Deep Learning Model for Screening Sepsis Using Electrocardiography

Joon-myoung Kwon (kwonjm@sejongh.co.kr)
Artificial Intelligence and Big Data Research Center, Sejong Medical Research Institute, Bucheon, South Korea https://orcid.org/0000-0001-6754-1010

Ye Rang Lee
Mediplex Sejong Hospital

Min-Seung Jung
Medical AI, Co.

Yoon-Ji Lee
Medical AI, Co.

Yong-Yeon Jo
Medical AI, Co.

Da-Young Kang
Mediplex Sejong Hospital

Soo Youn Lee
Mediplex Sejong Hospital

Yong-Hyeon Cho
Medical AI, Co.

Jae-Hyun Shin
Medical AI, Co.

Jang-Hyeon Ban
Bodyfriend, Co.

Kyung-Hee Kim
Mediplex Sejong Hospital

Original research

Keywords: Sepsis, Shock, Septic, Infections, Electrocardiography, Deep Learning, Artificial Intelligence

Posted Date: February 5th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-186976/v1

Abstract

Background: Sepsis is a life-threatening organ dysfunction and is a major healthcare burden worldwide. Although sepsis is a medical emergency that requires immediate management, it is difficult to screen the occurrence of sepsis. In this study, we propose an artificial intelligence based on deep learning-based model (DLM) for screening sepsis using electrocardiography (ECG).

Methods: This retrospective cohort study included 46,017 patients who admitted to two hospitals. 1,548 and 639 patients underwent sepsis and septic shock. The DLM was developed using 73,727 ECGs of 18,142 patients and internal validation was conducted using 7,774 ECGs of 7,774 patients. Furthermore, we conducted an external validation with 20,101 ECGs of 20,101 patients from another hospital to verify the applicability of the DLM across centers.

Results: During the internal and external validation, the area under the receiver operating characteristic curve (AUC) of an DLM using 12-lead ECG for screening sepsis were 0.901 (95% confidence interval 0.882–0.920) and 0.863 (0.846–0.879), respectively. During internal and external validation, AUC of an DLM for detecting septic shock were 0.906 (95% CI = 0.877–0.936) and 0.899 (95% CI = 0.872–0.925), respectively. The AUC of the DLM for detecting sepsis using 6-lead and single-lead ECGs were 0.845–0.882. A sensitivity map showed that the QRS complex and T wave was associated with sepsis. Subgroup analysis was conducted using ECGs from 4,609 patients who admitted with infectious disease, The AUC of the DLM for predicting in-hospital mortality was 0.817 (0.793–0.840). There was a significant difference in the prediction score of DLM using ECG according to the presence of infection in the validation dataset (0.277 vs 0.574, p<0.001), including severe acute respiratory syndrome coronavirus 2 (0.260 vs 0.725, p=0.018).

Conclusions: The DLM demonstrated reasonable performance for screening sepsis using 12-, 6-, and single-lead ECG. The results suggest that sepsis can be screened using not only conventional ECG devices, but also diverse life-type ECG machine employing the DLM, thereby preventing irreversible disease progression and mortality.

Introduction

Sepsis is a life-threatening organ dysfunction caused by a dysregulation of the host response to infection, and is a major healthcare problem worldwide [1, 2]. A total of 48.9 million incident cases of sepsis were recorded worldwide and 11.0 million sepsis-related deaths were reported, representing 19.7% of all global deaths [2, 3]. Although the mortality rate has decreased by 52.8% over the past 20 years, the incidence of sepsis is increasing, likely reflecting aging populations with more comorbidities [3]. Because sepsis is a medical emergency that requires immediate treatment and resuscitation, early recognition is a cornerstone for preventing disease progression and death [2].

Vital signs and blood tests are needed to screen and diagnose sepsis [1]. Vital signs are measured by the medical staff with intervals, and blood tests require infrastructure for blood sampling and analysis. For

this reason, it is difficult to monitor the occurrence of sepsis in real time in hospitals. Sepsis has its highest burden in areas with a lower sociodemographic index, and these areas lack medical resources for screening and diagnosis of sepsis as well as its treatment.[4] Home monitoring for the deterioration of infected patients and screening for sepsis is critical for a proper allocation of scarce medical resources in a pandemic, such as the current Coronavirus Disease-19 (COVID-19) pandemic. However, the current method for the screening of sepsis using vital signs and laboratory exams is limited in daily living situations and remote monitoring.

A low-cost and widely available method for screening sepsis patients has important therapeutic implications. An ECG is a noninvasive test that can be monitored in real time, and diverse wearable and life-type devices have been developed for remote monitoring and transfer. In previous studies, approximately 50% of the patients who were diagnosed with sepsis exhibited signs of myocardial dysfunction, and a prolonged duration and decreased amplitude of the QRS complex were reported in sepsis patients [5–8]. Artificial intelligence based on deep learning have been used in diverse medical domains, and a deep learning model (DLM) has been applied to the diagnosis of heart failure, pulmonary hypertension, valvular heart disease, electrolyte imbalance, and anemia using ECG[9–13]. As the most important advantage of a DLM over conventional statistical methods, it can extract the implications from the data, and thus it can diagnose or predict diseases by capturing non-linear and subtle changes in an ECG[14]. Thus, we hypothesized that we could develop a DLM to screen sepsis patients using an ECG. In this study, we developed and validated a DLM for screening sepsis using electrocardiography (ECG).

Methods

Study design and population

We conducted a retrospective multicenter cohort study in two hospitals. The study population applied adult patients who were admitted to two hospitals and underwent at least one standard 10-s 12-lead ECG during the study period. We excluded individuals who had missing ECG values. Data from Sejong General Hospital (SGH) were used to develop and validate the DLM. The patients admitted to SGH during the study period (October 2016 to November 2020) were randomly split into development (70%) and internal validation (30%) datasets (Fig. 1). Data from Mediplex Sejong Hospital (MSH) during the study period (March 2017 to November 2020) were only used for external validation, which confirmed that the developed DLA was robust across different hospitals. Patients with SGH and MSH were also divided exclusively. As the purpose of the validation dataset was to assess the accuracy of the DLM, we used only one ECG from each patient for the internal and external validation datasets, the time-closest to the sepsis time which was confirmed by critical care medicine physicians.

This study was approved by the institutional review board (IRB) of SGH (2020 – 0541) and MSH (2020 – 149). Clinical data, including an ECG, age, sex, admission note, vital signs, and laboratory exam results, were extracted from electronic health records of both hospitals after anonymization. Both IRBs waived

the need for informed consent because this was a retrospective study using fully anonymized data and the possibility of minimal harm.

Predictor variable

The predictor variable was only an ECG. Digitally stored 12-lead ECG data had 500 data points per second (500 Hz) at each lead for 10 s. In other words, one ECG data has 60,000 values. Because there were more artifacts at the beginning and end of the ECG, we removed 1 s of data at the beginning and end of the ECG. We created a dataset using 12-, 6-, and single-lead ECG datasets. We created a 12-lead ECG dataset using the entire 12-lead ECG data (12 × 4000). We also created a 6- and single-lead ECG dataset from partial datasets from the 12-lead ECG data. The 6-lead ECG dataset was made from limb 6-lead (I, II, III, aVL, aVR, aVF), and a single-lead ECG dataset was made from lead I. We selected these leads because they can be measured using diverse wearable and life-type ECG devices.

Endpoints

The primary endpoint of this study was the presence of sepsis. Sepsis was defined by the Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). Three critical care medicine physicians reviewed the medical records of the study population, including the admission notes, laboratory exam results, vital signs, and drug administration data, rapid response team's progression note, for labeling the presence and time range of sepsis. The secondary endpoint was the presence of septic shock. Septic shock was defined based on Sepsis-3.

For the primary endpoint, sepsis in patients with suspected infectious disease, we labeled the ECG within the time range of the sepsis as *Sepsis*, and the ECG outside of the time range of sepsis as *non-Sepsis*. In patients who had no history of infectious disease during hospitalization, we also labeled all ECG as *non-Sepsis*. For the secondary endpoint, septic shock, we labeled the ECG within the time range of septic shock as *Septic shock* and the other ECG as *non-Septic shock*, in the same manner.

Development of DLM for detecting sepsis using ECG

We developed a DLM based on a convolutional neural network, particularly residual neural networks (Fig. 2). The residual neural network contained a skip connection to avoid the problem of vanishing gradients. In a residual block with four stages, two convolutional layers and two batch normalization layers were repeated. There were two flattened layers in the DLM. The last layer of the seventh residual block was connected to a flattened layer, which was fully connected to a one-dimensional (1D) layer composed of neural nodes. The second fully connected 1D layer was connected to the output node, which is composed of two nodes. The value of the output nodes of the DLM represents the probability of the endpoints. The output node of the DLM uses a softmax function as an activation function.

Statistical analysis

At each input ECG of the validation data, the DLM calculated the probability of sepsis within the range of zero (non-sepsis) to 1 (sepsis). To confirm the accuracy of the DLM, we compared the probability

calculated by the DLM with the presence of sepsis (ground truth) in the internal and external validation datasets. For this purpose, we used the area under the receiver operating characteristic curve (AUC), sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). We confirmed the cut-off point from Youden's J statistics in the development dataset and applied the point to validate the performance of internal and external validation.[15] As a comparative measure, we used C-reactive protein (CRP) and body temperature abnormality (difference between measured body temperature and 36.5°C) to screen sepsis and septic shock.

Continuous variables were presented as the mean values (standard deviation, SD) and compared using the unpaired Student's *t*-test or Mann-Whitney U test. Categorical variables were expressed as the frequencies and percentages and compared using the $\chi 2$ test. The exact 95% confidence intervals (CIs) were used for all measures of the diagnostic performance, except for the AUC. The CIs for the AUC were determined based on the Sun and Su optimization of the De-long method using the pROC package in R (R Foundation for Statistical Computing, Vienna, Austria). A significant difference in patient characteristics was defined as a two-sided p value of < 0.05. Statistical analyses were computed using R software, version 3.4. In addition, we used PyTorch's open-source software library as the backend and Python (version 3.6) for the analysis.[16]

Visualizing the developed DLM for interpretation

To compare the findings from the developed DLM with current medical knowledge, we used a sensitivity map using the saliency method.[17, 18] The map showed the region that had a significant effect on the decision of the DLM. The sensitivity map was computed based on the first-order gradients of the classifier probabilities with respect to the input signals; if the probability of a classifier is sensitive to a specific region of the signal, the region would be important in the decision of the DLM. Using this method, we verified the region of ECG that was correlated with sepsis. We used a gradient class activation map (Grad-CAM) as a sensitivity map. We also confirmed the variable importance of the ECG features in the conventional statistical method (logistic regression) and the machine learning methods (random forest and deep learning). We calculated the variable importance of logistic regression, random forest, and deep learning based on a difference in deviance, mean degreased Gini, and Garson's relative importance, respectively.

Verifying DLM performance to predict in-hospital mortality among infectious disease patients

We hypothesized that the ECGs could display severity in infectious disease and that the developed DLM will predict in-hospital mortality of infectious disease patients. In other word, we hypothesized high score of DLM was correlated with severe infectious disease. We conducted a subgroup analysis of patients who suspected infectious disease in internal and external validation dataset. Among these patients, we confirmed the performance of DLM to predict in-hospital mortality. To confirm the accuracy of the DLM, we compared the score calculated by the DLM with the presence of in-hospital mortality in the subgroup datasets. As a comparative measure, we used the Sequential Organ Failure Assessment (SOFA) score,

quick SOFA score, National Early Warning Score (NEWS), Modified Early Warning Score (MEWS), lactate, white blood cell (WBC) count, and C-reactive protein (CRP) to predict in-hospital morality among infectious disease patients.[19–22]

Results

The eligible study population included patients who were admitted to SGH and MSH. As shown in Fig. 1, we excluded 8 patients because of missing clinical information, including ECGs, admission notes, and laboratory exams. The study included a total of 46,017 patients, of which 1,548 and 639 patients underwent sepsis and septic shock, respectively. The DLM was developed using a development dataset of 73,727 ECGs of 18,142 patients from SGH. The internal validation of the performance of DLM was conducted using 7,774 ECGs of 7,774 patients from SGH. External validation of the DLM was conducted using 20,101 ECGs of 20,101 patients from MSH. The patients were divided into development, internal validation, and external validation, exclusively. In patients with sepsis, the ECG had a rightward P, R, and T-wave axis, prolonged QTc, and tachycardia (Table 1).

Table 1
Baseline characteristics table

Characteristics	Non-sepsis patients (n = 44,469)	Sepsis patients (n = 1,548)	*p*
Age, yr, mean (SD)	58.01 (19.83)	61.83 (24.93)	< 0.001
Male, n (%)	20,836 (46.9)	810 (52.3)	< 0.001
Systolic blood pressure, mmHg, mean (SD)	121.61 (33.43)	101.05 (39.53)	< 0.001
Heart rate, bpm, mean (SD)	76.92 (17.63)	103.59 (23.65)	< 0.001
Respiratory rate, bpm, mean (SD)	18.79 (4.34)	26.29 (8.72)	< 0.001
Peripheral oxygen saturation, %, mean (SD)	97.18 (16.97)	92.81 (28.20)	< 0.001
Mental change, n (%)	288 (0.6)	753 (48.6)	< 0.001
C-reactive protein, mg/dL, mean (SD)	13.30 (37.22)	49.80 (74.46)	< 0.001
Lactate, mmol/L, mean (SD)	1.87 (1.87)	4.64 (5.03)	< 0.001
White blood cell count, per uL, mean (SD)	8180 (4200)	13090 (6190)	< 0.001
Total bilirubin, mg/dL, mean (SD)	0.72 (0.83)	1.39 (2.58)	< 0.001
Creatinine, mg/dL, mean (SD)	0.98 (0.97)	1.47 (1.53)	< 0.001
PR interval, ms, mean (SD)	169.07 (31.53)	161.47 (41.24)	< 0.001
QRS duration, ms, mean (SD)	96.59 (18.91)	96.96 (23.49)	0.461
QT interval, ms, mean (SD)	401.65 (47.90)	372.11 (63.62)	< 0.001
QTc, ms, mean (SD)	442.41 (37.66)	469.16 (43.79)	< 0.001
P-wave axis, degree, mean (SD)	43.42 (30.85)	45.50 (38.72)	0.033
R-wave axis, degree, mean (SD)	38.40 (46.36)	47.11 (61.76)	< 0.001
T-wave axis, degree, mean (SD)	47.42 (53.10)	66.09 (81.47)	< 0.001
Legend: SD denotes standard deviation.			

During the internal and external validation, the AUC of the DLM for detecting sepsis, the primary outcome, using a 12-lead ECG, was 0.901 (95% CI = 0.882–0.920) and 0.863 (95% CI = 0.846–0.879), respectively (Fig. 3). The AUC of the DLM for detecting septic shock using 12-lead ECGs during internal and external validations were 0.906 (95% CI = 0.877–0.936) and 0.899 (95% CI = 0.872–0.925), respectively. The AUC of the DLM for detecting sepsis using 6-lead and single-lead ECGs were 0.845–0.882, and the AUC of the DLM for detecting septic shock using 6-lead and single-lead ECGs were 0.881–0.906.

A sensitivity map showed that the QT interval and T wave were associated with sepsis, and the variable importance of deep learning confirmed that the prolonged QTc was associated with sepsis (Fig. 4). The logistic regression and random forest had different variable importance and showed that the prolonged QTc, T axis, and QRS duration were important variables (Table 2).

Table 2
Variable importance for detecting sepsis

Rank	Logistic regression (Deviance difference)	Random forest (Mean decrease Gini)	Deep learning (Relative importance)
1	QTc (4,640)	QTc (545.3)	QT interval (0.192)
2	QT interval (3,187)	QT interval (535.6)	QTc (0.145)
3	Age (243)	T wave axis (507.6)	PR interval (0.118)
4	QRS duration (204)	R wave axis (493.8)	Age (0.111)
5	T wave axis (152)	P wave axis (458.7)	QRS duration (0.104)
6	P wave axis (21)	Age (444.7)	T-wave axis (0.091)
7	R wave axis (14)	PR interval (433.2)	Sex (0.084)
8	PR interval (2)	QRS duration (424.4)	R wave axis (0.082)
9	Sex (-1)	Sex (0)	P wave axis (0.073)
Legend: none			

Subgroup analysis was conducted using ECGs from 4,609 patients who admitted with infectious disease patients in validation dataset. There were 256 in-hospital mortality cases in subgroup study population. The AUC of the DLM using 12, 6, and single-lead ECG, SOFA, qSOFA, NEWS, MEWS, lactate, WBC, and CRP for predicting in-hospital mortality was 0.817 (0.793–0.840), 0.815 (0.794–0.836), 0.802 (0.780–0.825), 0.817 (0.786–0.847), 0.797 (0.767–0.828), 0.808 (0.777–0.839), 0.778 (0.747–0.808), 0.801 (0.758–0.844), 0.591 (0.552–0.630), and 0.541 (0.499–0.583) outperformed other predictive models (Fig. 5).

As shown in Fig. 6, there was a significant difference in the prediction score of DLM using ECG according to the presence of infection in the validation dataset (0.277 vs 0.574, $p < 0.001$). In patients with COVID-19, the same trend was also observed in prediction score of DLM using ECG before and after COVID-19 infection (0.260 vs 0.725, $p = 0.018$).

Discussion

We developed a DLM for screening sepsis and septic shock using 12-, 6-, and single-lead ECGs and demonstrated reasonable accuracies for internal and external validation. We confirmed the performance

of predicting in-hospital mortality in subgroup analysis of infectious disease patients. We also showed the ECG regions and features that were associated with sepsis. To the best of our knowledge, this study is the first to develop a DLM for screening sepsis using ECG.

Approximately 50% of sepsis patients have cardiac dysfunction, and cardiac dysfunction is a well-known risk factor associated with a significantly increased mortality of 20–50%.[23] Sepsis develops into cardiac dysfunction by decreasing the beta-adrenergic receptor components, which are mediated by inflammatory substances such as cytokines and nitric oxide.[24] Direct cardiomyocyte injury or death is caused by toxins and complements from sepsis. Cardiomyocyte apoptosis is the leading cause of cardiac dysfunction, followed by the downregulation of a beta-adrenoreceptor and impairment of myofibril function owing to the disruption of calcium liberation. Because sepsis affects the cardiac function through a direct or indirect pathophysiology, we hypothesized that an ECG has information for sepsis detection. In previous studies, Rich et al. showed that the QRS amplitude of sepsis is smaller than that of normal people.[8] However, conventional statistical methods, such as logistic regression, could not develop diagnostic criteria for using this subtle change and non-linear correlations. ECG is affected not only by cardiac function but also by other human factors. For example, fat patients with a larger body mass index have a lower ECG amplitude.[25] In another previous study, Madias et al. described that the loss of QRS amplitude in the ECGs in patients with sepsis is not due to cardiac dysfunction, but to an extracardiac reason, such as a reduction in the transfer impedance of the body volume conductor owing to water accumulation.[26] Recent studies have shown the possibility of artificial intelligence for interpreting an ECG. Based on artificial intelligence with DLM, we could diagnose diseases that could not be diagnosed based on previous medical knowledge, such as heart failure, valvular heart disease, pulmonary hypertension, anemia, and hyperkalemia.[10–13, 27–29] The most important aspect of deep learning is its ability to extract features and develop an algorithm using various types of data, such as images, 2D data, and waveforms.[14] In this study, we developed a DLM for detecting sepsis and validated its performance based on an external validation. DLM can also detect septic shock using a prediction score of DLM. Previous studies have shown that inflammatory markers and infection are closely correlated with cardiac disease and an ECG.[30]

There has been an enormous development in diverse wearable and life-style devices worldwide. There is already a base for remote diagnosis and treatment based on diverse biosensors and internet technologies. However, there are limitations in the interpretation of biosignals captured from various wearable devices. An ECG is an important biosignal for remote medical monitoring and treatment. This is because it can be measured with diverse wearable devices and transferred to remote medical sites in real time. As a conventional statistical limitation, an ECG is only used in the diagnosis of arrhythmia and myocardial infarction. Based on current studies, AI technologies have enabled the diagnosis and prediction of diverse diseases using an ECG. In the current COVID-19 pandemic, such technologies are important for screening infectious diseases, monitoring the patient status, and capturing the deterioration of patients. In this study, we showed that the possibility of DLM for screening infectious disease including COVID-19 as shown in Fig. 6. There is a need for several studies on the use of AI in sepsis and

septic shock. However, this study shows the possibility of applying an ECG for detecting and monitoring infectious patients.

There are several limitations to this study. First, we validated the DLM using retrospective data; however, it is necessary to validate the DLM with prospective studies and data from daily living. Studies related to the clinical significance of the new technology are required for application in clinical practice. In our next study, we will verify the DLM performance and significance through a prospective study on daily clinical practice. Second, this study was conducted in only two hospitals in Korea; hence, it is necessary to validate the DLM with patients in other countries.

Conclusion

The DLM demonstrated an accurate performance in detecting sepsis and septic shock using an ECG. The application of artificial intelligence technologies based on the DLM to an ECG could enable screening for diverse infectious diseases and predict the development of sepsis.

List Of Abbreviations

AUC, Area under the receiver operating characteristic curve

CI, Confidence interval

COVID-19 Coronavirus Disease-19

CRP C-reactive protein

DLM Deep learning-based model

ECG Electrocardiography

MEWS Modified Early Warning Score

NEWS National Early Warning Score

NPV Negative predictive value

PPV Positive predictive value

Sepsis-3 Third International Consensus Definitions for Sepsis and Septic Shock

SGH Sejong General Hospital

SOFA Sequential Organ Failure Assessment

MSH Mediplex Sejong Hospital

IRB Institutional review board

WBC White blood cell

1D one-dimensional

Declarations

Ethics approval and consent to participate: The institutional review boards of Sejong General Hospital (2020-0541) and Mediplex Sejong Hospital (2020-149) approved this study protocol and waived the need for informed consent because of the impracticality and minimal harm involved.

Consent for publication: Not applicable

Availability of data and materials: The data underlying this article will be shared on reasonable request to the corresponding author.

Competing interests: YRL, DYK, KHK, and SYL declare that they have no competing interests. JK, YYJ, MSJ, YJL, YHC, and JHS are researchers of Medical AI Co., a medical artificial intelligence company. JK and JHB are researchers of Body friend Co. There are no products in development or marketed products to declare. This does not alter our adherence to Journal.

Funding: This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2020R1F1A1073791)

Authors' contributions: YRL and DYK contributed to the study idea and revised manuscript. MSJ and YJL performed data analysis and verified the clinical coding. YYJ, YHC, HJS, and JHB contributed to the study idea and design as well as data collection, performed data analysis, and contributed to subsequent drafts. SYL and KHK contributed to data collection and revised the manuscript. JK is the principal investigator and contributed to the study idea and design, data analysis, verified the clinical coding, and contributed to subsequent drafts.

Acknowledgement: This research was results of a study on the "High Performance Computing Support" and "AI voucher" Project, supported by the 'Ministry of Science and ICT and National IT Industry Promotion Agency of South Korea.

Authors' information: None

Footnotes: None

References

1. Singer M, Deutschman CS, Seymour CW, Shankar-Hari M, Annane D, Bauer M, et al. The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). JAMA [Internet].

2016;315:801. Available from: http://jama.jamanetwork.com/article.aspx?doi=10.1001/jama.2016.0287

2. Rhodes A, Evans LE, Alhazzani W, Levy MM, Antonelli M, Ferrer R, et al. Surviving Sepsis Campaign. Crit Care Med [Internet]. 2017;45:486–552. Available from: http://journals.lww.com/00003246-201703000-00015
3. Rudd KE, Johnson SC, Agesa KM, Shackelford KA, Tsoi D, Kievlan DR, et al. Global, regional, and national sepsis incidence and mortality, 1990–2017: analysis for the Global Burden of Disease Study. Lancet [Internet]. 2020;395:200–11. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0140673619329897
4. Rudd KE, Kissoon N, Limmathurotsakul D, Bory S, Mutahunga B, Seymour CW, et al. The global burden of sepsis: barriers and potential solutions. Crit Care [Internet]. 2018;22:232. Available from: https://ccforum.biomedcentral.com/articles/10.1186/s13054-018-2157-z
5. Charpentier J, Luyt C-E, Fulla Y, Vinsonneau C, Cariou A, Grabar S, et al. Brain natriuretic peptide: A marker of myocardial dysfunction and prognosis during severe sepsis. Crit Care Med [Internet]. 2004;32:660–5. Available from: http://journals.lww.com/00003246-200403000-00007
6. Court O, Kumar A, Parrillo JE, Kumar A. Clinical review: Myocardial depression in sepsis and septic shock. Crit. Care. 2002.
7. Rudiger A, Singer M. Mechanisms of sepsis-induced cardiac dysfunction. Crit Care Med [Internet]. 2007;35:1599–608. Available from: http://journals.lww.com/00003246-200706000-00021
8. Rich MM, McGarvey ML, Teener JW, Frame LH. ECG Changes during Septic Shock. Cardiology [Internet]. 2002;97:187–96. Available from: https://www.karger.com/Article/FullText/63120
9. Attia ZI, Kapa S, Lopez-Jimenez F, McKie PM, Ladewig DJ, Satam G, et al. Screening for cardiac contractile dysfunction using an artificial intelligence–enabled electrocardiogram. Nat. Med. 2019.
10. Kwon J myoung, Kim KH, Medina-Inojosa J, Jeon KH, Park J, Oh BH. Artificial intelligence for early prediction of pulmonary hypertension using electrocardiography. J Hear Lung Transplant [Internet]. Elsevier Inc.; 2020;39:805–14. Available from: https://doi.org/10.1016/j.healun.2020.04.009
11. Kwon J, Lee SY, Jeon K, Lee Y, Kim K, Park J, et al. Deep Learning–Based Algorithm for Detecting Aortic Stenosis Using Electrocardiography. J Am Heart Assoc [Internet]. 2020;9. Available from: https://www.ahajournals.org/doi/10.1161/JAHA.119.014717
12. Galloway CD, Valys A V., Shreibati JB, Treiman DL, Petterson FL, Gundotra VP, et al. Development and Validation of a Deep-Learning Model to Screen for Hyperkalemia From the Electrocardiogram. JAMA Cardiol [Internet]. 2019;4:428. Available from: http://cardiology.jamanetwork.com/article.aspx?doi=10.1001/jamacardio.2019.0640
13. Kwon J, Cho Y, Jeon K-H, Cho S, Kim K-H, Baek SD, et al. A deep learning algorithm to detect anaemia with ECGs: a retrospective, multicentre study. Lancet Digit Heal [Internet]. 2020;2:e358–67. Available from: https://linkinghub.elsevier.com/retrieve/pii/S2589750020301084
14. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature [Internet]. 2015;521:436–44. Available from: http://arxiv.org/abs/1603.05691

15. Schisterman EF, Perkins NJ, Liu A, Bondell H. Optimal cut-point and its corresponding Youden index to discriminate individuals using pooled blood samples. Epidemiology. 2005;16:73–81.
16. Ketkar N. Introduction to PyTorch. Deep Learn with Python [Internet]. Berkeley, CA: Apress; 2017. p. 195–208. Available from: http://link.springer.com/10.1007/978-1-4842-2766-4_12
17. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. Proc IEEE Int Conf Comput Vis. 2017. p. 1;618-626.
18. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. Int J Comput Vis [Internet]. 2020;128:336–59. Available from: http://link.springer.com/10.1007/s11263-019-01228-7
19. Subbe CP, Kruger M, Rutherford P, Gemmel L. Validation of a modified Early Warning Score in medical admissions. QJM. 2001;94:521–6.
20. Smith GB, Prytherch DR, Meredith P, Schmidt PE, Featherstone PI. The ability of the National Early Warning Score (NEWS) to discriminate patients at risk of early cardiac arrest, unanticipated intensive care unit admission, and death. Resuscitation [Internet]. European Resuscitation Council, American Heart Association, Inc., and International Liaison Committee on Resuscitation.~Published by Elsevier Ireland Ltd; 2013;84:465–70. Available from: http://dx.doi.org/10.1016/j.resuscitation.2012.12.016
21. Vincent J-L, Moreno R, Takala J, Willatts S, De Mendonça A, Bruining H, et al. The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. Intensive Care Med [Internet]. 1996;22:707–10. Available from: http://link.springer.com/10.1007/BF01709751
22. Seymour CW, Liu VX, Iwashyna TJ, Brunkhorst FM, Rea TD, Scherag A, et al. Assessment of Clinical Criteria for Sepsis. JAMA [Internet]. 2016;315:762. Available from: http://jama.jamanetwork.com/article.aspx?doi=10.1001/jama.2016.0288
23. Habimana R, Choi I, Cho HJ, Kim D, Lee K, Jeong I. Sepsis-induced cardiac dysfunction: a review of pathophysiology. Acute Crit Care [Internet]. 2020;35:57–66. Available from: http://accjournal.org/journal/view.php?doi=10.4266/acc.2020.00248
24. Suzuki T, Suzuki Y, Okuda J, Kurazumi T, Suhara T, Ueda T, et al. Sepsis-induced cardiac dysfunction and β-adrenergic blockade therapy for sepsis. J Intensive Care [Internet]. 2017;5:22. Available from: http://jintensivecare.biomedcentral.com/articles/10.1186/s40560-017-0215-2
25. Hassing GJ, van der Wall HEC, van Westen GJP, Kemme MJB, Adiyaman A, Elvan A, et al. Body mass index related electrocardiographic findings in healthy young individuals with a normal body mass index. Netherlands Hear J [Internet]. 2019;27:506–12. Available from: http://link.springer.com/10.1007/s12471-019-1282-x
26. Madias JE, Bazaz R. On the Mechanism of the Reduction in the ECG QRS Amplitudes in Patients with Sepsis. Cardiology [Internet]. 2003;99:166–8. Available from: https://www.karger.com/Article/FullText/70674
27. Attia ZI, Kapa S, Lopez-Jimenez F, McKie PM, Ladewig DJ, Satam G, et al. Screening for cardiac contractile dysfunction using an artificial intelligence–enabled electrocardiogram. Nat Med [Internet].

Springer US; 2019;25:70–4. Available from: http://www.nature.com/articles/s41591-018-0240-2

28. Kwon J, Kim K-H, Akkus Z, Jeon K-H, Park J, Oh B-H. Artificial intelligence for detecting mitral regurgitation using electrocardiography. J Electrocardiol [Internet]. 2020;59:151–7. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0022073619307046
29. Kwon JM, Kim KH, Jeon KH, Lee SY, Park J, Oh BH. Artificial intelligence algorithm for predicting cardiac arrest using electrocardiography. Scand J Trauma Resusc Emerg Med. Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine; 2020;28:1–10.
30. Pearson TA, Mensah GA, Alexander RW, Anderson JL, Cannon RO, Criqui M, et al. Markers of Inflammation and Cardiovascular Disease. Circulation [Internet]. 2003;107:499–511. Available from: https://www.ahajournals.org/doi/10.1161/01.CIR.0000052939.59093.45

Figures

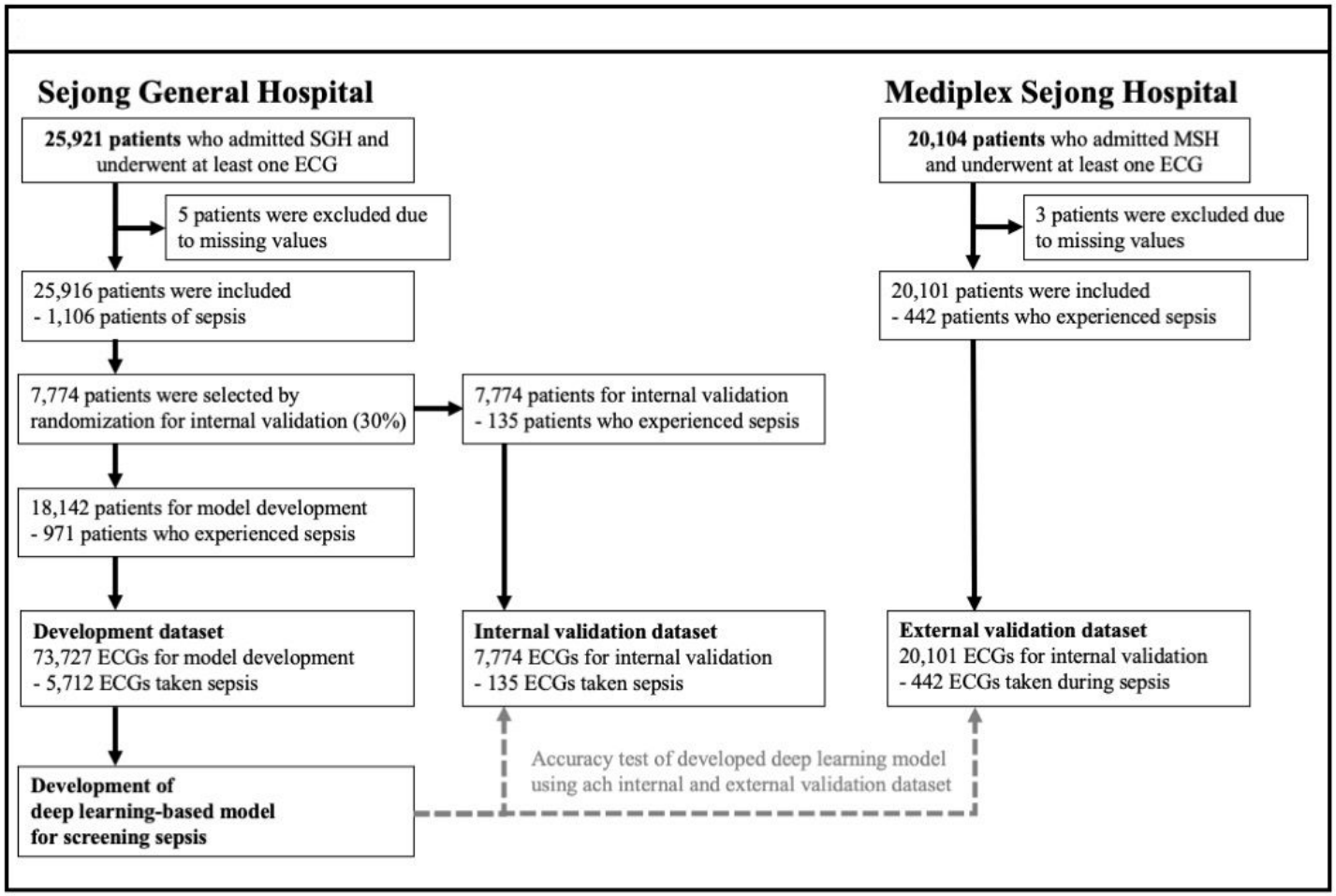


Figure 1

Study flowchart Legend: none

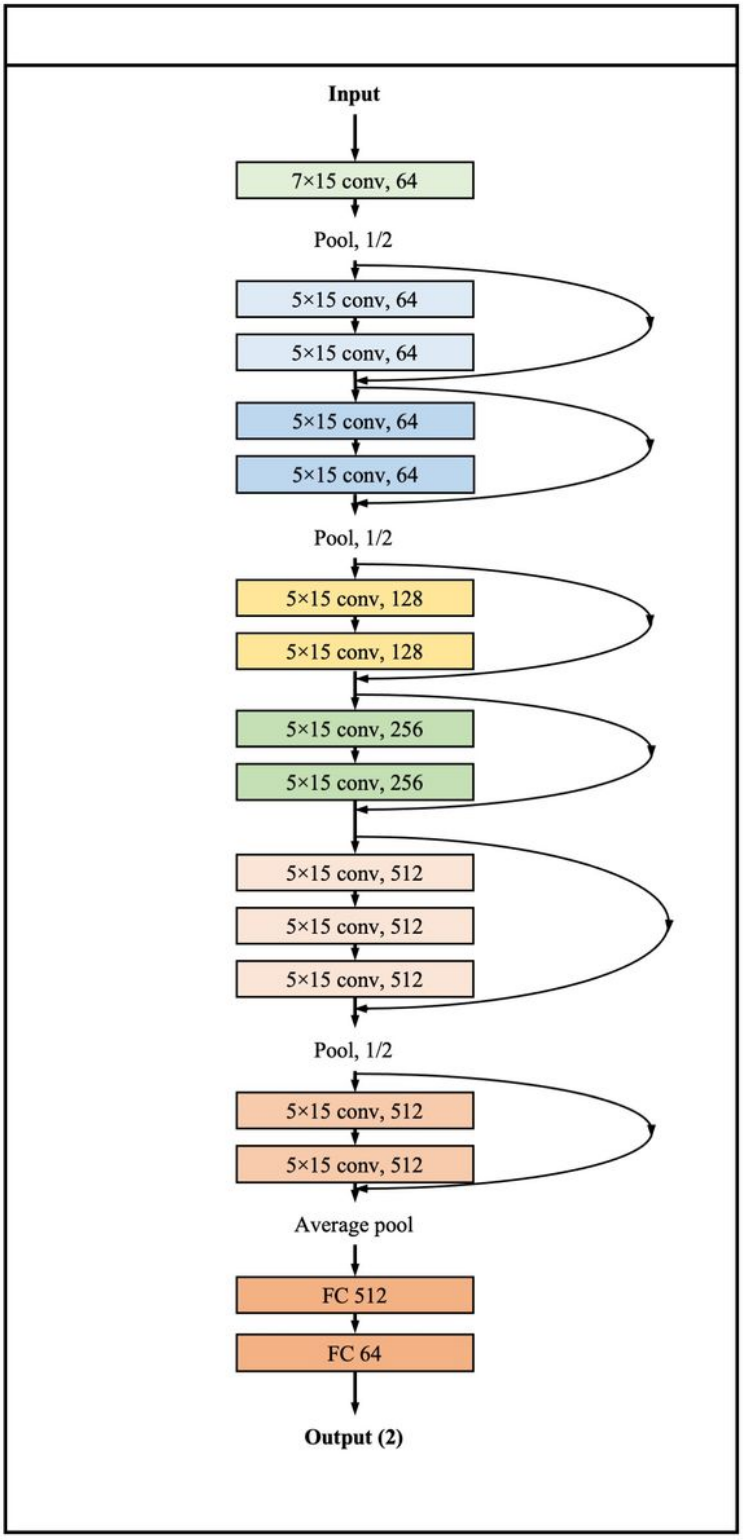
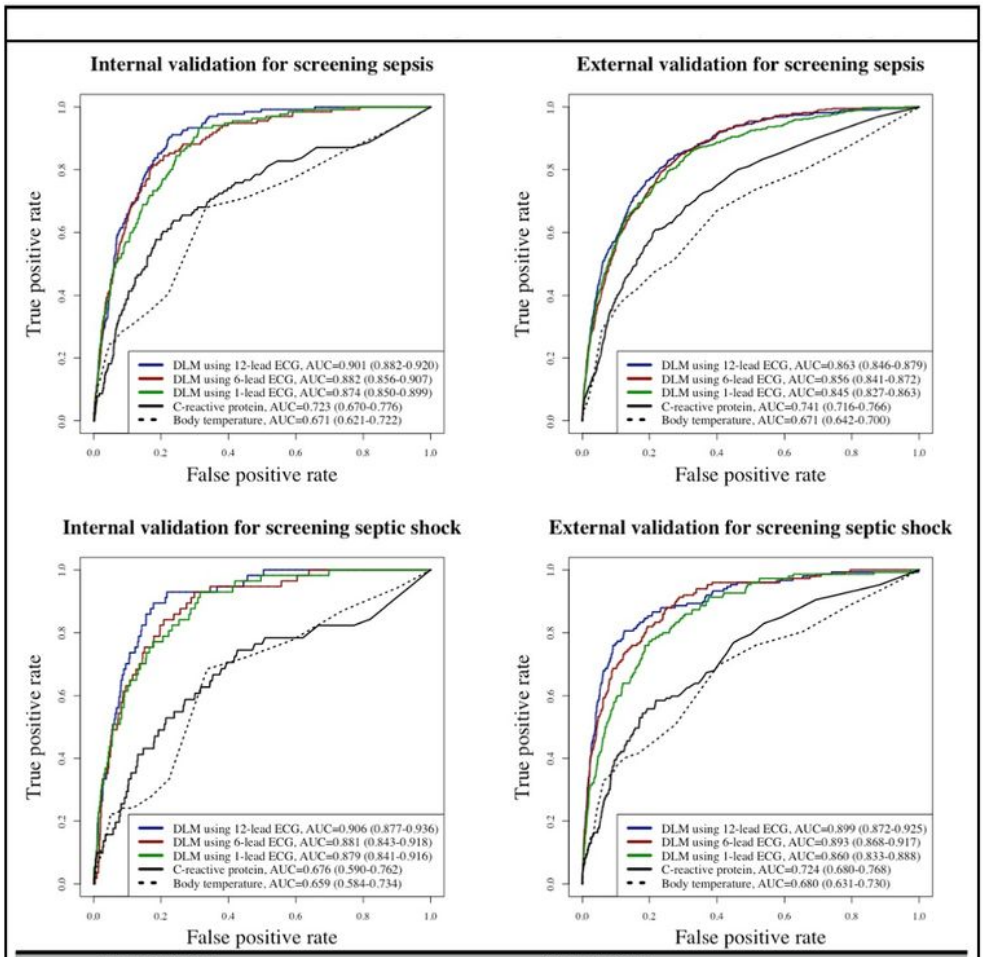


Figure 2

Architecture of DLM to screen sepsis using ECG Legend: conv denotes convolutional neural network layer, DLM deep learning based model, ECG electrocardiography, and FC fully connected layer.



	Internal validation					External validation				
Prediction model	AUC (95% CI)	SEN (95% CI)	SPE (95% CI)	PPV (95% CI)	NPV (95% CI)	AUC (95% CI)	SEN (95% CI)	SPE (95% CI)	PPV (95% CI)	NPV (95% CI)
Performance for screening sepsis										
DLM using 12-lead ECG	0.901 (0.882-0.920)	0.904 (0.854-0.953)	0.776 (0.767-0.785)	0.067 (0.055-0.078)	0.998 (0.997-0.999)	0.863 (0.846-0.879)	0.765 (0.725-0.804)	0.810 (0.805-0.816)	0.083 (0.075-0.091)	0.994 (0.992-0.995)
DLM using 6-lead ECG	0.882 (0.856-0.907)	0.815 (0.749-0.880)	0.824 (0.816-0.833)	0.076 (0.062-0.089)	0.996 (0.994-0.998)	0.856 (0.841-0.872)	0.794 (0.756-0.832)	0.766 (0.760-0.772)	0.071 (0.064-0.078)	0.994 (0.993-0.995)
DLM using 1-lead ECG	0.874 (0.850-0.899)	0.933 (0.891-0.975)	0.688 (0.678-0.698)	0.050 (0.042-0.059)	0.998 (0.997-0.999)	0.845 (0.827-0.863)	0.796 (0.759-0.834)	0.746 (0.739-0.752)	0.066 (0.059-0.072)	0.994 (0.993-0.995)
C-reactive protein	0.723 (0.670-0.776)	0.638 (0.550-0.725)	0.763 (0.748-0.778)	0.091 (0.071-0.111)	0.983 (0.977-0.988)	0.741 (0.716-0.766)	0.607 (0.561-0.654)	0.783 (0.776-0.791)	0.087 (0.077-0.097)	0.983 (0.981-0.986)
Body temperature	0.671 (0.621-0.722)	0.679 (0.599-0.759)	0.667 (0.655-0.679)	0.041 (0.032-0.049)	0.990 (0.987-0.993)	0.671 (0.642-0.700)	0.669 (0.624-0.714)	0.601 (0.593-0.608)	0.044 (0.039-0.049)	0.985 (0.983-0.987)
Performance for screening septic shock										
DLM using 12-lead ECG	0.906 (0.877-0.936)	0.895 (0.815-0.974)	0.822 (0.814-0.831)	0.036 (0.026-0.045)	0.999 (0.998-1.000)	0.899 (0.872-0.925)	0.807 (0.743-0.870)	0.875 (0.871-0.880)	0.046 (0.038-0.054)	0.998 (0.998-0.999)
DLM using 6-lead ECG	0.881 (0.843-0.918)	0.842 (0.747-0.937)	0.792 (0.783-0.801)	0.029 (0.021-0.037)	0.999 (0.998-0.999)	0.893 (0.868-0.917)	0.880 (0.828-0.932)	0.749 (0.743-0.755)	0.026 (0.021-0.030)	0.999 (0.998-0.999)
DLM using 1-lead ECG	0.879 (0.841-0.916)	0.930 (0.864-0.996)	0.684 (0.673-0.694)	0.021 (0.016-0.027)	0.999 (0.999-1.000)	0.860 (0.833-0.888)	0.773 (0.706-0.840)	0.801 (0.795-0.806)	0.028 (0.023-0.033)	0.998 (0.997-0.999)
C-reactive protein	0.676 (0.590-0.762)	0.745 (0.625-0.865)	0.573 (0.556-0.590)	0.027 (0.019-0.036)	0.993 (0.989-0.997)	0.724 (0.680-0.768)	0.585 (0.505-0.665)	0.781 (0.774-0.788)	0.030 (0.023-0.036)	0.994 (0.992-0.995)
Body temperature	0.659 (0.584-0.734)	0.685 (0.561-0.809)	0.663 (0.651-0.674)	0.017 (0.011-0.022)	0.996 (0.994-0.998)	0.680 (0.631-0.730)	0.697 (0.622-0.773)	0.596 (0.588-0.604)	0.016 (0.013-0.019)	0.995 (0.994-0.997)

Figure 3

Performance of DLM for screening sepsis and septic shock using electrocardiography Legend: AUC denotes area under the receiver operating characteristic curve, ECG electrocardiography, NPV negative predictive value, PPV positive predictive value, SEN sensitivity, and SPE specificity.

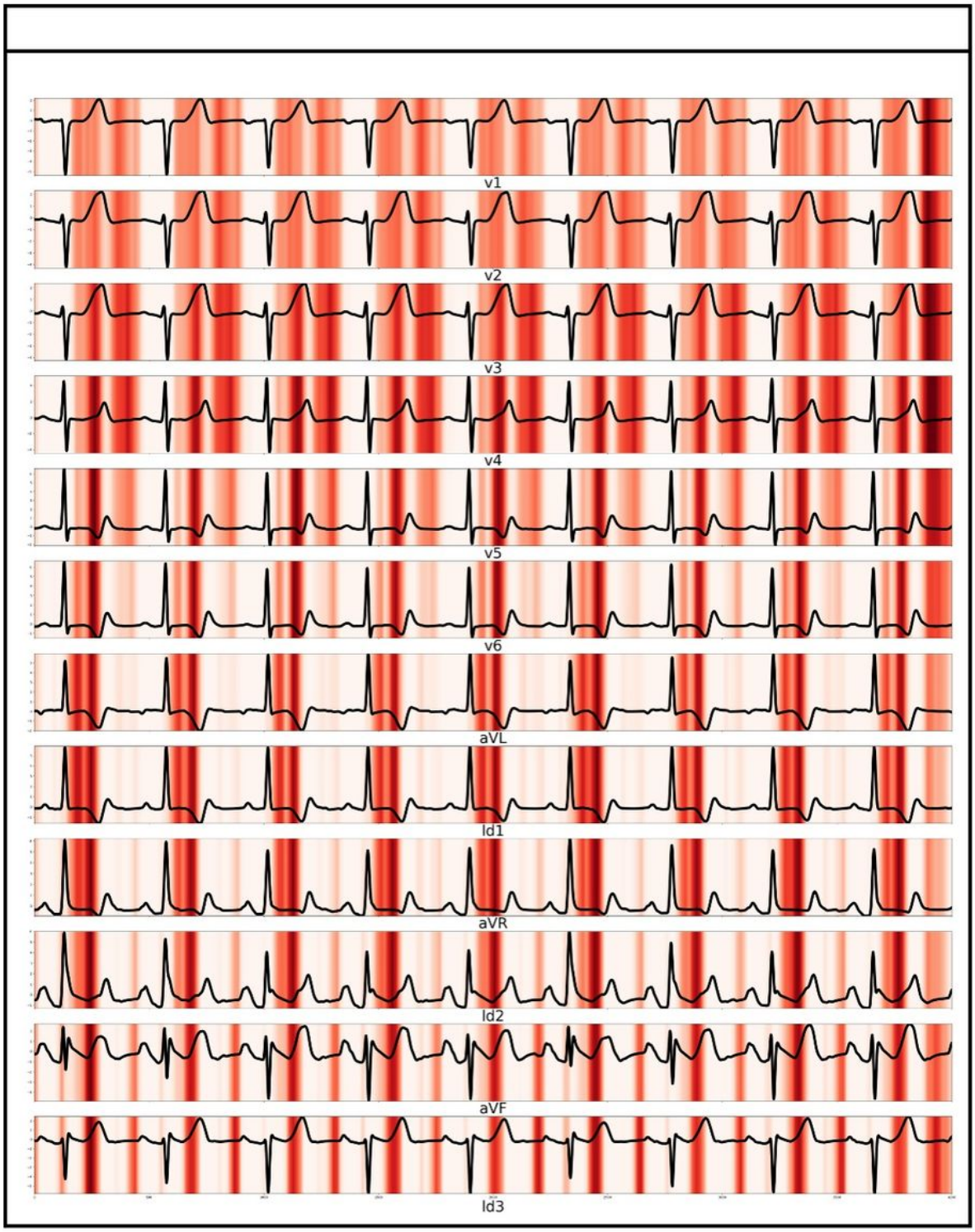
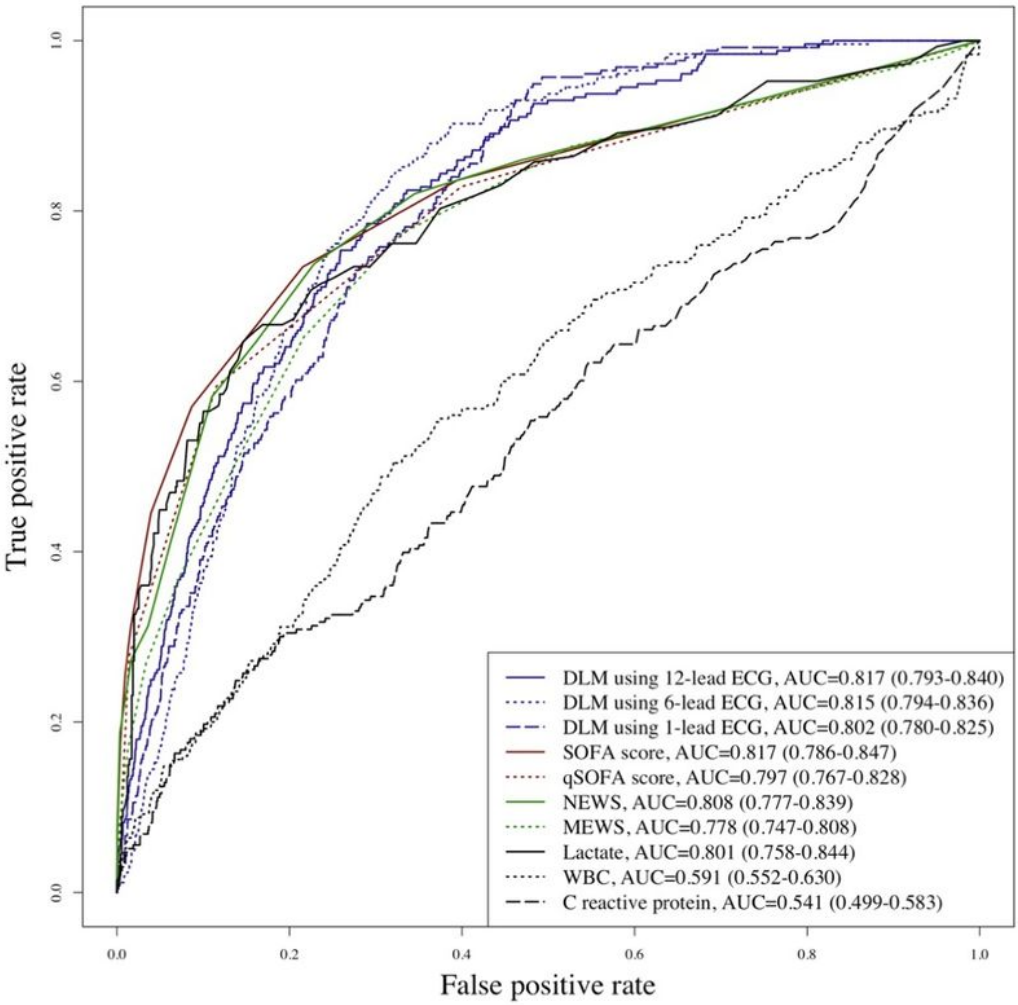


Figure 4

Sensitivity map of septic shock patients Legend: none



Predictive model	AUC (95% CI)	SEN (95% CI)	SPE (95% CI)	PPV (95% CI)	NPV (95% CI)
DLM using 12-lead ECG	0.817 (0.793-0.840)	0.785 (0.735-0.835)	0.710 (0.697-0.724)	0.137 (0.120-0.155)	0.983 (0.978-0.987)
DLM using 6-lead ECG	0.815 (0.794-0.836)	0.844 (0.799-0.888)	0.682 (0.668-0.695)	0.135 (0.118-0.152)	0.987 (0.983-0.991)
DLM using 1-lead ECG	0.802 (0.780-0.825)	0.945 (0.917-0.973)	0.524 (0.509-0.539)	0.105 (0.092-0.117)	0.994 (0.991-0.997)
SOFA score	0.817 (0.786-0.847)	0.734 (0.680-0.788)	0.785 (0.773-0.797)	0.167 (0.145-0.189)	0.980 (0.976-0.985)
quick SOFA score	0.797 (0.767-0.828)	0.594 (0.534-0.654)	0.885 (0.875-0.894)	0.233 (0.200-0.265)	0.974 (0.969-0.979)
NEWS	0.808 (0.777-0.839)	0.738 (0.684-0.792)	0.771 (0.759-0.784)	0.160 (0.139-0.180)	0.980 (0.976-0.985)
MEWS	0.778 (0.747-0.808)	0.773 (0.722-0.825)	0.670 (0.656-0.684)	0.121 (0.105-0.137)	0.981 (0.976-0.985)
Lactate	0.801 (0.758-0.844)	0.646 (0.569-0.724)	0.854 (0.829-0.879)	0.466 (0.397-0.534)	0.925 (0.905-0.944)
WBC	0.591 (0.552-0.630)	0.504 (0.442-0.566)	0.679 (0.665-0.694)	0.086 (0.072-0.101)	0.958 (0.951-0.965)
C-reactive protein	0.541 (0.499-0.583)	0.300 (0.242-0.359)	0.811 (0.799-0.823)	0.084 (0.065-0.103)	0.953 (0.946-0.960)

Figure 5

Performance of DLM for predicting in-hospital mortality of patients with infectious disease Legend: AUC denotes area under the receiver operating characteristic curve, ECG electrocardiography, MEWS Modified Early Warning Score, NEWS National Early Warning Score, NPV negative predictive value, PPV positive predictive value, SEN sensitivity, SOFA Sequential Organ Failure Assessment, and SPE specificity.

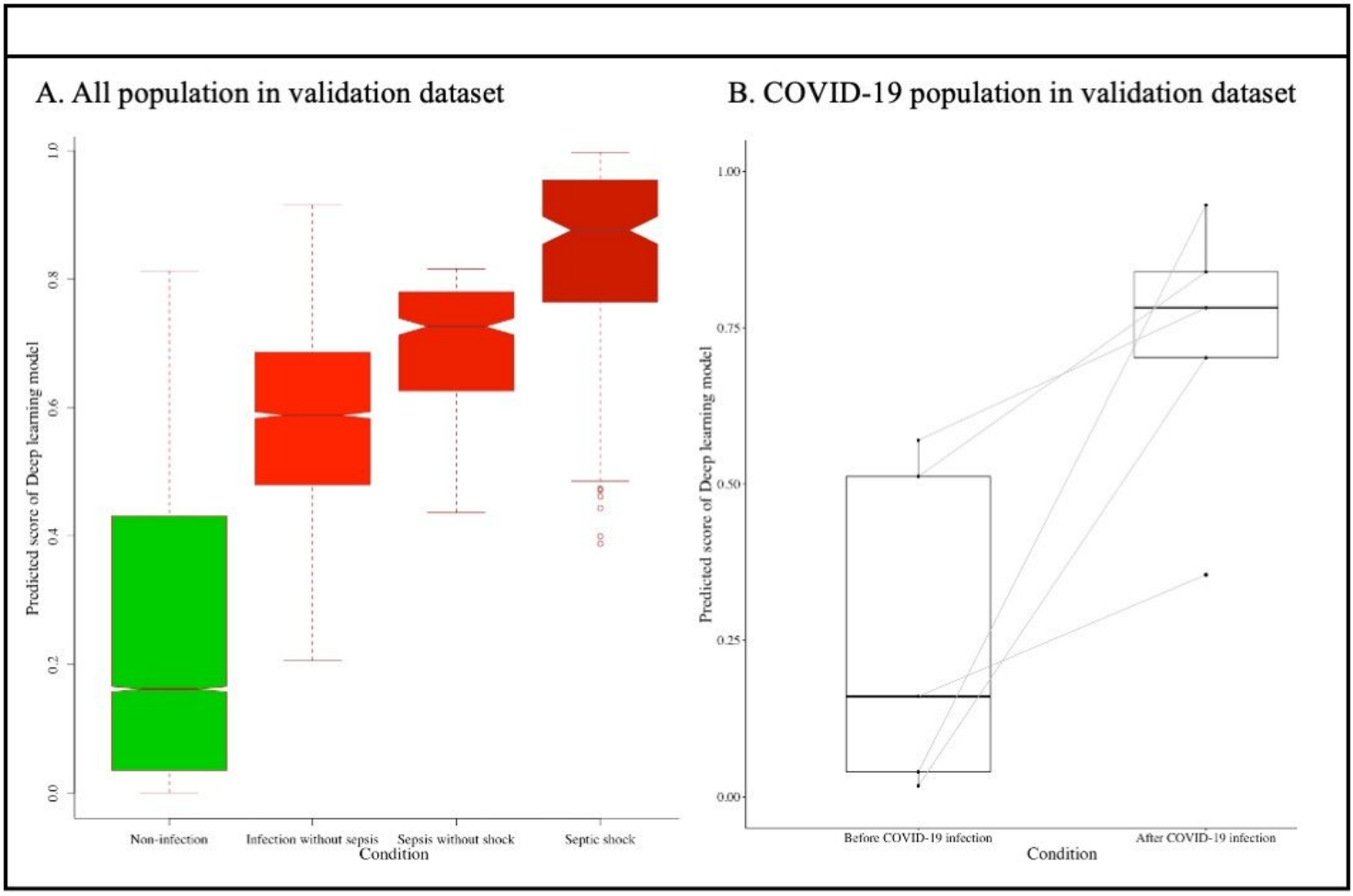


Figure 6

Change of DLM's prediction score according to infection Legend: COVID-19 Coronavirus Disease-19