A deep learning algorithm for detecting acute myocardial infarction

Wen-Cheng Liu[1], MD; Chin-Sheng Lin[1], MD, PhD; Chien-Sung Tsai[2], MD; Tien-Ping Tsao[3], MD; Cheng-Chung Cheng[1], MD; Jun-Ting Liou[1], MD; Wei-Shiang Lin[1], MD; Shu-Meng Cheng[1], MD, PhD; Yu-Sheng Lou[4], MS; Chia-Cheng Lee[5,6], MD; Chin Lin[4,7,8]*, PhD

1. Division of Cardiology, Department of Internal Medicine, Tri-Service General Hospital, National Defense Medical Center, Taipei, Taiwan, R.O.C.; 2. Division of Cardiovascular Surgery, Department of Surgery, Tri-Service General Hospital, National Defense Medical Center, Taipei, Taiwan, R.O.C.; 3. Division of Cardiology, Heart Centre, Cheng Hsin Hospital, Taipei, Taiwan, R.O.C.; 4. Graduate Institute of Life Sciences, National Defense Medical Center, Taipei, Taiwan, R.O.C.; 5. Planning and Management Office, Tri-Service General Hospital, National Defense Medical Center, Taipei, Taiwan, R.O.C.; 6. Division of Colorectal Surgery, Department of Surgery, Tri-Service General Hospital, National Defense Medical Center, Taipei, Taiwan, R.O.C.; 7. School of Medicine, National Defense Medical Center, Taipei, Taiwan, R.O.C.; 8. School of Public Health, National Defense Medical Center, Taipei, Taiwan, R.O.C.

This paper also includes supplementary data published online at: https://eurointervention.pcronline.com/doi/10.4244/EIJ-D-20-01155

Abstract

Background: Delayed diagnosis or misdiagnosis of acute myocardial infarction (AMI) is not unusual in daily practice. Since a 12-lead electrocardiogram (ECG) is crucial for the detection of AMI, a systematic algorithm to strengthen ECG interpretation may have important implications for improving diagnosis.

Aims: We aimed to develop a deep learning model (DLM) as a diagnostic support tool based on a 12-lead electrocardiogram.

Methods: This retrospective cohort study included 1,051/697 ECGs from 737/287 coronary angiogram (CAG)-validated STEMI/NSTEMI patients and 140,336 ECGs from 76,775 non-AMI patients at the emergency department. The DLM was trained and validated in 80% and 20% of these ECGs. A human-machine competition was conducted. The area under the receiver operating characteristic curve (AUC), sensitivity, and specificity were used to evaluate the performance of the DLM.

Results: The AUC of the DLM for STEMI detection was 0.976 in the human-machine competition, which was significantly better than that of the best physicians. Furthermore, the DLM independently demonstrated sufficient diagnostic capacity for STEMI detection (AUC=0.997; sensitivity, 98.4%; specificity, 96.9%). Regarding NSTEMI detection, the AUC of the combined DLM and conventional cardiac troponin I (cTnI) increased to 0.978, which was better than that of either the DLM (0.877) or cTnI (0.950).

Conclusions: The DLM may serve as a timely, objective and precise diagnostic decision support tool to assist emergency medical system-based networks and frontline physicians in detecting AMI and subsequently initiating reperfusion therapy.

Corresponding author: No.161 Min-Chun E. Rd, Sec. 6, Neihu, Taipei 114, Taiwan, R.O.C.
E-mail: xup6fup0629@gmail.com

Abbreviations

ALT	alanine aminotransferase
AMI	acute myocardial infarction
AST	aspartate aminotransferase
AUC	area under the ROC curve
BMI	body mass index
CAD	coronary artery disease
CAG	coronary angiogram
CK	creatine kinase
CKD	chronic kidney disease
COPD	chronic obstructive pulmonary disease
Cr	creatinine
cTnI	conventional cardiac troponin I
DLM	deep learning model
DM	diabetes mellitus
ECG	electrocardiogram
ED	emergency department
eGFR	estimated glomerular filtration rate
GLU	glucose
Hb	haemoglobin
HDL	high-density lipoprotein cholesterol
HF	heart failure
hsTnI	high-sensitivity cardiac troponin I
HTN	hypertension
IRA	infarct-related artery
K	potassium
LAD	left anterior descending artery
LCx	left circumflex artery
LDL	low-density lipoprotein
LMCA	left main coronary artery
Na	sodium
NSTE-ACS	non-ST-segment elevation acute coronary syndrome
NSTEMI	non-ST-segment elevation myocardial infarction
PLT	platelet count
PRROC	precision-recall receiver operating characteristic curve
PTB	Physikalisch-Technische Bundesanstalt
RCA	right coronary artery
ROC	receiver operating characteristic
STEMI	ST-segment elevation myocardial infarction
TC	total cholesterol
TG	triglycerides
WBC	white blood cell count

Introduction

Acute myocardial infarction (AMI) remains a major public health issue despite global advances in diagnosis and management[1]. AMI refers to the evidence of acute myocardial injury detected by abnormal cardiac biomarkers with necrosis in a clinical setting consistent with myocardial ischaemia. The categories of ST-segment elevation myocardial infarction (STEMI) and non-ST-segment elevation acute coronary syndrome (NSTE-ACS) based on the presentation of a 12-lead electrocardiogram (ECG) have customarily been included in the concept of acute coronary syndrome (ACS)[2]. Patients with

symptoms suggestive of myocardial ischaemia and ST-segment elevation on the ECG require timely reperfusion therapy to reduce cardiac morbidity and mortality[3]. Likewise, patients with non-ST-segment elevation myocardial infarction (NSTEMI) considered to be in the very high or high risk categories require an immediate/early invasive strategy to prevent a worse prognosis[4].

However, prompt management depends on rapid recognition and precise diagnosis. Despite the established criteria for the diagnosis of AMI, it remains a critical challenge for emergency physicians to recognise rapidly. Previous studies reported that the rate of misdiagnosis of AMI at first medical contact ranged from 2 to 30%[5-7]. Failure to identify high-risk ECG findings in patients with AMI results in lower quality care and higher adverse outcomes. One of the leading causes of missed identification in the diagnostic process was incorrect interpretation of a diagnostic test[8,9]. Systematic processes to improve ECG interpretation may therefore have important implications for improving diagnosis. Since the principal diagnostic tool for AMI is a 12-lead ECG, a more detailed analysis of the ECG may significantly speed up this process.

The current artificial intelligence revolution that started with a deep learning model (DLM) has provided us with an unprecedented opportunity to improve the healthcare system, and it has been proven to be effective in medical applications[10-12]. Additionally, DLMs were confirmed to be superior to cardiologists in ECG interpretation when they were trained by large annotated ECG data sets[13,14]. To our knowledge, the available and applicable ECG databases of AMI were relatively small. Our study aimed to develop a DLM to detect AMI in a timely, objective and precise manner by a 12-lead ECG. More than 100,000 AMI-associated ECGs were recruited and learned by the DLM. Facilitated by the system's powerful computing ability, the performance of the trained model was compared with that of physicians, including cardiologists and an emergency physician. The diagnostic power for STEMI and NSTEMI by the DLM and conventional cardiac troponin I (cTnI) was also evaluated.

Methods
STUDY DESIGN

This was a single-centre, case-control study. The data were provided by the Tri-Service General Hospital, Taipei, Taiwan, and the retrospective design was ethically approved by the institutional review board (IRB No. 2-107-05-168). An electronic health system was built for collecting ECGs and medical records. The study period was from January 2012 to December 2018.

STUDY POPULATION

AMI patients presenting to the emergency department (ED) who received a coronary angiogram (CAG) to rule in type I AMI and to confirm the infarct-related artery (IRA) of STEMI were recruited[2]. AMI patients with no electronic ECG available, right side ECG, posterior ECG and pacemaker rhythm were excluded. Non-AMI patients presenting to the ED during the same period were recruited, while excluding those with a history of AMI or

any elevated cTnI during their ED stay. The definitions of AMI, STEMI, NSTEMI, non-AMI, and non-STEMI in this study are provided in **Supplementary Table 1** and **Supplementary Appendix 1**. The AMI cases were divided into development (80%) and validation (20%) cohorts by date. The ECGs in the development cohort were excluded from the validation cohort. There was no overlap of patients between these two cohorts.

ADJUDICATED FINAL DIAGNOSIS

Adjudication of the final diagnosis was performed by three board-certified interventional cardiologists who did not participate in the human-machine competition and who retrospectively and independently reviewed the AMI cases according to the clinical presentations, serial ECGs, serial cTnI levels and angiographic findings to make the final diagnosis of STEMI and NSTEMI, as recommended in the current guidelines[2-4]. In situations of disagreement about the diagnosis, cases were reviewed and adjudicated in consensus meetings.

DATA COLLECTION AND IMPLEMENTATION OF THE DLM

Data collection and DLM implementation are shown in **Supplementary Appendix 1** and **Supplementary Figure 1**. ECG recordings were collected using a Philips 12-lead ECG machine (PH080A), and the DLM was based on ECG12Net, which had previously been developed[14]. The output of the DLM was the probability of STEMI, NSTEMI, and non-AMI.

HUMAN-MACHINE COMPETITION

We evaluated the performance of participanting physicians using a competition set of 450 ECGs, which included 174 STEMI, 138 NSTEMI, and 138 non-AMI ECGs. The STEMI ECGs, based on the IRA, were further classified into the left main coronary artery (LMCA), left anterior descending artery (LAD), left circumflex artery (LCx), or right coronary artery (RCA). Five cardiologists and one emergency attending physician participated in the competition. In addition, the Philips 12-lead algorithm was also included to detect AMI in the competition[15]. The physicians had no access to any patient information and no knowledge of the data. Their responses were entered into an online standardised data entry program. We calculated the sensitivities, specificities, and kappa values to compare their results with those of the DLM.

STATISTICAL ANALYSIS

The study cohort was divided into training, validation, and competition sets. We presented their characteristics as the means and standard deviations, numbers of patients, or percentages where appropriate. They were compared using either the Student's t-test or the chi-square test, as appropriate. The statistical analysis was performed using R software version 3.4.4 (R Foundation for Statistical Computing, Vienna, Austria).

All analyses were based on ECGs but not patients. A significance level of p<0.05 was used throughout the analysis. The primary analysis was to evaluate the performance of the DLM, the physicians and the Philips algorithm for STEMI detection in the human-machine competition. The receiver operating characteristic (ROC) curve and the area under the ROC curve (AUC) were applied to evaluate the competition results. We also used precision-recall ROC (PRROC) to evaluate the model performance in hypothetical real-world situations. Because the proportions of STEMI, NSTEMI, and non-AMI were distorted in the competition set, we re-weighted the samples based on the incidences in the real world (0.1%, 0.2%, and 99.7% of STEMI, NSTEMI, and non-AMI cases, respectively)[16-18]. The secondary analysis was performed on the whole validation cohort. We included more clinical information, such as patient characteristics and laboratory tests, to improve the model performance. A multivariable logistic regression model was used to integrate the DLM and clinical information. A series of logistic regression models identified the effects of different clinical information on the performance of STEMI and NSTEMI detection. The AUC was applied to evaluate the changes in model performance. The research interests, model comparison and statistical methods in this study are summarised in detail in **Supplementary Table 2**.

Results

BASELINE CHARACTERISTICS OF THE COHORTS

There were 1,051 ECGs before CAG from 737 STEMI patients, 697 ECGs before CAG from 287 NSTEMI patients and 140,336 ECGs from 76,775 non-AMI patients in this study. The development and validation cohorts included records from 58,056 and 19,743 patients, respectively. The characteristics and laboratory data are shown in **Supplementary Table 3**, and a detailed description is shown in **Supplementary Appendix 2**.

PREDICTION OF STEMI, NSTEMI AND NON-AMI

The results of the human-machine competition are summarised in **Figure 1**. The AUC of the DLM in the human-machine competition was 0.976 for STEMI detection, with a corresponding sensitivity and specificity of 89.7% and 94.6%, respectively. In contrast, the sensitivities and specificities for STEMI detection among the physicians and the Philips algorithm ranged from 60.5-92.6% and 76.0-97.5%, respectively, which were lower than those of the DLM. The PRROC analysis demonstrated the feasibility of an automatic ECG screening system, which revealed that the AUC of the DLM for STEMI detection was 0.586 in the hypothetical real world. The DLM achieved 63.2% precision and 50.3% recall using the appropriate cut-off point. These values were significantly better than those of all the physicians and the Philips algorithm.

Performance rankings and consistency analysis of STEMI detection among the DLM, the physicians and the Philips algorithm in the human-machine competition were carried out **(Figure 2)**. The DLM achieved the best global performance (kappa=0.645) **(Figure 2A)**, whereas the physicians had relatively better STEMI detection but poor discrimination of NSTEMI and non-AMI. The consistency analysis of AMI detection among the DLM, the physicians and the Philips algorithm is shown in the heatmap **(Figure 2B)**.
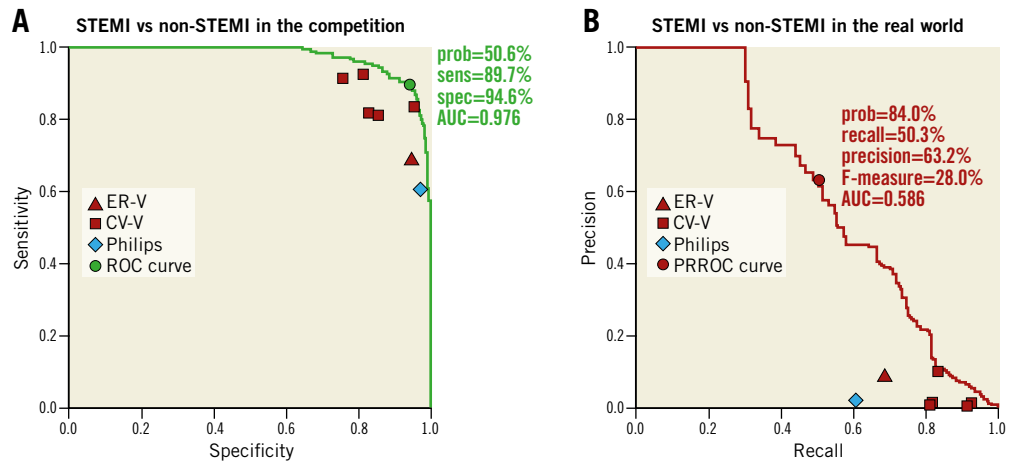
Figure 1. *Performance comparison for STEMI detection in the human-machine competition. The area under the receiver operating characteristic curve (AUC) was generated by the prediction of the DLM. The triangles, the square and the diamond denote the cardiologists, the emergency physician and the Philips algorithm, respectively. A) The ROC curve in the competition set (STEMI=174, NSTEMI=138, and non-AMI=138). B) The precision-recall ROC curve in the revised proportion of the hypothetical real world (STEMI=0.1%, NSTEMI=0.2%, and non-AMI=99.7%).*

ANALYSIS OF IRA OF STEMI

The DLM achieved the best global performance (kappa=0.629) for the IRA detection of STEMI **(Supplementary Figure 2)**. As shown in **Supplementary Figure 3**, after exclusion of LMCA and LCx, the AUC of the DLM for anterior STEMI detection was 0.975, with a corresponding sensitivity of 92.6%, which outperformed all participating physicians. Moreover, the AUC of the DLM in inferior STEMI detection was 0.974, with
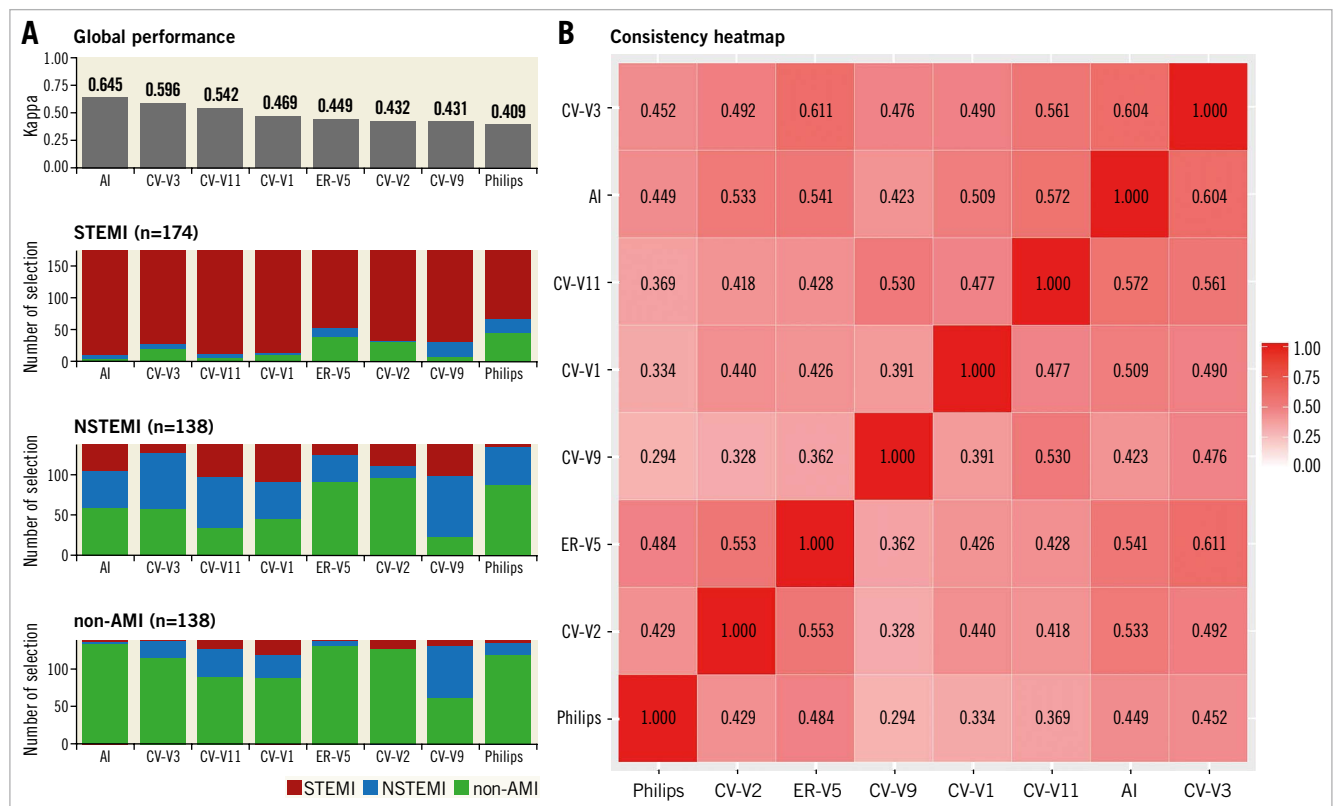
Figure 2. *Performance rankings and consistency analysis of STEMI detection among the DLM, the physicians and the Philips algorithm in the human-machine competition. A) Global performance rankings based on the class-3 kappa values. V(X) denotes (V) visiting staff with (X) years of experience. B) Consistency analysis as a heatmap coloured based on the values; the values in each cell were the kappa values of each pair.*

a corresponding sensitivity of 84.8%, which was better than all but one best physician. In the combined detection of anterior and inferior STEMI, the DLM had better performance than all physicians (AUC, 0.975; sensitivity, 89.4%).

INTERPRETATIONS OF STEMI ECGs BY THE DLM AND PHYSICIANS

Selected examples of STEMI ECGs in the human-machine competition are shown in **Figure 3**. A typical STEMI ECG with an IRA of the LAD **(Figure 3A)** was consistently detected by both the DLM and the physicians. One STEMI ECG with an IRA of the RCA **(Figure 3B)** was misdiagnosed by the DLM but correctly recognised by the best cardiologists. One STEMI ECG with an IRA of the RCA **(Figure 3C)** was misdiagnosed by both the DLM and the best cardiologists. The DLM correctly detected the ECG **(Figure 3D)** as STEMI with an IRA of the LAD, which was misdiagnosed by the best cardiologists.

Among the 138 NSTEMI ECGs, 58 ECGs were detected as non-AMI by the DLM, with an accuracy of 58.0%, which was worse than that of the best cardiologist (75.4%). This discrepancy was due to a more conservative AMI diagnostic strategy by the DLM. In contrast, among 138 non-AMI ECGs, the specificity of 96.4% of the DLM was much better than those of the two best cardiologists (82.6% and 64.5%). After adjustment for the specificities, the misdiagnosis of NSTEMI by the DLM was obviously less than that by the best cardiologists **(Table 1)**. Nevertheless, the DLM offered the best performance in AMI detection under the standardisation of the best cardiologists. The ECG lead-specific analyses for the detection of STEMI and the corresponding IRA are shown in **Supplementary Figure 4**, and a detailed description is shown in **Supplementary Appendix 2**.

LOGISTIC REGRESSION ANALYSIS OF STEMI AND NSTEMI
Univariate and multivariate logistic regression analyses in the development cohort revealed that male sex, prior CAD, cTnI, haemoglobin, total cholesterol and low-density lipoprotein (LDL) levels were independent risk factors for STEMI and NSTEMI detection **(Supplementary Figure 5)**.

DIAGNOSTIC VALUE ANALYSIS
We evaluated the performance of the DLM after adjusting for significant patient characteristics, disease histories, and laboratory data
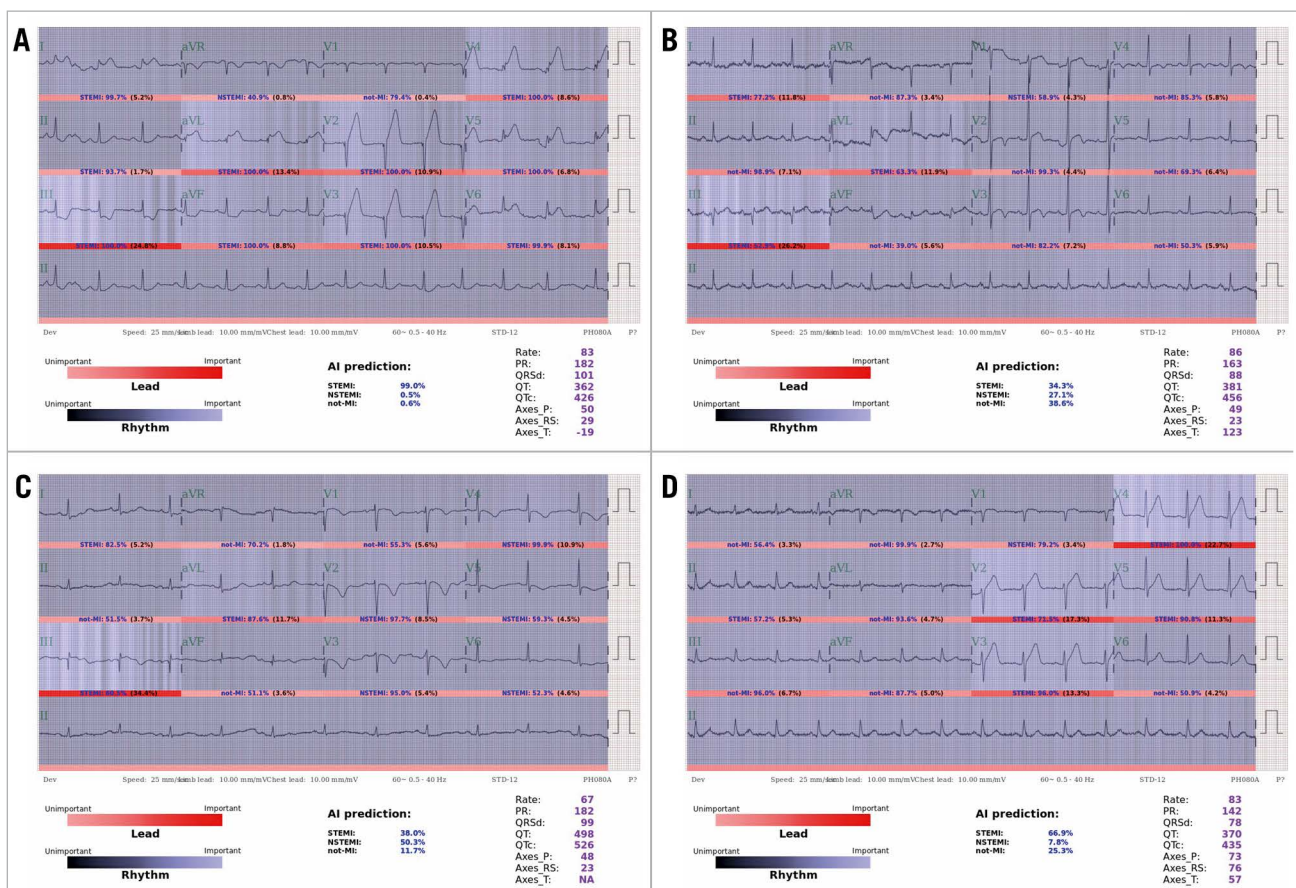
Figure 3. *Interpretations of selected STEMI ECGs by the DLM and physicians in the human-machine competition. A) Both the DLM and the best cardiologists consistently detected STEMI. B) The DLM misdetected STEMI, which was correctly detected by the best cardiologists. C) Both the DLM and the best cardiologists misdetected STEMI. D) The DLM correctly detected STEMI, which was misdetected by the best cardiologists.*

Table 1. Maximum sensitivity of the DLM for a specific specificity.

	Revised item[a]	Sensitivity[b] (STEMI)	Sensitivity[c] (NSTEMI)	Specificity[d]
DLM (original)	0.000	164/174 (94.3%)	80/138 (58.0%)	133/138 (96.4%)
CV-V3		146/174 (83.9%)	80/138 (58.0%)	114/138 (82.6%)
DLM (specificity=82.6%)	0.450	166/174 (95.4%)	108/138 (78.3%)	114/138 (82.6%)
CV-V11		162/174 (93.1%)	104/138 (75.4%)	89/138 (64.5%)
DLM (specificity=64.5%)	0.612	166/174 (95.4%)	123/138 (89.1%)	89/138 (64.5%)

[a]The revised item was used to modify the probability of non-AMI given by DLM. For example, if an original probability of STEMI/NSTEMI/non-AMI was 0.220/0.310/0.470, then the prediction was defined as non-AMI according to the largest probability. However, the revised item was used to make DLM become more sensitive, which was used to modify the probability of non-AMI as 0.470-0.450=0.020 as the first situation. Therefore, the new prediction of this case was defined as NSTEMI according to the largest revised probability (0.220/0.310/0.020). [b]The sensitivity of STEMI was defined as the percentage of STEMI that was correctly identified as STEMI. [c]The sensitivity of NSTEMI was defined as the percentage of NSTEMI that was correctly identified as NSTEMI. [d]The specificity is defined as the percentage of non-AMI cases that was correctly identified as non-AMI.

to ensure consistency across a wide range of putative confounding variables in the validation cohort. The DLM had significantly better performance than cTnI in detecting STEMI, with an AUC of 0.997 with a corresponding sensitivity and specificity of 98.4% and 96.9%, respectively **(Figure 4A)**. However, cTnI had significantly better performance than the DLM in detecting NSTEMI. The AUC for NSTEMI detection by the combination of the DLM and the first recorded cTnI increased to 0.978, with a corresponding sensitivity and specificity of 91.6% and 96.7%, respectively, which was better than that of the DLM (0.877) or cTnI (0.950) individually **(Figure 4B)**. Using the DLM independently was sufficient to detect STEMI, and the addition of patient characteristics did not significantly improve its performance. However, cTnI was found to improve the diagnostic accuracy for NSTEMI better than any additional characteristics **(Supplementary Figure 6)**.

Discussion

In this study, we established a DLM to detect STEMI precisely through ECG analysis, which applied a deep convolutional network to extract notable ECG features with a development cohort of more than 100,000 ECGs. All AMI patients were validated by CAG, and the corresponding IRA of STEMI was identified. Most importantly, our DLM performed better than the physicians in STEMI detection, with a high sensitivity of 89.7% and specificity of 94.6%.

The application of deep learning technology in the cardiovascular field for arrhythmias, dyskalaemia, and valvular heart disease has recently grown in popularity[13,14,19-21]. However, no large-scale study has been designed for AMI detection. Previous DLMs for AMI detection by a 12-lead ECG mainly used the Physikalisch-Technische Bundesanstalt (PTB) diagnostic ECG database[22,23]. These studies may be limited because they were not further validated. Moreover, comparisons between the DLM and physicians were lacking. In comparison with previous studies, we enrolled the largest number of clinically validated ECGs for development and validation. Additionally, we further confirmed the role of cTnI in assisting with NSTEMI detection by our DLM. All these results highlight the strengths of the current study.

The sensitivity and specificity for STEMI detection by the DLM were better than those of the physicians. ECG is the timeliest tool among all objective detection methods for AMI. However,
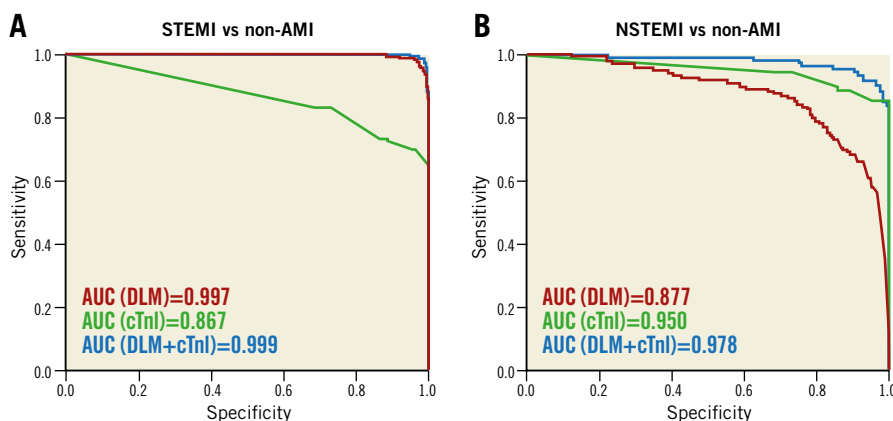
Figure 4. *Comparison of the diagnostic value between the DLM and cTnI in the validation cohort. The area under the receiver operating characteristic curve (AUC) was generated from the logistic regression analysis using the validation cohort. The p-values represent the comparison among the DLM, cTnI and the DLM plus cTnI. A) Regarding STEMI detection: DLM vs cTnI, p<0.001; DLM vs DLM+cTnI, p=ns. B) Regarding NSTEMI detection: cTnI vs DLM, p<0.01; cTnI+DLM vs cTnI, p<0.05.*

the low sensitivity and the disagreement in interpreting ECGs between physicians remain issues. The sensitivity of subjective interpretation for AMI detection using a 12-lead ECG ranged only from 61 to 74%, with a specificity ranging from 72 to 89%[24-26]. In contrast, previous prehospital computer algorithm interpretations for STEMI had a sensitivity of approximately 69%[27,28]. Our DLM provided extraordinary performance that could support decision-making systems in clinical practice.

The DLM could objectively identify STEMI based on analysing and learning a large number of ECGs. Moreover, subtle ECG changes in the earliest phase of STEMI, which are easily missed by physicians, could be correctly recognised by the DLM. Nevertheless, prior MI or cardiomyopathy might mislead the DLM owing to baseline ST-T changes. Therefore, information regarding previously available ECGs or the history of cardiovascular disease may be needed to strengthen the capacity of the DLM for STEMI detection further.

The performance of our DLM on the detection of STEMI equivalents and STEMI mimics was further evaluated. STEMI equivalents, including de Winter sign, Wellens' syndrome, hyperacute T-waves, ST elevation in the lead aVR with diffuse ST depression, ST elevation in the presence of bundle branch block, and posterior wall AMI, representing coronary occlusion without meeting the traditional ST elevation criteria, were crucial for timely recognition[29,30]. Additionally, high take-off T presentations, such as hyperkalaemia,

benign early repolarisation, left ventricular hypertrophy and Brugada syndrome, which mimick STEMI, were usually misdiagnosed, leading to false initiation of primary PCI[31,32]. Our study demonstrated that the DLM exhibited excellent diagnostic power in the detection of STEMI equivalents (except for type 1 Wellens' syndrome) and provided extraordinary differentiating capacity in the detection of high take-off T (**Supplementary Figure 7, Supplementary Figure 8**). Further prospective and large ECG validation data sets are needed to confirm the discriminating abilities of the DLM.

Our DLM has several potential clinical applications. First, the DLM could provide decision support and a high-risk alarm system for AMI that could help to reduce medical errors in the ED resulting from intense time pressures or heavy workloads and harried staff during busy working hours. Second, the DLM could be incorporated into ECG machines in ambulances to facilitate telemedicine and shorten the decision time before initiation of reperfusion therapies. Third, our DLM could be applied in rural and remote areas and places lacking experts to facilitate ECG interpretation and promote diagnostic accuracy, thereby initiating timely management and improving the prognosis of STEMI patients. Finally, the DLM could be incorporated into a wearable device for AMI detection, especially for patients with an extremely high risk of atherosclerotic cardiovascular disease. Accordingly, our DLM exhibits diagnostic benefits and may improve the quality of health care in the near future (**Central illustration**).
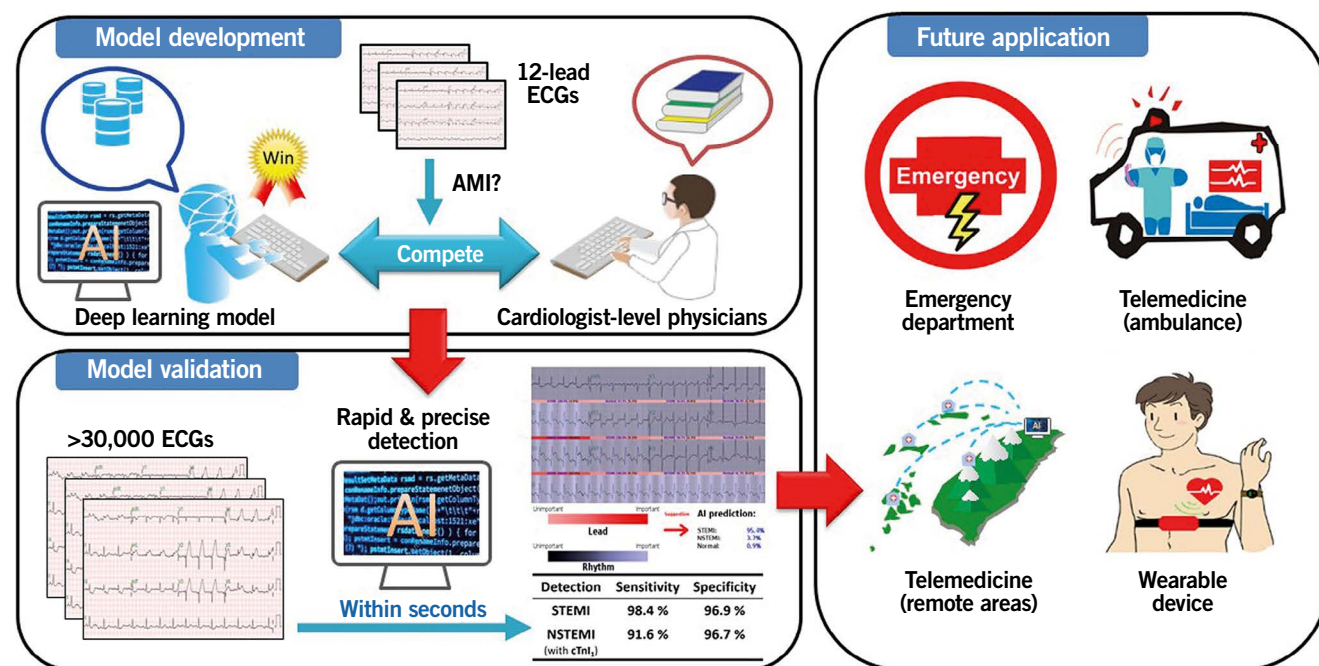
Central illustration. *Schematic diagram of the development, validation and future application of the current deep learning model for detecting AMI. The DLM learned from more than 100,000 ECGs was developed and trained. Compared with cardiologist-level physicians. The DLM exhibited the best performance in the detection of STEMI. The validated model achieved excellent diagnostic power with a sensitivity of 98.4% and a specificity of 96.9% for STEMI detection. With the ability of real-time detection, precise diagnosis and early alarm, the application of DLM for STEMI detection, including in-hospital, pre-hospital settings, telemedicine and wearable devices, would improve the quality of health care of cardiovascular disease in the near future.*

Limitations

Some limitations of this study should be noted. First, the human-machine competition was based on a well-designed retrospective study. A real-world prospective study should be conducted to verify the clinical impact of the DLM. Second, only six attending physicians participated in the competition with the DLM. Although their performance in AMI detection was relatively consistent with that in previous studies, comparisons should be made with more physicians to confirm the superiority of the DLM[33]. Third, the studied patients were enrolled from only one academic medical centre, although the diagnosis and management of AMI was based on the guidelines. Multicentre validation is needed to confirm the value and application of this study. Fourth, there were fewer NSTEMI cases than STEMI cases, which may limit the capacity for NSTEMI detection by our DLM. Fifth, during the study period, cTnI rather than hsTnI was used for AMI diagnosis. Sixth, information on prior ECGs to improve diagnostic performance was not available in our DLM system. Seventh, the impacts of coronary collateral flow on ST-T changes during AMI and the performance of the DLM in the detection of STEMI were not analysed. Finally, only patients in the ED were enrolled, which may have led to selection bias and constrained the generalisability of the results.

Conclusions

We established an optimal DLM to detect STEMI based on a 12-lead ECG with better accuracy than physicians. Integration of a DLM may assist frontline physicians in recognising AMI in a timely and precise manner to prevent delayed diagnosis or misdiagnosis of AMI and thereby provide prompt reperfusion therapy. Further prospective validation with pre-hospital and in-hospital ECG tests is needed to confirm the performance of our DLM.

Impact on daily practice

STEMI can now be recognised using this cardiologist-level algorithm, achieving real-time STEMI diagnosis and early alarms. A comprehensive ecosystem has been established including in-hospital, pre-hospital and wearable devices, improving the quality of care in AMI.

Funding

Conflict of interest statement

The authors have no conflicts of interest to declare.

References

1. GBD 2017 Congenital Heart Disease Collaborators. Global, regional, and national burden of congenital heart disease, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet Child Adolesc Health.* 2020;4:185-200.

2. Thygesen K, Alpert JS, Jaffe AS, Chaitman BR, Bax JJ, Morrow DA, White HD, ESC Scientific Document Group. Fourth universal definition of myocardial infarction (2018). *Eur Heart J.* 2019;40:237-269.

3. Ibanez B, James S, Agewall S, Antunes MJ, Bucciarelli-Ducci C, Bueno H, Caforio ALP, Crea F, Goudevenos JA, Halvorsen S, Hindricks G, Kastrati A, Lenzen MJ, Prescott E, Roffi M, Valgimigli M, Varenhorst C, Vranckx P, Widimsky P; ESC Scientific Document Group. 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.

4. Collet JP, Thiele H, Barbato E, Barthélémy O, Bauersachs J, Bhatt DL, Dendale P, Dorobantu M, Edvardsen T, Folliguet T, Gale CP, Gilard M, Jobs A, Jüni P, Lambrinou E, Lewis BS, Mehilli J, Meliga E, Merkely B, Mueller C, Roffi M, Rutten FH, Sibbing D, Siontis GCM; ESC Scientific Document Group. 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.

5. McCarthy BD, Beshansky JR, D'Agostino RB, Selker HP. Missed diagnoses of acute myocardial infarction in the emergency department: results from a multicenter study. *Ann Emerg Med.* 1993;22:579-82.

6. Masoudi FA, Magid DJ, Vinson DR, Tricomi AJ, Lyons EE, Crounse L, Ho PM, Peterson PN, Rumsfeld JS; Emergency Department Quality in Myocardial Infarction Study Investigators. Implications of the failure to identify high-risk electrocardiogram findings for the quality of care of patients with acute myocardial infarction: results of the Emergency Department Quality in Myocardial Infarction (EDQMI) study. *Circulation.* 2006;114:1565-71.

7. Wu J, Gale CP, Hall M, Dondo TB, Metcalfe E, Oliver G, Batin PD, Hemingway H, Timmis A, West RM. Editor's Choice - Impact of initial hospital diagnosis on mortality for acute myocardial infarction: A national cohort study. *Eur Heart J Acute Cardiovasc Care.* 2018;7:139-48.

8. Kachalia A, Gandhi TK, Puopolo AL, Yoon C, Thomas EJ, Griffey R, Brennan TA, Studdert DM. Missed and delayed diagnoses in the emergency department: a study of closed malpractice claims from 4 liability insurers. *Ann Emerg Med.* 2007;49: 196-205.

9. Schenkel S. Promoting patient safety and preventing medical error in emergency departments. *Acad Emerg Med.* 2000;7:1204-22.

10. Estava A, Kuprel B, Novoa R, Ko J, Swetter SM, Blau HM, Thrun S. Dermatologist level classification of skin cancer with deep neural networks. *Nature.* 2017;542: 115-8.

11. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, van der Laak JAWM, van Ginneken B, Sánchez CI. A survey on deep learning in medical image analysis. *Med Image Anal.* 2017;42:60-88.

12. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, Venugopalan S, Widner K, Madams T, Cuadros J, Kim R, Raman R, Nelson PC, Mega JL, Webster DR. Development and validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. *JAMA.* 2016;316: 2402-10.

13. Hannun AY, Rajpurkar P, Haghpanahi M, Tison GH, Bourn C, Turakhia MP, Ng AY. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nat Med.* 2019;25:65-9.

14. Lin CS, Lin C, Fang WH, Hsu CJ, Chen SJ, Huang KH, Lin WS, Tsai CS, Kuo CC, Chau T, Yang SJ, Lin SH. A Deep-Learning Algorithm (ECG12Net) for Detecting Hypokalemia and Hyperkalemia by Electrocardiography: Algorithm Development. *JMIR Med Inform.* 2020;8:e15931.

15. The Philips 12-lead Algorithm Physician's Guide. Available at: http://incenter. medical.philips.com/doclib/enc/fetch/577817/577818/12-Lead_Algorithm_ Physician_s_Guide_for_Algorithm_Verion_PH080A%2C_(ENG).pdf%3Fnodeid%3 D3325283%26vernum%3D-2. Last accessed 17/08/2021.

16. Widimsky P, Wijns W, Fajadet J, De Belder M, Knot J, Aaberge L, Andrikopoulos G, Baz JA, Betriu A, Claeys M, Danchin N, Djambazov S, Erne P, Hartikainen J, Huber K, Kala P, Klinceva M, Kristensen SD, Ludman P, Ferre JM, Merkely B, Milicic D, Morais J, Noc M, Opolski G, Ostojic M, Radovanovic D, De Servi S, Stenestrand U, Studencan M, Tubaro M, Vasiljevic Z, Weidinger F, Witkowski A, Zeymer U; European Association for Percutaneous Cardiovascular Interventions. Reperfusion therapy for ST elevation acute myocardial infarction in Europe: description of the current situation in 30 countries. *Eur Heart J.* 2010;31:943-57.

17. O'Gara PT, Kushner FG, Ascheim DD, Casey DE Jr, Chung MK, de Lemos JA, Ettinger SM, Fang JC, Fesmire FM, Franklin BA, Granger CB, Krumholz HM,

Linderbaum JA, Morrow DA, Newby LK, Ornato JP, Ou N, Radford MJ, Tamis-Holland JE, Tommaso CL, Tracy CM, Woo YJ, Zhao DX. 2013 ACCF/AHA guideline for the management of ST-elevation myocardial infarction: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. *J Am Coll Cardiol.* 2013;61:e78-140.

18. Jernberg T. Swedeheart annual report 2015. Karolinska University Hospital, Huddinge 2016. Available at: https://www.ucr.uu.se/swedeheart/dokument-sh/arsrapporter-sh/aeldre-arsrapporter-older-reports/arsrapport-2015/swedeheart-annual-report-2015

19. Attia ZI, Kapa S, Lopez-Jimenez F, McKie PM, Ladewig DJ, Satam G, Pellikka PA, Enriquez-Sarano M, Noseworthy PA, Munger TM, Asirvatham SJ, Scott CG, Carter RE, Friedman PA. Screening for cardiac contractile dysfunction using an artificial intelligence–enabled electrocardiogram. *Nat Med.* 2019;25:70-4.

20. Kwon JM, Lee SY, Jeon KH, Lee Y, Kim KH, Park J, Oh BH, Lee MM. Deep Learning–Based Algorithm for Detecting Aortic Stenosis Using Electrocardiography. *J Am Heart Assoc.* 2020;9:e014717.

21. Kwon JM, Kim KH, Akkus Z, Jeon KH, Park J, Oh BH. Artificial intelligence for detecting mitral regurgitation using electrocardiography. *J Electrocardiol.* 2020;59:151-7.

22. Gupta A, Huerta E, Zhao Z, Moussa I. Deep Learning for Cardiologist-Level Myocardial Infarction Detection in Electrocardiograms. In: Jarm T, Cvetkoska A, Mahnic-Kalamiza S, Miklavcic D (eds). 8th European Medical and Biological Engineering Conference. EMBEC 2020. IFMBE Proceedings, vol 80. Springer, Cham. https://doi.org/10.1007/978-3-030-64610-3_40

23. Strodthoff N, Strodthoff C. Detecting and interpreting myocardial infarction using fully convolutional neural networks. *Physiol Meas.* 2019;40:015001.

24. Asch FM, Shah S, Rattin C, Swaminathan S, Fuisz A, Lindsay J. Lack of sensitivity of the electrocardiogram for detection of old myocardial infarction: a cardiac magnetic resonance imaging study. *Am Heart J.* 2006;152:742-8.

25. Trägårdh E, Claesson M, Wagner GS, Zhou S, Pahlm O. Detection of acute myocardial infarction using the 12-lead ECG plus inverted leads versus the 16-lead ECG (with additional posterior and right-sided chest electrodes). *Clin Physiol Funct Imaging.* 2007;27:368-74.

26. McCabe JM, Armstrong EJ, Ku I, Kulkarni A, Hoffmayer KS, Bhave PD, Waldo SW, Hsue P, Stein JC, Marcus GM, Kinlay S, Ganz P. Physician accuracy in interpreting potential ST-segment elevation myocardial infarction electrocardiograms. *J Am Heart Assoc.* 2013;2:e000268.

27. Garvey JL, Zegre-Hemsey J, Gregg R, Studnek JR. Electrocardiographic diagnosis of ST segment elevation myocardial infarction: an evaluation of three automated interpretation algorithms. *J Electrocardiol.* 2016;49:728-32.

28. de Champlain F, Boothroyd LJ, Vadeboncoeur A, Huynh T, Nguyen V, Eisenberg MJ, Joseph L, Boivin JF, Segal E. Computerized interpretation of the prehospital electrocardiogram: predictive value for ST segment elevation myocardial infarction and impact on on-scene time. *CJEM.* 2014;16:94-105.

29. Lawner BJ, Nable JV, Mattu A. Novel patterns of ischemia and STEMI equivalents. *Cardiol Clin.* 2012;30:591-9.

30. Tzimas G, Antiochos P, Monney P, Eeckhout E, Meier D, Fournier S, Harbaoui B, Muller O, Schläpfer J. Atypical Electrocardiographic Presentations in Need of Primary Percutaneous Coronary Intervention. *Am J Cardiol.* 2019;124:1305-14.

31. Somers MP, Brady WJ, Perron AD, Mattu A. The prominent T wave: electrocardiographic differential diagnosis. *Am J Emerg Med.* 2002;20:243-51.

32. Wang K, Asinger RW, Marriott HJ. ST-segment elevation in conditions other than acute myocardial infarction. *N Engl J Med.* 2003;349:2128-35.

33. Willems JL, Abreu-Lima C, Arnaud P, van Bemmel JH, Brohet C, Degani R, Denis B, Gehring J, Graham I, van Herpen G, et al. The diagnostic performance of computer programs for the interpretation of electrocardiograms. *N Engl J Med.* 1991;325:1767-73.

Supplementary data

Supplementary Appendix 1. Methods including: the definition of AMI, STEMI, NSTEMI, non-AMI and non-STEMI; data collection; timelines of door-to-balloon/CAG, door-to-ECG, and ECG-to-balloon/CAG; implementation of the deep learning model (DLM); implementation details of the DLM; training details; data augmentation; model visualisation and summary of the research interests, model comparison and statistical methods.

Supplementary Appendix 2. Results including: the baseline characteristics of the cohorts, ECG lead-specific analysis, and discussion.

Supplementary Figure 1. Architecture of the DLM.

Supplementary Figure 2. Performance rankings of infarct-related artery detection of STEMI among DLM, physicians and the Philips algorithm in the human-machine competition.

Supplementary Figure 3. Performance comparison for anterior (LAD), inferior (RCA), and combined anterior and inferior (LAD+RCA) STEMI detection in the human-machine competition.

Supplementary Figure 4. ECG lead-specific analyses for the detection of STEMI, STEMI-LAD and STEMI-RCA.

Supplementary Figure 5. Univariate and multivariate logistic regression analysis of STEMI, and NSTEMI in the development cohort.

Supplementary Figure 6. Comparison of the diagnostic value among additional demographic variables, cTnI and DLM in the validation cohort.

Supplementary Figure 7. The test examples of the detection of STEMI equivalents by the DLM.

Supplementary Figure 8. The test examples of the detection of high take-off T ECG by the DLM.

Supplementary Table 1. The definitions of AMI, STEMI, NSTEMI, non-AMI and non-STEMI.

Supplementary Table 2. The research interests, model comparison and statistical methods.

Supplementary Table 3. Corresponding patient characteristics and laboratory results of STEMI, NSTEMI, and non-AMI ECGs in the development and validation cohorts.

The supplementary data are published online at:
https://eurointervention.pcronline.com/
doi/10.4244/EIJ-D-20-01155