A Deep-learning Algorithm With the Real World Validation for Detecting Acute Myocardial Infarction

Wen-Cheng Liu  
Tri-Service General Hospital

Chin-Sheng Lin  
Tri-Service General Hospital

Chien-Sung Tsai  
Tri-Service General Hospital

Tien-Ping Tsao  
Cheng Hsin General Hospital

Cheng-Chung Cheng  
Tri-Service General Hospital

Jun-Ting Liou  
Tri-Service General Hospital

Wei-Shiang Lin  
Tri-Service General Hospital

Shu-Meng Cheng  
Tri-Service General Hospital

Yu-Sheng Lou  
National Defense Medical Center

CHIN LIN ( xu6fup@gmail.com )  
National Defense Medical Center  https://orcid.org/0000-0003-2337-2096

Research

Keywords: Artificial intelligence, electrocardiogram, deep learning, myocardial infarction

DOI: https://doi.org/10.21203/rs.3.rs-83284/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.  Read Full License
Abstract

Background

The initial detection and diagnosis of ST-segment or non-ST-segment elevation myocardial infarction (STEMI or NSTEMI) definitely rely on a 12-lead electrocardiogram (ECG). Delay or misdiagnosis is not unusual by subjective interpretation. Our aim is to develop a DLM as a diagnostic support tool to detect MI based on a 12-lead ECG and to evaluate the performance of this model.

Methods

This study included 1,051 ECGs from 737 coronary angiography (CAG)-validated STEMI patients, 697 ECGs from 287 CAG-validated NSTEMI patients, and 140,336 not-MI ECGs from 76,775 patients at emergency departments. DLM was trained and validated for the performance using 80% and 20% of the ECGs, respectively. 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 DLM and experts. STEMI versus not-STEMI, and MI versus not-MI were evaluated by DLM.

Results

The AUCs of DLM for identifying STEMI and MI were 0.976 and 0.944 in the human-machine competition, respectively, which were significantly better than those of our best clinicians. In the real world setting, DLM presented with AUC of 0.995/0.916 with corresponding sensitivities of 96.9%/77.0%, and specificities of 96.2%/92.9% in the identification of STEMI and MI, respectively. Furthermore, DLM demonstrated sufficient diagnostic capacity for STEMI without the aid of troponin I (TnI) (AUC= 0.996) with corresponding sensitivity and specificity of 98.4% and 96.9%. The AUC of combined DLM and the first recorded TnI for the detection of NSTEMI were increased to 0.978 with corresponding sensitivity and specificity of 91.6% and 96.7%, which was better than that of DLM (0.877) or TnI (0.949) alone.

Conclusions

DLM may serve as a diagnostic decision tool to assist intensive or emergency medical system-based networks and frontline physicians in identifying STEMI and NSTEMI in a timely and precise manner to prevent delay or misdiagnosis, and thereby to facilitate subsequent reperfusion therapy.

Background

Acute myocardial infarction (AMI) remains a major public health issue despite advances in diagnosis and management globally.[1] AMI refers to an abrupt cause of an unmet need of coronary blood supply to the myocardium. Based on the presentations of electrocardiogram (ECG), it is categorized mainly into two distinct populations: ST-segment elevation myocardial infarction (STEMI) and non-ST-segment elevation myocardial infarction (NSTEMI).[2] STEMI with ECG presentations of ST-segment elevation over infarcted areas indicates acute complete coronary occlusion that warrants prompt aggressive therapeutic strategies for coronary reperfusion of an occluded infarct-related artery (IRA) to prevent a cardiac disaster.[3] A delay in reperfusion therapy is significantly associated with an increase in subsequent mortality.[4-6] Similarly, as for NSTEMI with a high risk profile of unstable condition, an invasive reperfusion strategy should be adopted to prevent a worse outcome.[2, 7]

However, prompt management depends on rapid recognition and precise diagnosis. The diagnosis of AMI requires a syndrome indicative of myocardial ischemia with some extents of myocardial necrosis detected by ECG and cardiac biomarkers. Even though the established criteria for the diagnosis of AMI, it is still a critical challenge for emergent physicians to rapidly recognize.[8] Previous studies have reported a missed rate of diagnosis of AMI at first medical contact that ranges from 2 to 30%.[9-12] The failure to identify high-risk ECG findings in patients with AMI results in lower quality care and higher adverse prognosis. One of the most leading causes in the diagnostic process was incorrect interpretation of a diagnostic test.[13, 14] Systematic processes to improve ECG interpretation may have important implications for treatment and outcomes. Since the principal diagnostic tool for patients with suspected AMI is a 12-lead ECG, a more detailed analysis of the ECG could speed up this process significantly.

The current artificial intelligence revolution started by deep learning model (DLM) has provided us with an unprecedented opportunity to improve the health care system, and DLMs have been proven to be effective in medical applications.[15-18] DLM have been confirmed to surpass the cardiologist level on ECG interpretation when they are trained by large annotated ECG datasets.[19-21] To our knowledge, the available ECG databases of AMI are relative small.[22] Our study aimed to develop a DLM to timely, objectively and precisely diagnose AMI by ECG. DLM which learned more than 100,000 ECGs associated MI exhibited excellent diagnosis power in the detection of MI by ECG. Facilitated by the system's powerful computing ability, the performance of the trained model was compared with that of different levels of participants including cardiologists, emergency physicians, residents and medical students. We also evaluate the diagnostic power for STEMI and NSTEMI by DLM and conventional cardiac troponin I (TnI).
Methods

Study design and setting

This was a single center, retrospective and case-control study. The data were provided from the Tri-Service General Hospital, Taipei, Taiwan, and the retrospective research was ethically approved by the institutional review board (IRB NO. 2-107-05-168). Our hospital has built an electronic health system for collecting ECGs and the records from January 1, 2012, to December 31, 2018.

Study population

The MI patients were collected from records at the cardiac catheterization lab; they had received coronary angiography (CAG) to rule in type I MI and to confirm the IRA in STEMI.[23] There were 1051 ECGs before primary percutaneous coronary intervention from 737 STEMI cases. For NSTEMI cases, ECG records were collected before CAG, and 697 ECGs from 287 NSTEMI cases were included in this study. Right side or posterior ECG records were excluded. Not-MI ECGs were collected from patients in the ED during the same period. Patients with a history of AMI or elevated TnI were excluded in the not-MI population. A total of 76,775 patients with 140,336 ECGs were defined as not-MI in this study. We divided these cases into development (80%) and validation (20%) cohorts by date. The ECGs in the development cohort were excluded in the validation cohorts. There were no overlapping patients between these two cohorts.

Data collection

ECG recordings were collected using a Philips 12-lead ECG machine (PH080A). The ECG signal was recorded in a digital format. The sampling frequency was 500 Hz with 10 seconds recorded in each lead. Patient characteristics and laboratory tests were collected from our electronic medical records. The timely nearest laboratory data were assigned for each ECG record. Because the ECG records were sometimes conducted in a related short time period, some ECGs from the same patients shared the same patient characteristics and laboratory data.

Implementation of the DLM

We have developed a DLM with 82-layer convolutional layers and an attention mechanism. The technology details, such as the model architecture, data augmentation, and model visualization, were described previously.[21] We used the same architecture to train two new deep learning models for MI detection and location analysis of STEMI. The first deep learning model was trained via full samples with 3 categories, including STEMI, NSTEMI, and not-MI, and the output of this model was a 3-class softmax output. The second deep learning model was trained via STEMI ECGs, and the output of this model was a 4-class softmax output for location analysis.

The standard input format of the DLM is a length of 1,024 numeric sequences, but the original length of our 12-lead ECG signals is 5,000. In the training process, we randomly cropped a length of 1,024 sequences as input. For the inference stage, 9 overlapping lengths of 1,024 sequences based on interval sampling were used to generate a prediction and averaged as the final prediction. Due to the scarcity of MI cases in our study, an oversampling process was implemented to ensure that rare samples were adequately recognized. The settings for the training model were as follows: (1) Adam optimizer with standard parameters ($\beta_1 = 0.9$ and $\beta_2 = 0.999$) and a batch size of 36 for optimization; (2) a learning rate of 0.001; and (3) a weight decay of $10^{-4}$. The 100th epoch model was used as the final model, and the presented performance in the validation set was only evaluated once.

Human-machine competition

We evaluated the performance of participant physicians using a sub-validation set. This sub-dataset included 174 STEMI, 138 NSTEMI, and 138 not-MI ECGs. In STEMI, based on the IRA, it was further classified into the left main coronary artery (LMCA), left anterior descending artery (LAD), left circumflex artery (LCx), or right coronary artery (RCA). There were six visiting staff, five residents, and six medical students who participated in the competition. The physicians had no possible access to patient information for further diagnosis. The responses they provided were entered into an online standardized data entry program. We calculated their sensitivities, specificities, and kappa values to compare their results with those of the DLM.

Statistical analysis

We presented their characteristics as the means and standard deviations, numbers of patients, or percentages, where appropriate. They were compared using either Student's t-test or the chi-square test, as appropriate. The statistical analysis was carried out using the software environment R version 3.4.4.

All analyses were based on ECGs but not patients. The described statistical analyses are shown in the Supplement, and we used a significance level of $p < 0.05$ throughout the analysis. The primary analysis was to evaluate the performance of the DLM and clinicians in MI and STEMI identification in a human-machine competition. Receiver operating characteristic (ROC) curve analysis and the area under curve (AUC) were applied to evaluate the competition results. Because the proportions of STEMI, NSTEMI, and not-MI cases were distorted in the competition set,
we reweighted the samples via the proportions in the hypothetical real world (0.1%, 0.2%, and 99.7% of STEMI, NSTEMI, and not-MI cases, respectively).

The baseline characteristics of cohorts.

The development and validation cohorts included records from 58,056 and 19,743 patients, respectively, and the characteristics and laboratory results are shown in the Table 1. Patients in the validation cohort were significantly older, had more comorbidities, an impaired estimated glomerular filtration rate, impaired alanine aminotransferase, lower TnI, higher glucose and low-density lipoprotein cholesterol than those in the development cohort. The development/validation cohorts consisted of 860/191, 559/138, and 109,904/30,432 ECGs from STEMI, NSTEMI, and not-MI, respectively. The LAD and RCA were the most commonly identified IRA in STEMI. Patients with STEMI were more likely to be male, more overweight, had more prior coronary artery disease (CAD), had higher TnI, and more impaired lipid profiles than those in the not-MI group.

The analysis of infarct related artery of STEMI

DLM also achieved the best global performance (kappa = 0.629) in the IRA of STEMI (Additional file 1: Figure S1). DLM achieved the best global performance (kappa = 0.629) for the IRA detection of STEMI. Both the LAD and RCA were easily detected by the DLM and clinicians. The LCx had troublesome interpretation. The LMCA was only correctly detected by medical students.

Consistency assessments of MI ECGs

Selected STEMI ECGs in the human-machine competition were shown in Figure 3. A typical ECG of STEMI was consistently detected as STEMI with an IRA of the LAD by both DLM and the clinicians (Figure 3 Case A). A total of 10 ECGs were detected as not-STEMI by DLM. Five of ten misdiagnosed by DLM were correctly recognized as STEMI by the best cardiologists (Figure 3 Case B), and the remainder were misdiagnosed by both DLM and the best cardiologists (Figure 3 Case C). DLM could identify ECGs as STEMI that expert cardiologists had misdiagnosed (Figure 3 Case D). Among 138 NSTEMI ECGs in the human-machine competition, 58 cases were detected as not-MI by DLM, with an accuracy of 58.0%, which was worse than the 75.4% accuracy of the best cardiologists. This was due to a more conservative MI diagnostic strategy by DLM. The specificity of 96.4% of DLM in 138 not-MI cases was much better than that of 82.6% and 64.5% of the two best cardiologists. After adjustment of the specificity, the misdiagnosis of NSTEMI cases by DLM was obviously less than that by cardiologists (Table 2). Nevertheless, DLM still offered the best performance in the detection of MI by ECG under the standardization of the best cardiologists.

ECG lead-specific analysis
ECG leads were specifically analyzed for the detection of STEMI and MI in the hypothetical real world (Additional file 1: Figure S2). Lead III, V2, aVL, and V3 demonstrated better performance than other leads for the detection of STEMI, with the AUC of 0.913, 0.913, 0.911, and 0.908, respectively. For the detection of MI, V4, Lead I, and V3 demonstrated better performance, with the AUC of 0.841, 0.825 and 0.825, respectively. Lead-specific PRROC curve on the detection of MI and STEMI in the hypothetical real world (Additional file 1: Figure S3), and on the IRA of STEMI (Additional file 1: Figure S4) were analyzed. Lead-specific PRROC curve analysis demonstrated the best performance for the detection STEMI with the AUC of 0.300 on aVL. Moreover, lead-specific PRROC curve analysis on the IRA of STEMI demonstrated the best performance for the LAD with the AUC of 0.970, 0.955, and 0.953 on V4, V2, and V3, respectively; that for the RCA yielded the AUC of 0.995, 0.978, and 0.966 on aVL, Lead III, and aVF, respectively.

Logistic regression analysis of MI, STEMI, and NSTEMI

The univariate and multivariate logistic regression analyses in the development cohort revealed that male, prior CAD, troponin I, hemoglobin, total cholesterol and low density lipoprotein were independent risk factors for the detection of MI, STEMI and NSTEMI (Additional file 1: Figure S5).

Diagnostic value analysis

We evaluated the algorithm performance after adjusting for significant patient characteristics, disease histories, and laboratory results to ensure consistency across a wide range of putative confounding variables in the validation cohort. DLM had significantly better performance than the use of troponin I alone to detect STEMI with the AUC of 0.996. The corresponding sensitivity and specificity are 98.4% and 96.9%, respectively. However, the use of troponin I alone had significantly better performance than DLM to detect NSTEMI. The AUC of combined DLM and the first recorded TnI for the detection of NSTEMI were increased to 0.978, with the corresponding sensitivity and specificity are 91.6% and 96.7%, respectively, which was better than that of DLM (0.877) or TnI (0.949) alone (Figure 4). It is enough to detect STEMI using the DLM alone, and the addition of patient characteristics did not significantly improve the performance. However, troponin I was found to improve the diagnostic accuracy for NSTEMI, and the improvement was better than the combination of all additional characteristics (Additional file 1: Figure S6).

Discussion

In this study, we established a DLM to precisely detect STEMI and MI through ECG analysis, which applied a deep convolutional network to extract notable ECG features with a development cohort of more than 110,000 ECGs. All MI cases were validated by coronary angiography with the identification of the corresponding IRA in patients with STEMI. Most importantly, our DLM demonstrated better performance than that of clinicians for STEMI and MI detection with high sensitivities of 89.7%/83.7% and specificities of 94.6%/95.7%.

The application of deep learning technology in the cardiovascular field for arrhythmias, dyskalemia, and valvular heart disease had become popularized recently.[19-21, 27-29] However, no large scale study has been designed to apply deep learning technology for MI detection. Previous DLMs for MI detection by ECG were analyzed mainly from the Physikalisch-Technische Bundesanstalt (PTB) diagnostic ECG Database.[30, 31] These studies may be limited because they did not have further validation. Moreover, the comparison between DLM and human experts was lacking. In comparison with previous studies, we enrolled the largest clinically validated ECG records for the development and validation processes. Additionally, we further confirmed the role of TnI in assisting with NSTEMI detection by our DLM. All these results point out the strengths of the current study.

The sensitivities and specificities for STEMI/MI detection by DLM were better than those of the participating experts. ECG is the timeliest tool among all objective detection methods for MI. However, the low sensitivity and disagreement in interpreting ECGs between physicians are issues for detecting STEMI and NSTEMI. The sensitivity of manual interpretation for MI detection using a 12-lead ECG is only from 61 to 74% with the specificity ranging from 72 to 89.0%.[32-35] In contrast to previously prehospital computer algorithm interpretation for STEMI with the sensitivity of approximately 69%. [36-38] Our DLM provides extraordinary performance, which supports decision-making systems in clinical practice.

With the aid of the first recorded TnI, DLM exhibited excellent diagnostic yield with an AUC of 0.978 for NSTEMI detection, which was significantly better than those of DLM or TnI alone, with the AUC of 0.877 and 0.949, respectively. The universal diagnosis of NSTEMI is derived from the clinical presentation, 12-lead ECG, and cardiac troponin.[2] To date, biomarker measurement for cardiomyocyte injury, preferably high-sensitivity cardiac troponin (hsTnI), is mandatory in all patients with suspected NSTEMI due to the high sensitivity and specificity.[2, 39, 40] However, several concerns should be considered in current practice. First, the guidelines suggest the second cardiac troponin assessment to be performed 1-3 hours after the first blood test in unconfirmed cases. Repeated time-costly laboratory tests might delay the diagnosis. Second, cardiac troponin might be perturbed in some clinical conditions other than MI. Combined with the information of the first recorded TnI, DLM allows rapid and powerful NSTEMI detection in high-risk patients.

DLM can objectively conclude highly suspected STEMI based on analyzing and learning a large amount of ECG data. Moreover, subtle changes in the ECG presentation in the acute and early phases of STEMI that were easily missed by clinicians could be correctly recognized. Interestingly, there were two main characteristics of 10 STEMI ECGs that were unrecognized by DLM, including an infarct Q wave with ST elevation in indicated
leads and an atypical ST-T change in reciprocal leads related to old MI. Thus, information regarding previously available ECGs and the history of old MI may be needed to further strengthen the capacity of DLM in STEMI detection.

Regarding NSTEMI detection, DLM showed less sensitivity than cardiologists. Several points should be clarified. Among 58 NSTEMI ECGs unrecognized by DLM, there were several atypical ECG presentations, including intraventricular conduction disorders, ventricular hypertrophy, poor R wave progression, or baseline variant.[41] Even experienced cardiologists could not identify some of these ECGs. Moreover, overdiagnosis of NSTEMI by ECG is commonplace in clinical practice, which may partly explain the high sensitivity and low specificity of the performance of the physicians in this study.[42, 43] With the aid of DLM with high specificity in the detection of NSTEMI, clinicians could easily exclude NSTEMI, which reduces subsequent lab tests and ED observation time and guides clinicians to differentiate it from other diagnosis with clinical presentations at that time. As a result, it is worthwhile to increase the ECG training data along with the first-record cardiac biomarkers to enhance the capacity of DLM in NSTEMI detection in the future.

Our novel DLM has several potential clinical and educational applications. First, DLM could be incorporated into ECG machines in ambulances or remote areas to facilitate telemedicine and shorten the decision time for reperfusion therapies. Second, the developed model can be applied to a wearable device for MI detection, especially for patients with an extremely high risk of atherosclerotic cardiovascular disease. Third, DLM provides decision support and a high-risk alarm system for MI and will help to reduce medical errors in the ICU or ED resulting from intense time pressure or heavy workload and harried staff during the busy working hours. Finally, the application of a DLM in medical education is probably a future trend. [44] Young physicians and medical students could be trained and tested for the detection of MI with currently developed explainable DLM. Accordingly, our DLM exhibits diagnostic and educational benefits and promotes healthcare for cardiovascular disease in the near future.

4.1 Limitations

Some limitations of this study should be mentioned. First, the human-machine competition was based on a well-design retrospective study. A real-world prospective study should be conducted to verify the clinical impact of DLM. Moreover, only eleven clinicians participated in the competition with DLM.[45] Although their performance in MI detection was relative consistent with that of the previous studies, comparisons should be made with more experts to confirm the superiority of DLM. Second, the studied patients were only enrolled from one academic medical center although the diagnosis and management of MI was followed up according to the guidelines. Multicenter validation is needed to confirm the value and application of this study. Third, the number of NSTEMI cases was not as large as that of STEMI cases, which may limit the capacity for NSTEMI detection with our DLM. Finally, patients only in the ED with both an ECG and a diagnosis of MI were enrolled in this study, which may have led to selection bias and constrained the generalizability of the results.

Conclusion

We established an optimal DLM to detect STEMI and discriminate between MI and not-MI based on 12 lead ECG with an accuracy better than that of clinicians. Integration of a DLM may assist frontline physicians to recognize MI, especially STEMI, in a timely and precise manner to prevent delay or misdiagnosis and thereby provide prompt reperfusion therapy. Further prospective validation with prehospital and in-hospital ECG tests are needed to confirm the performance of our DLM.

Abbreviations

AMI: Acute myocardial infarction; ECG: Electrocardiogram; STEMI: ST-segment elevation myocardial infarction; NSTEMI: non-ST-segment elevation myocardial infarction; IRA: infarct-related artery; DLM: Deep learning model; TnI: conventional cardiac troponin I; CAG: Coronary angiography; LMCA: Left main coronary artery; LAD: Left anterior descending artery; LCx: Left circumflex artery; RCA: Right coronary artery; CAD: Coronary artery disease; ROC: Receiver operating characteristic; AUC: the area under curve; PRROC: Precision-recall receiver operating characteristic; PTB: Physikalisch-Technische Bundesanstalt; hsTnI: High-sensitivity cardiac troponin I.

Declarations

Ethics approval and consent to participate

The institutional review boards of the Tri-Service General Hospital, Taipei, Taiwan granted a waiver of consent for this study (IRB NO. 2-107-05-168).

Consent for publication

Not applicable.

Availability of data and materials
The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available because they contain information that could compromise research participant privacy.

**Competing interests**

The authors declare that they have no competing interests.

**Funding**

The work was supported by the Ministry of Science and Technology, Taiwan (MOST 108-2314-B-016-001 to C. Lin, MOST 109-2314-B-016-026 to C. Lin), the National Science and Technology Development Fund Management Association, Taiwan (MOST 108-3111-Y-016-009 and MOST 109-3111-Y-016-002 to C. Lin), and the Cheng Hsin General Hospital, Taiwan (CHNDMC-109-19 to C. Lin).

**Authors’ contributions**

All authors made substantial contributions to the study, were involved in critically revising it for intellectual content and accuracy, and approved the final version of the article submitted for publication. WCL, CSL CST and TPT conceived and designed the study and drafted the article. CCC, JTL, WSL and SMC acquired and analyzed results. YSL and CCL was responsible for the statistical analyses. CL takes responsibility for the paper as a whole.

**Funding**

The work was supported by the Ministry of Science and Technology, Taiwan (MOST 108-2314-B-016-001 to C. Lin, MOST 109-2314-B-016-026 to C. Lin), the National Science and Technology Development Fund Management Association, Taiwan (MOST 108-3111-Y-016-009 and MOST 109-3111-Y-016-002 to C. Lin), and the Cheng Hsin General Hospital, Taiwan (CHNDMC-109-19 to C. Lin).

**Acknowledgments**

None

**Author details**

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

**References**

1. Roth GA, Huffman MD, Moran AE, Feigin V, Mensah GA, Naghavi M, Murray CJ: Global and regional patterns in cardiovascular mortality from 1990 to 2013. *Circulation* 2015, 132(17):1667-1678.
2. Collet J-P, Thiele H, Barbato E, Barthélémy O, Bauersachs J, Bhatt DL, Dendale P, Dorobantu M, Edvardsen T, Folliguet T: 2020 ESC Guidelines for the management of acute coronary syndromes in patients presenting without persistent ST-segment elevation. *European Heart Journal* 2020.
3. Ibanez B, James S, Agewall S, Antunes MJ, Bucciarelli-Ducci C, Bueno H, Caforio ALP, Crea F, Goudevenos JA, Halvorsen S 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). *European Heart Journal* 2018, 39(2):119-177.
4. Mullvain R, Saman DM, Rostvedt A, Landgren P: ECG-to-Decision Time Impact on 30-Day Mortality and Reperfusion Times for STEMI Care. *Critical pathways in cardiology* 2018, 17(1):19-24.
5. Scholz KH, Maier SKG, Maier LS, Lengenfelder B, Jacobshagen C, Jung J, Fleischmann C, Werner GS, Olbrich HG, Ott R et al: Impact of treatment delay on mortality in ST-segment elevation myocardial infarction (STEMI) patients presenting with and without haemodynamic...
6. Tsukii T, Sakakura K, Taniguchi Y, Yamamoto K, Wada H, Momomura SI, Fujita H: Determinants of short and long door-to-balloon time in current primary percutaneous coronary interventions. Heart and vessels 2018, 33(5):498-506.

7. Li YH, Wang YC, Wang YC, Liu JC, Lee CH, Chen CC, Hsieh IC, Kuo FY, Huang WC, Sung SH et al: 2018 Guidelines of the Taiwan Society of Cardiology, Taiwan Society of Emergency Medicine and Taiwan Society of Cardiovascular Interventions for the management of non ST-segment elevation acute coronary syndrome. Journal of the Formosan Medical Association = Taiwan yi zhi 2018, 117(9):766-790.

8. Thygesen K, Alpert JS, Jaffe AS, Simoons ML, Chaitman BR, White HD, Thygesen K, Alpert JS, White HD, Jaffe AS et al: Third universal definition of myocardial infarction. Journal of the American College of Cardiology 2012, 60(16):1581-1598.

9. 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(3):579-582.

10. Schull MJ, Vermeulen MJ, Stukel TA: The risk of missed diagnosis of acute myocardial infarction associated with emergency department volume. Ann Emerg Med 2006, 48(6):647-655.

11. Masoudi FA, Magid DJ, Vinson DR, Tricomi AJ, Lyons EE, Crouse L, Ho PM, Peterson PN, Rumsfeld JS: 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(15):1565-1571.

12. Wu J, Gale CR, 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. European heart journal Acute cardiovascular care 2018, 7(2):139-148.

13. 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(2):196-205.

14. Schenkel S: Promoting patient safety and preventing medical error in emergency departments. Academic Emergency Medicine 2000, 7(11):1204-1222.

15. Estava A, Kuprel B, Novoa R: Dermatologist level classification of skin cancer with deep neural networks [J]. Nature 2017, 542:115.

16. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, Van Der Laak JA, Van Ginneken B, Sánchez CI: A survey on deep learning in medical image analysis. Medical image analysis 2017, 42:60-88.

17. Bejnordi BE, Veta M, Van Diest PJ, Van Ginneken B, Karssemeijer N, Litjens G, Van Der Laak JA, Hermans M, Manson QF, Balkenhol M: Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. Jama 2017, 318(22):2199-2210.

18. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, Venugopalan S, Widner K, Madams T, Cuadros J: Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. Jama 2016, 316(22):2402-2410.

19. 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. Nature medicine 2019, 25(1):65.

20. Galloway CD, Valys AV, Shreibati JB, Treiman DL, Petterson FL, Gundotra VP, Albert DE, Attia ZI, Carter RE, Asirvatham SJ: Development and validation of a deep-learning model to screen for hyperkalemia from the electrocardiogram. JAMA cardiology 2019, 4(5):428-436.

21. Lin C-S, Lin C, Fang W-H, Hsu C-J, Chen S-J, Huang K-H, Lin W-S, Tsai C-S, Kuo C-C, Chau T: A Deep-Learning Algorithm (ECG12Net) for Detecting Hypokalemia and Hyperkalemia by Electrocardiography: Algorithm Development. JMIR medical informatics 2020, 8(3):e15931.

22. Goldberger AL, Amaral LA, Glass L, Hausdorff JM, Ivanov PC, Mark RG, Mietus JE, Moody GB, Peng CK, Stanley HE: PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. Circulation 2000, 101(23):e215-e220.

23. Bax JJ, Baumgartner H, Cecconi M, Dean V, Fagard R, Funck-Brentano C, Hasdai D, Hoes A, Kirchhof P, Knuuti J: Third universal definition of myocardial infarction. Journal of the American College of Cardiology 2012, 60(16):1581-1598.

24. Jernberg T: Swedeheart annual report 2015. Karolinska university hospital, Huddinge 2016, 14186.

25. McManus DD, Gore J, Yarzebski J, Spencer F, Lessard D, Goldberg RJ: Recent trends in the incidence, treatment, and outcomes of patients with STEMI and NSTEMI. The American journal of medicine 2011, 124(1):40-47.

26. Yin WH, Lu TH, Chen KC, Cheng CF, Lee JC, Liang FW, Huang YT, Yang LT: The temporal trends of incidence, treatment, and in-hospital mortality of acute myocardial infarction over 15years in a Taiwanese population. International journal of cardiology 2016, 209:103-113.

27. Attia ZI, Kapa S, Lopez-Jimenez F, McKie PM, Ladewig DJ, Satam G, Pellikka PA, Enriquez-Sarano M, Noseworthy PA, Munger TM: Screening for cardiac contractile dysfunction using an artificial intelligence—enabled electrocardiogram. Nature medicine 2019, 25(1):70-74.

28. Kwon J, Lee SY, Jeon KH, Lee Y, Kim KH, Park J, Oh BH, Lee MM: Deep Learning—Based Algorithm for Detecting Aortic Stenosis Using Electrocardiography. Journal of the American Heart Association 2020, 9(7):e014717.

29. Kwon J-M, Kim K-H, Akkus Z, Jeon K-H, Park J, Oh B-H: Artificial intelligence for detecting mitral regurgitation using electrocardiography. Journal of Electrocardiology 2020.
30. Gupta A, Huerta E, Zhao Z, Moussa I: Deep Learning for Cardiologist-level Myocardial Infarction Detection in Electrocardiograms. arXiv preprint arXiv:191207618 2019.

31. Strodthoff N, Strodthoff C: Detecting and interpreting myocardial infarction using fully convolutional neural networks. Physiological measurement 2019, 40(1):015001.

32. 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. American heart journal 2006, 152(4):742-748.

33. McClelland AJ, Owens CG, Menown IB, Lown M, Adgey AA: Comparison of the 80-lead body surface map to physician and to 12-lead electrocardiogram in detection of acute myocardial infarction. The American journal of cardiology 2003, 92(3):252-257.

34. McCabe JM, Armstrong EJ, Ku I, Kulkarni A, Hoffmayer KS, Bhave PD, Waldo SW, Hsue P, Stein JC, Marcus GM et al: Physician accuracy in interpreting potential ST-segment elevation myocardial infarction electrocardiograms. Journal of the American Heart Association 2013, 2(5):e000268.

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

36. Pollehn T, Brady W, Perron A, Morris F: The electrocardiographic differential diagnosis of ST segment depression. Emergency medicine journal: EMJ 2002, 19(2):129.

37. Moynihan R, Doust J, Henry D: Preventing overdiagnosis: how to stop harming the healthy. Bmj 2012, 344.

38. Spence D: Bad medicine: non-ST elevation myocardial infarction. Bmj 2013, 347:f5967.

39. Chartrand G, Cheng PM, Vorontsov E, Drozdal M, Turcotte S, Pal CJ, Kadoury S, Tang A: Deep learning: a primer for radiologists. Radiographics 2017, 37(7):2113-2131.

40. Willems JL, Abreu-Lima C, Arnaud P, van Bemmel JH, Brochet C, Degani R, Denis B, Gehring J, Graham I, van Herpen G: The diagnostic performance of computer programs for the interpretation of electrocardiograms. New England Journal of Medicine 1991, 325(25):1767-1773.

Tables

Table 1 Corresponding patient characteristics and laboratory results of STEMI, NSTEMI, and not-MI ECGs in development cohort and validation cohort.
|                      | Development cohort | Validation cohort | p-value# |
|----------------------|--------------------|-------------------|----------|
|                      | STEMI (n = 860)    | NSTEMI (n = 559)  | not-MI (n = 109904) | p-value | STEMI (n = 191) | NSTEMI (n = 138) | not-MI (n = 30432) | p-value |
| **STEMI location**   |                    |                   |                      |         |                |                  |                      |         |
| STEMI-LMCA           | 21(2.4%)           |                   |                      |         | 3(1.6%)        |                   |                      |         |
| STEMI-LAD            | 420(48.8%)         |                   |                      |         | 105(55.0%)     |                   |                      |         |
| STEMI-LCx            | 87(10.1%)          |                   |                      |         | 11(5.8%)       |                   |                      |         |
| STEMI-RCA            | 332(38.6%)         |                   |                      |         | 72(37.7%)      |                   |                      |         |
| **Gender (Male)**    | 688(83.8%)         | 420(76.2%)        | 55453(50.5%)         | <0.001  | 150(82.9%)     | 84(62.2%)         | 15484(50.9%)        | <0.001  |
| **Age (years)**      | 61.8±13.8          | 64.3±13.8         | 60.9±19.6            | <0.001  | 62.9±14.6      | 65.9±13.7         | 62.6±20.2           | 0.165   |
| **BMI (kg/m²)**      | 25.9±4.5           | 24.4±3.9          | 24.5±8.8             | 0.009   | 26.9±4.7       | 25.0±4.9          | 24.5±6.0            | 0.043   |
| **Disease history**  |                    |                   |                      |         |                |                  |                      |         |
| CAD                  | 197(24.0%)         | 188(34.1%)        | 20275(18.4%)         | <0.001  | 133(73.5%)     | 95(70.4%)         | 7439(24.4%)         | <0.001  |
| HF                   | 50(6.1%)           | 66(12.0%)         | 8099(7.4%)           | <0.001  | 21(11.6%)      | 33(24.4%)         | 2972(9.8%)          | <0.001  |
| DM                   | 176(21.4%)         | 187(33.9%)        | 25429(23.1%)         | <0.001  | 39(21.5%)      | 50(37.0%)         | 7675(25.2%)         | 0.004   |
| HTN                  | 249(30.3%)         | 243(44.1%)        | 42081(38.3%)         | <0.001  | 67(37.0%)      | 83(61.5%)         | 14177(46.6%)        | <0.001  |
| CKD                  | 68(8.3%)           | 101(18.3%)        | 9929(9.0%)           | <0.001  | 8(4.4%)        | 26(19.3%)         | 2332(7.7%)          | <0.001  |
| Lipidemia            | 198(24.1%)         | 219(39.7%)        | 30087(27.4%)         | <0.001  | 34(18.8%)      | 53(39.3%)         | 8579(28.2%)         | <0.001  |
| COPD                 | 85(10.4%)          | 62(11.3%)         | 21600(19.7%)         | <0.001  | 24(13.3%)      | 19(14.1%)         | 7090(23.3%)         | <0.001  |
| **Laboratory test**  |                    |                   |                      |         |                |                  |                      |         |
| Na (mEq/L)           | 137.3±3.2          | 136.9±3.6         | 136.6±4.5            | <0.001  | 137.1±2.7      | 135.9±3.4         | 135.8±4.7           | 0.005   |
| K (mEq/L)            | 3.9±0.6            | 4.0±0.6           | 3.9±0.5              | 0.006   | 3.8±0.5        | 4.0±0.6           | 3.9±0.5             | 0.008   |
| eGFR (mL/min)        | 74.2±26.3          | 63.8±30.7         | 82.5±37.0            | <0.001  | 74.2±26.5      | 64.3±37.4         | 81.0±35.0           | <0.001  |
| Cr (mg/dl)           | 1.3±1.3            | 1.9±2.2           | 1.3±1.6              | <0.001  | 1.3±0.9        | 2.3±2.6           | 1.2±1.3             | <0.001  |
| CK (ng/ml)           | 389.8±650.7        | 296.1±325.4       | 131.7±409.0          | <0.001  | 348.9±597.0    | 252.5±310.7       | 122.5±306.9         | <0.001  |
| TnI (ng/ml)          | 60.6±598.7         | 224.8±1121.7      | 0.0±0.0              | <0.001  | 4.8±16.6       | 2.7±6.5           | 0.0±0.0             | <0.001  |
| WBC (10⁹/ul)         | 11.1±3.6           | 8.8±3.0           | 8.9±4.5              | <0.001  | 11.2±3.2       | 9.3±2.8           | 8.8±4.6             | <0.001  |
| Hb (gm/dl)           | 14.6±1.9           | 13.2±2.4          | 12.9±2.3             | <0.001  | 14.7±1.7       | 13.2±2.7          | 12.9±2.3            | <0.001  |
| PLT (10⁹/ul)         | 228.5±64.0         | 221.0±74.6        | 227.0±81.9           | 0.425   | 228.4±90.7     | 216.5±52.9        | 210.1±74.9          | 0.015   |
| GLU (mg/dl)          | 193.9±85.3         | 219.4±126.3       | 198.7±114.8          | 0.631   | 166.0±13.1     | 215.8±85.5        | 241.1±128.5         | 0.462   |
| AST (U/L)            | 54.0±85.3          | 45.6±104.5        | 32.6±81.3            | <0.001  | 51.3±65.0      | 36.4±37.1         | 33.0±91.3           | 0.075   |
### Table 2 Maximum sensitivity of the AI system for a specific specificity.

| Revise item          | Sensitivity<sup>b</sup> (STEMI) | Sensitivity<sup>c</sup> (NSTEMI) | Specificity<sup>d</sup> |
|----------------------|---------------------------------|---------------------------------|-------------------------|
| AI system (original) | 0.000                           | 164/174 (94.3%)                 | 133/138 (96.4%)         |
| CV-V3                | 0.450                           | 166/174 (95.4%)                 | 114/138 (82.6%)         |
| AI system (Specificity = 82.6%) | 0.450                           | 166/174 (95.4%)                 | 114/138 (82.6%)         |
| CV-V11               | 0.612                           | 166/174 (95.4%)                 | 89/138 (64.5%)          |

<sup>a</sup>: The revised item is used to modify the probability of non-MI given by DLM. For example, if an original probability of STEMI/NSTEMI/non-MI is 0.220/0.310/0.470, then the prediction is defined as not-MI according to the largest probability. However, the revised item is used to let DLM become more sensitive, which is used to modify the probability of not-MI as 0.470 - 0.450 = 0.020 as the first situation. Therefore, the new prediction of this case is defined as NSTEMI according to the largest revised probability (0.220/0.310/0.020).

<sup>b</sup>: The sensitivity of STEMI is defined as the percentage of STEMI cases that are correctly identified as STEMI.

<sup>c</sup>: The sensitivity of NSTEMI is defined as the percentage of NSTEMI cases that are correctly identified as STEMI/NSTEMI.

<sup>d</sup>: The specificity is defined as the percentage of not-MI cases that are correctly identified as not-MI.

TEMI classified into the LMCA, LAD, RCA and LCx are shown.

### Figures
Figure 1

Performance comparison of MI and STEMI recognition in the human-machine competition. The curves were made by the predictions of DLM. The top and bottom panels show the performance for MI and STEMI detection, respectively. The left and middle panels show the ROC curves in the competition set (STEMI = 174, NSTEMI = 138, and not-MI = 138) and the hypothetical real world (STEMI = 0.1%, NSTEMI = 0.2%, and not-MI = 99.7%), respectively. The right panels show the PRROC curves in the hypothetical real world. The red and blue points represent visiting staffs and residents, respectively. The triangle and square marks represent the emergency physicians and cardiologists, respectively.
Figure 2

Consistency analysis of the DLM and human experts and their performance rankings in the human-machine competition. Global performance rankings based on the 3-class kappa values. The abbreviations V(X), R(X), and M(X) denote the visiting staff, residents, and medical students with (X) years of experience, respectively. Number consistency heatmap colored according to the values. The values in each cell are the kappa values of each pair.
Figure 3

Selected STEMI ECGs of consistent and inconsistent assessments given by DLM and human experts. Case A shows a typical ECG of STEMI with ST segment elevation and reciprocal lead depression that was consistently detected as anterior wall STEMI with an IRA related to the LAD by both DLM and the clinicians; Case B was misdiagnosed by DLM but was correctly detected as STEMI by cardiologists; Case C revealed an old anterior wall MI that caused misdetection by both DLM and clinicians; Case D was an early-phase STEMI misdiagnosed by cardiologists but correctly detected by DLM.
Figure 4

Comparison of the diagnostic value between DLM based on ECGs and cardiac enzymes in the validation cohort. The ROC curves were generated from the logistic regression analysis using the validation cohort. The p values were the comparison between each AUC and the AUC of DLM.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementarymaterial.docx
- Graphicalabstract.png
- FigS1supp.pdf
- FigS2supp.pdf
- FigS3supp.pdf
- FigS4supp.pdf
- FigS5supp.pdf
- FigS6supp.pdf