14. Aslanger EK, Meyers HP, Bracey A, Smith SW. The STEMI/NonSTEMI Dichotomy needs to be replaced by Occlusion MI vs. Non-Occlusion MI. *International Journal of Cardiology.* 2021;330:15.

Formatted: Font: (Default) Cambria Math
Formatted: Font: (Default) Arial
Formatted: Font: (Default) Cambria Math
Formatted: Font: (Default) Arial
Formatted: Font: (Default) Cambria Math
Formatted: Font: (Default) Arial
Formatted: Font: (Default) Cambria Math
Formatted: Font: (Default) Arial
Formatted: Font: (Default) Cambria Math
Formatted: Font: (Default) Arial

15. Avdikos G, Michas G, Smith SW. From Q/Non-Q Myocardial Infarction to STEMI/NSTEMI: Why It's Time to Consider Another Simplified Dichotomy; a Narrative Literature Review. *Archives of Academic Emergency Medicine.* 2022;10(1):e78-e78.

16. Aslanger EK, Meyers PH, Smith SW. STEMI: A transitional fossil in MI classification? *Journal of Electrocardiology.* 2021;65:163-169.

17. Meyers HP, Bracey A, Lee D, et al. Comparison of the ST-elevation myocardial infarction (STEMI) vs. NSTEMI and occlusion MI (OMI) vs. NOMI paradigms of acute MI. *The Journal of emergency medicine.* 2021;60(3):273-284.

18. Bouzid Z, Faramand Z, Gregg RE, et al. In search of an optimal subset of ECG features to augment the diagnosis of acute coronary syndrome at the emergency department. *Journal of the American Heart Association.* 2021;10(3):e017871.

19. Meyers HP, Bracey A, Lee D, et al. Ischemic ST-Segment Depression Maximal in V1–V4 (Versus V5–V6) of Any Amplitude Is Specific for Occlusion Myocardial Infarction (Versus Nonocclusive Ischemia). *Journal of the American Heart Association.* 2021;10(23):e022866.

20. Birnbaum Y, de Luna AB, Fiol M, et al. Common pitfalls in the interpretation of electrocardiograms from patients with acute coronary syndromes with narrow QRS: a consensus report. *Journal of Electrocardiology.* 2012;45(5):463-475.

21. Meyers HP, Bracey A, Lee D, et al. Accuracy of OMI ECG findings versus STEMI criteria for diagnosis of acute coronary occlusion myocardial infarction. *IJC Heart & Vasculature.* 2021;33:100767.

22. Al-Zaiti S, Callaway CW, Kozik TM, Carey M, Pelter M. Clinical Utility of Ventricular Repolarization Dispersion for Real-Time Detection of Non-ST Elevation Myocardial Infarction in Emergency Departments. *Journal of the American Heart Association.* 2015;4(7):e002057.

Formatted: Font: (Default) Cambria Math

Formatted: Font: (Default) Arial

23. Al-Zaiti S, Alrawashdeh M, Martin-Gill C, Callaway C, Mortara D, Nemec J. Evaluation of Beat-to-Beat Ventricular Repolarization Lability from Standard 12-Lead ECG During Acute Myocardial Ischemia. *Journal of Electrocardiology.* 2017;50(6):717-724.

24. Al-Zaiti S, Sejdic E, Nemec J, Callaway C, Soman P, Lux RL. Spatial Indices of Repolarization Correlate with Non-ST Elevation Myocardial Ischemia in Patients with Chest Pain. *Medical & Biological Engineering & Computing* 2018;56(1):1-12.

25. Sharma A, Miranda DF, Rodin H, Bart BA, Smith SW, Shroff GR. Interobserver variability among experienced electrocardiogram readers to diagnose acute thrombotic coronary occlusion in patients with out of hospital cardiac arrest: Impact of metabolic milieu and angiographic culprit. *Resuscitation.* 2022;172:24-31.

26. Gregg RE, Yang T, Smith SW, Babaeizadeh S. ECG reading differences demonstrated on two databases. *Journal of Electrocardiology.* 2021;69:75-78.

27. Cook DA, Oh S-Y, Pusic MV. Accuracy of physicians' electrocardiogram interpretations: a systematic review and meta-analysis. *JAMA internal medicine.* 2020;180(11):1461-1471.

28. McRae AD, Innes G, Graham M, et al. Undetectable concentrations of an FDA-approved high-sensitivity cardiac Troponin T assay to rule out acute myocardial infarction at emergency department arrival. *Academic Emergency Medicine.* 2017;24:DOI: 10.1111/acem.13229.

29. Body R, Mahler S. Welcome to the real world: Do the conditions of FDA approval devalue high sensitivity troponin? *Academic Emergency Medicine.* 2017;24:DOI: 10.1111/acem.13256.

30. Wereski R, Chapman AR, Lee KK, et al. High-sensitivity cardiac troponin concentrations at presentation in patients with ST-segment elevation myocardial infarction. *JAMA cardiology.* 2020;5(11):1302-1304.

Formatted: Font: (Default) Cambria Math

Formatted: Font: (Default) Arial

Formatted: Font: (Default) Cambria Math

Formatted: Font: (Default) Arial

31. Cotterill PG, Deb P, Shrank WH, Pines JM. Variation in chest pain emergency department admission rates and acute myocardial infarction and death within 30 days in the Medicare population. *Academic Emergency Medicine.* 2015;22(8):955-964.

32. Kang MG, Kang Y, Kim K, et al. Cardiac mortality benefit of direct admission to percutaneous coronary intervention-capable hospital in acute myocardial infarction: Community registry-based study. *Medicine (Baltimore).* 2021;100(10):e25058-e25058.

33. Quinn T, Johnsen S, Gale CP, et al. Effects of prehospital 12-lead ECG on processes of care and mortality in acute coronary syndrome: a linked cohort study from the Myocardial Ischaemia National Audit Project. *Heart.* 2014;100(12):944-950.

34. Bouzid Z, Faramand Z, Martin-Gill C, et al. Incorporation of Serial 12-Lead Electrocardiogram With Machine Learning to Augment the Out-of-Hospital Diagnosis of Non-ST Elevation Acute Coronary Syndrome. *Annals of Emergency Medicine.* 2023;81(1):57-69.

35. Al-Zaiti S, Besomi L, Bouzid Z, et al. Machine learning-based prediction of acute coronary syndrome using only the pre-hospital 12-lead electrocardiogram. *Nature communications.* 2020;11(3966):https://doi.org/10.1038/s41467-41020-17804-41462.

36. Stopyra JP, Harper WS, Higgins TJ, et al. Prehospital modified HEART score predictive of 30-day adverse cardiac events. *Prehospital and disaster medicine.* 2018;33(1):58-62.

37. Ashburn NP, Snavely AC, O'Neill JC, et al. Performance of the European Society of Cardiology 0/1-Hour Algorithm With High-Sensitivity Cardiac Troponin T Among Patients With Known Coronary Artery Disease. *JAMA Cardiology.* 2023.

38. Sabatine MS, Morrow DA, McCabe CH, Antman EM, Gibson CM, Cannon CP. Combination of quantitative ST deviation and troponin elevation provides independent prognostic and therapeutic information in unstable angina and non–ST-elevation myocardial infarction. *American heart journal.* 2006;151(1):25-31.

Formatted: Font: (Default) Arial

Formatted: Font: (Default) Arial

39. Lux RL. Non-ST-Segment Elevation Myocardial Infarction: A Novel and Robust Approach for Early Detection of Patients at Risk. *Journal of the American Heart Association.* 2015;4(7):e002279.

40. Marrusa S, Zhangc M, Arthurb M. Identification of Acute Coronary Syndrome via Activation and Recovery Times in Body-Surface Mapping and Inverse Electrocardiography. *International Journal of Bioelectromagnetism.* 2019;21(1-6).

41. Lux RL. Basis and ECG measurement of global ventricular repolarization. *Journal of Electrocardiology.* 2017;50(6):792-797.

42. Al-Zaiti S, Runco K, Carey M. Increased T-Wave Complexity Can Indicate Subclinical Myocardial Ischemia in Asymptomatic Adults. *Journal of Electrocardiology.* 2011;44(6):684-688.

43. Forberg JL, Green M, Björk J, et al. In search of the best method to predict acute coronary syndrome using only the electrocardiogram from the emergency department. *Journal of Electrocardiology.* 2009;42:58-63.

44. Green M, Bjork J, Forberg J, Ekelund U, Edenbrandt L, Ohlsson M. Comparison between neural networks and multiple logistic regression to predict acute coronary syndrome in the emergency room. *Artificial Intelligence in Medicine.* 2006;38:305-318.

45. Hong S, Zhou Y, Shang J, Xiao C, Sun J. Opportunities and challenges of deep learning methods for electrocardiogram data: A systematic review. *Computers in Biology and Medicine.* 2020;122:103801.

46. Baxt WG, Skora J. Prospective validation of artificial neural network trained to identify acute myocardial infarction. *The Lancet.* 1996;347(8993):12-15.

47. Tsien CL, Fraser HS, Long WJ, Kennedy RL. Using classification tree and logistic regression methods to diagnose myocardial infarction. *Studies in health technology and informatics.* 1998;52 Pt 1:493-497.

Formatted: Font: (Default) Cambria Math

Formatted: Font: (Default) Arial

Formatted: Font: (Default) Cambria Math

Formatted: Font: (Default) Arial

48. Berikol GB, Yildiz O, Özcan İT. Diagnosis of Acute Coronary Syndrome with a Support Vector Machine. *Journal of Medical Systems.* 2016;40(84).

49. Wu C-C, Hsu W-D, Islam M, et al. An artificial intelligence approach to early predict non-ST-elevation myocardial infarction patients with chest pain. *Computer Methods and Programs in Biomedicine.* 2019;173:109-117.

50. Brisk R, Bond R, Finlay D, et al. Neural networks for ischaemia detection: Revolution or red herring? A systematic review and meta-analysis. *Journal of Electrocardiology.* 2021;69:79.

51. Bond R, Finlay D, Al-Zaiti SS, Macfarlane P. Machine learning with electrocardiograms: A call for guidelines and best practices for 'stress testing'algorithms. *Journal of Electrocardiology.* 2021;69S:1-6.

52. Elul Y, Rosenberg AA, Schuster A, Bronstein AM, Yaniv Y. Meeting the unmet needs of clinicians from AI systems showcased for cardiology with deep-learning–based ECG analysis. *Proceedings of the National Academy of Sciences.* 2021;118(24):e2020620118.

53. Cohen MV, Downey JM. What are optimal P2Y12 inhibitor and schedule of administration in patients with acute coronary syndrome? *Journal of Cardiovascular Pharmacology and Therapeutics.* 2020;25(2):121-130.

54. Tziakas D, Chalikias G, Al-Lamee R, Kaski JC. Total coronary occlusion in non ST elevation myocardial infarction: Time to change our practice? *International Journal of Cardiology.* 2021;329:1-8.

55. Udelson JE, Selker HP, Braunwald E. Glucose–Insulin–Potassium Therapy for Acute Myocardial Infarction: 50 Years On and Time for a Relook. *Circulation.* 2022;146(7):503-505.

56. Zvuloni E, Read J, Ribeiro AH, Ribeiro ALP, Behar JA. On Merging Feature Engineering and Deep Learning for Diagnosis, Risk-Prediction and Age Estimation Based on the 12-Lead ECG. *arXiv preprint arXiv:220706096.* 2022.

57. Al-Zaiti SS, Martin-Gill C, Sejdic E, Alrawashdeh M, Callaway C. Rationale, development, and implementation of the Electrocardiographic Methods for the Prehospital Identification of Non-ST Elevation Myocardial Infarction Events (EMPIRE). *J Electrocardiol.* 2015;48(6):921-926.

58. Zègre-Hemsey JK, Hogg M, Crandell J, et al. Prehospital ECG with ST-depression and T-wave inversion are associated with new onset heart failure in individuals transported by ambulance for suspected acute coronary syndrome. *Journal of Electrocardiology.* 2021.

59. Al-Zaiti SS, Alghwiri AA, Hu X, et al. A clinician's guide to understanding and critically appraising machine learning studies: a checklist for Ruling Out Bias Using Standard Tools in Machine Learning (ROBUST-ML). *European Heart Journal-Digital Health.* 2022;3(2):125-140.

60. Al-Zaiti SS, Faramand Z, Alrawashdeh MO, Sereika SM, Martin-Gill C, Callaway C. Comparison of clinical risk scores for triaging high-risk chest pain patients at the emergency department. *The American journal of emergency medicine.* 2019;37(3):461-467.

Table 1. Baseline demographic and clinical characteristics

	DERIVATION & TESTING COHORT (N=4,026)	EXTERNAL VALIDATION COHORT (N=3,287)
AGE (YEARS)	59±16 (18−102)	60±15 (21−100)
SEX		
Male	2,122 (53%)	1,814 (55%)
Female	1,904 (47%)	1,473 (45%)
RACE		
White	1,698 (42%)	1,326 (40%)
Black	1,328 (33%)	1,544 (47%)
Others	52 (1.3%)	40 (1%)
Unknown	948 (24%)	377 (12%)
ETHNICITY		
Not Hispanic	3,043 (76%)	2,850 (87%)
Hispanic / Latino	19 (1%)	116 (3.5%)
Unknown	964 (23%)	321 (9.5%)
PAST MEDICAL HISTORY		
Hypertension	2,767 (69%)	2,090 (64%)
Diabetes	1,146 (29%)	1,067 (33%)
High cholesterol	1,520 (38%)	1,376 (42%)
Current smoker	1,244 (31%)	802 (25%)
Known CAD	1,388 (35%)	964 (30%)
Prior MI	930 (23%)	929 (29%)
Prior PCI	963 (24%)	134 (4%)
Prior CABG	357 (10%)	470 (14%)
ECG & LAB FINDINGS		
Sinus rhythm	3,496 (87%)	2,614 (80%)
Atrial fibrillation	354 (9%)	352 (11%)
Left BBB	94 (2.3%)	114 (3.5%)
Right BBB	237 (5.9%)	215 (6.6%)
ECG-LVH	383 (9.5%)	467 (14.2%)
cTnI positive (initial)	330 (8.2%)	736 (22.4%)
cTnI positive (serial testing)	729 (18.1%)	1,177 (35.8%)
MEDICAL THERAPY		
PCI (ANY STENT)	300 (7.5%)	245 (7.5%)
Emergent PCI (<90 MIN)	144 (3.6%)	157 (4.8%)
Total LAD occlusion	91 (2.3%)	94 (2.9%)
Total LCX occlusion	63 (1.6%)	88 (2.7%)
Total RCA occlusion	101 (2.5%)	102 (3.1%)
CABG	34 (0.8%)	30 (0.9%)
STUDY OUTCOMES		
CONFIRMED ACS	550 (13.7%)	537 (16.3%)
OMI	210 (5.2%)	209 (6.4%)
Other Acute MI (NOMI)	240 (6.0%)	220 (6.7%)
Unstable Angina	100 (2.5%)	108 (3.3%)
30-DAY CV DEATH	137 (3.4%)	111 (3.4%)

Values are mean ± SD (min-max) or n (%); CAD: coronary artery disease; MI: myocardial infarction; BBB: bundle branch block; LVH: left ventricular hypertrophy; PCI: percutaneous coronary intervention; LAD: left anterior descending artery; LCX: left circumflex artery; RCA: right coronary artery; CABG: coronary artery bypass graft; OMI: occlusion MI; NOMI: non-occlusion MI; CV: cardiovascular.

FIGURE LEGENDS

Fig. 1. Cohort and sample selection

This flow diagram shows patient inclusion and exclusion in each cohort, as well as the dataset partition for training, internal testing, and external validation. Exclusions are not mutually exclusive.

Fig. 2. Algorithm derivation and testing

This figure shows (A) the classification performance of the machine learning model against other reference standards for detecting occlusion myocardial infarction (OMI), (B) the probability density plots of OMI(+) and OMI(-) classes as denoted by the machine learning model, along with optimal cutoffs of low-risk, intermediate, and high-risk, and (C) distribution of patients in low-risk (+), intermediate risk (++) and high-risk (+++) as per the machine learning model and HEART score.

Fig. 3. Model explainability for OMI detection

This figure shows (A) SHAP values for the 25 most important features driving the predictions of the machine learning classifier in the derivation cohort, and (B) the aggregate median beats of ECGs with occlusion myocardial infarction (OMI) class (red) and the aggregate median beats of ECGs with normal sinus rhythm and no OMI (blue).

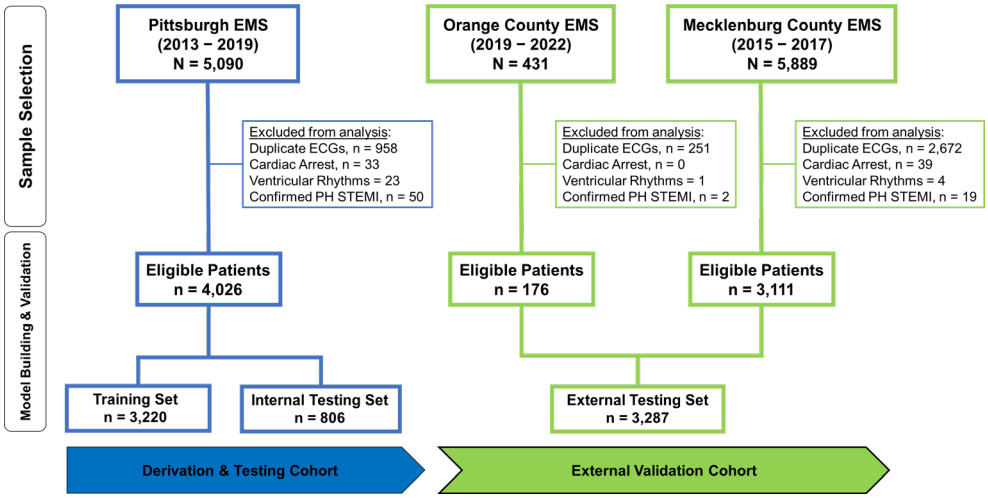
Fig. 4. External validation of ECG-SMART algorithm

This figure shows (A) the classification performance of the machine learning model against other reference standards for detecting occlusion myocardial infarction (OMI), (B) the independent clinical predictors of OMI on multivariate logistic regression testing, and (C) the overall sensitivity and specificity (95% confidence interval [CI]) of the derived OMI score, along with breakdown across subgroups based on age, sex, comorbidities, and baseline ECG findings. The size of markers denotes the sample size of the respective subgroup.

Fig. 5. Net reclassification improvement of OMI risk score when integrated in the clinical workflow at first medical contact ~~Decision analysis for the incremental gain of OMI risk score in reclassifying patients~~

779 This figure simulates the incremental gain of the derived risk score in reclassifying the initial triage
780 decisions by emergency personnel at first medical contact.

781 **Fig. 1. Cohort and sample selection**



782

783

784

Fig. 2. Algorithm derivation and testing for OMI detection

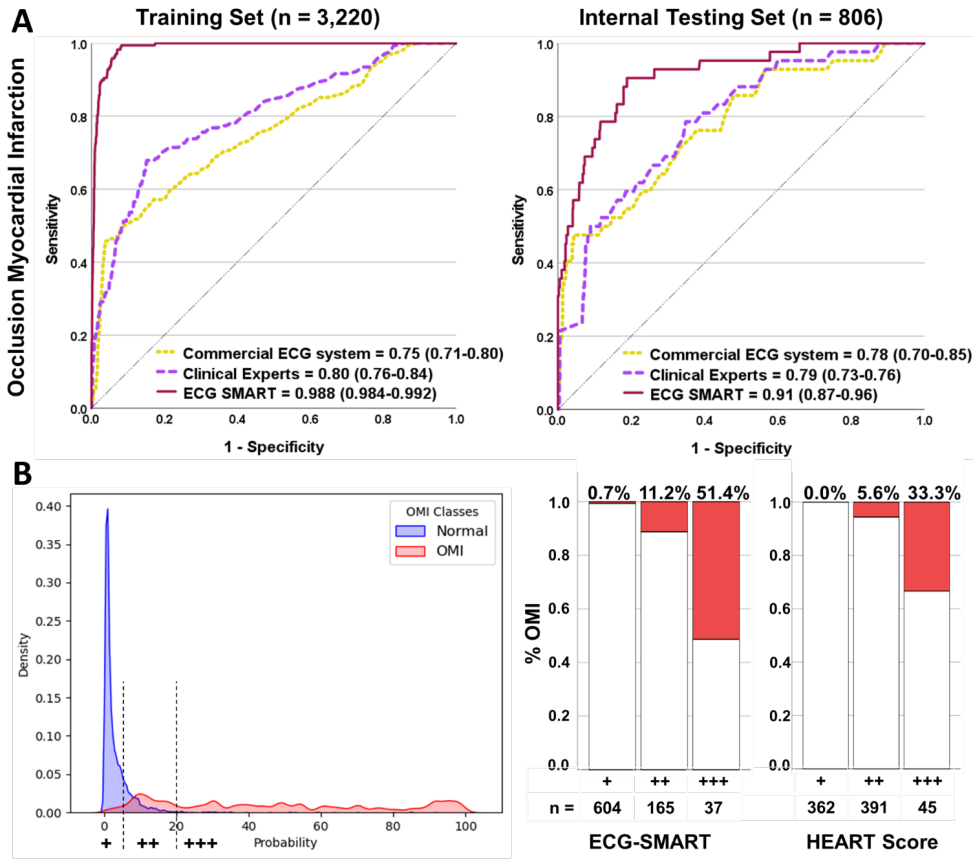
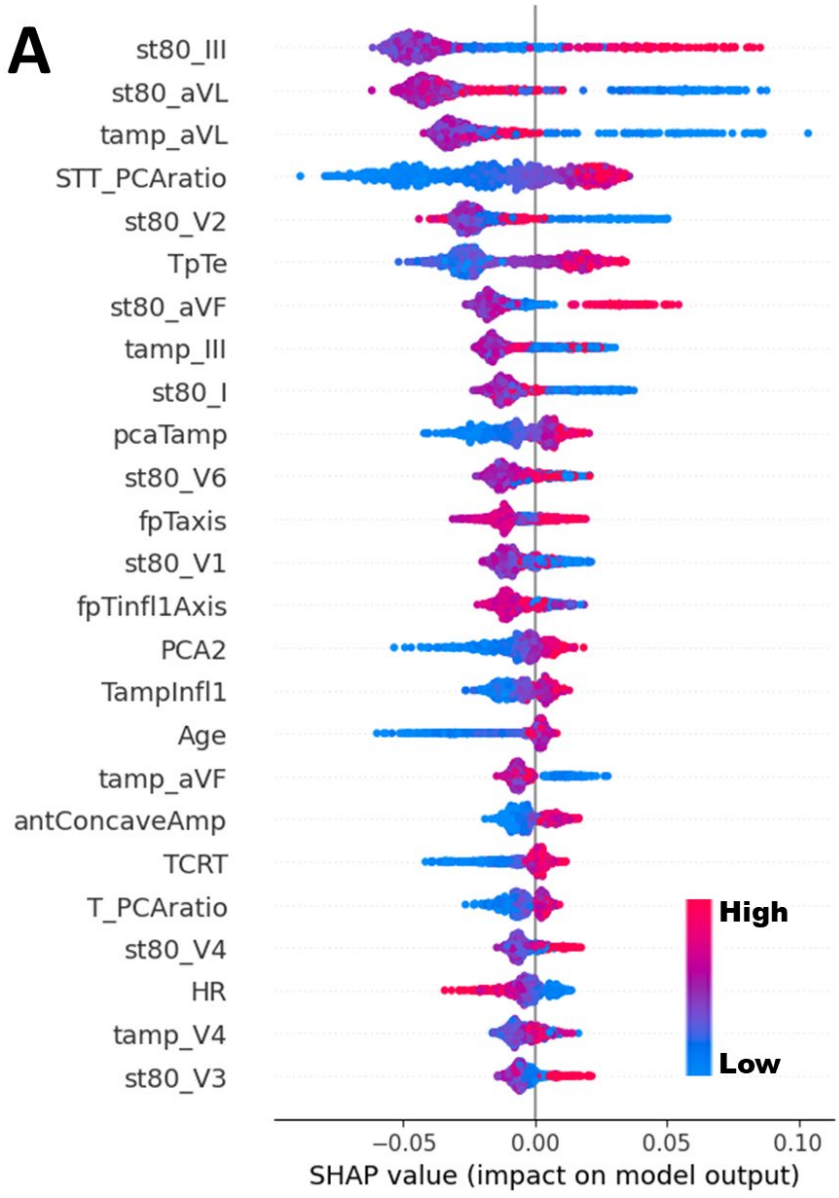


Fig. 3. Model explainability for OMI detection



790 **(Fig. 3B)**

791

Fig. 4. External validation of ECG-SMART algorithm for OMI detection

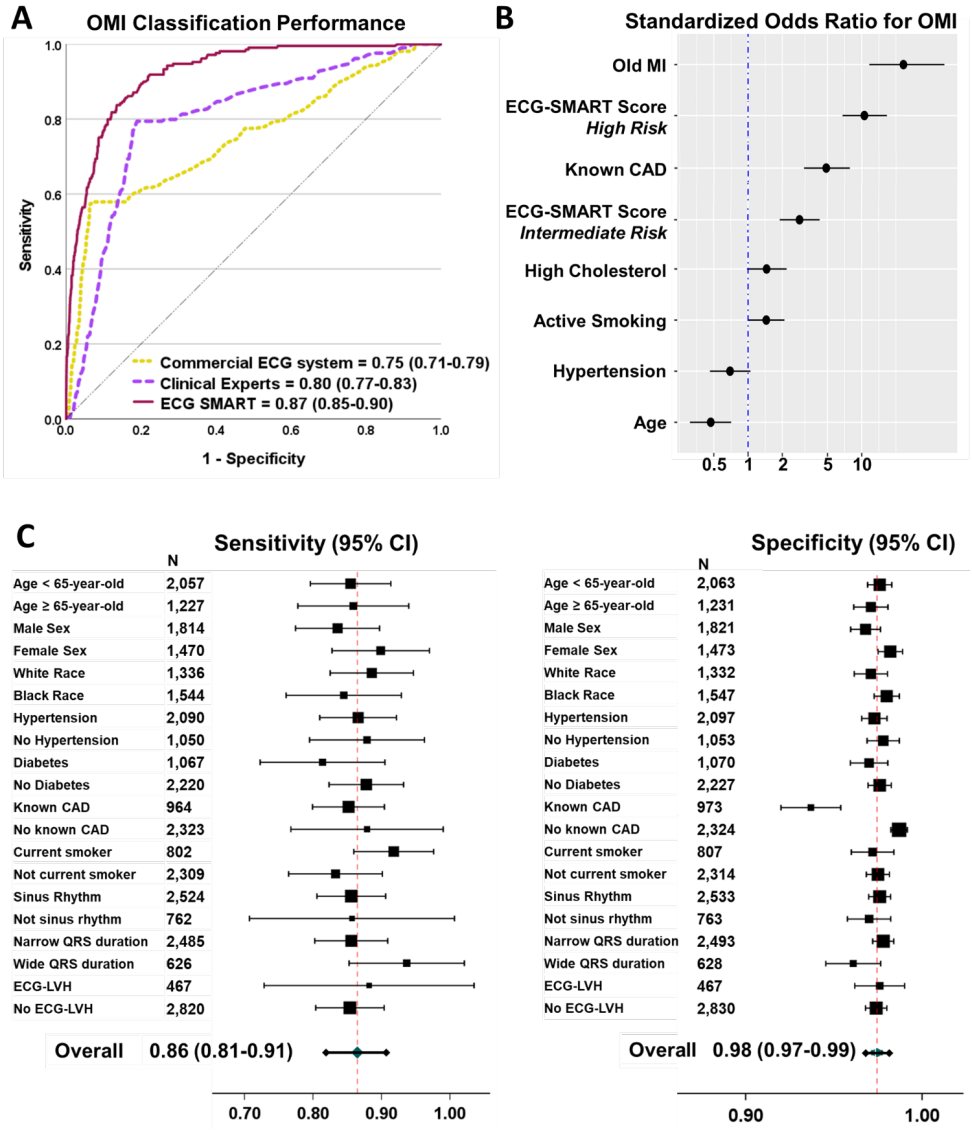
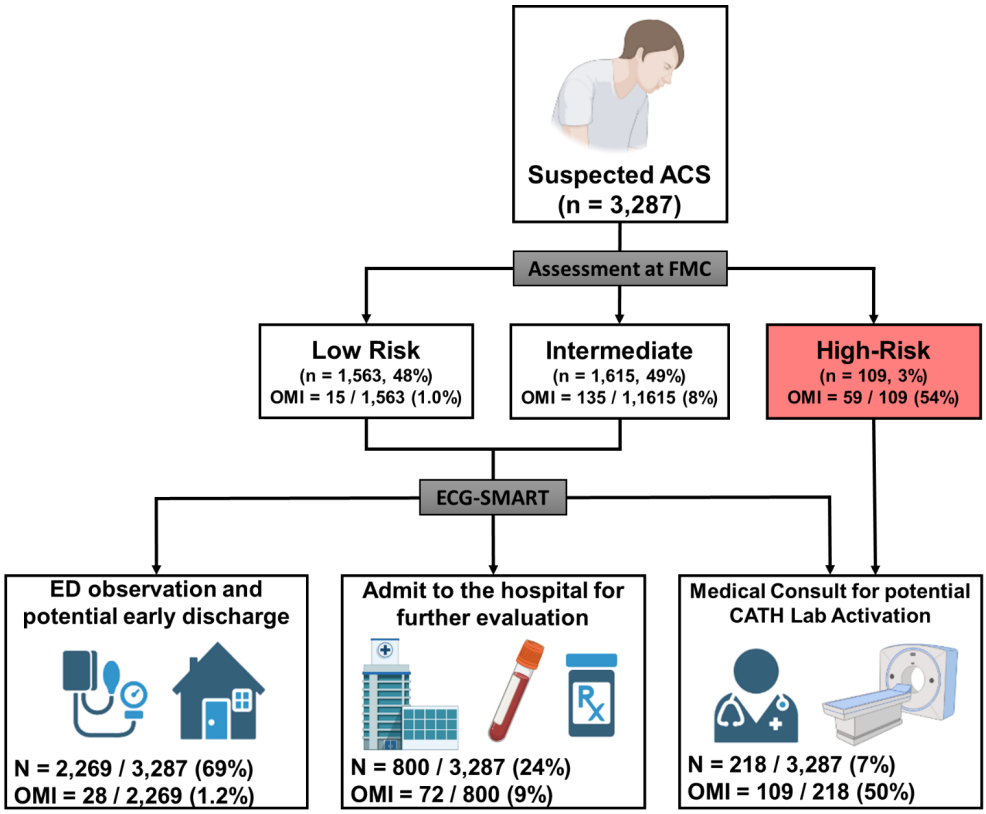


Fig. 5. ~~Decision analysis for the incremental gain~~ Net reclassification improvement of OMI risk score when integrated in the clinical workflow ~~re-classifying patients~~ at first medical contact



DATA SUPPLEMENT

Legend for Extended Data Figures

Extended Data Fig. 1. The relationship between the magnitude of vessel occlusion and the classification of acute coronary events

This figure shows the spectrum of coronary artery disease (CAD) as a function of severity and extent of atherosclerosis plaque progression, ranging from patent coronary artery (far left) to total coronary occlusion (far right). Among patients who develop symptomatic CAD, including those evaluated for chest pain or angina-like symptoms, a subset is diagnosed with acute coronary syndrome (ACS). This group is subclassified as either acute myocardial infarction (MI) or unstable angina (UA). Those with acute MI can be further subclassified, based on the presence of ST-elevation on the ECG, as either ST-elevation myocardial infarction (STEMI) or without ST-elevation (NSTEMI). The STEMI and NSTEMI patients overlap in terms of the presence or absence of total occlusion (depicted as triangles across the continuum in the figure). Alternatively, the same group with acute MI can be subclassified, based on angiographic TIMI flow criteria, as either occlusion (OMI) or non-occlusion (non-OMI) myocardial infarction. Unlike STEMI, OMI classification better aligns with focal angiographic findings since this group exclusively contains patients with total coronary occlusion. The color gradient indicates the severity of disease. This Figure was created with BioRender.com. Reproduced with permission from Al-Zaiti et. al.[1] (permission number 5471421247333, Licensed content publisher: Elsevier).

Extended Data Fig. 2. Graphical abstract summarizing the flow of study and main findings

This figure provides a graphical summary of the study flow and main findings.

Extended Data Fig. 3. Local explainability of feature importance on a selected example

This figure shows the ~~B~~baseline ECG of a 50-year-old female with a past medical history of hypertension, high cholesterol, prior myocardial infarction, and current smoking. The ECG was documented as benign with isolated non-specific T wave changes, and the patient was triaged as

intermediate risk. ~~The OMI score was 62 indicating the need to up-triage.~~ The patient was later sent to the catheterization lab where she had complete occlusion of the right coronary artery. The OMI score on this baseline ECG was 62 indicating high risk designation. The force plot identified the five most important ECG features that met the contribution threshold of the random forest model: negative T wave in aVL, slight ST depression in aVL and V2, and slight ST elevation in aVF and III.

Extended Data Fig. ~~3~~4. Selected example~~s~~ of a patient~~s~~ correctly reclassified as OMI

This figure shows ~~two examples of~~ an ECG that was ~~patients who were~~ correctly reclassified as occlusion myocardial infarction by the machine learning model. ~~(A)~~ This ~~Baseline~~ baseline ECG was for ~~of~~ a 67-year-old male with a past medical history of high cholesterol and a prior myocardial infarction. The ST-depression in anterior-lateral leads were noted, and the patient was triaged as intermediate risk. The OMI score was 49 indicating the need to up-triage. The patient was later sent to the catheterization lab where he had complete occlusion of the right coronary artery. ~~(B) Baseline ECG of a 50-year-old female with a past medical history of hypertension, high cholesterol, prior myocardial infarction, and current smoking. The ECG was documented as benign with isolated non-specific T-wave changes, and the patient was triaged as intermediate risk. The OMI score was 62 indicating the need to up-triage. The patient was later sent to the catheterization lab where she had complete occlusion of the right coronary artery.~~

Extended Data Fig. ~~4~~5. Selected example of a missed OMI by our model

This figure provides a selected example of a patient with occlusion myocardial infarction that was missed by the machine learning model and other reference standards. This ECG was obtained on a 70-year-old female with a past medical history of hypertension, high cholesterol, prior myocardial infarction, and current smoking. The baseline clinical interpretation suggests normal sinus rhythm with benign findings. There are isolated Q waves in inferior leads, low ECG voltage, and some baseline wander and high frequency noise in few leads. The OMI risk score was 2 indicating a low risk. The patient was later sent to the catheterization lab, which showed significant left main occlusion and had

many stents placed. The patient developed new-onset HF during hospitalization. A closer look at this ECG by experienced ECG readers suggests that this ECG could resemble the “precordial swirl pattern”, a rightward ST-elevation vector, with STE in V1 and aVR and reciprocal ST-depression in V5 and V6. This pattern was found to correlate with LAD occlusion.

Extended Data Fig. ~~5~~6. Development and validation of an algorithm to screen for any ACS event

This figure shows the classification performance of the machine learning model against other reference standards for detecting any acute coronary syndrome event (ACS). The figure also shows the distribution of patients in low-risk, intermediate risk, and high-risk groups as per our derived risk score. There is a notable gain in precision (rule-in) but a significant loss in recall (rule-out).

Extended Data Fig. ~~6~~7. Limitations of ST amplitude on surface ECG as a sole marker of myocardial ischemia

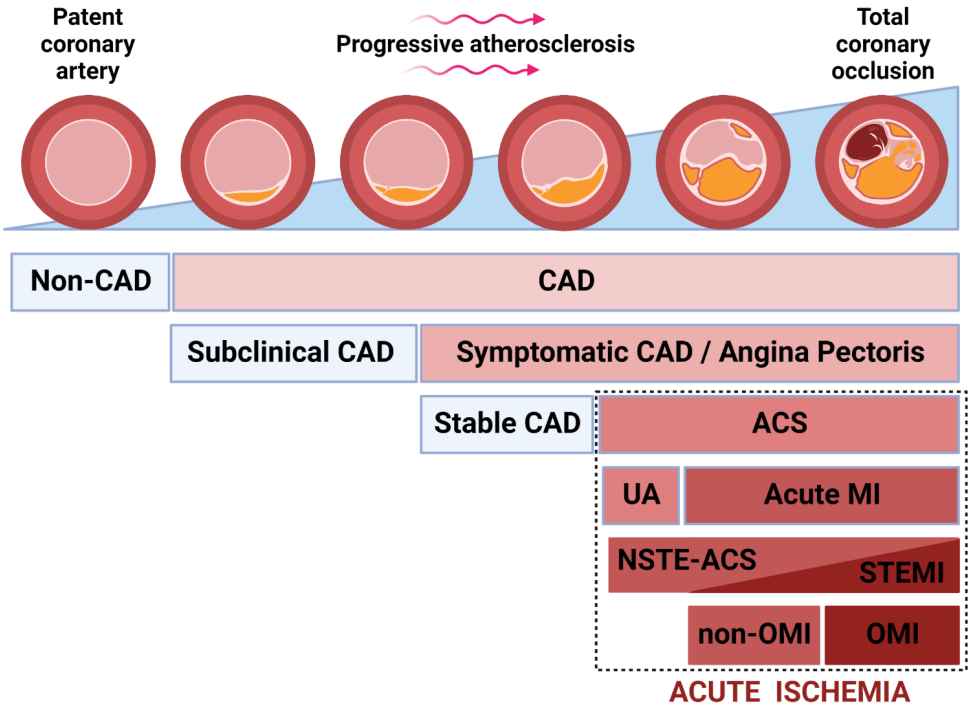
This figure shows: (A) cardiac model of anterior wall epicardial ischemia with corresponding ST-elevation on V3 to V5 of the 12-lead ECG. (B) cardiac model of anterolateral and inferior-apical epicardial ischemia with corresponding attenuation of ST changes on the 12-lead ECG. This figure was generated using ECGSIM (www.ecgsim.org). Reproduced with permission from Al-Zaiti et. al.[1] (permission number 5471421247333, Licensed content publisher: Elsevier).

Extended Data Fig. ~~7~~8. Comparison between 10 algorithms trained on the derivation cohort to classify OMI

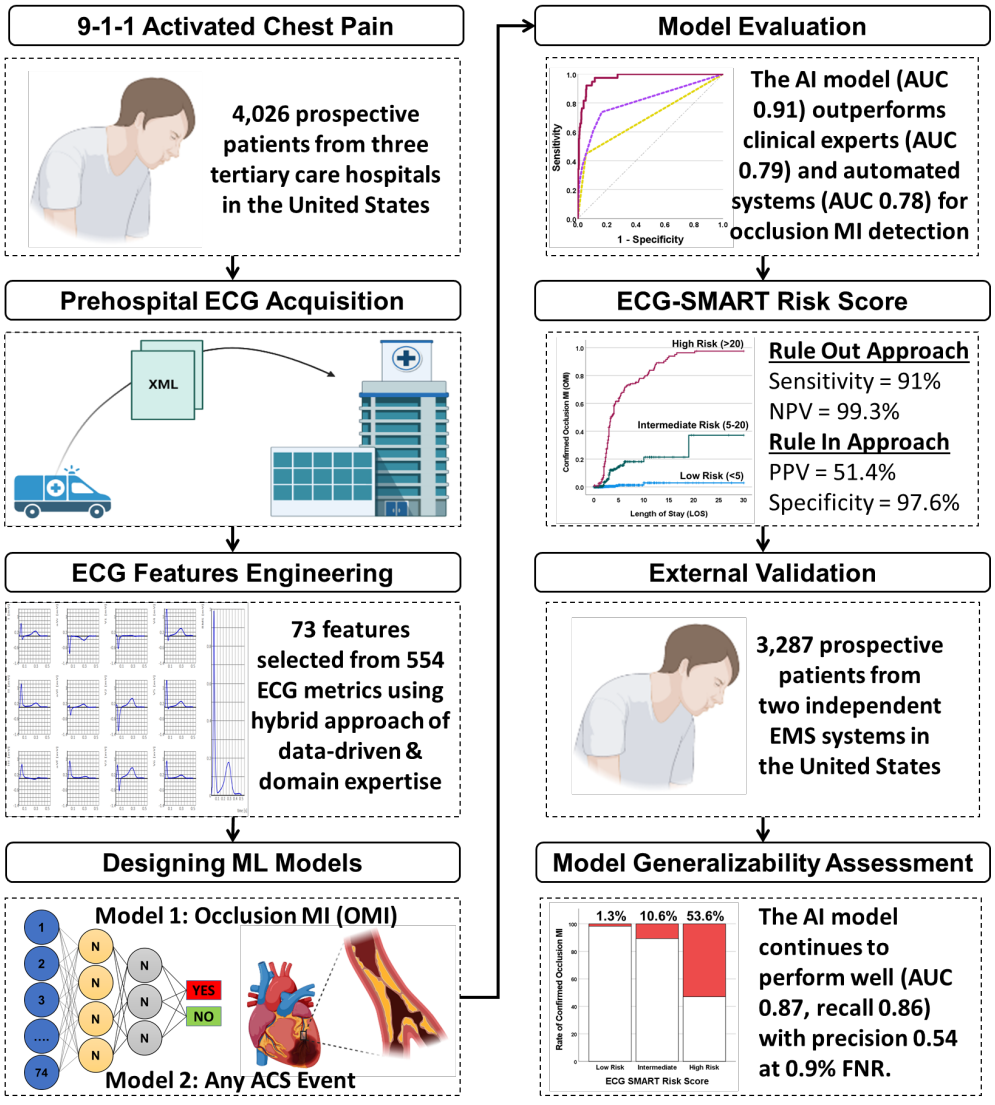
This figure compares the area under the receiver operator characteristics curves (95% confidence interval) of 10 classifiers during training (left) and testing (right) on the derivation cohort. RF: random forest; KNN: K-nearest neighbors; GBM: gradient boosting machine; XGB: extreme gradient boosting; SVM: support vector machine; ANN: artificial neural networks; LogReg: regularized logistic regression; LDA: linear discriminant analysis; SGD_LogReg: stochastic gradient descent logistic regression; G_NB: Gaussian Naïve Bayes.

871

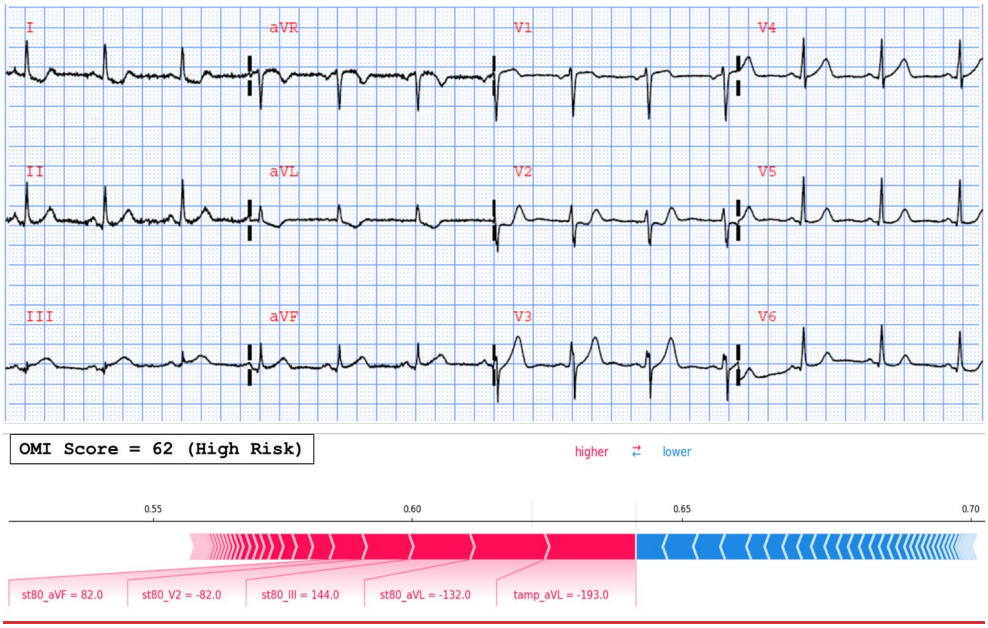
Extended Data Fig 1. The relationship between magnitude of coronary occlusion and coronary artery disease and acute events classification



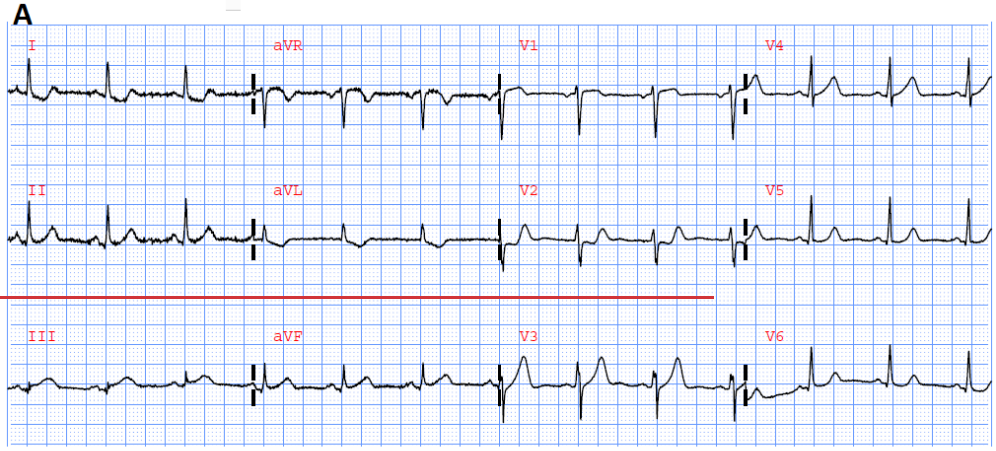
Extended Data Fig 2. Graphical abstract summarizing the flow of study and main findings

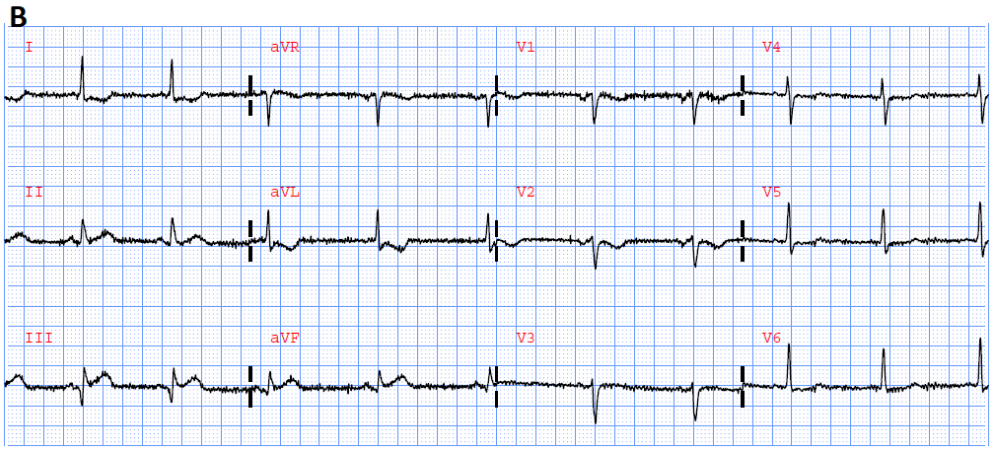


Extended Data Fig. 3. Local explainability of feature importance on a selected example

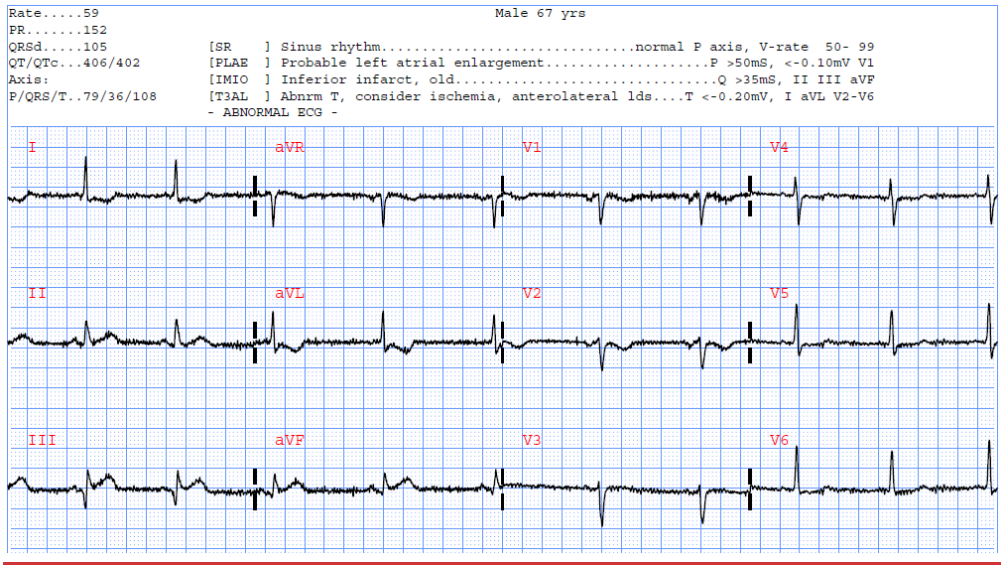


Extended Data Fig ~~3~~4. Selected example~~s~~ of a patient~~s~~ correctly reclassified as OMI

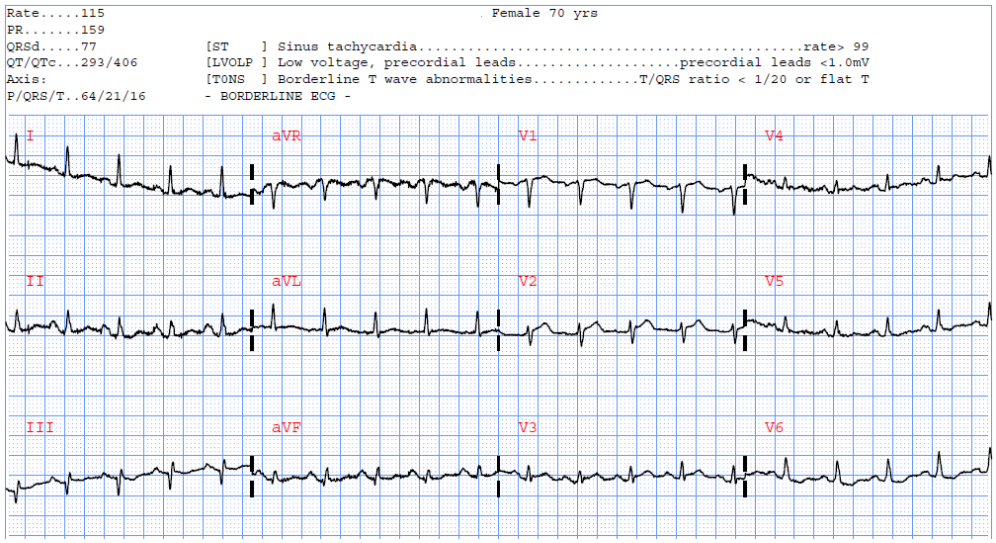

A
I
aVR
V1
V4
II
aVL
V2
V5
III
aVF
V3
V6


B
I
aVR
V1
V4
II
aVL
V2
V5
III
aVF
V3
V6

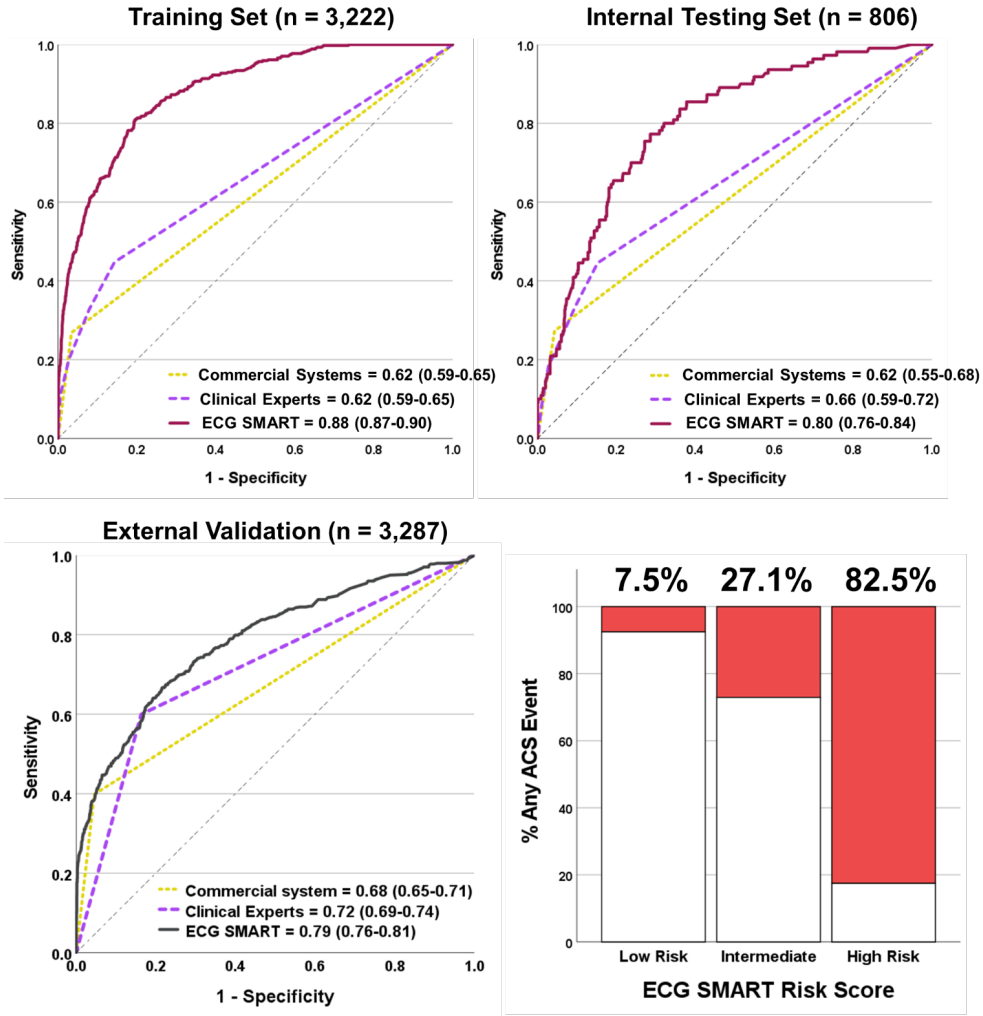
882


Rate.....59
Male 67 yrs
PR.....152
QRSd.....105
[SR] Sinus rhythm.....normal P axis, V-rate 50- 99
QT/QTc...406/402
[PLAE] Probable left atrial enlargement.....P >50mS, <-0.10mV V1
Axis:
[IMIO] Inferior infarct, old.....Q >35mS, II III aVF
P/QRS/T..79/36/108
[T3AL] Abnrm T, consider ischemia, anterolateral lds....T <-0.20mV, I aVL V2-V6
- ABNORMAL ECG -
I
aVR
V1
V4
II
aVL
V2
V5
III
aVF
V3
V6

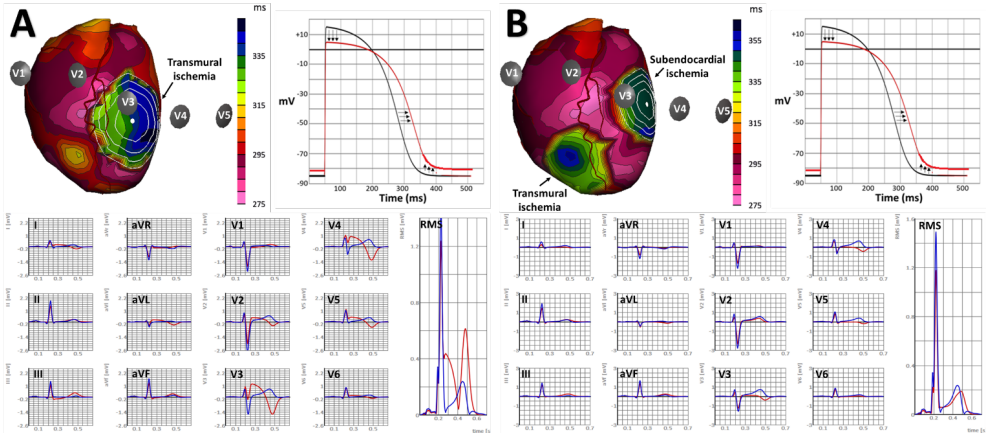
Extended Data Fig 45. Selected example of a missed OMI by our model



Extended Data Fig. ~~56~~. Development and validation of an algorithm to screen for any ACS event



Extended Data Fig. ~~6~~7: Limitations of ST amplitude on surface ECG as a sole marker of myocardial ischemia



Extended Data Fig ~~7~~8. Comparison between 10 algorithms trained on the derivation cohort to classify OMI