

ORIGINAL RESEARCH

ISCHEMIC HEART DISEASE

Sex-Consistent Performance of an AI-Enabled ECG for Acute Myocardial Infarction

The ROMIAE Study



Hak Seung Lee, MD,^{a,b,*} Sora Kang, MS,^{a,b,*} Joon-myoung Kwon, MD, MS,^a Tae Gun Shin, MD, PhD,^c Youngjoo Lee, MD, PhD,^d Dong Hoon Kim, MD,^e Sung Hyuk Choi, MD, PhD,^f Hanjin Cho, MD, PhD,^g Mi Jin Lee, MD, PhD,^h Ki Young Jeong, MD, PhD,ⁱ Won Young Kim, MD, PhD,^j Young Gi Min, MD,^k Chul Han, MD, PhD,^l Jae Chol Yoon, MD, PhD,^m Eugene Jung, MD, PhD,ⁿ Woo Jeong Kim, MD, PhD,^o Chiwon Ahn, MD, PhD,^p Jeong Yeol Seo, MD, PhD,^q Tae Ho Lim, MD, PhD,^r Jae Seong Kim, MD,^s Jeong Min Son, MD,^{a,b} Kyung Su Kim, MD, PhD,^{a,b} Kyuseok Kim, MD, PhD,^t Min Sung Lee, MD,^{a,b} the ROMIAE Study Group

ABSTRACT

BACKGROUND Women with suspected acute myocardial infarction (AMI) are at increased risk of delayed or missed diagnosis. Artificial intelligence-enabled electrocardiogram (AI-ECG) may support earlier AMI detection, but prospective evidence for sex-consistent performance and rule-out safety is limited.

OBJECTIVES The purpose of this study was to evaluate sex-stratified diagnostic performance, rule-out safety, and phenotype robustness of an AI-ECG for AMI in a prospective, multicenter emergency department cohort.

METHODS ROMIAE (Rule-Out Acute Myocardial Infarction Using Artificial Intelligence Electrocardiogram Analysis) was a prospective external validation study at 18 centers in South Korea. Adults with suspected AMI were enrolled. AI-ECG analyzed initial ECGs using prespecified risk thresholds. Diagnostic performance was assessed using the area under the receiver-operating characteristic curve. Rule-out safety was evaluated by sensitivity, negative predictive value, and missed AMI rate at the low-risk cutoff. All analyses were sex-stratified, with prespecified subgroup analyses.

RESULTS Among 8,493 patients, 3,186 were women. AI-ECG demonstrated comparable discrimination in women and men (area under the receiver-operating characteristic curve 0.875 [95% CI: 0.853-0.896] vs 0.871 [95% CI: 0.858-0.883]). At the prespecified low-risk cutoff, rule-out sensitivity was 98.8% (95% CI: 97.0-99.5) in women and 99.8% (95% CI: 99.4-100.0) in men, with high negative predictive value (99.2% [95% CI: 97.9-99.7] in women and 99.1% [95% CI: 96.7-99.7] in men) and low missed AMI rates (<1% in both sexes). Performance remained stable across ST-segment elevation and non-ST-segment elevation myocardial infarction, without degradation in women.

CONCLUSIONS In a prospective, multicenter emergency department cohort, AI-ECG demonstrated sex-consistent diagnostic performance and preserved rule-out safety for AMI. Further validation and implementation studies are warranted. (ROMIAE [Rule-Out Acute Myocardial Infarction Using Artificial Intelligence Electrocardiogram Analysis] Trial; [NCT05435391](https://clinicaltrials.gov/ct2/show/study/NCT05435391)) (JACC Adv. 2026;5:102813) © 2026 The Authors. Published by Elsevier on behalf of the American College of Cardiology Foundation. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

**ABBREVIATIONS
AND ACRONYMS****AI-ECG** = artificial
intelligence-enabled
electrocardiogram**AMI** = acute myocardial
infarction**AUROC** = area under the
receiver-operating
characteristic curve**ECG** = electrocardiogram**ED** = emergency department**NPV** = negative predictive
value**NSTEMI** = non-ST-segment
elevation myocardial infarction**PPV** = positive predictive value**STEMI** = ST-segment elevation
myocardial infarction

Acute myocardial infarction (AMI) remains a leading cause of morbidity and mortality worldwide.¹ Despite advances in diagnostic strategies, women remain at increased risk of delayed or missed AMI diagnoses, which has been associated with worse clinical outcomes.^{2,3} Compared with men, women more frequently present with atypical symptoms, non-ST-elevation phenotypes, and less overt electrocardiographic abnormalities at initial evaluation, particularly on early electrocardiogram (ECG) assessment, factors that increase vulnerability to underrecognition in acute care settings.

Current diagnostic pathways for suspected AMI rely largely on physician interpretation, rule-based clinical risk scores, and biomarker testing.^{4,5} Although effective in many clinical contexts, these approaches may be influenced by cognitive bias and delays inherent to biomarker-based confirmation. Such limitations may disproportionately affect populations already at risk of underdiagnosis, highlighting the need for diagnostic tools that are objective and rapidly deployable.^{6,7}

Artificial intelligence (AI)-ECG enables automated AMI risk assessment independent of presenting symptoms or provider judgment.⁸⁻¹¹ In particular, prospective evidence is limited regarding its ability to preserve rule-out safety—such as high sensitivity and negative predictive value (NPV)—and consistent

performance across sexes in clinically challenging phenotypes.

In this prespecified secondary analysis of the ROMIAE (Rule-Out Acute Myocardial Infarction Using Artificial Intelligence Electrocardiogram Analysis) cohort, we evaluated sex-stratified diagnostic performance and rule-out safety of AI-ECG for AMI, extending the findings of the primary ROMIAE publication.¹²

METHODS

STUDY DESIGN AND DATA SOURCE. The ROMIAE study was a prospective, multicenter external validation study designed to evaluate the diagnostic performance of an AI-ECG model for AMI.¹² The study was conducted across 18 emergency departments (EDs) in South Korea between March 2022 and October 2023. The study protocol was approved by the Institutional Review Board at each participating site, and all participants provided written informed consent prior to enrollment. The trial was registered at ClinicalTrials.gov (NCT05435391). The present study represents a prespecified secondary analysis of the ROMIAE cohort focusing on sex-stratified diagnostic performance and subgroup robustness of AI-ECG.¹³

STUDY POPULATION. Adult patients (age ≥18 years) presenting to the ED with clinically suspected AMI were eligible for inclusion. Patients were enrolled if they presented within 24 hours of symptom onset or experienced worsening or recurrent chest pain

From the ^aMedical AI Co., Ltd., Seoul, Republic of Korea; ^bDigital Healthcare Institute, Sejong Hospital, Republic of Korea; ^cDepartment of Emergency Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea; ^dDepartment of Emergency Medicine, Soonchunhyang University Seoul Hospital, Seoul, Republic of Korea; ^eDepartment of Emergency Medicine, Uijeongbu Eulji Medical Center, Uijeongbu, Republic of Korea; ^fDepartment of Emergency Medicine, Korea University Guro Hospital, Seoul, Republic of Korea; ^gDepartment of Emergency Medicine, Korea University Ansan Hospital, Ansan, Republic of Korea; ^hDepartment of Emergency Medicine, St. Vincent's Hospital, College of Medicine, The Catholic University of Korea, Seoul, Republic of Korea; ⁱDepartment of Emergency Medicine, Kyung Hee University College of Medicine, Kyung Hee University Hospital, Seoul, Republic of Korea; ^jDepartment of Emergency Medicine, Asan Medical Center, Seoul, Republic of Korea; ^kDepartment of Emergency Medicine, Ajou University School of Medicine, Suwon, Republic of Korea; ^lDepartment of Emergency Medicine, College of Medicine, Ewha Womans University, Seoul, Republic of Korea; ^mDepartment of Emergency Medicine, Research Institute of Clinical Medicine, Jeonbuk National University, Biomedical Research Institute, Jeonbuk National University Hospital, Jeonju, Republic of Korea; ⁿDepartment of Emergency Medicine, Chonnam National University Medical School, Gwangju, Republic of Korea; ^oCollege of Medicine, Jeju National University, Jeju, Republic of Korea; ^pDepartment of Emergency Medicine, College of Medicine, Chung-Ang University, Seoul, Republic of Korea; ^qDepartment of Emergency Medicine, Hallym University Chuncheon Sacred Heart Hospital, Chuncheon, Republic of Korea; ^rDepartment of Emergency Medicine, Hanyang University College of Medicine, Seoul, Republic of Korea; ^sDepartment of Emergency Medicine, Incheon Sejong Hospital, Incheon, Republic of Korea; and the ^tDepartment of Emergency Medicine, CHA Bundang Medical Center CHA University School of Medicine, Bundang, Republic of Korea. *These authors contributed equally to this work. The authors attest they are in compliance with human studies committees and animal welfare regulations of the authors' institutions and Food and Drug Administration guidelines, including patient consent where appropriate. For more information, visit the [Author Center](#).

Manuscript received January 5, 2026; revised manuscript received April 14, 2026, accepted April 15, 2026.

within 24 hours prior to or after ED arrival. Patients were excluded if they presented with out-of-hospital cardiac arrest, traumatic chest pain, declined participation, or had a clear noncardiac diagnosis such as pneumothorax. Inclusion and exclusion criteria were applied uniformly across sexes. Biological sex was defined according to electronic medical record documentation at the time of presentation.

AI-ECG MODEL. The AI-ECG model evaluated in this study was based on a previously developed deep learning architecture (AiTiAMI, version 1.00.00; Medical AI Co, Ltd). The model processes raw 12-lead ECG signals sampled at 500 Hz and generates a continuous probability score (0-100) representing the likelihood of AMI.¹²

The model was trained using multicenter ECG data sets from institutions not participating in the present validation cohort, ensuring complete separation between training and validation data. No clinical variables, laboratory values, or demographic characteristics—including sex—were incorporated as explicit model inputs. Detailed descriptions of the model architecture and training procedures have been reported previously in the original ROMIAE study.¹² Prespecified thresholds were used to categorize patients into low-, intermediate-, and high-risk groups. A score <3.0 defined the low-risk category, selected a priori to achieve approximately 99% sensitivity, while a score ≥ 48.5 defined the high-risk category, corresponding to approximately 90% specificity in the development data set.^{14,15}

DATA COLLECTION AND OUTCOME ADJUDICATION. Clinical care was not influenced by study participation. All participating centers followed contemporary international guidelines for the evaluation of suspected AMI using high-sensitivity troponin-based diagnostic algorithms.^{4,5} Prospectively collected data included the initial 12-lead ECG, high-sensitivity troponin levels, presenting symptoms, and physician-estimated clinical probability of AMI after initial evaluation (Physician AMI score, 0-10 scale). Follow-up ECGs and serial biomarker testing were performed at the treating team's discretion.

The primary outcome was a diagnosis of AMI during the index hospitalization, including type 1 and type 2 MI, according to the Fourth Universal Definition of Myocardial Infarction.¹⁵ Outcome adjudication was performed by 2 board-certified emergency physicians at each site, with a third adjudicator resolving disagreements. Adjudicators were blinded to AI-ECG results.

STATISTICAL ANALYSIS

Baseline characteristics were summarized using descriptive statistics. Continuous variables were reported as medians with IQRs and compared between groups using the Mann-Whitney *U* test. Categorical variables were reported as frequencies and percentages and compared using the chi-square test or Fisher exact test, as appropriate.

The diagnostic performance of the AI-ECG model was evaluated for the primary outcome of AMI diagnosis. All analyses were performed using prospectively collected data according to prespecified analytic plans.¹³ Discrimination was assessed using the area under the receiver-operating characteristic curve (AUROC) and the area under the precision-recall curve. Sensitivity, specificity, positive predictive value (PPV), and NPV were calculated at prespecified thresholds or optimal cutoffs determined by Youden's *J* statistic. CIs for all performance metrics were estimated using bootstrap resampling. Rule-out safety at the prespecified low-risk cutoff was assessed using sensitivity, NPV, and the missed AMI rate, defined as the proportion of patients with a final diagnosis of AMI among those classified as low risk by AI-ECG.

To contextualize AI-ECG performance within established diagnostic strategies, prespecified comparative analyses were performed against commonly used clinical assessment tools. These included the Physician AMI score, History, ECG, Age, Risk factors, and Troponin (HEART) score, GRACE (Global Registry of Acute Coronary Events) 2.0 score, and initial high-sensitivity troponin levels, reflecting routine ED practice.^{16,17} For ST-segment elevation myocardial infarction (STEMI), where early diagnostic decisions are primarily driven by electrocardiographic interpretation and physician assessment, AI-ECG performance was compared directly with the Physician AMI score. For non-STEMI (NSTEMI), a diagnostically more heterogeneous phenotype, AI-ECG performance was compared with the Physician AMI score, HEART score, GRACE score, and initial high-sensitivity troponin levels. Comparisons of AUROC between the AI-ECG model and each competing diagnostic tool were performed using DeLong's test for correlated receiver-operating characteristic curves, applied separately within each sex and MI phenotype subgroup. Comparisons of other performance metrics, including sensitivity, specificity, PPV, and NPV, were conducted using bootstrap resampling with 1,000 iterations.

TABLE 1 Baseline Characteristics by Sex

	ROMIAE		P Value
	Women (n = 3,186)	Men (n = 5,307)	
Demographics			
Age, y	64.0 (52.0-75.0)	61.0 (50.0-70.0)	<0.001
Past history			
Hypertension	1,456 (45.7)	2,467 (46.5)	0.496
Diabetes mellitus	740 (23.2)	1,377 (25.9)	0.005
Hyperlipidemia	931 (29.2)	1,382 (26.0)	0.002
Coronary artery disease	576 (18.1)	1,186 (22.3)	<0.001
Acute myocardial infarction	198 (6.2)	628 (11.8)	<0.001
Congestive heart failure	146 (4.6)	156 (2.9)	<0.001
Chronic kidney disease	201 (6.3)	347 (6.5)	0.710
ER chief complaints			<0.001
Typical chest pain	946 (29.7)	2,192 (41.3)	
Atypical chest pain	1,118 (35.1)	1,632 (30.8)	
Noncardiac chest pain	318 (10.0)	441 (8.3)	
Other symptoms	804 (25.2)	1,042 (19.6)	
Chest pain onset ≤ 3 h	1,596 (50.1)	2,760 (52.0)	0.092
Physician ECG score	2.0 (2.0-4.0)	3.0 (2.0-6.0)	<0.001
HEART score	4.0 (3.0-5.0)	5.0 (3.0-6.0)	<0.001
GRACE score	84.0 (60.0-111.0)	84.0 (61.0-107.0)	0.202
AI-ECG score	12.1 (4.5-31.8)	22.7 (9.9-50.4)	<0.001
Type 1	65.2 (39.6-86.2)	77.0 (46.6-95.0)	<0.001
Type 2	20.1 (11.2-66.9)	47.3 (16.1-72.4)	0.287
STEMI	85.3 (67.6-97.1)	95.7 (79.1-98.0)	<0.001
NSTEMI	53.8 (26.5-72.6)	54.6 (27.6-78.5)	0.168
AMI type			<0.001
Non-AMI	2,851 (89.5)	4,056 (76.4)	
Type 1	312 (9.8)	1,192 (22.5)	
Type 2	23 (0.7)	59 (1.1)	
ACS type			<0.001
Non-ACS	2,685 (84.3)	3,646 (68.7)	
Unstable angina	166 (5.2)	410 (7.7)	
STEMI	104 (3.3)	538 (10.1)	
NSTEMI	231 (7.3)	713 (13.4)	

Values are median (IQR) or n (%).

ACS = acute coronary syndrome; AI-ECG = artificial intelligence-enabled electrocardiogram; AMI = acute myocardial infarction; ER = emergency room; HEART = History, ECG, Age, Risk factors, and Troponin; GRACE = Global Registry of Acute Coronary Events; NSTEMI = non-ST-segment elevation myocardial infarction; ROMIAE = Rule-Out Acute Myocardial Infarction Using Artificial Intelligence Electrocardiogram Analysis; STEMI = ST-segment elevation myocardial infarction.

To assess diagnostic equity and robustness, all primary analyses were conducted in a sex-stratified manner. Prespecified subgroup analyses were performed across MI phenotype, symptom type, ECG rhythm, MI subtype, and culprit coronary artery, as detailed in the study protocol. Within each subgroup, diagnostic performance metrics were estimated using identical analytic methods. Comparisons of AUROC between women and men within each prespecified clinical subgroup were performed using DeLong's test. For MI subtype and culprit artery subgroups,

heterogeneity of AUROC across categories was assessed using likelihood ratio tests comparing models with and without subgroup indicator variables.

To formally assess risk prediction performance beyond discrimination, sex-stratified calibration analyses and decision curve analysis were performed. Calibration was evaluated using smoothed calibration curves estimated by locally weighted scatterplot smoothing with 95% CIs. Calibration-in-the-large (intercept) and calibration slope were estimated by fitting a logistic regression model of observed outcomes on the logit-transformed predicted probabilities, with the intercept assessed after fixing the slope to 1.0 and the slope estimated in an unconstrained model. Overall probabilistic accuracy was quantified using the Brier score. Decision curve analysis was performed to assess the net clinical benefit of AI-ECG across a range of threshold probabilities, with comparators including physician-estimated AMI probability, HEART score, GRACE score, and initial high-sensitivity troponin. All calibration and decision curve analyses were conducted sex-stratified.

All statistical analyses were performed using R software (version 4.3.2) and Python (version 3.9.7). Statistical significance was assessed using 2-sided tests with a threshold of $P < 0.05$. Multiple-comparison adjustment was not performed because subgroup analyses were intended to evaluate performance robustness rather than formal hypothesis testing. This study was reported in accordance with the European Heart Rhythm Association AI reporting checklist for AI-based medical research and the TRIPOD + AI (Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis + Artificial Intelligence) reporting guidelines¹⁸ (Supplemental Tables 1 and 2).

RESULTS

STUDY POPULATION AND BASELINE CHARACTERISTICS BY SEX.

Baseline demographic and clinical characteristics stratified by sex are summarized in Table 1. Women were older at presentation and more frequently reported atypical chest pain compared with men (35.1% [1,118/3,186] vs 30.8% [1,632/5,307]). Men more commonly had diabetes mellitus, dyslipidemia, coronary artery disease, and prior MI, whereas congestive heart failure was more prevalent among women. The index AMI diagnosis was less prevalent among women than men (10.5% [335/3,186] vs 23.6% [1,251/5,307]). Among AMI cases, STEMI was more common in men, whereas NSTEMI accounted

for a relatively larger proportion of cases among women.

SEX-STRATIFIED DISCRIMINATION AND RISK STRATIFICATION PERFORMANCE OF AI-ECG FOR AMI. AI-ECG demonstrated comparable discrimination performance for AMI in women and men (Table 2, Figure 1). The AUROC was 0.875 (95% CI: 0.853-0.896) in women and 0.871 (95% CI: 0.858-0.883) in men, with no significant difference between groups (DeLong's test, $P = 0.71$). The area under the precision-recall curve was 0.593 (95% CI: 0.542-0.646) in women and 0.763 (95% CI: 0.741-0.784) in men.

Beyond discrimination, AI-ECG provided effective risk stratification across predefined risk categories (Figure 2, Central Illustration). In women, the prevalence of AMI increased stepwise across AI-ECG risk categories (0.8%, 4.8%, and 47.5% in the low-, intermediate-, and high-risk groups, respectively). A similar enrichment pattern was observed in men (0.9%, 9.4%, and 64.7%). Among AMI-positive cases, NSTEMI accounted for a larger proportion of total AMI in women than in men in both the intermediate- (93.5% vs 88.4%) and high-risk (56.5% vs 44.8%) categories.

To formally assess risk prediction performance beyond discrimination, calibration analyses and decision curve analysis were performed sex-stratified. Calibration results and decision curve analysis are presented in Supplemental Figures 1 and 2.

RULE-OUT SAFETY OF AI-ECG AT THE PRESPECIFIED LOW-RISK CUTOFF. As shown in Table 2, at the prespecified low-risk cutoff, AI-ECG demonstrated preserved rule-out safety in both women and men. Sensitivity was 98.8% (95% CI: 97.0-99.5) in women and 99.8% (95% CI: 99.4-100.0) in men, indicating a low false-negative rate. NPV was similarly high in both sexes, at 99.2% (95% CI: 97.9-99.7) in women and 99.1% (95% CI: 96.7-99.7) in men. In addition, at the same low-risk cutoff, AI-ECG classified a larger proportion of women as low risk compared with men (15.1% [95% CI: 13.9%-16.4%; 480/3186] vs 4.1% [95% CI: 3.6%-4.7%; 217/5307]).

The miss rate of AI-ECG was <1% in both women and men, occurring in 0.83% (95% CI: 0.23-2.12%; 4/480) of women and 0.92% (95% CI: 0.11-3.29%; 2/217) of men. These results demonstrate rule-out safety of AI-ECG at the prespecified low-risk threshold, supporting its potential utility as an early triage adjunct (Figure 3).

ROBUSTNESS OF AI-ECG PERFORMANCE ACROSS PRESPECIFIED CLINICAL SUBGROUPS

The robustness of AI-ECG performance was evaluated across AMI phenotypes, with a particular focus on

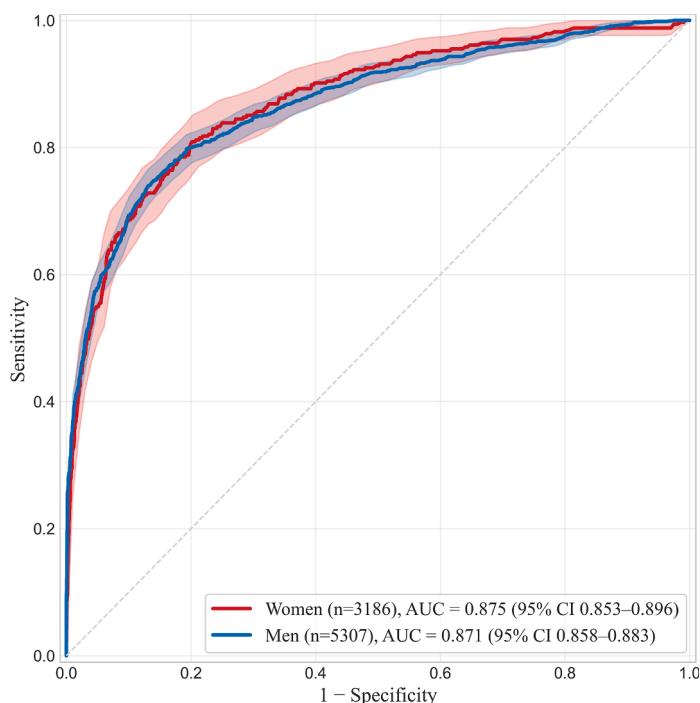
TABLE 2 Sex-Stratified Diagnostic Performance of AI-ECG

	Women (n = 3,186)	Men (n = 5,307)
AMI events, n (%)	335 (10.5%)	1,251 (23.6%)
Discrimination		
AUROC (95% CI)	0.875 (0.853-0.896)	0.871 (0.858-0.883)
AUPRC (95% CI)	0.593 (0.542-0.646)	0.763 (0.741-0.784)
Rule-out safety (low-risk cutoff) ^a		
Sensitivity (%)	98.8% (97.0-99.5)	99.8% (99.4-100.0)
Negative predictive value (%)	99.2% (97.9-99.7)	99.1% (96.7-99.7)
Missed AMI rate (%) ^b	0.83% (4/480)	0.92% (2/217)
Rule-in context (high-risk cutoff) ^a		
Specificity (%)	91.3% (90.3-92.3)	87.9% (86.9-88.8)
Positive predictive value (%)	47.6% (43.3-52.1)	64.8% (62.4-67.2)

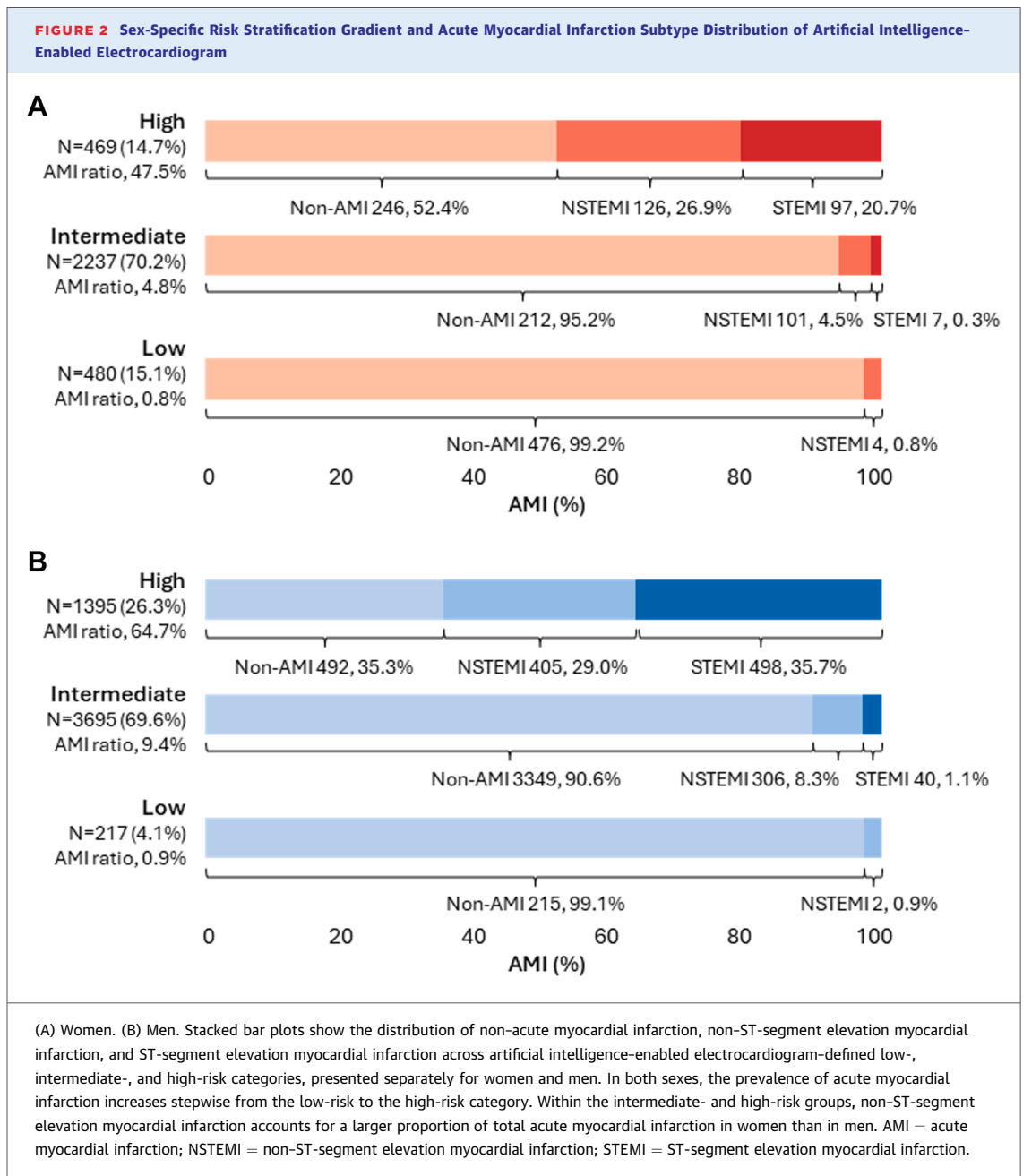
Performance metrics were calculated at the prespecified low-risk cutoff defined in the study protocol. ^aThe prespecified low- and high-risk cutoffs (<3.0 and ≥48.5) were defined a priori during model development to achieve approximately 99% sensitivity and 90% specificity, respectively, in the development data set. ^bMissed AMI rate was defined as the proportion of patients with a final diagnosis of acute myocardial infarction among those classified as low risk by AI-ECG. The denominator represents the total number of patients classified as low risk within each sex.

AUPRC = area under the precision-recall curve; AUROC = area under the receiver operating characteristic curve; other abbreviations as in Table 1.

FIGURE 1 Sex-Stratified Discrimination Performance of Artificial Intelligence-Enabled Electrocardiogram for Acute Myocardial Infarction



Receiver-operating characteristic curves with 95% CIs comparing the discrimination performance of artificial intelligence-enabled electrocardiogram for acute myocardial infarction in women and men. AUC = area under the receiver operating characteristic curve.

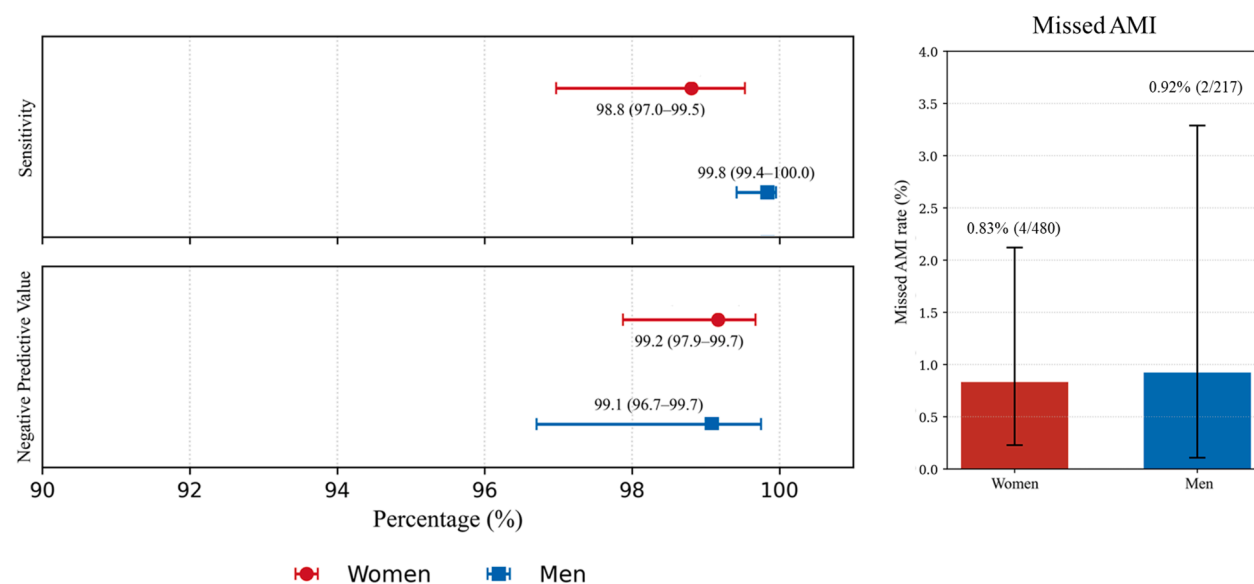


differences between STEMI and NSTEMI, as summarized in [Table 3](#). In patients with STEMI, AI-ECG demonstrated high discrimination performance and preserved safety metrics and showed numerically comparable or higher performance across multiple metrics compared with the physician-estimated clinical probability of AMI. In contrast, within patients with NSTEMI, AI-ECG maintained stable discrimination and demonstrated overall comparable performance relative to the HEART score, physician-

estimated probability of AMI, and initial high-sensitivity troponin I or T. No systematic degradation of AI-ECG performance was observed across discrimination, sensitivity, specificity, or NPV.

Beyond MI phenotype, AI-ECG performance remained generally consistent across prespecified clinical subgroups, including demographic characteristics, symptom presentation, electrocardiographic rhythm, MI subtype, and culprit coronary artery, as shown in [Supplemental Tables 3 and 4](#).

FIGURE 3 Rule-Out Safety of Artificial Intelligence-Enabled Electrocardiogram by Sex at the Prespecified Low-Risk Cutoff



The figure shows sensitivity and negative predictive value for women and men at the prespecified low-risk cutoff. Points indicate point estimates and horizontal lines represent 95% CIs. The right panel shows the missed acute myocardial infarction rate by sex, with bars representing point estimates and error bars indicating 95% CIs. Missed acute myocardial infarction rates were below 1% in both women and men, indicating preserved rule-out safety across sexes. Abbreviation as in Figure 2.

Although numerical differences were observed in selected subgroups, overall discrimination remained stable across most subgroups in both sexes. Heterogeneity analyses demonstrated statistically significant differences across selected MI phenotypes and culprit artery subgroups, with a general trend toward lower discrimination in more complex or non-ST-elevation presentations. Additionally, performance was attenuated in patients with prior coronary artery disease, prior AMI, or congestive heart failure, consistent with the known impact of underlying cardiac disease on ECG-based interpretation.

DISCUSSION

In this prespecified secondary analysis of the prospective multicenter ROMIAE cohort, we demonstrate that AI-ECG provides sex-consistent diagnostic performance for AMI, extending the findings of the primary ROMIAE study. AI-ECG showed comparable discrimination in women and men and preserved rule-out safety in both sexes. In addition, diagnostic performance remained robust across myocardial infarction phenotypes, including STEMI and NSTEMI, supporting a phenotype-agnostic diagnostic capability. Together, these findings provide prospective

evidence that AI-ECG may support stable diagnostic performance across sexes and could enable more equitable early triage in patients presenting with suspected AMI.

SEX-CONSISTENT AND SAFE DIAGNOSTIC PERFORMANCE. Women with suspected AMI have historically faced structural challenges in timely and accurate diagnosis, driven by higher rates of atypical symptom presentation and less overt ECG findings.^{2,19,20} Diagnostic pathways relying heavily on symptom interpretation, clinical risk scores, and biomarker confirmation may contribute to delayed recognition in women. In this context, our findings demonstrate that AI-ECG provides comparable discrimination and safety in women and men, with preserved rule-out safety and similarly low missed AMI rates. In women, this was accompanied by meaningful rule-out efficiency, allowing a larger proportion to be safely classified as low risk at presentation. Rule-in performance showed sex-related numerical differences, which can be interpreted in the context of differing AMI prevalence.

Notably, the ROMIAE cohort comprised a higher proportion of male patients (62%), with a correspondingly higher AMI prevalence in men. This sex-related difference in prevalence is the primary

TABLE 3 Performance for STEMI and NSTEMI

	AUROC (95% CI)	P Value	Sensitivity (95% CI)	P Value	Specificity (95% CI)	P Value
Women						
STEMI						
AI-ECG	0.973 (0.960-0.983)	–	0.931 (0.877-0.974)	–	0.914 (0.904-0.924)	–
Physician AMI score	0.971 (0.949-0.986)	0.834	0.874 (0.804-0.938)	0.157	0.969 (0.963-0.976)	<0.001
NSTEMI						
AI-ECG	0.831 (0.801-0.859)	–	0.547 (0.480-0.609)	–	0.914 (0.904-0.924)	–
HEART score	0.880 (0.858-0.898)	0.001	0.568 (0.502-0.628)	0.622	0.927 (0.918-0.936)	0.037
Physician AMI score	0.781 (0.748-0.814)	0.01	0.192 (0.141-0.244)	<0.001	0.969 (0.963-0.975)	<0.001
hs-Troponin I or T	0.851 (0.827-0.873)	0.231	0.857 (0.810-0.900)	<0.001	0.844 (0.831-0.857)	<0.001
GRACE score	0.755 (0.728-0.782)	<0.001	0.200 (0.148-0.250)	<0.001	0.931 (0.921-0.940)	0.007
Men						
STEMI						
AI-ECG	0.968 (0.960-0.975)	–	0.926 (0.904-0.947)	–	0.879 (0.869-0.888)	–
Physician AMI score	0.960 (0.950-0.969)	0.159	0.862 (0.832-0.890)	<0.001	0.945 (0.938-0.952)	<0.001
NSTEMI						
AI-ECG	0.797 (0.778-0.815)	–	0.568 (0.531-0.606)	–	0.879 (0.868-0.889)	–
HEART score	0.835 (0.821-0.850)	<0.001	0.483 (0.447-0.519)	<0.001	0.922 (0.913-0.929)	<0.001
Physician AMI score	0.751 (0.731-0.769)	<0.001	0.271 (0.238-0.303)	<0.001	0.945 (0.938-0.952)	<0.001
hs-Troponin I or T	0.804 (0.787-0.819)	0.545	0.772 (0.740-0.802)	<0.001	0.835 (0.825-0.846)	<0.001
GRACE score	0.667 (0.646-0.688)	<0.001	0.098 (0.078-0.121)	<0.001	0.959 (0.953-0.965)	<0.001

TABLE 3 Continued

	PPV (95% CI)	P Value	NPV (95% CI)	P Value	AUPRC (95% CI)	P Value
Women						
STEMI						
AI-ECG	0.281 (0.233-0.330)	–	0.997 (0.995-0.999)	–	0.691 (0.596-0.777)	–
Physician AMI score	0.506 (0.430-0.585)	<0.001	0.995 (0.992-0.998)	0.198	0.753 (0.677-0.830)	0.654
NSTEMI						
AI-ECG	0.339 (0.290-0.389)	–	0.961 (0.954-0.968)	–	0.372 (0.311-0.435)	–
HEART score	0.386 (0.333-0.436)	0.067	0.964 (0.956-0.970)	0.524	0.387 (0.324-0.455)	0.840
Physician AMI score	0.334 (0.252-0.415)	0.892	0.937 (0.928-0.945)	<0.001	0.264 (0.210-0.318)	<0.001
hs-Troponin I or T	0.308 (0.271-0.343)	0.130	0.987 (0.982-0.991)	<0.001	0.588 (0.555-0.619)	<0.001
GRACE score	0.189 (0.141-0.235)	<0.001	0.935 (0.926-0.944)	<0.001	0.196 (0.152-0.240)	<0.001
Men						
STEMI						
AI-ECG	0.503 (0.470-0.535)	–	0.989 (0.985-0.992)	–	0.871 (0.845-0.893)	–
Physician AMI score	0.675 (0.640-0.710)	<0.001	0.981 (0.977-0.985)	0.001	0.845 (0.816-0.871)	0.003
NSTEMI						
AI-ECG	0.451 (0.417-0.485)	–	0.921 (0.912-0.929)	–	0.469 (0.426-0.510)	–
HEART score	0.519 (0.481-0.557)	<0.001	0.910 (0.902-0.919)	0.014	0.511 (0.471-0.550)	0.252
Physician AMI score	0.465 (0.417-0.509)	0.602	0.881 (0.871-0.891)	<0.001	0.373 (0.338-0.411)	<0.001
hs-Troponin I or T	0.452 (0.424-0.479)	0.955	0.954 (0.947-0.961)	<0.001	0.629 (0.605-0.650)	<0.001
GRACE score	0.297 (0.238-0.355)	<0.001	0.858 (0.848-0.869)	<0.001	0.242 (0.217-0.269)	<0.001

Threshold-dependent metrics were calculated at predefined high-risk cutoffs: AI-ECG at the prespecified cutoff; HEART score, physician AMI score, and GRACE score at the high-risk cutoffs used in the ROMIAE study; and hs-troponin at the 99th-percentile URL. Performance metrics are shown for completeness; interpretation focuses on discrimination and rule-out safety. *P* values indicate comparisons with AI-ECG within the same sex and myocardial infarction phenotype.

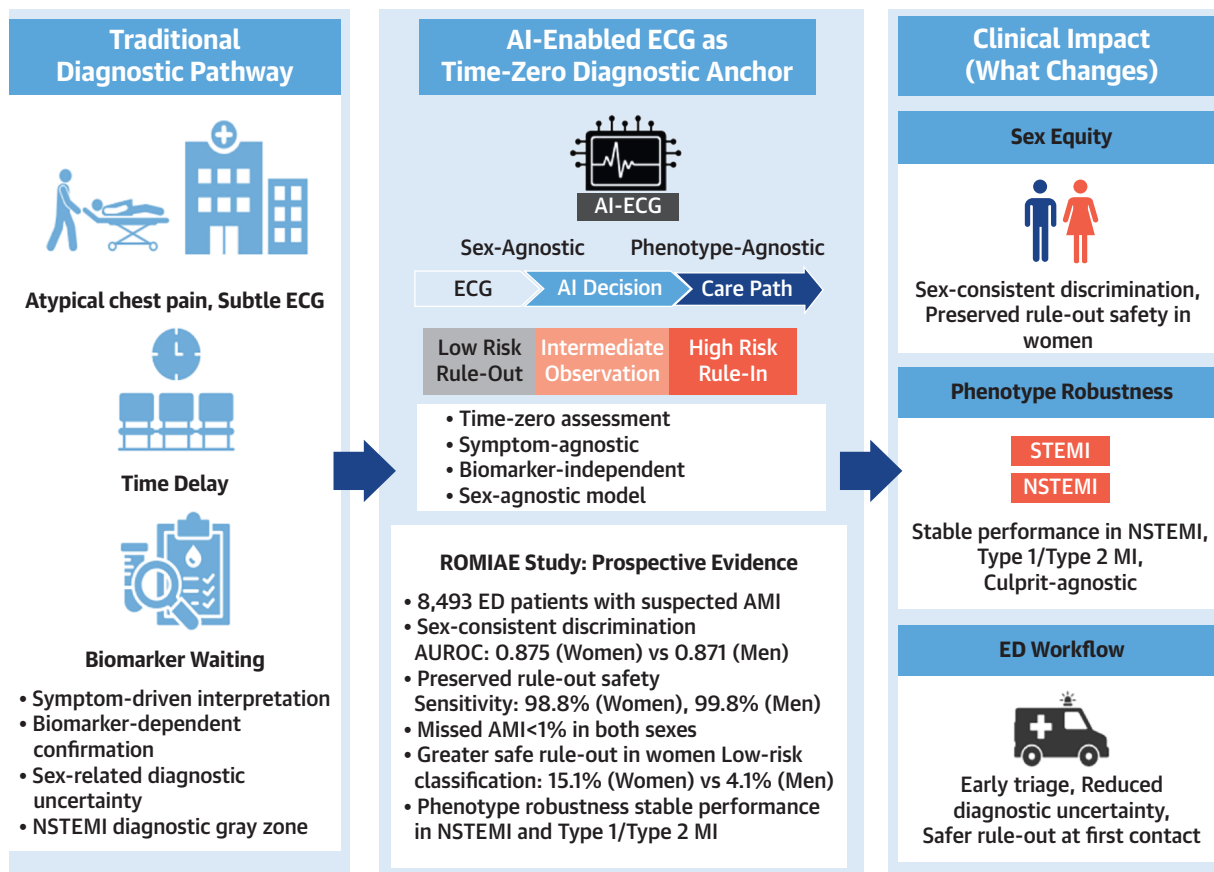
NPV = negative predictive value; PPV = positive predictive value; other abbreviations as in [Tables 1 and 2](#).

driver of the observed PPV disparity between sexes, as PPV is inherently dependent on disease prevalence rather than model performance per se. As a result, the safety estimates derived in women—including sensitivity and NPV—should be interpreted with this context in mind, as they should be interpreted in the

context of the cohort's sex composition, as performance metrics are inherently influenced by AMI prevalence within the study population.

From a workflow perspective, these results suggest that AI-ECG may serve as an early triage tool in the ED, enabling risk stratification based on ECG alone at an

CENTRAL ILLUSTRATION Artificial Intelligence-Enabled Electrocardiogram for Early Triage of Suspected Acute Myocardial Infarction



Lee HS, et al. JACC Adv. 2026;5(6):102813.

This schematic contrasts the conventional diagnostic pathway for suspected acute myocardial infarction with an artificial intelligence-enabled electrocardiogram-based triage approach. Traditional evaluation relies on symptom interpretation, clinician assessment, and serial cardiac biomarker testing, which can introduce diagnostic delay and uncertainty—particularly in women, who more often present with atypical symptoms or less overt electrocardiographic findings. In contrast, artificial intelligence-enabled electrocardiogram provides an immediate, objective, time-zero assessment based solely on electrocardiographic signals at the time of presentation, independent of symptom interpretation or biomarker availability. In the prospective multicenter ROMIAE cohort, artificial intelligence-enabled electrocardiogram demonstrated sex-consistent discrimination and preserved rule-out safety, with low missed acute myocardial infarction rates at the prespecified low-risk threshold. Performance remained stable across myocardial infarction phenotypes, including ST-segment elevation myocardial infarction and non-ST-segment elevation myocardial infarction. These findings support artificial intelligence-enabled electrocardiogram as a phenotype-agnostic early triage adjunct that may improve diagnostic efficiency and equity at first clinical contact. AI-ECG = artificial intelligence-enabled electrocardiogram; AUROC = area under the receiver operating characteristic curve; ED = emergency department; MI = myocardial infarction; other abbreviations as in Figure 2.

early stage of evaluation.⁸ By operating independently of biomarker availability, AI-ECG has the potential to support safe rule-out decisions before laboratory confirmation, which may be particularly relevant in women presenting with suspected AMI.²¹

When evaluating rule-out strategies in the ED, the missed AMI rate is an important safety metric. In the present study, AI-ECG demonstrated a missed AMI rate below 1% in both sexes at the prespecified low-risk threshold. However, certain presentations—

such as ischemic injury without overt biomarker elevation, including mental stress-induced myocardial ischemia, which has been reported to be more prevalent in women—may not be reliably captured by ECG-based assessment alone.²² These findings underscore that AI-ECG should be interpreted as a supportive early triage tool, and that clinical judgment and serial biomarker assessment remain essential. Furthermore, the relatively low prevalence of type 2 MI in this cohort, representing

approximately 5% of AMI cases, should be considered when interpreting safety estimates, as the observed results are largely driven by type 1 MI events.

ROBUSTNESS ACROSS PHENOTYPES IN THE CONTEXT OF PRIOR DIAGNOSTIC STRATEGIES. NSTEMI represents a diagnostically challenging phenotype, especially in women, given their more frequent atypical presentations and less overt ECG findings.^{19,23,24} Importantly, AI-ECG performance in women with NSTEMI remained clinically robust, with meaningful discrimination and high NPV. In addition, AI-ECG performance remained consistent across MI subtypes and symptom presentation patterns, including typical vs atypical symptoms and timing of symptom onset, as well as across different culprit coronary arteries.

High-sensitivity troponin-based algorithms and clinical risk scores such as GRACE and HEART have been shown to exhibit sex-related classification differences.²⁵⁻²⁷ In contrast, AI-ECG in the present study operated independently of biomarker thresholds and clinical scoring systems, while maintaining consistent performance across sexes and MI phenotypes. Notably, this sex-consistent performance was achieved without sex-specific model optimization, suggesting that objective ECG-based assessment may reduce interpretive variability inherent to symptom-driven diagnostic pathways.

CLINICAL IMPLICATIONS. The present analysis extends the primary ROMIAE study by providing several clinically relevant insights. First, AI-ECG preserved rule-out safety specifically in women, with a sensitivity of 98.8% and a missed AMI rate below 1%, supporting its safety as an early triage adjunct in a population historically at risk for underdiagnosis. Second, AI-ECG classified a substantially larger proportion of women as low risk compared with men (15.1% vs 4.1%), suggesting potential workflow efficiency gains in this population. Third, diagnostic performance remained robust across MI phenotypes, including NSTEMI—a clinically challenging presentation disproportionately affecting women—without systematic degradation.

Despite these findings, AI-ECG should be interpreted as a supportive triage tool requiring integration with the full clinical context, not as a standalone diagnostic criterion. Importantly, ischemic presentations from nonatherosclerotic etiologies—including myocardial infarction with nonobstructive coronary arteries, spontaneous coronary artery dissection, and vasospastic angina—may not be reliably captured by ECG-based algorithms, as these entities frequently lack

characteristic electrocardiographic abnormalities.²⁸ Women are disproportionately represented among these etiologies, and clinicians should maintain a high index of suspicion for such phenotypes, even when AI-ECG risk scores are low.

Acknowledging these contextual limitations, several prior studies have demonstrated that AI-ECG can meaningfully augment acute coronary syndrome workflows. The ARISE (Artificial Intelligence-Powered Rapid Identification of ST-Elevation Myocardial Infarction via electrocardiogram) trial evaluated AI-ECG as a real-time clinical decision support system for STEMI and showed that AI-ECG-assisted triage reduced treatment delays.²⁹ In parallel, occlusion MI-focused approaches have emphasized detecting acute coronary occlusion among patients without classic ST elevation and have highlighted the value of ECG-only risk scores for both rule-in and rule-out classification and workflow reclassification.⁹⁻¹¹

When considered alongside prior studies, the present analysis complements existing evidence by providing prospective multicenter evidence of sex-consistent diagnostic performance and rule-out safety across MI phenotypes in ED settings.

STUDY LIMITATIONS. Several limitations of this study should be acknowledged. First, this investigation was conducted in a single country, which may limit the generalizability of the findings to health care systems with different patient populations, clinical workflows, or ECG acquisition practices. Second, biological sex was defined based on electronic medical record documentation and does not capture gender identity or sociocultural factors, which may also influence symptom presentation and health care-seeking behavior. Third, the physician-estimated probability of AMI was inherently subjective and may vary across clinicians and institutions, potentially affecting comparative performance assessments. Fourth, prespecified AI-ECG risk thresholds were fixed based on prior model development and were not optimized separately for each subgroup, which may limit flexibility across different clinical contexts. Finally, although outcome adjudication was performed during the index hospitalization, certain clinical presentations without overt biomarker elevation may remain diagnostically challenging. Therefore, the observed missed AMI rate should be interpreted with caution, particularly in women who more frequently present with atypical phenotypes. These findings should be interpreted as supportive evidence, and further external validation across diverse health care systems and prospective

implementation studies will be important to determine the real-world clinical impact of AI-ECG-guided triage strategies.³⁰

CONCLUSIONS

In this prospective, multicenter external validation study, AI-ECG demonstrated sex-consistent diagnostic performance for AMI, with preserved discrimination and rule-out safety in women. AI-ECG performance remained robust across MI phenotypes, including STEMI and NSTEMI, supporting a phenotype-agnostic diagnostic capability. These findings suggest that AI-ECG may serve as an early triage adjunct in ED settings, with potential to improve diagnostic equity during the initial evaluation of suspected AMI, particularly in women.

FUNDING SUPPORT AND AUTHOR DISCLOSURES

This work has been supported by a nonconditional 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). Drs HS Lee, S Kang, JM Son, KS Kim, JM Kwon, and MS Lee are employees of Medical AI Co, Ltd and hold stocks in the company. Dr Kyuseok Kim is a stakeholder in Medical AI Co., Ltd. All other authors have

reported that they have no relationships relevant to the contents of this paper to disclose.

ADDRESS FOR CORRESPONDENCE: Dr Min Sung Lee, Medical AI Co., Ltd., Seoul, Republic of Korea, Digital Healthcare Institute, Sejong Hospital, 38, Yeongdong-daero 85-gil, Gangnam-gu, Seoul 06180, Republic of Korea. E-mail: lylm85@gmail.com.

PERSPECTIVES

COMPETENCY IN MEDICAL KNOWLEDGE: Women with suspected AMI face an increased risk of diagnostic delay and underrecognition due to atypical presentations and less overt ECG findings. In a prospective, multicenter ED cohort, AI-ECG demonstrated sex-consistent diagnostic performance and preserved rule-out safety.

TRANSLATIONAL OUTLOOK: AI-ECG may serve as an early triage adjunct when biomarker results are pending, with sex-consistent performance supporting integration into ED workflows to improve diagnostic equity. Future studies should evaluate real-world implementation across diverse health care systems and assess impacts on clinical decision-making, downstream testing, and outcomes—particularly among women and other high-risk populations.

REFERENCES

- Martin SS, Aday AW, Almarzooq ZI, et al. 2024 Heart disease and stroke statistics: a report of US and Global data from the American Heart Association. *Circulation*. 2024;149:e347–e913.
- van Oosterhout RE, de Boer AR, Maas AH, Rutten FH, Bots ML, Peters SA. Sex differences in symptom presentation in acute coronary syndromes: a systematic review and meta-analysis. *J Am Heart Assoc*. 2020;9:e014733.
- DeFilippis EM, Collins BL, Singh A, et al. Women who experience a myocardial infarction at a young age have worse outcomes compared with men: the Mass General Brigham YOUNG-MI registry. *Eur Heart J*. 2020;41:4127–4137.
- Rao SV, O'Donoghue ML, Ruel M, et al. 2025 ACC/AHA/ACEP/NAEMSP/SCAI guideline for the management of patients with acute coronary syndromes: a report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. *J Am Coll Cardiol*. 2025;85:2135–2237.
- Byrne RA, Rossello X, Coughlan J, et al. 2023 ESC guidelines for the management of acute coronary syndromes. *Eur Heart J*. 2023;44:3720–3826.
- Kwok CS, Bennett S, Azam Z, et al. Misdiagnosis of acute myocardial infarction: a systematic review of the literature. *Crit Pathw Cardiol*. 2021;20:155–162.
- Than M, Herbert M, Flaws D, 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;166:752–754.
- Khera R, Oikonomou EK, Nadkarni GN, et al. Transforming cardiovascular care with artificial intelligence: from discovery to practice: JACC state-of-the-art review. *J Am Coll Cardiol*. 2024;84:97–114.
- Herman R, Meyers HP, Smith SW, et al. International evaluation of an artificial intelligence-powered electrocardiogram model detecting acute coronary occlusion myocardial infarction. *Eur Heart J Digit Health*. 2024;5:123–133.
- Al-Zaiti SS, Martin-Gill C, Zègre-Hemsey JK, et al. Machine learning for ECG diagnosis and risk stratification of occlusion myocardial infarction. *Nat Med*. 2023;29:1804–1813.
- Liu W-C, Lin C-S, Tsai C-S, et al. A deep learning algorithm for detecting acute myocardial infarction: deep learning model to detect AMI. *EuroIntervention*. 2021;17:765.
- Lee MS, Shin TG, Lee Y, et al. Artificial intelligence applied to electrocardiogram to rule out acute myocardial infarction: the ROMIAE multicentre study. *Eur Heart J*. 2025;46:1917–1929.
- Shin TG, Lee Y, Kim K, Lee MS, Kwon JM, ROMIAE study group. 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.
- Doudesis D, Lee KK, Yang J, et al. Validation of the myocardial-ischaemic-injury-index machine learning algorithm to guide the diagnosis of myocardial infarction in a heterogeneous population: a prespecified exploratory analysis. *Lancet Digit Health*. 2022;4:e300–e308.
- Lopez-Ayala P, Boeddinghaus J, Nestelberger T, et al. External validation of the myocardial-ischaemic-injury-index machine learning algorithm for the early diagnosis of myocardial infarction: a multicentre cohort study. *Lancet Digit Health*. 2024;6:e480–e488.
- Backus B, Six A, Kelder J, et al. A prospective validation of the HEART score for chest pain patients at the emergency department. *Int J Cardiol*. 2013;168:2153–2158.

17. Hung J, Roos A, Kadesjö E, et al. Performance of the GRACE 2.0 score in patients with type 1 and type 2 myocardial infarction. *Eur Heart J*. 2021;42:2552-2561.
18. Svennberg E, Han JK, Caiani EG, et al. State of the art of artificial intelligence in clinical electrophysiology in 2025: a scientific statement of the European Heart Rhythm Association (EHRA) of the ESC, the Heart Rhythm Society (HRS), and the ESC Working Group on E-Cardiology. *Europace*. 2025;27:euaf071.
19. Khraishah H, Daher R, Garelnabi M, Karere G, Welty FK. Sex, racial, and ethnic disparities in acute coronary syndrome: novel risk factors and recommendations for earlier diagnosis to improve outcomes. *Arterioscler Thromb Vasc Biol*. 2023;43:1369-1383.
20. Redfors B, Angerås O, RAamunddal T, et al. Trends in gender differences in cardiac care and outcome after acute myocardial infarction in Western Sweden: a report from the Swedish web system for enhancement of evidence-based care in heart disease evaluated according to recommended therapies (SWEDEHEART). *J Am Heart Assoc*. 2015;4:e001995.
21. Giordano V, Guillari A, Sansone V, Catone M, Rea T. Women acute myocardial infarction—identifying and understanding the gender gap (WAMy-GAP): a study protocol. *Healthcare MDPI*. 2024;12:972.
22. Vaccarino V, Almuwaqqat Z, Kim JH, et al. Association of mental stress-induced myocardial ischemia with cardiovascular events in patients with coronary heart disease. *JAMA*. 2021;326:1818-1828.
23. Lichtman JH, Leifheit EC, Safdar B, et al. Sex differences in the presentation and perception of symptoms among young patients with myocardial infarction: evidence from the VIRGO study (variation in recovery: role of gender on outcomes of young AMI patients). *Circulation*. 2018;137:781-790.
24. Safdar B, Spatz ES, Dreyer RP, et al. Presentation, clinical profile, and prognosis of young patients with myocardial infarction with non-obstructive coronary arteries (MINOCA): results from the VIRGO study. *J Am Heart Assoc*. 2018;7:e009174.
25. Supples MW, Snavely AC, O'Neill JC, et al. Sex and race differences in the performance of the European Society of Cardiology 0/1-h algorithm with high-sensitivity troponin T. *Clin Cardiol*. 2024;47:e24199.
26. Wenzl FA, Kraler S, Ambler G, et al. Sex-specific evaluation and redevelopment of the GRACE score in non-ST-segment elevation acute coronary syndromes in populations from the UK and Switzerland: a multinational analysis with external cohort validation. *Lancet*. 2022;400:744-756.
27. Lee KK, Ferry AV, Anand A, et al. Sex-specific thresholds of high-sensitivity troponin in patients with suspected acute coronary syndrome. *J Am Coll Cardiol*. 2019;74:2032-2043.
28. Tamis-Holland JE, Jneid H, Reynolds HR, et al. Contemporary diagnosis and management of patients with myocardial infarction in the absence of obstructive coronary artery disease: a scientific statement from the American Heart Association. *Circulation*. 2019;139:e891-e908.
29. Lin C, Liu W-T, Chang C-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:A0a2400190.
30. Cha H, Kim J. Ethical considerations of artificial intelligence in emergency medicine for triage and resource allocation: a scoping review. *Clin Exp Emerg Med*. 2025;12:306.

KEY WORDS acute myocardial infarction, AI-ECG, artificial intelligence, electrocardiogram

APPENDIX For a list of the ROMIAE Study Group as well as supplemental tables and figures, please see the online version of this paper.