ESC European Society of Cardiology

European Heart Journal (2025) **46**, 1917–1929
https://doi.org/10.1093/eurheartj/ehaf004

CLINICAL RESEARCH
Digital health and innovation

Artificial intelligence applied to electrocardiogram to rule out acute myocardial infarction: the ROMIAE multicentre study

Min Sung Lee [1,2,†], Tae Gun Shin[3,†], Youngjoo Lee[4,†], Dong Hoon Kim[5], Sung Hyuk Choi[6], Hanjin Cho[7], Mi Jin Lee[8], Ki Young Jeong[9], Won Young Kim[10], Young Gi Min[11], Chul Han[12], Jae Chol Yoon[13], Eujene Jung[14], Woo Jeong Kim[15], Chiwon Ahn[16], Jeong Yeol Seo[17], Tae Ho Lim[18], Jae Seong Kim[19], Jeff Choi[20], Joon-myoung Kwon [1,2,*], and Kyuseok Kim[21,*], ROMIAE study group

[1]Medical AI Co., Ltd., 13F, 38, Yeongdong-daero 85-gil, Gangnam-gu, Seoul 06180, Republic of Korea; [2]Digital Healthcare Institute, Sejong Hospital, 28, Hohyeon-ro 489beon-gil, Sosa-gu, Bucheon-si, Gyeonggi-do 14754, Republic of Korea; [3]Department of Emergency Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea; [4]Department of Emergency Medicine, Soonchunhyang University Seoul Hospital, Seoul, Republic of Korea; [5]Department of Emergency Medicine, Gyeongsang National University College of Medicine, Jinju, Republic of Korea; [6]Department of Emergency Medicine, Korea University Guro Hospital, Seoul, Republic of Korea; [7]Department of Emergency Medicine, Korea University Ansan Hospital, Ansan, Republic of Korea; [8]Department of Emergency Medicine, School of Medicine, Kyungpook National University, Daegu, Republic of Korea; [9]Department of Emergency Medicine, Kyung Hee University College of Medicine, Kyung Hee University Hospital, Seoul, Republic of Korea; [10]Department of Emergency Medicine, Asan Medical Center, Seoul, Republic of Korea; [11]Department of Emergency Medicine, Ajou University School of Medicine, Suwon, Republic of Korea; [12]Department of Emergency Medicine, College of Medicine, Ewha Womans University, Seoul, Republic of Korea; [13]Department of Emergency Medicine, Research Institute of Clinical Medicine, Jeonbuk National University, Biomedical Research Institute, Jeonbuk National University Hospital, Jeonju, Republic of Korea; [14]Department of Emergency Medicine, Chonnam National University Medical School, Gwangju, Republic of Korea; [15]College of Medicine, Jeju National University, Jeju, Republic of Korea; [16]Department of Emergency Medicine, College of Medicine, Chung-Ang University, Seoul, Republic of Korea; [17]Department of Emergency Medicine, Hallym University Chuncheon Sacred Heart Hospital, Chuncheon, Republic of Korea; [18]Department of Emergency Medicine, Hanyang University College of Medicine, Seoul, Republic of Korea; [19]Department of Emergency Medicine, Incheon Sejong Hospital, Incheon, Republic of Korea; [20]Department of Surgery, Stanford University, Stanford, CA, USA; and [21]Department of Emergency Medicine, CHA Bundang Medical Center CHA University School of Medicine, 155, Bundangnaegok-ro, Bundang-gu, Seongnam-si, Gyeonggi-do 13524, Republic of Korea

Received 23 July 2024; revised 12 October 2024; accepted 1 January 2025; online publish-ahead-of-print 24 February 2025

See the editorial comment for this article 'Acute myocardial infarction diagnosis: impact of artificial intelligence-powered ECG interpretation', by I.-M Chiu and D. Ouyang, https://doi.org/10.1093/eurheartj/ehaf142.

Abstract

Background and Aims Emerging evidence supports artificial intelligence–enhanced electrocardiogram (AI-ECG) for detecting acute myocardial infarction (AMI), but real-world validation is needed. The aim of this study was to evaluate the performance of AI-ECG in detecting AMI in the emergency department (ED).

Methods The Rule-Out acute Myocardial Infarction using Artificial intelligence Electrocardiogram analysis (ROMIAE) study is a prospective cohort study conducted in the Republic of Korea from March 2022 to October 2023, involving 18 university-level teaching hospitals. Adult patients presenting to the ED within 24 h of symptom onset concerning for AMI were assessed. Exposure included AI-ECG score, HEART score, GRACE 2.0 score, high-sensitivity troponin level, and Physician AMI score. The primary outcome was diagnosis of AMI during index admission, and the secondary outcome was 30 day major adverse cardiovascular event (MACE).

Results The study population comprised 8493 adults, of whom 1586 (18.6%) were diagnosed with AMI. The area under the receiver operating characteristic curve for AI-ECG was 0.878 (95% CI, 0.868–0.888), comparable with the HEART score (0.877; 95% CI, 0.869–0.886) and superior to the GRACE 2.0 score, high-sensitivity troponin level, and Physician AMI score. For

* Corresponding authors. Tel: +82 10 4780 8321, Email: dreinstein70@gmail.com (K.K.); Tel: +82 10 7592 0533, Email: happywithhj@gmail.com (J.-m.K.)
† The first three authors contributed equally to this work.

predicting 30 day MACE, AI-ECG (area under the receiver operating characteristic, 0.866; 95% CI, 0.856–0.877) performed comparably with the HEART score (0.858; 95% CI, 0.848–0.868). The integration of the AI-ECG improved risk stratification and AMI discrimination, with a net reclassification improvement of 19.6% (95% CI, 17.38–21.89) and a *C*-index of 0.926 (95% CI, 0.919–0.933), compared with the HEART score alone.

Conclusions In this multicentre prospective study, the AI-ECG demonstrated diagnostic accuracy and predictive power for AMI and 30 day MACE, which was similar to or better than that of traditional risk stratification methods and ED physicians.

Structured Graphical Abstract

Key Question

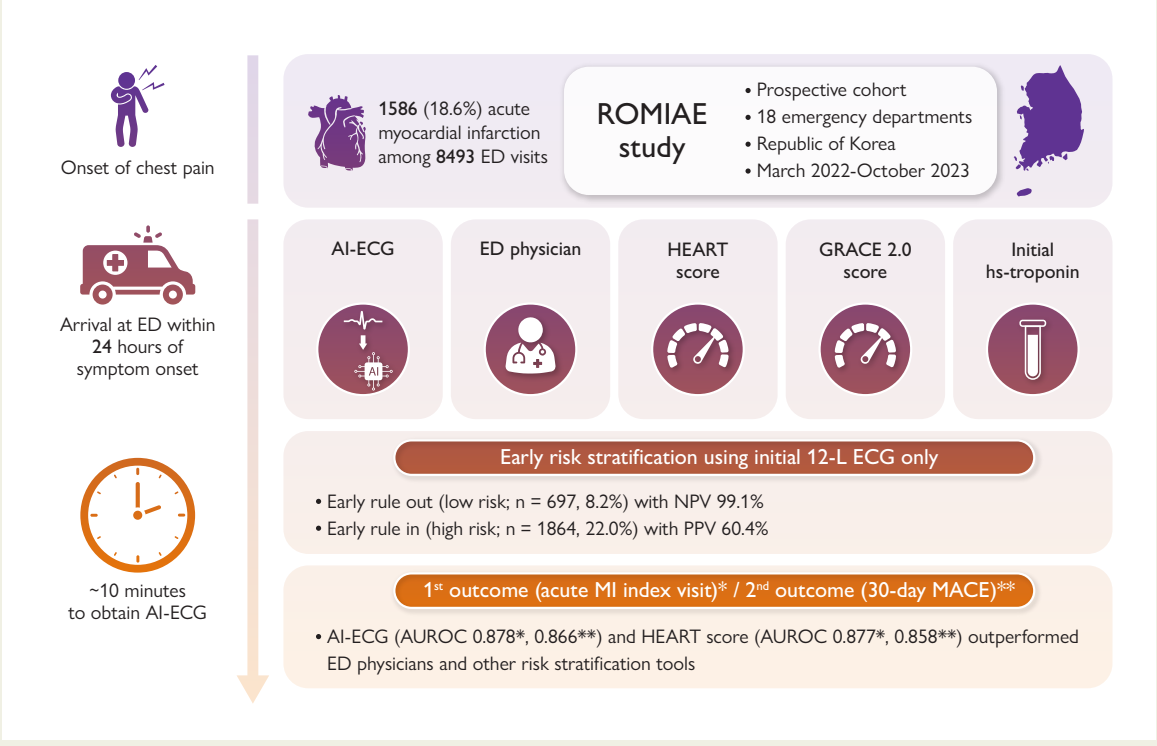
Does artificial intelligence-enhanced electrocardiography (AI-ECG) provide more effective risk stratification of acute myocardial infarction (AMI) in the emergency department (ED) setting compared to existing clinical scores?

Key Finding

AI-ECG demonstrated effective discrimination and risk stratification for AMI, performing comparably to the HEART score, and surpassing other tools. Its integration with HEART score further enhanced risk stratification.

Take Home Message

AI-ECG can rapidly and accurately stratify AMI risk in EDs, matching or surpassing risk stratification based on traditional methods. These findings could impact on the management of patients with suspected AMI.



An overview of ROMIAE study including design and main findings. ROMIAE, Rule-Out acute Myocardial Infarction using Artificial intelligence Electrocardiogram analysis; AI-ECG, artificial intelligence–enhanced electrocardiogram; AMI, acute myocardial infarction; AUROC, area under the receiver operating characteristic curve; ED, emergency department; MACE, major adverse cardiovascular event; NPV, negative predictive value; PPV, positive predictive value.

Keywords Acute myocardial infarction • Artificial intelligence • Electrocardiogram • Emergency department • Acute coronary syndrome • AI/ML-enabled SaMD

Introduction

Cardiovascular disease remains the leading cause of mortality worldwide, with acute myocardial infarction (AMI) being a significant contributor.[1,2] The conventional protocol for patients presenting to the emergency department (ED) with signs suggestive of AMI includes immediate acquisition of an electrocardiogram (ECG), with a target timeframe of within 10 min of arrival.[3] The interpretation of these ECGs by a specialist is crucial for the early risk stratification of ST-elevation myocardial infarction (STEMI) and non–ST-elevation acute coronary

syndromes (NSTE-ACS). Nevertheless, the prompt and precise analysis of ECG findings is a complex task in practical settings. Although the presence of STEMI might be discernible on the initial ECG, discrepant interpretations can occur among even well-trained physicians.[4,5] Additionally, ECGs can present as normal upon expert review in 17%–33% of patients who are ultimately diagnosed with NSTE-ACS.[3,6] Examinations can be delayed by ED overcrowding or during pivotal healthcare challenges, such as the COVID-19 pandemic.[7]

Clinical decision support using artificial intelligence (AI) has recently been introduced into numerous medical fields, including emergency care.[8,9] Despite advances, the translation of AI for practical clinical use remains limited,[9–12] partly due to a dearth of rigorous clinical validation studies. Several AI algorithms for ECG interpretation have been proposed, including for AMI detection,[13,14] but prospective studies are needed to validate the utility of such AI systems for AMI diagnosis in clinical settings. This prospective multicentre study aimed to externally validate the predictive performance of an AI-enhanced 12-lead ECG (AI-ECG) analytic model and compare its performance with existing AMI risk stratification models. We hypothesized that AI-ECG, a low-cost and easily-implementable model, would have superior discrimination in comparison with existing models, warranting larger-scale implementation and subsequent clinical efficacy studies. This study involved comparing initial AI-ECG interpretations upon ED admission with established risk stratification tools, namely the HEART and GRACE scores, and clinical assessments that consider the initial ECG manifestations and patient presentation.

Methods

This study was a prospective multicentre external validation study conducted in 18 EDs (1 certified cardiovascular hospital and 17 university-level hospitals) in Korea from March 2022 to October 2023. The institutional review board at each of the 18 hospitals approved the study protocol. All participants provided written informed consent. A comprehensive study protocol, encompassing the eligibility criteria, AI-ECG model, and variable descriptions, was published previously[15]; we highlight key study design elements according to the TRIPOD-AI checklist (see Supplementary data online, supplementary file). This study was registered on ClinicalTrials.gov (NCT05435391).

Artificial intelligence–enhanced electrocardiogram model to estimate probability of acute myocardial infarction

This study evaluated an advanced algorithm based on the previously reported AI-ECG model, AiTiAMI version 1.00.00 (Medical AI Co., Ltd, Seoul, Republic of Korea; see Supplementary data online, *Figure S1*). AiTiAMI is built on a residual neural network and was trained using multicentre 12-lead ECG raw data from Korea. Notably, the hospitals included in this validation study were not part of the training dataset. The sole input for the model is 500 Hz 12-lead ECG data, and it is agnostic to the ECG machine manufacturer. Quality filtering is performed through an internal noise assessment module. In the initial retrospective external validation study, the model demonstrated an area under the receiver operating characteristic (AUROC) curve of 0.951 for STEMI and 0.901 for AMI [STEMI and non-STEMI (NSTEMI)].[13] AiTiAMI ultimately derives AMI probability scores ranging from 0 to 100, stratifying patients into low- (score <3.0, corresponding to a sensitivity of 99%), intermediate-, and high-risk (score ≥48.5, corresponding to a specificity of 90%) categories for having AMI. In this prospective observational study, physicians in the ED were blinded from AI-ECG output.

Study population

This study cohort encompassed adult patients (aged >18 years) who presented to the ED with clinically suspected AMI. Patients who arrived at the ED within 24 h of their initial chest pain, those reporting worsening chest pain within 24 h prior to their ED admission, and those encountering recurrent symptoms within 24 h after ED admission were eligible for inclusion. Patients presenting with out-of-hospital cardiac arrest upon ED arrival, those who declined participation in the study, those experiencing traumatic chest pain, and those diagnosed with conditions clearly distinct from myocardial infarction (MI) (such as pneumothorax) were excluded. As described in the protocol paper, assuming a 10% dropout rate, the estimated sample size to detect totalled 8814 participants.

Data collection

Participation in this study did not impose any treatment restrictions on enrolled patients; each ED provided standard care according to international guidelines. The 18 EDs adhered to the 0/1 h algorithm or the 0/3 h algorithm based on the guidelines.[16,17] In accordance with sanctioned protocol,[15] we collected data prospectively without interfering with patients' ED processes. Mandatory collection data comprised the initial 12-lead ECG, initial high-sensitivity troponin (hs-troponin) I or T, the 12-lead ECG machine manufacturer, cardiac biomarker assay manufacturer, centre-specific 99th percentile troponin upper reference limit, chief complaints (classified as typical, atypical, and non-typical), and ED physician-estimated AMI probability score after initial patient examinations with manual ECG review (Physician AMI score, ranging from 0 to 10). Follow-up ECG and cardiac biomarker assessments were not mandatory.

Outcomes

The primary outcome was an AMI diagnosis during the index admission, encompassing both Type 1 and Type 2 MIs. We used the fourth universal definition of AMI.[18] The types of AMI (Types 1, 2, 3, 4, and 5) and labelling of acute coronary syndrome (STEMI, NSTEMI, and unstable angina) were determined based on clinical information at the time of the ED visits and from further examinations, such as coronary angiography and echocardiography.[3] Two board-certified emergency medicine specialists from each emergency centre were responsible for this task, and when their opinions did not align, a third board-certified emergency medicine specialist was consulted to reach a final decision. The secondary outcome was major adverse cardiovascular events (MACEs), defined as any instance of death, MI (index and recurrent), stroke, target vessel revascularization, or stent thrombosis within 30 days after the index admission.

Statistical analysis

Our primary analysis evaluated the predictive performance of the AI-ECG against the Physician AMI score, HEART score, GRACE 2.0 score, and initial hs-troponin level (AMI diagnosis threshold: 99% upper reference limit) for both the primary and secondary outcomes. We evaluated discrimination using accuracy, the AUROC, area under the precision-recall curve (AUPRC), sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) at the best cut-off value based on Youden's J statistic. The confidence intervals (CIs) for the thresholds are calculated using bootstrap resampling and the averaging methods.[19]

Furthermore, to evaluate the risk stratification potentials in clinical workflow, the performance of the AI-ECG model and each conventional clinical risk stratification tool (HEART score and GRACE 2.0 score) were assessed at their pre-specified low- to intermediate-risk and intermediate- to high-risk thresholds. For the AI-ECG model, the thresholds were <3.0 for low risk and ≥48.5 for high risk. The HEART score thresholds were 0–3 for low risk and 7–10 for high risk, while the GRACE 2.0 score thresholds were <109 for low risk and >140 for high risk.[20,21] The Physician AMI score and hs-troponin lack reference values for stratifying patients into three risk levels, so we only compared the AI-ECG model with the HEART score and GRACE 2.0 score in this evaluation. In addition to comparing the

performance of AI-ECG, HEART score, and GRACE 2.0 score individually, we also conducted clinical simulation to determine how patient clinical flow is affected when the hs-troponin or HEART score is applied following the AI-ECG assessment. We calculated *C*-index and the net reclassification improvement (NRI) to evaluate whether the addition of AI-ECG to the HEART score enhanced its discriminatory capability.[22] We developed a logistic regression model incorporating both the AI-ECG and HEART scores to evaluate the incremental value in risk stratification. The clinical utility of the AI-ECG model was evaluated using decision curve analysis (DCA) to compare its net benefit against other modalities.[23]

We conducted pre-defined subgroup analyses[15] according to the following items: demographics, past history, onset of chest pain, type of chest pain, ECG type and manufacturer, hospitals, type of MI (STEMI, NSTEMI, Type 1 MI, and Type 2 MI), and culprit coronary artery. In the subgroup analyses, AI-ECG's AMI discrimination performances were estimated using AUROC, AUPRC, sensitivity, specificity, PPV, and NPV at the best cut-off value based on Youden's J statistic. For STEMI, NSTEMI, Type 1, and Type 2 MI, the performance of the AI-ECG model was compared with each conventional clinical risk stratification tool. For STEMI, since the initial risk stratification is typically made by the ED physician based on a single ECG and clinical judgement, the AI-ECG score was compared directly with the Physician AMI score alone.

The DeLong test was used to compare two AUROCs. We used the PRROC R package to compare AUPRC values with bootstrap resampling.[24] Methods for comparing sensitivity, specificity, PPV, and NPV were chosen as appropriate.[25,26] We compared demographic characteristics of patients who were and were not ultimately diagnosed with AMI (AMI and non-AMI groups, respectively) using Student's *t*-test, Mann–Whitney *U* test, χ^2 test, or Fisher's exact test. For DCA, dcurves 0.5.0 R package was used. Statistical significance was assessed using a two-sided threshold of $P < .05$. During comparison analyses, we did not adjust the *P*-values for multiple comparisons because our primary objective was to explore and identify differences between each conventional risk stratification tool and the AI-ECG. We used R software version 4.3.2 and Python 3.9.7

Results

Study population

Between March 2022 and October 2023, 25,935 patients presented at the 18 EDs with suspected AMI, among whom 8493 comprised our study cohort (*Figure 1*). The median age of these patients was 62 (51–72) years (male, 62.5%). Among the 1586 patients (18.6%) diagnosed with AMI (*Table 1*), Type 1 and Type 2 MI comprised 94.8% and 5.1%, respectively, and STEMI and NSTEMI comprised 40.4% and 59.5%, respectively. The baseline characteristics are presented in *Table 1* and Supplementary data online, *Tables S1* and *S2*. Initial 12-lead ECGs and clinical risk score of the study population were all analysed and calculated using AiTiAMI and various risk stratification tools, with no missing data. In the AMI group, the Physician AMI score, HEART score, GRACE 2.0 score, and AI-ECG score (6.0, 7.0, 104.0, and 72.7, respectively) were significantly higher compared with the non-AMI group (2.0, 4.0, 79.0, and 14.0, respectively). Myocardial injury was present in the non-AMI group and AMI group at 1111 (16.0%) and 1199 (75.6%), respectively. In the non-AMI group, unstable angina was present in 8.3%. Within 30 days, MACE occurred in 7.1% of the AMI group and 1.8% of the non-AMI group. Supplementary data online, *Tables S1* and *S2* compare other variables described in the study protocol. Notably, more than 99% of AMI patients' AI-ECG score were distributed in the intermediate to high risk zone, and 92.6% of STEMI patients' AI-ECG score were distributed in the high-risk zone. The AI-ECG score distribution is detailed in Supplementary data online, *Figure S2*.

Discrimination performance for diagnosing acute myocardial infarction

The performance metrics of AI-ECG score, Physician AMI score, HEART score, GRACE 2.0 score, and initial hs-troponin level were compared (*Table 2*; *Figure 2*). The difference in the AUROC between AI-ECG score and HEART score was not statistically significant (*P*-value .944), with values of 0.878 (95% CI 0.868–0.888) and 0.877 (95% CI 0.869–0.886), respectively. The AUROC values for Physician AMI score, GRACE 2.0 score, and initial hs-troponin level were 0.846 (95% CI 0.834–0.857), 0.711 (95% CI 0.698–0.724), and 0.798 (95% CI 0.783–0.812), respectively, all significantly lower than that of AI-ECG score (*P*-value <.001). The HEART score demonstrated slightly higher sensitivity than AI-ECG score (0.794 vs. 0.767; $P = .035$), but AI-ECG had significantly higher specificity (0.848 vs. 0.814; $P < .001$) and PPV (0.536 vs. 0.495; $P < .001$). The AUPRC for AI-ECG score was significantly higher than that for HEART score, with values of 0.727 (95% CI 0.707–0.748) and 0.641 (95% CI 0.615–0.667), respectively. The precision-recall curves for AI-ECG, HEART, and Physician AMI scores are presented in Supplementary data online, *Figure S2*.

Discrimination performance for 30 day major adverse cardiovascular event

The difference in the AUROC between AI-ECG and HEART scores was again insignificant. The AUROC values for Physician AMI score, GRACE 2.0 score, and initial hs-troponin level were significantly lower than that of AI-ECG score ($P < .001$) (*Table 2*; *Figure 2*). Sensitivity did not differ significantly between AI-ECG score and the four comparison tests. The specificity of AI-ECG score was 0.852 (95% CI 0.844–0.861), which was significantly higher than that of HEART, Physician AMI, and GRACE 2.0 scores ($P < .001$). Among the five tests, the AUPRC for AI-ECG score was significantly higher than that of the others (0.720, 95% CI 0.698–0.740).

Comparison of acute myocardial infarction risk stratification methods

For the primary outcome, AI-ECG, HEART, and GRACE 2.0 scores classified 697 (8.2%), 3106 (36.6%), and 6308 (74.3%) of the 8493 patients, respectively, into the low-risk group (*Figure 3*). At the low to intermediate risk cut-off, the sensitivity of AI-ECG, HEART, and GRACE 2.0 scores was 99.6% (95% CI 99.3–99.9), 97.0% (95% CI 96.2–97.9), and 46.0% (95% CI 43.6–48.5), respectively, and the NPV was 99.1% (95% CI, 98.5–99.8), 98.5% (95% CI 98.1–98.9), and 86.4% (95% CI 85.6–87.3), respectively (see Supplementary data online, *Figures S3* and *S4* and *Table S3*). Artificial intelligence–enhanced electrocardiogram was the only model that met the accepted threshold of a missed AMI rate of <1%.[27] AI-ECG, HEART, and GRACE 2.0 scores categorized 1864 (22.0%), 1459 (17.2%), an d 595 (7.0%) of the 8493 patients, respectively, into the high-risk group (*Figure 3*). At the intermediate to high-risk cut-off, the specificity of AI-ECG, HEART, and GRACE 2.0 scores was 89.3% (95% CI 88.6–90.0), 92.4% (95% CI 91.7–93.0), and 94.7% (95% CI 94.2–95.3), respectively, and the PPV was 60.4% (95% CI 58.2–62.6), 63.9% (95% CI 61.4–66.3), and 38.8% (95% CI 34.9–42.7), respectively (see Supplementary data online, *Figures S3* and *S4* and *Table S3*). The DCA demonstrated a mostly positive and higher net benefit of the AI-ECG model compared with other scores for decision thresholds below 0.086 and above 0.385 (see Supplementary data online, *Figure S9*).

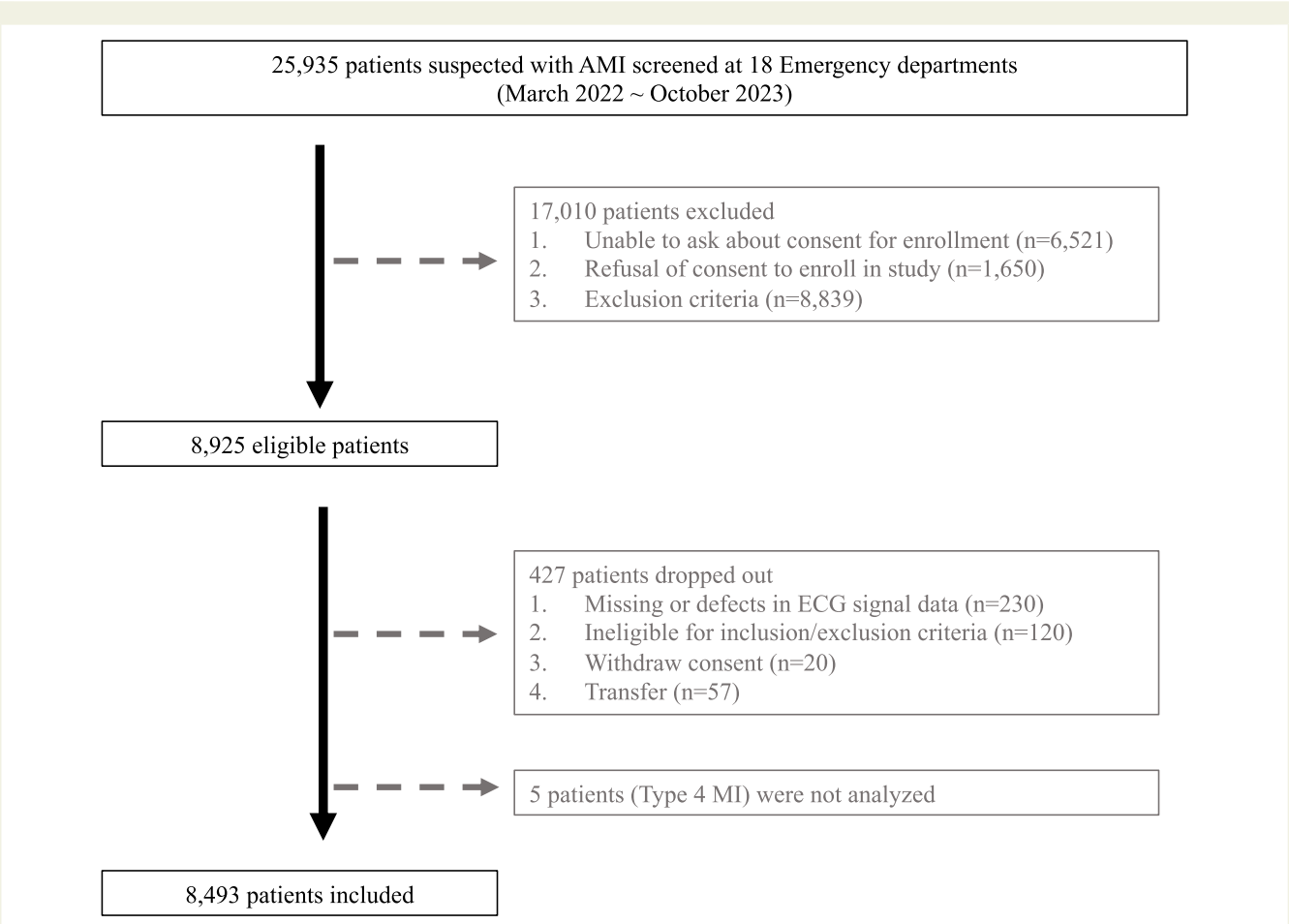


Figure 1 Study flow. This illustrates the selection process for patient inclusion in this study, which was conducted in 18 emergency departments from March 2022 to October 2023

Integration of an artificial intelligence–enhanced electrocardiogram with high-sensitivity troponin and HEART scores for enhanced clinical decision-making

The AI-ECG model categorized 8.2% of patients ($n = 697$) as low risk, yielding an AMI rate of 0.86%. This rate is significantly lower than that observed with the HEART score, which classified a larger cohort of patients ($n = 3016$) as low risk with an AMI rate of 1.5%. Similarly, the GRACE 2.0 score identified 74.3% of patients ($n = 6309$) as low risk, but with a substantially higher AMI rate of 13.6% (*Figure 3A*). Furthermore, the AI-ECG model demonstrated superior performance in identifying high-risk AMI patients, recognizing 1126 cases, while the HEART score identified 932 cases and the GRACE 2.0 score identified just 231 cases.

In the clinical simulation where the AI-ECG was applied to the initial ECG in the ED and subsequently combined with the HEART score (designated as AI-ECG + HEART), the low-risk group identified by the AI-ECG remained classified as low risk. The intermediate- and high-risk groups identified by the AI-ECG were then reclassified based on the HEART score (*Figure 3B*). Artificial intelligence–enhanced electrocardiogram + HEART resulted in a greater number of patients being stratified as low risk, yielding a reduced AMI rate of 0.77% among 3112 patients (37.0%), compared with the HEART score alone, which encompassed 3016 patients (36.6%) with an AMI rate of 1.5% (*Figure 3B*). Notably, the AI-ECG + HEART approach identified an additional 187 AMI patients classified as high-risk ($n = 596$, AMI rate 42.7%) or very high-risk ($n = 717$, AMI rate 82.4%) compared with the high-risk group identified solely by the HEART score. Moreover, the intermediate-risk group within the AI-ECG + HEART cohort ($n = 3,116$, 36.6%; AMI rate 7.9%) demonstrated a decrease in both patient numbers and AMI risk. The integration of AI-ECG resulted in a 19.6% (95% CI 17.38–21.89) increase in the NRI compared with the HEART score alone (*Table 3*). The combined AI-ECG + HEART model demonstrated superior AMI discrimination with a *C*-index of 0.926 (95% CI 0.919–0.933) vs. the HEART score alone (AUROC 0.877, 95% CI 0.869–0.886).

In the clinical simulation of applying AI-ECG and hs-troponin measurement in the ED, the combined assessment further stratified patients identified as low or intermediate risk by the AI-ECG into subgroups characterized by differing AMI rates. Specifically, those with no hs-troponin elevation exhibited lower AMI rates of 0.5% in the low-risk group and 2.4% in the intermediate-risk group. Conversely, patients with hs-troponin elevation exhibited significantly higher AMI rates, with 8.6% in the low-risk group and 29.4% in the intermediate-risk group (*Figure 3B*). Notably, the high-risk group identified by the AI-ECG group consistently showed an elevated AMI rate, regardless of the hs-troponin results (*Figure 3B*).

Table 1 Baseline characteristics

	Non-AMI (*N* = 6907)	AMI (*N* = 1586)
Demographics		
Age, year	61 (49–71)	65 (56–74)
Male	4056 (58.7)	1251 (78.8)
History		
Hypertension	3104 (44.9)	819 (51.6)
Diabetes mellitus	1603 (23.2)	514 (32.4)
Hyperlipidaemia	1902 (27.5)	411 (25.9)
Coronary artery disease	1432 (20.7)	330 (20.8)
AMI	605 (8.7)	221 (13.9)
Congestive heart failure	246 (3.5)	56 (3.5)
Chronic kidney disease	420 (6.0)	128 (8.0)
ECG device		
GE	5991 (86.7)	1434 (90.4)
Philips	916 (13.2)	152 (9.5)
ED chief complaints		
Typical chest pain	2060 (29.8)	1078 (67.9)
Atypical chest pain	2456 (35.5)	294 (18.5)
Non-cardiac chest pain	695 (10.0)	64 (4.0)
Other symptoms	1695 (24.5)	151 (9.5)
Chest pain onset to ED visit ≤3h	3529 (51.0)	827 (52.1)
Physician AMI score, mean (SD)	2.0 (1.0–3.0)	6.0 (4.0–9.0)
HEART score, mean (SD)	4.0 (2.0–5.0)	7.0 (6.0–8.0)
GRACE 2.0 score, mean (SD)	79.0 (56.0–103.0)	104.0 (84.0–127.0)
AI-ECG score, mean (SD)	14.0 (5.9–29.8)	72.7 (43.0–92.2)
Type 1		74.4 (44.7–93.0)
Type 2		43.9 (13.6–69.7)
STEMI		94.8 (77.8–98.0)
NSTEMI		54.1 (22.0–77.3)
Myocardial injury[a]	1111 (16.0)	1199 (75.6)
AMI type		
Type 1		1504 (94.8)
Type 2		82 (5.1)
ACS type		
STEMI		642 (40.4)
NSTEMI		944 (59.5)
Unstable angina	576 (8.3)	
30 day MACE		
Total MACE	125 (1.8)	113 (7.1)

Values are expressed as *n* (%) or median with interquartile range.
ACS, acute coronary syndrome.
[a]Myocardial injury is defined as an elevation of hs-troponin above the 99th percentile upper reference limit specific to each ED.

Table 2 Comparison of performance detecting acute myocardial infarction and predicting 30 day major cardiovascular adverse event

	AUROC	*P*	Sensitivity	*P*	Specificity	*P*	PPV	*P*	NPV	*P*	AUPRC	*P*
AMI												
AI-ECG	**0.878 (0.868–0.888)**		0.767 (0.746–0.788)		**0.848 (0.839–0.856)**		**0.536 (0.516–0.557)**		**0.941 (0.935–0.946)**		**0.727 (0.707–0.748)**	
HEART score	**0.877 (0.869–0.886)**	.944	**0.794 (0.775–0.814)**	.035	0.814 (0.805–0.823)	<.001	0.495 (0.476–0.515)	<.001	**0.945 (0.939–0.951)**	.154	0.641 (0.615–0.667)	<.001
Physician AMI score	0.846 (0.834– 0.857)	<.001	0.751 (0.730–0.772)	.205	0.790 (0.780–0.799)	<.001	0.450 (0.431–0.469)	<.001	0.932 (0.926–0.939)	.009	0.657 (0.633–0.679)	<.001
hs-troponin level	0.798 (0.783–0.812)	<.001	0.756 (0.735–0.777)	.471	**0.839 (0.830–0.848)**	.116	**0.519 (0.499–0.539)**	.085	**0.937 (0.931–0.943)**	.378	0.462 (0.442–0.482)	<.001
GRACE score	0.711 (0.698–0.724)	<.001	0.745 (0.723–0.766)	.122	0.565 (0.554–0.577)	<.001	0.282 (0.269–0.296)	<.001	0.906 (0.897–0.915)	<.001	0.339 (0.318–0.362)	<.001
30 day major cardiovascular adverse event												
AI-ECG	**0.866 (0.856–0.877)**		0.736 (0.716–0.757)		**0.852 (0.844–0.861)**		**0.557 (0.537–0.578)**		**0.928 (0.921–0.934)**		**0.720 (0.698–0.740)**	
HEART score	**0.858 (0.848–0.868)**	.181	0.756 (0.736–0.777)	.136	0.816 (0.807–0.825)	<.001	0.509 (0.489–0.528)	<.001	**0.930 (0.923–0.936)**	.519	0.634 (0.610–0.658)	<.001
Physician AMI score	0.828 (0.817–0.840)	<.001	0.718 (0.696–0.739)	.115	0.791 (0.782–0.801)	<.001	0.464 (0.445–0.483)	<.001	0.917 (0.910–0.924)	.002	0.645 (0.623–0.667)	<.001
hs-troponin level	0.786 (0.771–0.801)	<.001	0.728 (0.707–0.749)	.543	**0.843 (0.834–0.852)**	.122	0.539 (0.519–0.560)	.010	**0.925 (0.918–0.931)**	.430	0.475 (0.455–0.493)	<.001
GRACE score	0.717 (0.705–0.730)	<.001	0.746 (0.725–0.766)	.532	0.571 (0.560–0.583)	<.001	0.305 (0.291–0.319)	<.001	0.899 (0.890–0.908)	<.001	0.375 (0.351–0.397)	<.001

The values demonstrating the highest performance in each column are highlighted in bold. Each *P*-value represents the result of a comparative analysis with the AI-ECG score.

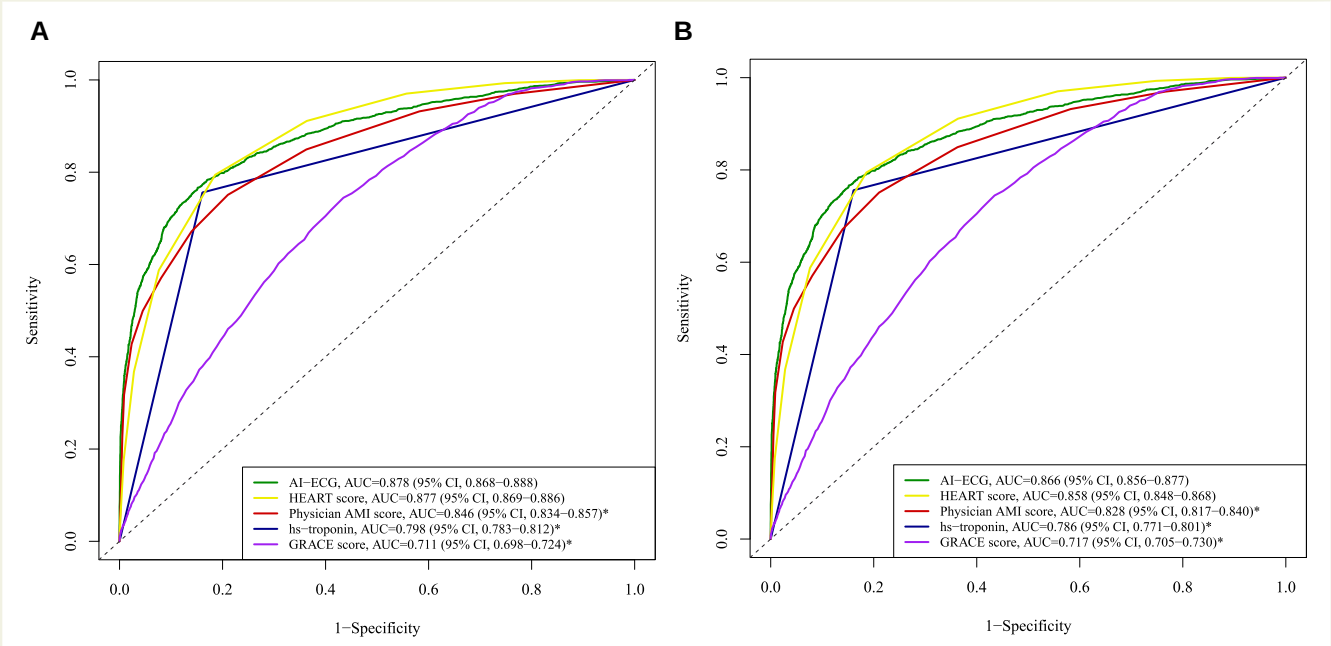


Figure 2 Primary and secondary outcome analyses using receiver operating characteristic curves. (*A*) Receiver operating characteristic curve for the primary outcome (diagnosis of acute myocardial infarction at the index visit). (*B*) Receiver operating characteristic curve for the secondary outcome (prediction of 30 day major adverse cardiovascular events). Each curve represents a different diagnostic tool or score: Artificial Intelligence–Enhanced Electrocardiogram score, HEART score, Physician Acute Myocardial Infarction score, initial high-sensitivity troponin level, and GRACE score. The area under the receiver operating characteristic curve for each diagnostic measure is displayed with the corresponding 95% confidence intervals, indicating the performance of each test in terms of its sensitivity (true positive rate) vs. 1−specificity (false positive rate). *$P < .05$

Pre-defined subgroup analysis

The findings of pre-defined subgroup analysis indicate that the detection capability of AI-ECG score diminished among individuals aged 65 and older or those with coronary artery disease, heart failure, chronic kidney disease, or peripheral artery disease (see Supplementary data online, *Table S4*). There was no difference in performance based on biological sex. Discrepancies were noted when comparing different types of ECGs, their features, and the institutions where they were performed. However, no significant differences in performance were noted based on the ECG device, ECG features, type, or onset of chest pain (see Supplementary data online, *Table S4*). The location of the culprit artery did not significantly affect the performance of AI-ECG (see Supplementary data online, *Table S5*).

When we evaluated the performance of AI-ECG score for STEMI and NSTEMI, the metrics were as follows. For STEMI, the AUROC was 0.971 (95% CI 0.965–0.977), with a sensitivity of 92.5% (95% CI 90.5–94.6), specificity of 89.9% (95% CI 89.2–90.6), PPV of 45.9% (95% CI 43.2–48.6), and NPV of 99.2% (95% CI 99.0–99.4). For NSTEMI, the values were an AUROC of 0.814 (95% CI 0.799–0.830), sensitivity of 65.5% (95% CI 62.4–68.5), specificity of 82.9% (95% CI 82.1–83.8), PPV of 34.4% (95% CI 32.2–36.6), and NPV of 94.6% (95% CI 94.0–95.2) (see Supplementary data online, *Table S5*). When comparing the risk stratification tools for the detection of STEMI (see Supplementary data online, *Table S6*), we found no significant difference between the AI-ECG score and the Physician AMI score in terms of AUROC (0.971 vs. 0.965). However, the AI-ECG score demonstrated higher sensitivity (92.5% vs. 86.4%) and NPV (99.2% vs. 98.7%). Artificial intelligence–enhanced electrocardiogram score's performance for NSTEMI had lower sensitivity than HEART score and blood biomarker. Nonetheless, AI-ECG score exceeded Physician AMI score in terms of both AUROC and AUPRC. Notably, no STEMI patients were identified among those classified as low-risk by AI-ECG score. Consequently, the NPV for STEMI based on the low to intermediate risk cut-off was 100%. The NPV for NSTEMI based on the low-intermediate risk cut-off was 99.1% (95% CI 98.5%–99.8%) (see Supplementary data online, *Table S7*).

The performance of AI-ECG for Type 1 MI and Type 2 MI was evaluated, and the results were described in Supplementary data online, *Table S5*. When comparing the risk stratification performances, AI-ECG and HEART scores had the highest AUROC (0.887, 0.883) to detect Type 1 MI. Notably, AI-ECG demonstrated higher specificity, PPV, and AUPRC for Type 1 MI, while its sensitivity was comparable with other risk stratification tools. For Type 2 MI, all risk stratification tools showed lower discrimination performance. Detailed results are presented in Supplementary data online, *Table S8*.

Discussion

To our knowledge, this is the first prospective multicentre observational study to validate an AI-ECG model for AMI detection. The results demonstrate that the AI-ECG outperformed existing risk stratification tools, showing superior or similar accuracy for both early rule-out and rule-in of AMI. The overall diagnostic performance of AI-ECG score, as manifested by the AUROC, was comparable with that of the HEART score (*Structured Graphical Abstract*). Considering that HEART score includes hs-troponin level, this is a noteworthy finding. The AI-ECG model's performance suggests its potential as a reliable digital biomarker for assisting clinicians in making timely decisions about patient management in emergency settings.

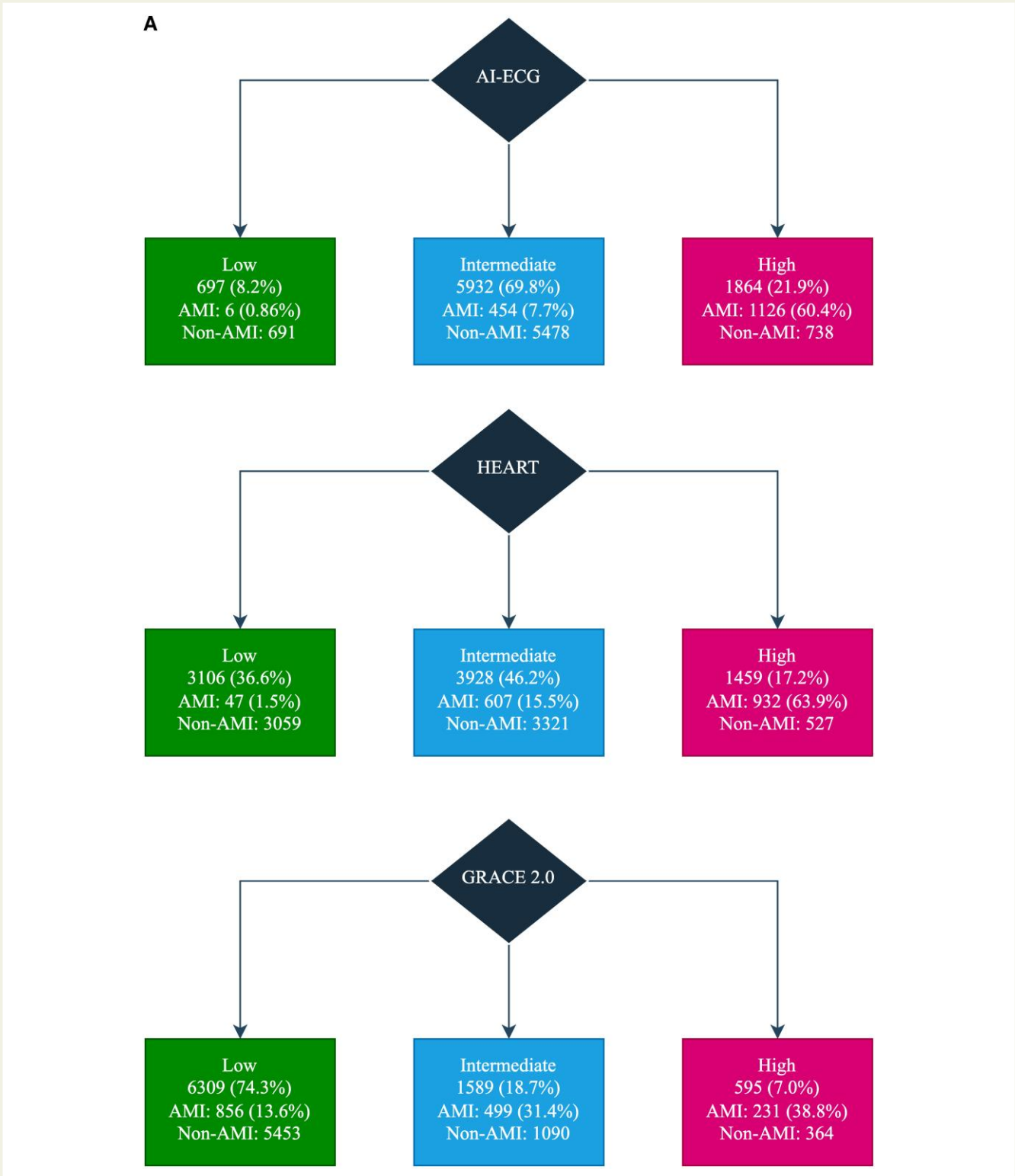


Figure 3 Risk stratification of acute myocardial infarction using Artificial Intelligence–Enhanced Electrocardiogram, hs-troponin, and HEART scores. (*A*) Classification of patients based on three acute myocardial infarction risk stratification tools: Artificial Intelligence–Enhanced Electrocardiogram, HEART score, and GRACE 2.0 score. Patients are categorized into low-, intermediate-, and high-risk groups for each tool. (*B*) Implementation scenario of Artificial Intelligence–Enhanced Electrocardiogram, evaluating patients based on initial Artificial Intelligence–Enhanced Electrocardiogram results, further stratified using high-sensitivity troponin levels and HEART scores. The HEART score is applied to patients initially classified as intermediate- or high-risk by Artificial Intelligence–Enhanced Electrocardiogram, further categorizing them into low-, intermediate-, high-, and very high-risk groups as depicted in the figure

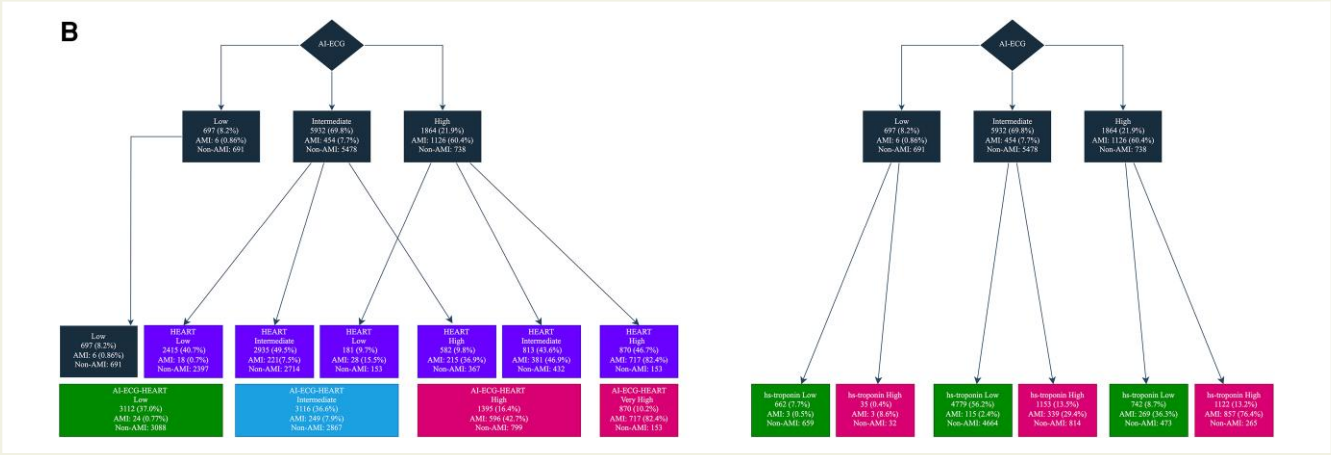


Figure 3 Continued

Table 3 Net reclassification improvement by integrating artificial intelligence–enhanced electrocardiogram with HEART score for acute myocardial infarction risk stratification

	HEART	AI-ECG + HEART			
		Low	Intermediate	High	Total
AMI	Low	19	28	0	47
	Intermediate	5	221	381	607
	High	0	0	932	932
	Total	24	249	1313	1586
Non-AMI	Low	2906	153	0	3059
	Intermediate	175	2714	432	3321
	High	7	0	520	527
	Total	3088	2867	952	6907

Reclassified					
[a]Higher		[a]Lower		Net	[b]NRI
409	25.79 (P up)	5	0.32 (P down)	25.47	19.64
585	8.47 (N up)	182	2.64 (N down)	−5.83	

Displays the NRI achieved by combining the AI-ECG score with the HEART score for risk stratification in patients with and without AMI. The values in the table represent the number of patients reclassified into different risk categories (low, intermediate, or high) before and after incorporating the AI-ECG score. Rows correspond to the initial HEART score classification, while columns indicate the reclassification after integrating AI-ECG.

[a]For AMI (top row groups): 'Higher' refers to patients correctly reclassified into a higher-risk category, while 'Lower' refers to patients incorrectly reclassified into a lower-risk category. For non-AMI (bottom row groups): 'Higher' refers to patients incorrectly reclassified into a higher-risk category, while 'Lower' refers to patients correctly reclassified into a lower-risk category.

[b]The NRI is calculated using the formula: NRI = (P up−P down) + (N down−N up).

Prior research on AI-ECG for AMI detection has primarily focused on retrospective model development and the assessment of diagnostic performance. Those studies often use comparisons with traditional ECG device diagnosis and interpretations by physicians in their validation process.[13,14] In this study, we performed a multi-institutional prospective observational investigation of a pre-existing AI-ECG model[13] and compared its performance with that of traditional risk stratification tools. This approach might facilitate the integration of AI-based clinical decision support into real practice.

Concern about the consequences of missing an AMI diagnosis is an important issue in emergency medicine. A systematic review has reported that the rates of missed AMI are ~1%–2%.[28] Furthermore, a large survey indicated that clinicians consider an AMI miss rate of 1% or less to be acceptable.[27] When employing 0/1 h or 0/3 h algorithms, the occurrence of AMI ratio and MACE in low-risk groups has been reported to be below 1%.[3,29] This implies that the sensitivity of any diagnostic tool should ideally be 99% or higher. In our study, various diagnostic tools were evaluated, and the AI-ECG scoring system was

the only model to meet the 1% threshold. The pre-specified cut-off of the AI-ECG score (<3.0 from AiTiAMI version 1.00.00) demonstrated a sensitivity of 99.6% and a NPV of 99.1% for the primary outcome in the low-risk cohort. Consequently, we think that the NPV of the AI-ECG tool is sufficiently high for clinical application. Additionally, it is important to note that the HEART score requires troponin level assessment, which necessitates ~1 h for result availability in real-world practice. In contrast, the diagnostic performance of AI-ECG, achieved through non-invasive analysis of 12-lead ECGs within a few minutes, underscores the potential value of AI-ECG for the early rule out of chest pain in ED patients.

In a recent pragmatic randomized controlled trial utilizing an AI-ECG model for detecting STEMI demonstrated a significant reduction in time to treatment.[30] This finding suggests that real-time application of the AI-ECG could aid physician's decision with identifying patients with STEMI. In this study, the AI-ECG model exhibited a performance detecting STEMI similar to Physician AMI score, but a higher sensitivity. In addition, previous studies demonstrated the ability of AI-ECG to successfully detect occlusive NSTEMI.[27,28] In our study, the diagnostic performance of AI-ECG was higher than Physician AMI score in detecting NSTEMI. Non–ST-elevation myocardial infarction is considered to be harder than STEMI to diagnose with ECG, so this finding shows the potential merit of AI-ECG score. Interestingly, for Type 1 MI and any culprit artery MI, the AI-ECG demonstrated similar levels of performance compared with previous studies, highlighting the potential utility of AI-ECG in detecting occlusive NSTEMI.[31,32] In the current Rule-Out acute Myocardial Infarction using Artificial intelligence Electrocardiogram analysis (ROMIAE) cohort, we additionally defined occlusive NSTEMI as NSTEMI patients who underwent percutaneous coronary intervention (PCI) or coronary artery bypass graft (CABG), achieving an AUROC of 0.832, a sensitivity of 71.6%, and a specificity of 80.4%, consistent with previous research findings.

To assess the value of early rule-in, we need to focus on the results that the AI-ECG score had a notably higher AUPRC than all other tools tested in this study. When considering a low-incidence variable, AUPRC is often more informative than AUROC because it focuses on the classifier's performance on the positive (less prevalent) class.[33,34] Moreover, considering the consequences of a false negative (patient's life at risk) and false positive (unnecessary admission or invasive tests), AUPRC might be a more appropriate metric if the primary concern is to ensure that all cases of MI are caught (high recall) while also maintaining a reasonable level of precision to avoid over-treating patients (high precision). Indeed, both AI-ECG and AI-ECG + HEART identified more AMI patients as high risk compared with using the HEART score alone. This was true not only for STEMI but also for NSTEMI. Consequently, the AMI ratio in the intermediate-risk group for AI-ECG and AI-ECG + HEART was lower than that for the HEART score and GRACE 2.0 score. In summary, AI-ECG demonstrated the potential for early rule-out and early rule-in capabilities.

The comprehensive discrimination performance (AUROC and AUPRC) of the AI-ECG, as well as the results of risk stratification at the pre-specified cut-off, is crucial. This is because the clinical needs addressed by the AI-ECG model can vary depending on the healthcare setting. The current reliance on physician-dependent practices for AMI detection might be less appropriate in regions without 24/7 access to experienced physicians. In those areas, the value of the AI-ECG model could be higher than in our participating institutions. In centres where full-time coverage by emergency medicine specialists is not available, the use of AI-ECG for early rule-in can help ensure that high-risk patients with STEMI and NSTEMI are not missed. Conversely, AI-ECG can also be employed for early rule-out to efficiently allocate resources and alleviate ED overcrowding by quickly identifying low-risk patients.

In patients aged 65 years and older, those with obesity, and those with chronic illnesses, the performance of AI-ECG showed a tendency to decrease. In addition, it decreased in the presence of left ventricular hypertrophy, bundle branch block, atrial fibrillation, and pacemaker rhythm. The variation in AI-ECG performance based on demographics, comorbidities, and ECG characteristics was similar to reports on AI-ECG models predicting left ventricular systolic dysfunction.[35,36] When healthcare professionals use AI-ECG for diagnostic assistance and screening for AMI, they should consider the individual patient's baseline ECG and underlying medical conditions.

We implemented an explainable AI (XAI) framework to incorporate heatmap-based methodologies such as saliency mapping[9] and a generative counterfactual-based ECG XAI that was previously used in research[37] to confirm the AI-ECG's ability to detect ECG alterations associated with the coronary territory implicated in pathologic incidents. The application of XAI will assist physicians and researchers in the future by fostering transparent decision-making and providing new insights.

Future technical advancements aimed at further reducing the missed rate of AMI may be necessary, potentially involving the use of AI-ECG alone or in combination with other clinical scoring systems, such as the HEART score. This study demonstrates that the implementation of a protocol wherein AI-ECG was utilized prior to assessing the HEART score resulted in improved rule-out performance compared with reliance on the HEART score alone. Furthermore, the AI-ECG system has the potential for application in the pre-hospital phase, enabling patients with suspected occlusive AMI to be directly transferred to the angio-room, similar to the protocols for patients with STEMI employed in some developed countries. This approach could enhance the timely intervention for patients at risk of AMI and ultimately improve patient outcomes.

Limitations

This study has limitations to discuss. First, its external validity might be confined to the Republic of Korea, necessitating further validation in diverse populations and healthcare settings through international collaborations. We plan to conduct external validation both domestically (using data from Sejong Hospital and other participating hospitals that were unable to obtain consent) and internationally. A demo page is available for reference (https://aitia-demo.medicalai.com/). Second, this study evaluated short-term outcomes (AMI diagnosis during index ED admission and 30 day MACE), so an assessment of long-term outcomes, including recurrent cardiovascular events and long-term mortality, is essential for a comprehensive understanding of the AI-ECG model's prognostic value. Third, this study did not investigate the clinical impacts of AI-ECG model implementation. Future research should focus on user experience and clinical workflows to evaluate the AI-ECG's influence on decision-making, time efficiency, and patient outcomes. Specifically, we found that the performance of the AI-ECG combined with the HEART score was superior in discrimination compared with conventional tools. Even when used alone, the AI-ECG was comparable with the HEART score, which could significantly reduce the 1–2 h waiting time for results. Additionally, its strong discrimination performance for Type 1 MI and occlusive NSTEMI patients who underwent PCI or CABG was confirmed. These findings will help design future randomized controlled trials using AiTiAMI and evaluate patient safety and outcomes. Fourth, the GRACE score was

originally developed to predict in-hospital and 6-month mortality in patients with acute coronary syndromes,[38] rather than in patients presenting with chest pain in the ED. The use of the GRACE score as a comparator in the primary outcome may be inherently inferior in terms of purpose when compared with AI-ECG to detect index AMI. Nonetheless, we employed the GRACE score in this context, as there was no objective scoring system available for the detection of AMI in the ED other than the HEART score. Moreover, reasoning or causality research regarding high-risk AI is still incomplete. Even with explainable AI-ECG, it remains challenging to intuitively trust or fully grasp the 'black box' nature of AiTiAMI version 1.00.00, which is a key feature of this product. Therefore, we believed that conducting explorative comparisons across various clinical proven risk stratifying tools would help in understanding the characteristics of AI-ECG. Additionally, previous studies[39,40] have applied the GRACE score with similar objectives, providing some precedent for its use in this analysis. Fifth, our dropout criteria and exclusion due to missing or noisy data might have introduced bias, so their effects on result generalizability should be carefully considered.

Conclusions

The ROMIAE study contributes significant evidence supporting the potential of AI-ECG for early AMI detection. The AI-ECG model's superior performance suggests a promising future for AI applications in EDs. Addressing the limitations of this study, validating our results in diverse populations, and ensuring ethical considerations are crucial for the successful integration of AI technologies into routine clinical practice.

Supplementary data

Supplementary data are available at *European Heart Journal* online.

Declarations

Disclosure of Interest

Joon-myoung Kwon is the founder and stakeholder in Medical AI Co., Ltd., a medical artificial intelligence company. Min Sung Lee, Jeong Min Son, Sora Kang, Hak Seung Lee, Min-yeong Kim, Nuri Shin, Jong-Hwan Jang, Yong-Yeon Jo are employees and stakeholders in Medical AI Co., Ltd. Kyuseok Kim is a stakeholder in Medical AI Co., Ltd. No other potential conflict of interest relevant to this article was reported.

Model Transparency and Data Availability

The data will be shared upon reasonable request to the corresponding author for research purposes. Our researchers are committed to ensuring the transparency and credibility of the models developed. We will actively share AI-ECG model (AiTiAMI version 1.00.00) for external validation and further development with the global research community. This process aims to achieve extensive clinical validity across diverse populations and settings, enhancing the model's robustness and applicability for real-world clinical use.

Funding

This work has been supported by a non-conditional grant from Medical AI Co., Ltd. This research was supported by a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (RS-2024-00335935).

Ethical Approval

The institutional review board at each of the 18 hospitals approved the study protocol. All participants provided written informed consent.

Pre-registered Clinical Trial Number

This study was registered on ClinicalTrials.gov: NCT05435391.

Appendix

The ROMIAE study group members are as follows: Jong Eun Park (Department of Emergency Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea), Heajin Chung (Department of Emergency Medicine, Soonchunhyang University Seoul Hospital, Seoul, Republic of Korea), Jin Hee Jeong (Department of Emergency Medicine, Gyeongsang National University College of Medicine, Jinju, Korea), Sung-Jun Park (Department of Emergency Medicine, Korea Uiversity Guro Hospital, Seoul, Korea), Sejoong Ahn (Department of Emergency Medicine, Korea University Ansan Hospital, Ansan, Korea), Haewon Jung (Department of Emergency Medicine, School of Medicine, Kyungpook National University, Daegu, South Korea), Seok Hoon Ko (Department of Emergency Medicine, Kyung Hee University Hospital, Seoul, Korea), Han Sung Choi (Department of Emergency Medicine, Kyung Hee University College of Medicine, Kyung Hee University Hospital, Seoul, Korea), June-sung Kim (Department of Emergency Medicine, Asan Medical Center, Seoul, Korea), Bangshill Rhee (Department of Emergency Medicine, Ajou University School of Medicine, Suwon, Korea), Keon Kim (Department of Emergency Medicine, College of Medicine, Ewha Womans University, Seoul, Korea), So Eun Kim (Department of Emergency Medicine, Research Institute of Clinical Medicine, Jeonbuk National University, Biomedical Research Institute, Jeonbuk National University Hospital, South Korea), Ji Hwan Bu (Department of Emergency Medicine, Jeju National University Hospital), Jae Hwan Kim (Department of Emergency Medicine, College of Medicine, Chung-Ang University, Seoul, Korea), Dong Won Kim (Department of Emergency Medicine, Hallym University Chuncheon Sacred Heart Hospital, Chuncheon, Korea), Yongil Cho (Department of Emergency Medicine, Hanyang University College of Medicine, Seoul, Korea), Yonghee Lee (Department of Emergency Medicine, CHA Bundang Medical Center CHA University School of Medicine), Kihwan Choi (Department of Emergency Medicine, CHA Bundang Medical Center CHA University School of Medicine), Sangchun Choi (Emergency Department, Soonchunhyang University College of Medicine, Gyeonggi-Do, Republic of Korea), Jeong Min Son (Medical AI Co., Ltd, Seoul, Republic of Korea; Digital Healthcare Institute, Sejong Hospital, Republic of Korea), Sora Kang (Medical AI Co., Ltd, Seoul, Republic of Korea; Digital Healthcare Institute, Sejong Hospital, Republic of Korea), Hak Seung Lee (Medical AI Co., Ltd, Seoul, Republic of Korea; Digital Healthcare Institute, Sejong Hospital, Republic of Korea), Min-yeong Kim (Medical AI Co., Ltd, Seoul, Republic of Korea; Digital Healthcare Institute, Sejong Hospital, Republic of Korea), Nuri Shin (Medical AI Co., Ltd, Seoul, Republic of Korea; Digital Healthcare Institute, Sejong Hospital, Republic of Korea), Jong-Hwan Jang (Medical AI Co., Ltd, Seoul, Republic of Korea), and Yong-Yeon Jo (Medical AI Co., Ltd, Seoul, Republic of Korea).

References

1. Roth GA, Johnson C, Abajobir A, Abd-Allah F, Abera SF, Abyu G, *et al.* Global, regional, and national burden of cardiovascular diseases for 10 causes, 1990 to 2015. *J Am Coll Cardiol* 2017;**70**:1–25. https://doi.org/10.1016/j.jacc.2017.04.052
2. Ahn S, Ko E, Ro YS. Acute myocardial infarction diagnosed in emergency departments: a report from the National Emergency Department Information System (NEDIS) of Korea, 2018–2022. *Clin Exp Emerg Med* 2023;**10**:S42–7. https://doi.org/10.15441/ceem.23.140
3. Byrne RA, Rossello X, Coughlan JJ, Barbato E, Berry C, Chieffo A, *et al.* 2023 ESC guidelines for the management of acute coronary syndromes. *Eur Heart J* 2023;**44**:3720–826. https://doi.org/10.1093/eurheartj/ehad191
4. Cook DA, Oh SY, Pusic MV. Accuracy of physicians' electrocardiogram interpretations: a systematic review and meta-analysis. *JAMA Intern Med* 2020;**180**:1461–71. https://doi.org/10.1001/jamainternmed.2020.3989
5. McCabe JM, Armstrong EJ, Ku I, Kulkarni A, Hoffmayer KS, Bhave PD, *et al.* Physician accuracy in interpreting potential ST-segment elevation myocardial infarction electrocardiograms. *J Am Heart Assoc* 2013;**2**:e000268. https://doi.org/10.1161/JAHA.113.000268
6. Turnipseed SD, Trythall WS, Diercks DB, Laurin EG, Kirk JD, Smith DS, *et al.* Frequency of acute coronary syndrome in patients with normal electrocardiogram performed during presence or absence of chest pain. *Acad Emerg Med* 2009;**16**:495–9. https://doi.org/10.1111/j.1553-2712.2009.00420.x
7. Emanuel EJ, Persad G, Upshur R, Thome B, Parker M, Glickman A, *et al.* Fair allocation of scarce medical resources in the time of covid-19. *N Engl J Med* 2020;**382**:2049–55. https://doi.org/10.1056/NEJMsb2005114
8. Chang H, Cha WC. Artificial intelligence decision points in an emergency department. *Clin Exp Emerg Med* 2022;**9**:165–8. https://doi.org/10.15441/ceem.22.366
9. Okada Y, Ning Y, Ong MEH. Explainable artificial intelligence in emergency medicine: an overview. *Clin Exp Emerg Med* 2023;**10**:354–62. https://doi.org/10.15441/ceem.23.145
10. Moulik SK, Kotter N, Fishman EK. Applications of artificial intelligence in the emergency department. *Emerg Radiol* 2020;**27**:355–8. https://doi.org/10.1007/s10140-020-01794-1
11. Tang KJW, Ang CKE, Constantinides T, Rajinikanth V, Acharya UR, Cheong KH. Artificial intelligence and machine learning in emergency medicine. *Biocybern Biomed Eng* 2021;**41**:156–72. https://doi.org/10.1016/j.bbe.2020.12.002
12. Mueller B, Kinoshita T, Peebles A, Graber MA, Lee S. Artificial intelligence and machine learning in emergency medicine: a narrative review. *Acute Med Surg* 2022;**9**:e740. https://doi.org/10.1002/ams2.740
13. Cho Y, Kwon JM, Kim KH, Medina-Inojosa JR, Jeon KH, Cho S, *et al.* Artificial intelligence algorithm for detecting myocardial infarction using six-lead electrocardiography. *Sci Rep* 2020;**10**:20495. https://doi.org/10.1038/s41598-020-77599-6
14. Lee BT, Kwon J-m, Cho J, Bae W, Park H, Seo W-W, *et al.* Usefulness of deep-learning algorithm for detecting acute myocardial infarction using electrocardiogram alone in patients with chest pain at emergency department: DAMI-ECG study. *J Cardiovasc Interv* 2023;**2**:100–12. https://doi.org/10.54912/jci.2022.0028
15. Shin TG, Lee Y, Kim K, Lee MS, Kwon JM, group R. ROMIAE (rule-out acute myocardial infarction using artificial intelligence electrocardiogram analysis) trial study protocol: a prospective multicenter observational study for validation of a deep learning-based 12-lead electrocardiogram analysis model for detecting acute myocardial infarction in patients visiting the emergency department. *Clin Exp Emerg Med* 2023;**10**:438–45. https://doi.org/10.15441/ceem.22.360
16. Collet JP, Thiele H, Barbato E, Barthelemy O, Bauersachs J, Bhatt DL, *et al.* 2020 ESC guidelines for the management of acute coronary syndromes in patients presenting without persistent ST-segment elevation. *Eur Heart J* 2021;**42**:1289–367. https://doi.org/10.1093/eurheartj/ehaa575
17. Ibanez B, James S, Agewall S, Antunes MJ, Bucciarelli-Ducci C, Bueno H, *et al.* 2017 ESC guidelines for the management of acute myocardial infarction in patients presenting with ST-segment elevation: the task force for the management of acute myocardial infarction in patients presenting with ST-segment elevation of the European Society of Cardiology (ESC). *Eur Heart J* 2018;**39**:119–77. https://doi.org/10.1093/eurheartj/ehx393
18. Thygesen K, Alpert JS, Jaffe AS, Chaitman BR, Bax JJ, Morrow DA, *et al.* Fourth universal definition of myocardial infarction (2018). *J Am Coll Cardiol* 2018;**72**:2231–64. https://doi.org/10.1016/j.jacc.2018.08.1038
19. Fawcett T. An introduction to ROC analysis. *Pattern Recognit Lett* 2006;**27**:861–74. https://doi.org/10.1016/j.patrec.2005.10.010
20. Backus BE, Six AJ, Kelder JC, Bosschaert MA, Mast EG, Mosterd A, *et al.* A prospective validation of the HEART score for chest pain patients at the emergency department. *Int J Cardiol* 2013;**168**:2153–8. https://doi.org/10.1016/j.ijcard.2013.01.255
21. Hung J, Roos A, Kadesjo E, McAllister DA, Kimenai DM, Shah ASV, *et al.* Performance of the GRACE 2.0 score in patients with type 1 and type 2 myocardial infarction. *Eur Heart J* 2021;**42**:2552–61. https://doi.org/10.1093/eurheartj/ehaa375
22. Cook NR, Ridker PM. Advances in measuring the effect of individual predictors of cardiovascular risk: the role of reclassification measures. *Ann Intern Med* 2009;**150**:795–802. https://doi.org/10.7326/0003-4819-150-11-200906020-00007
23. Vickers AJ, Cronin AM, Elkin EB, Gonen M. Extensions to decision curve analysis, a novel method for evaluating diagnostic tests, prediction models and molecular markers. *BMC Med Inform Decis Mak* 2008;**8**:53. https://doi.org/10.1186/1472-6947-8-53
24. Davis J, Goadrich M. The relationship between precision-recall and ROC xurves. In: *Proceedings of the 23rd International Conference on Machine Learning, Pittsburgh, PA, USA*: Association for Computing Machinery, 2006, 233–40.
25. Mc NQ. Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika* 1947;**12**:153–7. https://doi.org/10.1007/BF02295996
26. Kosinski AS. A weighted generalized score statistic for comparison of predictive values of diagnostic tests. *Stat Med* 2013;**32**:964–77. https://doi.org/10.1002/sim.5587
27. Than M, Herbert M, Flaws D, Cullen L, Hess E, Hollander JE, *et al.* What is an acceptable risk of major adverse cardiac event in chest pain patients soon after discharge from the emergency department?: a clinical survey. *Int J Cardiol* 2013;**166**:752–4. https://doi.org/10.1016/j.ijcard.2012.09.171
28. Kwok CS, Bennett S, Azam Z, Welsh V, Potluri R, Loke YK, *et al.* Misdiagnosis of acute myocardial infarction: a systematic review of the literature. *Crit Pathw Cardiol* 2021;**20**:155–62. https://doi.org/10.1097/HPC.0000000000000256
29. Twerenbold R, Costabel JP, Nestelberger T, Campos R, Wussler D, Arbucci R, *et al.* Outcome of applying the ESC 0/1-hour algorithm in patients with suspected myocardial infarction. *J Am Coll Cardiol* 2019;**74**:483–94. https://doi.org/10.1016/j.jacc.2019.05.046
30. Lin C, Liu W-T, Chang C-H, Lee C-C, Hsing S-C, Fang W-H, *et al.* Artificial intelligence–powered rapid identification of ST-elevation myocardial infarction via electrocardiogram (ARISE)—a pragmatic randomized controlled trial. *NEJM AI* 2024;**1**:AIoa2400190. https://doi.org/10.1056/AIoa2400190
31. Herman R, Meyers HP, Smith SW, Bertolone DT, Leone A, Bermpeis K, *et al.* International evaluation of an artificial intelligence-powered electrocardiogram model detecting acute coronary occlusion myocardial infarction. *Eur Heart J Digit Health* 2024;**5**:123–33. https://doi.org/10.1093/ehjdh/ztad074
32. Al-Zaiti SS, Martin-Gill C, Zegre-Hemsey JK, Bouzid Z, Faramand Z, Alrawashdeh MO, *et al.* Machine learning for ECG diagnosis and risk stratification of occlusion myocardial infarction. *Nat Med* 2023;**29**:1804–13. https://doi.org/10.1038/s41591-023-02396-3
33. Saito T, Rehmsmeier M. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLoS One* 2015;**10**:e0118432. https://doi.org/10.1371/journal.pone.0118432
34. Pinker E. Reporting accuracy of rare event classifiers. *NPJ Digit Med* 2018;**1**:56. https://doi.org/10.1038/s41746-018-0062-0
35. Sangha V, Nargesi AA, Dhingra LS, Khunte A, Mortazavi BJ, Ribeiro AH, *et al.* Detection of left ventricular systolic dysfunction from electrocardiographic images. *Circulation* 2023;**148**:765–77. https://doi.org/10.1161/CIRCULATIONAHA.122.062646
36. Harmon DM, Carter RE, Cohen-Shelly M, Svatikova A, Adedinsewo DA, Noseworthy PA, *et al.* Real-world performance, long-term efficacy, and absence of bias in the artificial intelligence enhanced electrocardiogram to detect left ventricular systolic dysfunction. *Eur Heart J Digit Health* 2022;**3**:238–44. https://doi.org/10.1093/ehjdh/ztac028
37. Jeon KH, Jang JH, Kang S, Lee HS, Lee MS, Son JM, *et al.* Identifying atrial fibrillation with Sinus rhythm electrocardiogram in embolic stroke of undetermined source: a validation study with insertable cardiac monitors. *Korean Circ J* 2023;**53**:758–71. https://doi.org/10.4070/kcj.2023.0009
38. Granger CB, Goldberg RJ, Dabbous O, Pieper KS, Eagle KA, Cannon CP, *et al.* Predictors of hospital mortality in the global registry of acute coronary events. *Arch Intern Med* 2003;**163**:2345–53. https://doi.org/10.1001/archinte.163.19.2345
39. Zheng W, Wang G, Ma J, Wu S, Zhang H, Zheng J, *et al.* Evaluation and comparison of six GRACE models for the stratification of undifferentiated chest pain in the emergency department. *BMC Cardiovasc Disord* 2020;**20**:199. https://doi.org/10.1186/s12872-020-01476-3
40. Huang Z, Wang K, Yang D, Gu Q, Wei Q, Yang Z, *et al.* The predictive value of the HEART and GRACE scores for major adverse cardiac events in patients with acute chest pain. *Intern Emerg Med* 2021;**16**:193–200. https://doi.org/10.1007/s11739-020-02378-0