ESC European Society of Cardiology

European Heart Journal - Digital Health (2025) **6**, 170–179
https://doi.org/10.1093/ehjdh/ztae088

ORIGINAL ARTICLE
Artificial intelligence (machine learning, deep learning)

Sudden cardiac arrest prediction via deep learning electrocardiogram analysis

Matt T. Oberdier[1,†], Luca Neri [1,2,*,†], Alessandro Orro[3,†], Richard T. Carrick[1], Marco S. Nobile[4], Sujai Jaipalli[5], Mariam Khan[1], Stefano Diciotti[6,7], Claudio Borghi [2,8], and Henry R. Halperin[1,5,9]

[1]Department of Medicine, Division of Cardiology, Johns Hopkins University, Baltimore, MD 21205, USA; [2]Department of Medical and Surgical Sciences, University of Bologna, 40138 Bologna, Italy; [3]Institute of Biomedical Technologies, Department of Biomedical Sciences, National Research Council (ITB-CNR), 20054 Segrate, Italy; [4]Department of Environmental Sciences, Informatics and Statistics, Ca' Foscari University of Venice, 30172 Mestre (Venice), Italy; [5]Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD 21205, USA; [6]Department of Electrical, Electronic, and Information Engineering 'Guglielmo Marconi', University of Bologna, 47521 Cesena, Italy; [7]Alma Mater Research Institute for Human-Centred Artificial Intelligence, University of Bologna, 40121 Bologna, Italy; [8]Cardiovascular Medicine Unit, IRCCS Azienda Ospedaliero-Universitaria di Bologna, 40138 Bologna, Italy; and [9]Department of Radiology, Johns Hopkins University, Baltimore, MD 21205, USA

Received 22 July 2024; revised 24 September 2024; accepted 16 October 2024; online publish-ahead-of-print 25 February 2025

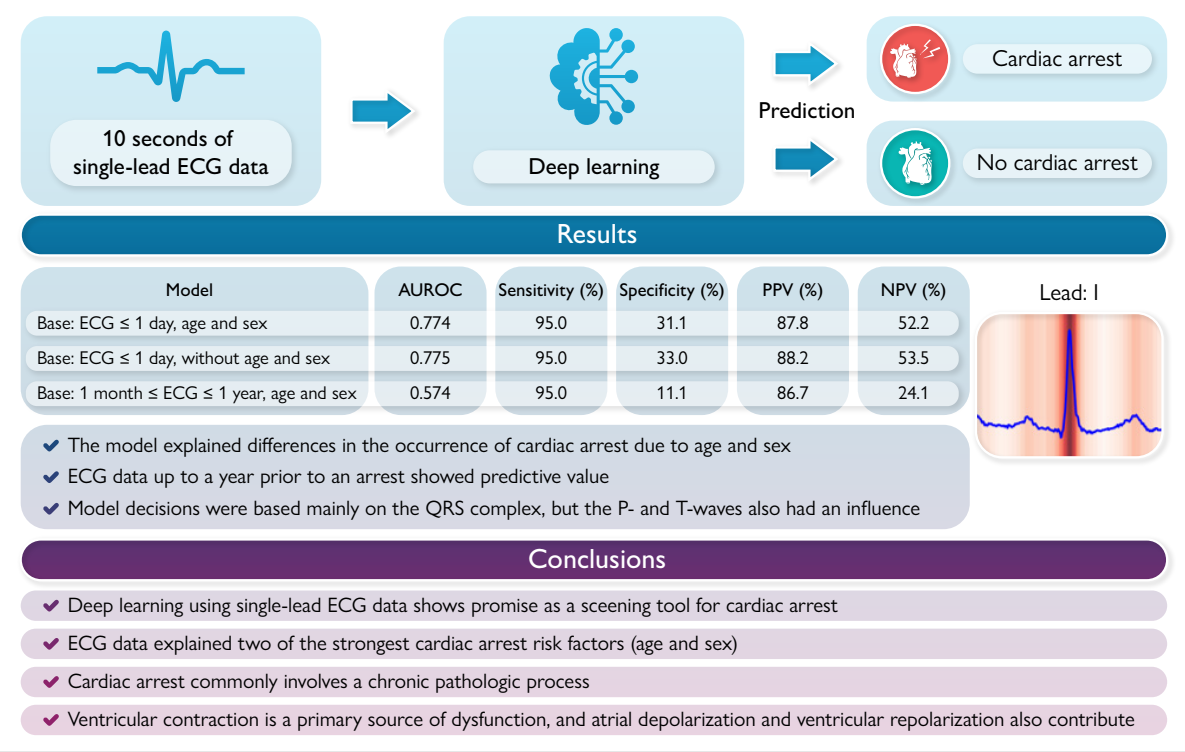
Aims Sudden cardiac arrest (SCA) is a commonly fatal event that often occurs without prior indications. To improve outcomes and enable preventative strategies, the electrocardiogram (ECG) in conjunction with deep learning was explored as a potential screening tool.

Methods and results A publicly available data set containing 10 s of 12-lead ECGs from individuals who did and did not have an SCA, information about time from ECG to arrest, and age and sex was utilized for analysis to individually predict SCA or not using deep convolution neural network models. The base model that included age and sex, ECGs within 1 day prior to arrest, and data sampled from windows of 720 ms around the R-waves from 221 individuals with SCA and 1046 controls had an area under the receiver operating characteristic curve of 0.77. With sensitivity set at 95%, base model specificity was 31%, which is not clinically applicable. Gradient-weighted class activation mapping showed that the model mostly relied on the QRS complex to make predictions. However, models with ECGs recorded between 1 day to 1 month and 1 month to 1 year prior to arrest demonstrated predictive capabilities.

Conclusion Deep learning models processing ECG data are a promising means of screening for SCA, and this method explains differences in SCAs due to age and sex. Model performance improved when ECGs were nearer in time to SCAs, although ECG data up to a year prior had predictive value. Sudden cardiac arrest prediction was more dependent upon QRS complex data compared to other ECG segments.

* Corresponding author. Tel: +1 410 955 3330, Fax: +1 410 614 1980, Email: lneri1@jhmi.edu
†The first three authors are co-first authors.

Graphical Abstract



Keywords Electrocardiogram • Cardiac arrest • Convolutional neural network • Deep learning • Prediction • Artificial intelligence

Introduction

Sudden cardiac arrest (SCA) affects about 650 000 people per year in the USA[1] with successful cardiopulmonary resuscitation (CPR) rates only around 10%.[2] Major limitations in the prevention and treatment of SCA are that at least 20%[3–5] and up to 50%[6] of events are unexpected, roughly 50% of cardiac arrests occur in persons with undiagnosed heart disease,[3,7–9] and 50% of SCAs are first cardiac events.[10] That is, many of those who are most likely to need emergency care are the most unlikely to expect such an event or to be aware of the need to take precautionary or preventative measures. In addition, 60% of SCAs occur outside the hospital[11] where immediate intervention is not available.[12,13]

Towards improving outcomes and enabling proactions, risk factor analysis has been used to predict SCA.[14] Age and sex are among the most strongly predictive SCA factors[14]; however, age and sex are generic cardiovascular risk factors and are thus not valuable in distinguishing SCA from other potential cardiovascular events. Beyond risk factors, the electrocardiogram (ECG) has strong potential as a biomarker to predict SCA, especially in the general population via users of wearable devices,[15] because it is inexpensive, non-invasive, provides real-time data, and is particularly valuable for detecting cardiac anomalies.[16,17] Summary features extracted from the ECG (e.g. QRS duration and QT interval) have been investigated for predicting SCA[18–21] and are most valuable in combination[22] but were found to have limited utility.[23,24]

Instead of using the summary feature extraction approach involving predefined clusters of data, an emerging opportunity exists to predict SCA using all ECG data and their relations via artificial intelligence. Both machine learning[25] and deep learning (DL)[26,27] have been used to analyse patients' data for SCA prediction; however, DL has even more profound potential beyond machine learning because of its ability to use large data sets, learn independently, determine non-linear relationships, and be more accurate.[27,28]

This study aimed to explore the degree to which ECGs collected within 24 h prior to an event, in conjunction with age and sex, could predict arrest in a specific population with the vision that a broadly refined future DL model could utilize ECG data obtained from wearable devices to screen the public. Therefore, it was hypothesized that ECG data could be used in conjunction with DL algorithms to predict the occurrence of SCAs. Secondary objectives included determining the following: (i) contributions of age and sex, (ii) influence of shockable vs. non-shockable rhythms, (iii) impact of ECG data collected beyond 24 h prior to arrest, (iv) dependency on the amount of data around the R-wave, and (v) value of physiologic features beyond the QRS complex.

Methods

Data set

This study was performed with the Nightingale Open Science—Subtyping Cardiac Arrest data set[29,30] because it is the only publicly available data

set to contain ECGs prior to arrests in a large enough volume to enable DL. Electrocardiogram recordings consisted of 10-s recordings at a sampling rate of 500 Hz across 12 leads from individuals who had their ECGs recorded between 1 day and 10 years prior to having a cardiac arrest and who, upon having SCA, presented in the emergency department (ED) (see Supplementary material online, *Figure S1*). A control group of patients consisted of individuals who visited the ED on the same day as a case from the SCA group but who did not have a cardiac arrest and who had their ECGs recorded within 1 month and 2 years of their visit (see Supplementary material online, *Figure S2*). The data set features a mix of presenting SCA rhythms including asystole, pulseless electrical activity, and ventricular tachycardia/ventricular fibrillation (see Supplementary material online, *Figure S3*). There were 7015 total subjects in the original data set, of which 605 were SCA patients and 6410 were controls (*Figure 1*).

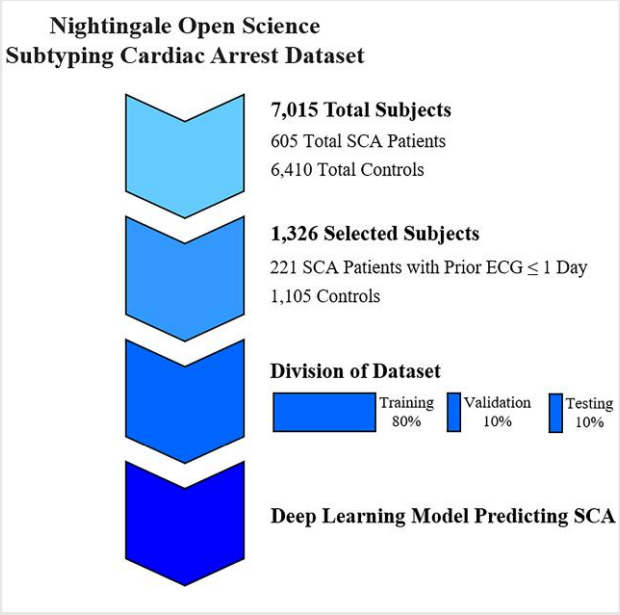


Figure 1. Overview of the selection of subjects and dataset division for the base deep learning model.

Pre-processing

Each signal was separated into segments (about 10 per signal depending on the heart rate) that were centred around the R-waves via high-quality detectors.[16,31] Intervals of 720 ms (360 ms on either side of the R-wave) were used for the base analysis. Therefore, for each individual, ECG data consisting of 12 leads and ~10 cardiac cycles were provided as independent inputs to the DL model.

Deep learning model

The DL model was implemented in Python (version 3.10) with TensorFlow (version 2.12.0)[32] and external libraries (biosppy 1.0.0, matplotlib 3.7.1, neurokit2 0.2.4, numpy 1.23.5, pandas 2.0.1, and scikit-learn 1.2.2). The model consisted of three modules: ECG, age, and sex data that were processed separately, and a final layer concatenated the three outputs (*Figure 2*). The ECG module is a sequence of four one-dimensional convolutional blocks followed by max pooling. The four blocks used 64, 128, 256, and 512 filters while the kernel and max pooling sizes were maintained at 7 and 2, respectively.

The modules for age and sex used normalized and binary values, respectively, and processed them through a dense layer with 16 outputs. An Adam optimizer,[33] a cross-entropy loss function, and a 0.001 learning rate were adopted.

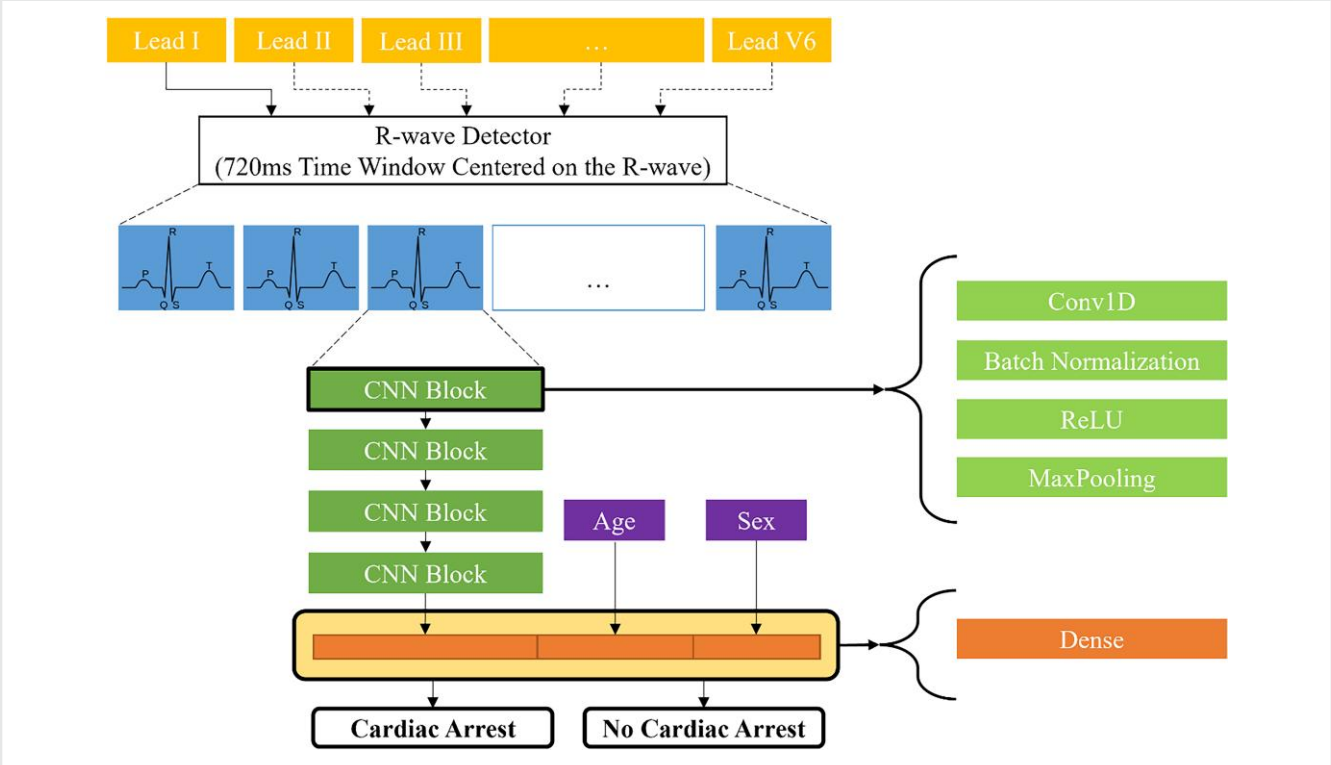


Figure 2. Architecture of the base deep learning model. CNN is Convolutional Neural Network, Conv1D is one dimensional convolutional layer, ReLu is Rectified Linear Unit.

Training and validation

The model was trained via supervision to classify ECGs as occurring prior to a SCA or not (i.e. data input for training included a binary variable for SCA/no-SCA). The data set was randomly split into training (80%), validation (10%), and testing (10%) sets while ensuring there was not an overlap of individuals among the sets. Model training was performed using nine-fold cross validation for each lead, and final predictions are the average for each lead and each cardiac cycle. For both the training and validation sets, the area under the receiver operating characteristic (AUROC) curve was evaluated.

Testing and post-processing

The model tested whether ECGs were classified as occurring prior to a SCA or not, and the classification was the output from the model. That is, the binary variable for SCA/no-SCA was not used as an input; instead, these data were used as the reference to determine model performance.

Among 50 epochs, the best performing experiment from the validation set was selected based on the highest AUROC. With sensitivity set at 95%, additional measures of performance including accuracy, F1, positive predictive value (PPV), negative predictive value (NPV), and specificity were computed.

To gain a visual understanding of the DL model's decision strategy when analysing the temporal evolution of the ECG signal, gradient-weighted class activation mapping (GRAD-CAM)[34] was utilized. For each lead and each SCA of the test set, activation maps were determined as the average of the four convolutional neural networks layers from the derivative of class outputs, and a final image was obtained by data point matching of the map to the ECG. One hundred GRAD-CAM images were then randomly selected, and for each lead, subjective tallies were made corresponding to the regions of the ECG (i.e. P-wave, P-Q interval, QRS complex, S-T interval, and T-wave) that were most influential in determining SCA.

Table 1 Characteristics of the subjects used as input for the base model

Variable	Controls ($n = 1046$)	Cardiac arrest ($n = 221$)
Age (years)	69.7 (12.3)	68.5 (15.5)
Sex (men)	549	146

Values are reported as mean (SD).

Models and their varying inputs

The base model, X, corresponds to the primary aim of the study and was generated by selecting only those SCA patients who had their ECGs recorded within 1 day of arrests, regardless of presenting rhythm. Additional inputs included age and sex and ECG data within a window of 720 ms around the R-wave. The controls were randomly selected to be age and sex balanced with the SCA patients (*Table 1*). There were 221 cardiac arrest cases, of which 218 were primary arrests, and 1046 controls used in the base model.

To explore secondary aims of the study, additional models were created in which the inputs were individually varied, resulting in experimental arms (*Figure 3*). Specifically, a model was created in which age and sex were not inputs (A1). Another series of models were generated in which the times between ECG recording and SCA were varied from <1 week, <1 month, between 1 day and 1 month, and between 1 month and 1 year (B1–B4, respectively). A third group of models was created by limiting the presenting rhythms to those that were non-shockable (C1) or only pulseless electrical activity (C2). A final set of models were generated with input data obtained

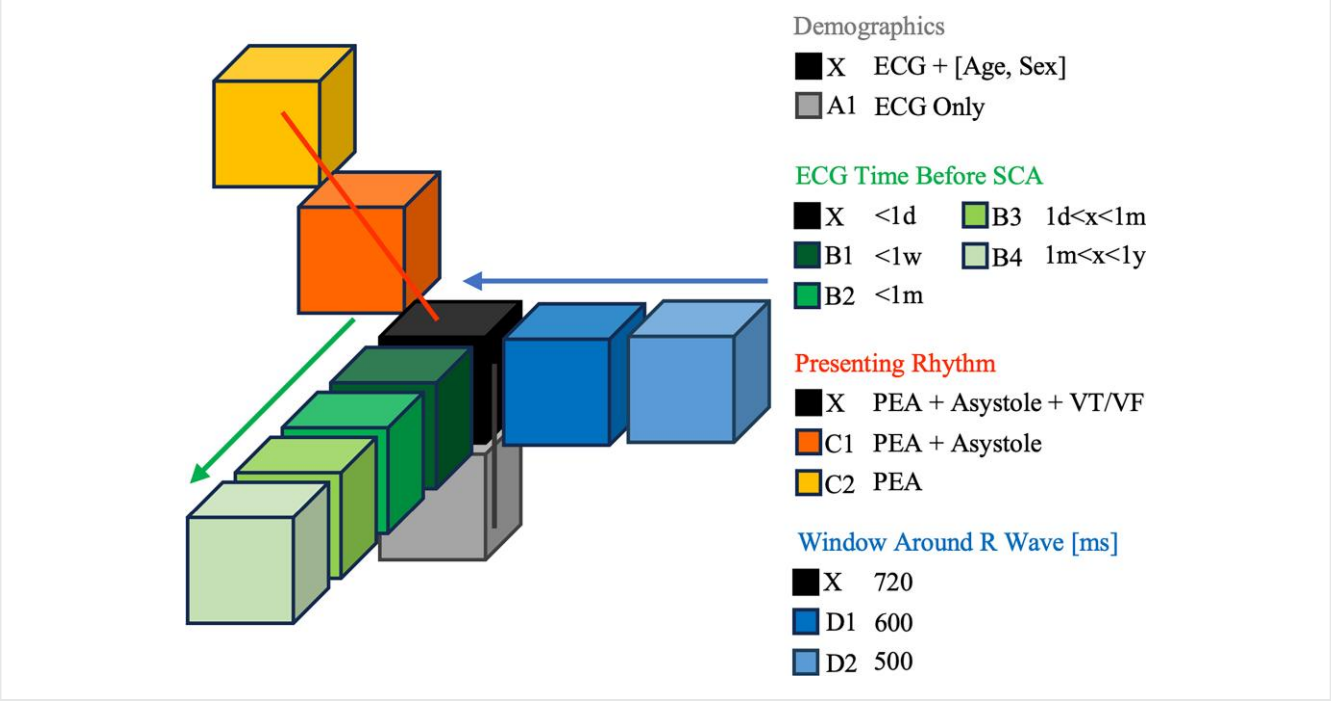


Figure 3. Conceptual representation of the base model (black), addressing the primary aim, and arms that signify secondary aims. In each arm, one input from the base model was changed (demographics, gray; ECG time before SCA, green; presenting rhythm, orange/yellow; window around R wave, blue). The base model included age and sex with ECG data less than one day prior to arrest and a window around the R wave of 720 ms (X) as input. d is day; w is week; m is month; y is year; ECG is electrocardiogram; SCA is sudden cardiac arrest; PEA is pulseless electrical activity; VT is ventricular tachycardia; VF is ventricular fibrillation.

when the window around the R-wave was 600 or 500 ms (D1 and D2, respectively).

Statistical analyses

Comparisons among groups were made using the Kruskal–Wallis test, and when differences were detected, the *post hoc* Dunn test was performed to determine significant groupings. The threshold for significance was 0.05.

Results

Base model performance

For the base model, X, average AUROC curves for training, validation, and testing are shown (*Figure 4*). With sensitivity fixed at 95% and across nine experiments, the base DL model had an average AUROC of 0.774, which was 84.8% accurate and 31.1% specific. Other obtained values of performance include 91.3% F1, 87.8% PPV, and 52.2% NPV (*Table 2* and *Figure 5*).

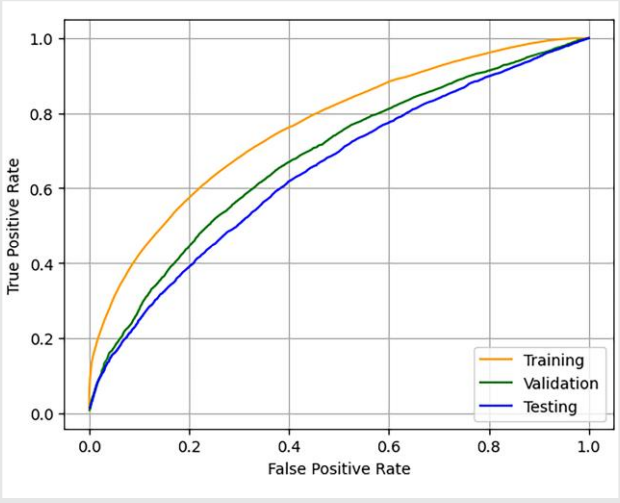


Figure 4. Average receiver operating characteristic curves for training, validation, and testing.

When considering model output by individual ECG leads, AUROC varied between 0.716 and 0.817 (III and V3, respectively), and the performance of the model with lead III was significantly lower than that of V3 ($P = 0.001$) (*Figure 6*). Further, accuracy was between 83.6 and 86.2% (aVF and V5, respectively), F1 was between 90.6 and 92.0% (aVF and V5, respectively), PPV was between 86.7 and 89.2% (aVF and V5, respectively), NPV was between 45.2 and 59.0% (aVF and V5, respectively), and specificity was between 23.4 and 40.1% (aVF and V5, respectively) (*Table 3*).

Base model explainability

Gradient-weighted class activation mapping visual analysis revealed a general emphasis on the QRS complex via all lead configurations (*Figure 7*). Tallies from 100 randomly selected GRAD-CAM images showed that all but 2 leads (V3 and V6) had P-wave contributions and there were as many as 14 from another lead (V1) (*Table 4*). All leads had T-wave contributions with as few as 3 (AVR) and as many as 22 (V1 and V4) while all leads involved the QRS complex from as few as 94 (AVL) to as many as 99 (V3 and V4). The P-Q interval and S-T interval were generally not influential in determining SCA.

Performance with alternative models

In a model in which age and sex were not considered (A1), the AUROC was 0.775 with a specificity of 33.0% (*Table 2*). Models created via ECGs recorded with increasing time from arrest, including <1 week, <1 month, between 1 day and 1 month, and between 1 month and 1 year (B1–B4, respectively) had AUROCs of 0.760, 0.734, 0.654, and 0.574, respectively, corresponding to specificities of 27.9, 25.4, 14.5, and 11.1%. When shockable rhythms were not included in an additional model (C1), the AUROC was 0.769 and the specificity was 33.1%, and when cases of asystole were further removed (C2), the AUROC was 0.773 with a specificity of 34.8%. When the temporal width around the QRS complex was reduced to 600 ms in another model (D1), the AUROC was 0.766 and the specificity was 30.8%, whereas when the temporal width was reduced to 500 ms (D2), the AUROC was 0.753 and the specificity was 28.5%. For each alternative model (A1–D2), averages across leads were generally consistent, although some were significantly different (see Supplementary material online, *Tables S1–S9*). For each alternative model, measures of performance varied across leads but were generally consistent with results obtained when all 12 leads were considered as an average; however, significant

Table 2 Measures of overall performance with sensitivity fixed at 95% for the base model, X, and its variations (A1, B1, B2, B3, B4, C1, C2, D1, and D2)

Model	Accuracy (%)	F1 (%)	PPV (%)	NPV (%)	Specificity (%)
X	84.8 (3.1)	91.3 (1.7)	87.8 (3.1)	52.2 (12.8)	31.1 (10.8)
A1	85.1 (2.8)	91.5 (1.5)	88.2 (2.8)	53.5 (13.3)	33.0 (11.4)
B1	82.2 (3.0)	89.6 (1.6)	84.7 (3.0)	54.7 (13.0)	27.9 (10.0)
B2	79.2 (2.8)	87.6 (1.6)	81.2 (2.8)	57.9 (11.3)	25.4 (8.7)
B3	87.6 (2.4)	93.3 (1.2)	91.6 (2.4)	21.7 (14.1)	14.5 (9.9)
B4	83.3 (2.4)	90.7 (1.3)	86.7 (2.4)	24.1 (15.4)	11.1 (8.1)
C1	87.4 (3.0)	93.0 (1.5)	91.0 (3.0)	46.2 (12.7)	33.1 (11.5)
C2	89.6 (2.7)	94.3 (1.4)	93.6 (2.7)	38.5 (12.4)	34.8 (14.0)
D1	84.7 (3.4)	91.2 (1.8)	87.7 (3.4)	51.4 (14.0)	30.8 (11.5)
D2	84.4 (3.2)	91.0 (1.7)	87.4 (3.2)	50.0 (13.0)	28.5 (10.4)

Values are reported as mean (SD).

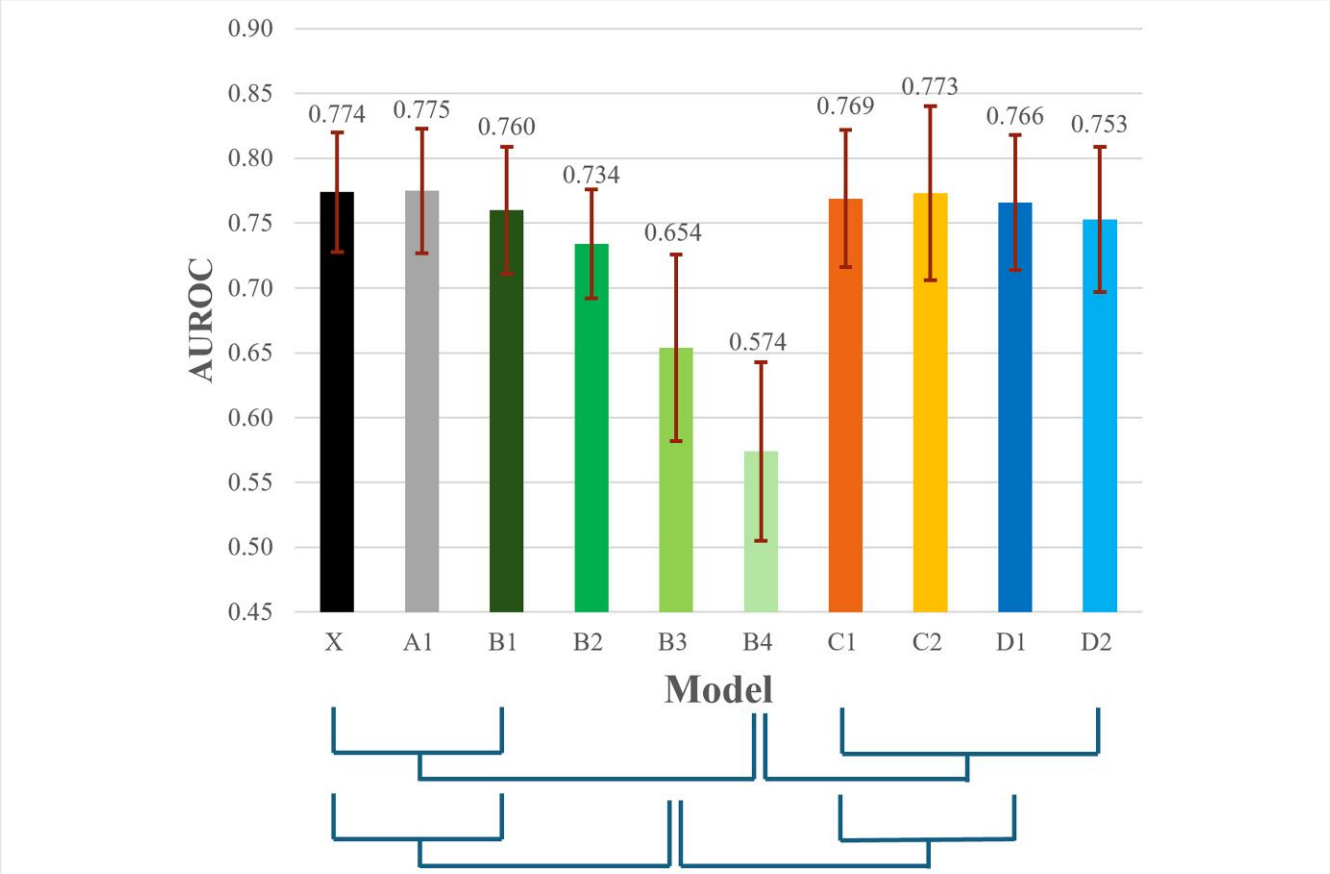


Figure 5. AUROC for each model with significant differences. B4 ≠ (X, A1, and B1); B4 ≠ (C1, C2, D1, and D2); B3 ≠ (X, A1, and B1); B3 ≠ (C1, C2, and D1).

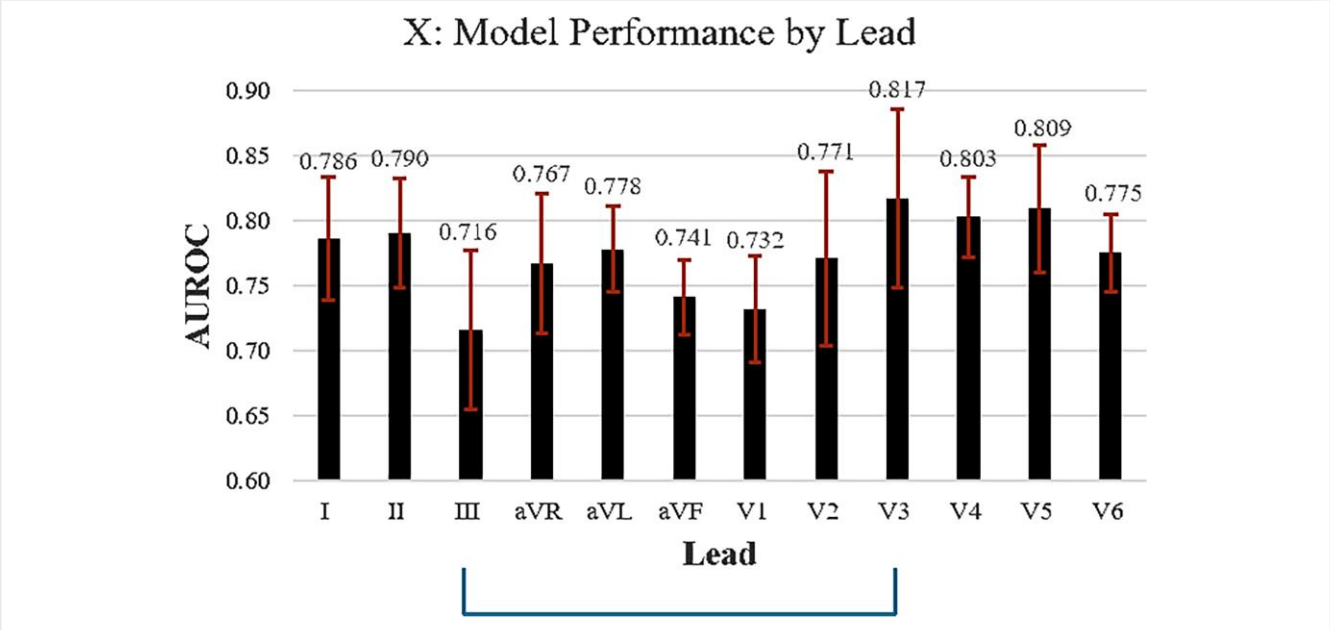


Figure 6. Model X performance by lead with significant differences. III ≠ V3.

Table 3 Measures of performance for the base case for each lead for the base model, X, with sensitivity fixed at 95%

	Lead	Accuracy (%)	F1 (%)	PPV (%)	NPV (%)	Specificity (%)
X	I	85.4 (3.0)	91.6 (1.8)	88.4 (3.4)	55.6 (13.8)	35.6 (13.4)
	II	84.9 (2.6)	91.3 (1.7)	87.9 (3.2)	56.2 (6.8)	33.7 (6.3)
	III	84.8 (1.6)	91.3 (1.1)	87.9 (2.0)	50.9 (16.0)	29.6 (11.2)
	aVR	84.4 (2.8)	91.0 (1.8)	87.4 (3.2)	52.9 (10.0)	30.1 (7.8)
	aVL	84.5 (1.4)	91.1 (1.0)	87.5 (1.9)	50.9 (15.1)	28.7 (9.8)
	aVF	83.6 (2.2)	90.6 (1.4)	86.7 (2.6)	45.2 (15.2)	23.4 (10.5)
	V1	84.0 (2.3)	90.9 (1.5)	87.1 (2.9)	48.5 (11.5)	25.7 (6.1)
	V2	84.5 (2.8)	91.1 (1.8)	87.6 (3.3)	50.0 (14.2)	28.6 (11.8)
	V3	85.6 (4.2)	91.7 (2.5)	88.7 (4.6)	55.0 (12.0)	36.9 (18.3)
	V4	84.7 (3.3)	91.3 (2.0)	87.8 (3.7)	51.4 (12.4)	30.8 (10.7)
	V5	86.2 (3.1)	92.0 (1.9)	89.2 (3.5)	59.0 (9.7)	40.1 (12.8)
	V6	84.6 (2.3)	91.2 (1.5)	87.6 (2.7)	50.8 (16.6)	29.4 (11.3)

Values are reported as mean (SD).

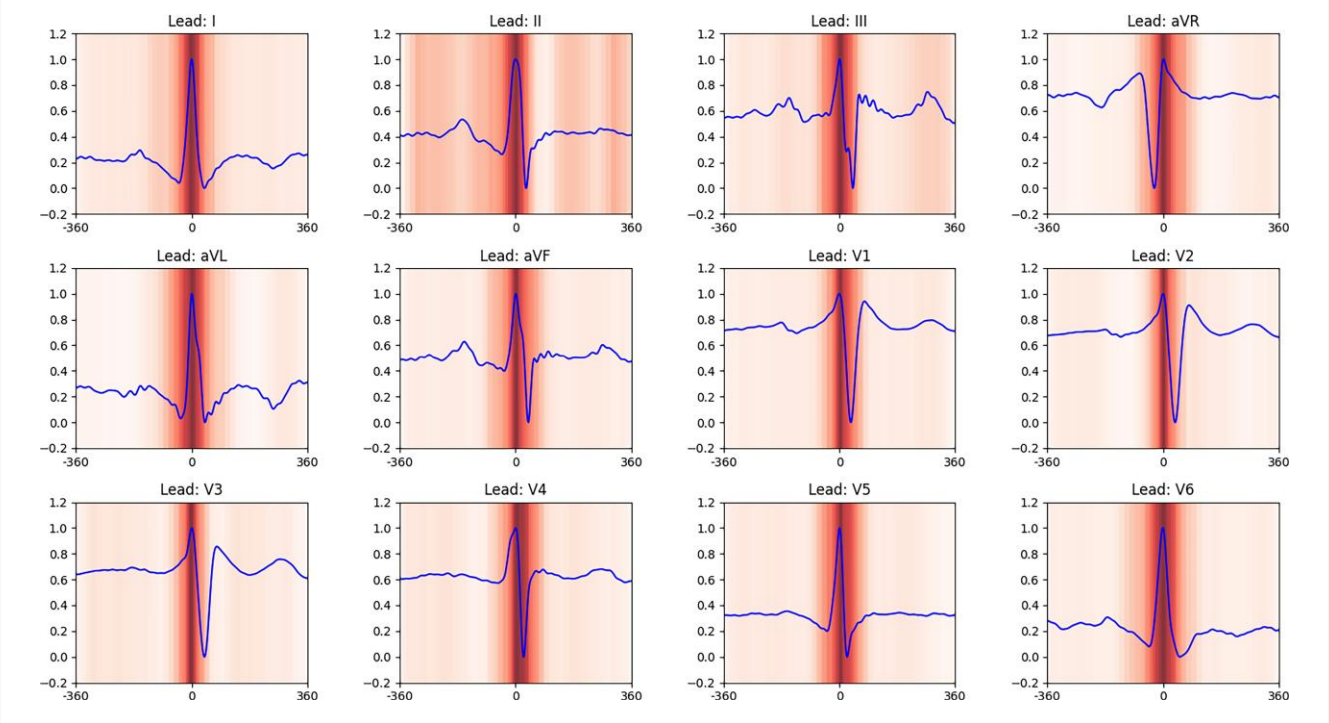


Figure 7. Representative GRAD-CAM images from the base case, X, in which only the QRS complex was highlighted. The color map is white (no activation) to red (full activation).

differences among leads were present (see Supplementary material online, *Tables S1–S9* and *Figures S4–S12*).

Discussion

Novelty

This is the first study to utilize the Nightingale Open Science—Subtyping Cardiac Arrest data set to predict SCA with DL. This is also the first DL application of ECG data to predict SCA in which determinations were made about the relative contributions of (i) risk factors such as age and sex, (ii) shockable vs. non-shockable rhythms, (iii) ECG data collected beyond 24 h prior to SCA, (iv) ECG data in varying temporal widths around the R-wave, and (v) regions of the ECG beyond the QRS complex.

Model implications

Output from the base model demonstrates that 10 s of ECG data collected 24 h prior to an SCA can predict 95% of SCA events. While the

Table 4 Tallies of the major contributing segments of the electrocardiogram to sudden cardiac arrest determination from 100 randomly selected gradient-weighted class activation mapping images from the base case, X

	P-Wave	P-Q interval	QRS complex	S-T interval	T-Wave
I	5	1	96	0	18
II	6	2	96	0	12
III	3	0	96	0	5
aVR	1	0	96	0	3
aVL	5	0	94	0	8
aVF	4	1	96	0	5
V1	14	0	95	0	22
V2	5	1	99	1	17
V3	0	0	99	0	16
V4	2	0	99	1	22
V5	4	0	98	0	20
V6	0	0	98	1	11

number of false positives is too high to enable its direct application, the model demonstrates that there is promise for a broadly refined future DL method that utilizes ECGs obtained from wearable devices to screen the public for SCA. Therefore, the current data justify exploration of prospective DL models built from larger and more diverse data sets.

Currently, some of the most influential factors in predicting SCA are age and sex.[14] In our models, differences in the ECG explained the contribution of age and sex in predicting SCA because the exclusion of age and sex did not change model performance. The explanation of age and sex is significant because the model is specific for SCA whereas age and sex are generic cardiovascular risk factors and thus do not differentiate SCA from other possible cardiovascular disorders. The finding is also important because age has a strongly positive and non-linear relationship with SCA,[35,36] and sudden cardiac death is two to three times more common in men than in women,[37–39] making SCA prediction via traditional tools such as linear regression more challenging whereas DL can explain these factors without adjustments. Finally, the model explaining age and sex is significant because these demographics are some of the few additional inputs beyond the ECG that would be commonly available to screen the public for SCA, and even they are not needed, which demonstrates the robustness and practicality of the approach.

The analyses were not able to determine differences among the arrhythmias associated with SCA. Removing the shockable rhythm arrests from the analysis did not influence model performance. This was also true when further removal excluded the asystole cases and left only pulseless electrical activity arrests for analysis. Nonetheless, the number of cases among the various presenting rhythms may not have been large enough to substantially influence model performance. In particular, a model based exclusively on shockable rhythms was not created because these arrests were <20% of those arrest rhythms in the data set and were thus not of sufficient quantity to independently consider.

Models that varied the time between ECG recording and SCA showed that physiologic differences indicative of an event are evident up to a year in advance, suggesting a chronic pathologic process leading to arrest in a large number of individuals. The finding is significant because it further encourages analysis of ECGs with DL as a potential screening tool in that it would allow more opportunity for prevention and intervention than if SCA were strictly an acute process. However, with longer time between ECG recording and arrest, model performance significantly diminished.

Model performance also had a diminishing trend with decreasing width around the QRS complex. The finding implies that ECG data relationships beyond the QRS complex confer information about an impending SCA. The varying width data are significant because they show that as much of the ECG should be included for SCA prediction as possible with the entirety of the signal being ideal.

Gradient-weighted class activation mapping analysis showed that the most relevant ECG features (in order of importance) are the QRS complex followed by the T-wave and then the P-wave. This result suggests that differences during ventricular depolarization and repolarization predict SCA and that atrial depolarization may also play a role. Whether this influence is due to electrical or mechanical dysfunction or both is yet to be determined. Gradient-weighted class activation mapping data in conjunction with analyses that isolate presenting rhythms could give prospective insight into pathologic mechanisms and anatomic sources of dysfunction that lead to SCA.

Comparison to prior sudden cardiac arrest prediction tools

The only other similar study utilizing DL with ECG data to predict SCA was performed by Kwon *et al.*[27] Consistent with some of the data from our study, ECG data were collected 24 h prior to arrest. However, beyond our scope, Kwon *et al.* performed both internal and external validation and reported AUROCs of 0.91 and 0.95, respectively. While these levels of performance are notably higher than ours, the data utilized for training and validation by this group are unbalanced, featuring relatively few cardiac arrests, which we believe facilitates the high AUROC and simultaneously results in their PPV of only around 8% as compared to our 88% PPV. Consistent with our findings, Kwon *et al.* showed that model performance among leads was comparable and described that the DL models' decision strategy primarily involved the QRS complex but also had contributions from other unspecified and not quantified temporal regions of the ECG.

Another study by the same group utilized DL to predict in-hospital SCA using heart rate, systolic blood pressure, respiratory rate, and body temperature and showed an AUROC of 0.85 and up to 76% sensitivity with 76% specificity, which outperformed logistic regression.[26] Of note, this DL algorithm used summary features of multiple physiologic signals rather than entire waveforms, and it is not clear if the ECG was utilized; heart rate could have been derived manually via auscultation or from analysis of the blood pressure waveform. While this approach is notable for its in-hospital capabilities, due to the reliance on up to four signals, it is not practical to employ for ECG prediction in the out-of-hospital setting, which is where 60% of SCAs occur.[11]

Even though the analysis here did not examine the relative contributions of ECG summary features in predicting SCA, it is reasonable to expect that the DL models captured their influence because they are weaker risk factors than age and sex,[14] which were explained. In theory, ECG analysis with DL is likely superior to other methods for SCA prediction because DL models can utilize characteristics that are beyond the threshold of human perception whereas the ECG summary features are only comprised of discrete portions of the ECG and result in a loss of information.

Other studies involving SCA prediction utilize either summary features extracted from the ECG[18–21] and/or other physiologically based metrics such as left ventricular ejection fraction[7,22,40] as input to statistical models and then report findings via odds ratios or relative risks. The endpoint of our study was the classification of an ECG as either a SCA or a control, and thus, the only appropriate way to express model performance is via measures of performance (e.g. AUROC, sensitivity, and

specificity). Therefore, the results of our DL model cannot be resolved with studies that utilized traditional approaches.

Clinical considerations

In our study, sensitivity was set at 95% while all other performance metrics were dependent. In general, sensitivity and specificity are negatively correlated, and therefore, the lower the sensitivity, the higher the specificity. However, it is uncertain as to whether decreasing sensitivity to increase specificity could result in a combination that is clinically relevant because such thresholds have not been defined. Nonetheless, in the theoretical application of the current model, having more than half false positives when the annual individual risk of SCA is 0.03–0.10%[14] would create a combination of apathy among users and an unrealistic and unnecessary burden on EDs. Nonetheless, output from the current models shows that age and sex do not improve prediction and do not need to be included, ECGs up to a year prior to an SCA event have predictive value and should be used as input, and, if possible, the entirety of the ECG signal should be used as input, despite the decision strategy relying most on the QRS complex and T- and P-waves, in that order.

The use of a future DL model to screen for SCA via ECG data could be highly impactful because 20–50% of SCA cases have no previously identified heart disease,[3–6] 50% of SCAs are first cardiac events,[10] and over 60% of cardiac arrests occur outside the hospital[41] where delays in initiating CPR and defibrillation are common, bystander CPR is suboptimal,[42] and there is a lack of epinephrine until paramedics arrive. Also, the proliferation of ECG recording wearable devices makes such an application readily available as a potentially inexpensive yet practical SCA screening tool for the general public. Widespread adoption could improve outcomes, enhance proactions, and thereby decrease dependency on paramedics for emergency transport and resuscitation.

Limitations

This study was limited in that all data were provided via the National Taiwan University Hospital; thus, the study population is not diverse enough to make conclusions that are generalizable to western populations. A second limitation is that this was a retrospective, case–control study in which the incidence of SCA is greatly inflated relative to the general population. Also, the SCA group here had a disproportionately large number of men. Thus, a more sex-balanced prospective cohort study is necessary to allow more confidence in the findings. Third, other DL models were not explored, and as a result, there may be alternative methods available that are more predictive of SCA. Fourth, comorbidity and other clinical data are not available for the control group, thus preventing the use of that information in the overall model. Fifth, with further regard to the control group, precise times for each individual follow-up from ECG recording to ED visit are not specifically known, and it cannot be guaranteed that members of the control group did not have cardiac arrests after their follow-up periods. Sixth, only those data certain widths around the QRS complex were utilized because the model would not converge when the entire signal was input, and it cannot be guaranteed that the entire T-wave was consistently captured. Finally, DL is highly dependent on large data sets, and a few hundred cases may not be enough to reveal the full potential of its algorithms to predict SCA exclusively from ECG waveforms. Overcoming these issues is necessary to create a generalized screening tool for SCA.

Conclusions

This study shows that 10 s of single-lead ECG data enables the prediction of an impending SCA within 24 h with 95% sensitivity and 31% specificity, which demonstrates potential for screening via DL. This method also explains differences in SCAs due to age and sex. Further, model performance improved when ECGs were nearer in time to SCAs, although ECG data up to a year prior had predictive value, suggesting that SCA commonly involves a chronic pathologic process. Sudden cardiac arrest prediction is also mostly dependent upon the QRS complex, implicating ventricular contraction as the primary source of dysfunction; however, the T- and P-waves had influence, suggesting that ventricular repolarization and atrial depolarization also contribute to dysfunction. Studies with larger data sets and more diverse patients are necessary to confirm these findings and to potentially increase specificity.

Supplementary material

Supplementary material is available at *European Heart Journal – Digital Health*.

Author contribution

Conceptualization: L.N., M.T.O., A.O., R.T.C. and H.R.H. Methodology: L.N., M.T.O, A.O., R.T.C., S.D., and H.R.H. Software: A.O. and L.N. Validation: A.O. and L.N. Formal analysis: L.N., A.O., and M.T.O. Investigation: L.N., A.O., and M.T.O. Resources: L.N., A.O., and M.T.O. Data curation: L.N., A.O., and M.T.O. Writing—original draft preparation: M.T.O., L.N., and A.O. Writing—review & editing: M.T.O., L.N., H.R.H., A.O., R.T.C., M.S.N., S.D., C.B., and S.J. Visualization: L.N., M.T.O., A.O., M.K., and R.T.C. Supervision: H.R.H., C.B., R.T.C., and S.D. Project administration: L.N., M.T.O., A.O., C.B., and H.R.H. Funding acquisition: N/A.

Funding

RTC is funded by the National Institutes of Health (T32HL007227, L30HL165535) and is a recipient of the Semyon and Janna Friedman Fellowship award. AO is funded by the Italian Ministry of Education, University and Research (project CNRBIOMICS PON R&I PIR01_00017). SD is funded by the European Union - NextGenerationEU through the Italian Ministry of University and Research under PNRR - M4C2-I1.3 Project PE_00000019 "HEAL ITALIA" - CUP J33C22002920006. The views and opinions expressed are those of the authors only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them.

Conflict of interest: none declared.

Data availability

The data used in this study are openly available via the Nightingale Open Science—Subtyping Cardiac Arrest data set at http://docs.ngsci.org/datasets/arrest-ntuh-ecg/.

References

1. Benjamin EJ, Muntner P, Alonso A, Bittencourt MS, Callaway CW, Carson AP, *et al.* Heart disease and stroke statistics—2019 update: a report from the American Heart Association. *Circulation* 2019;**139**:e56–e528.
2. Meaney PA, Bobrow BJ, Mancini ME, Christenson J, De Caen AR, Bhanji F, *et al.* Cardiopulmonary resuscitation quality: improving cardiac resuscitation outcomes both inside and outside the hospital: a consensus statement from the American Heart Association. *Circulation* 2013;**128**:417–435.
3. Kannel WB, Doyle JT, McNamara PM, Quickenton P, Gordon T. Precursors of sudden coronary death. Factors related to the incidence of sudden death. *Circulation* 1975;**51**:606–613.
4. Kuller LH. Sudden death—definition and epidemiologic considerations. *Prog Cardiovasc Dis* 1980;**23**:1–12.
5. Goldstein S. The necessity of a uniform definition of sudden coronary death: witnessed death within 1 hour of the onset of acute symptoms. *Am Heart J* 1982;**103**:156–159.
6. Myerburg RJ, Interian A, Mitrani RM, Kessler KM, Castellanos A. Frequency of sudden cardiac death and profiles of risk. *Am J Cardiol* 1997;**80**:10F–19F.

7. Stecker EC, Vickers C, Waltz J, Socoteanu C, John BT, Mariani R, *et al.* Population-based analysis of sudden cardiac death with and without left ventricular systolic dysfunction: two-year findings from the Oregon Sudden Unexpected Death Study. *J Am Coll Cardiol* 2006;**47**:1161–1166.
8. Chiuve SE, Fung TT, Rexrode KM, Spiegelman D, Manson JE, Stampfer MJ, *et al.* Adherence to a low-risk, healthy lifestyle and risk of sudden cardiac death among women. *JAMA* 2011;**306**:62–69.
9. Modi S, Krahn AD. Sudden cardiac arrest without overt heart disease. *Circulation* 2011; **123**:2994–3008.
10. Myerburg RJ. Sudden cardiac death: interface between pathophysiology and epidemiology. *Card Electrophysiol Clin* 2017;**9**:515–524.
11. Berdowski J, Berg RA, Tijssen JGP, Koster RW. Global incidences of out-of-hospital cardiac arrest and survival rates: systematic review of 67 prospective studies. *Resuscitation* 2010;**81**:1479–1487.
12. Andersson A, Arctaedius I, Cronberg T, Levin H, Nielsen N, Friberg H, *et al.* In-hospital versus out-of-hospital cardiac arrest: characteristics and outcomes in patients admitted to intensive care after return of spontaneous circulation. *Resuscitation* 2022;**176**:1–8.
13. Merchant RM, Topjian AA, Panchal AR, Cheng A, Aziz K, Berg KM, *et al.* Part 1: executive summary: 2020 American Heart Association guidelines for cardiopulmonary resuscitation and emergency cardiovascular care. *Circulation* 2020;**142**:S337–S357.
14. Ha ACT, Doumouras BS, Wang CN, Tranmer J, Lee DS. Prediction of sudden cardiac arrest in the general population: review of traditional and emerging risk factors. *Can J Cardiol* 2022;**38**:465–478.
15. Neri L, Oberdier MT, van Abeelen KCJ, Menghini L, Tumarkin E, Tripathi H, *et al.* Electrocardiogram monitoring wearable devices and artificial-intelligence-enabled diagnostic capabilities: a review. *Sensors (Basel)* 2023;**23**:4805.
16. Neri L, Oberdier MT, Augello A, Suzuki M, Tumarkin E, Jaipalli S, *et al.* Algorithm for mobile platform-based real-time QRS detection. *Sensors (Basel)* 2023;**23**:1625.
17. Neri L, Corazza I, Oberdier MT, Lago J, Gallelli I, Cicero AFG, *et al.* Comparison between a single-lead ECG garment device and a Holter monitor: a signal quality assessment. *J Med Syst* 2024;**48**:57.
18. Chugh SS, Reinier K, Singh T, Uy-Evanado A, Socoteanu C, Peters D, *et al.* Determinants of prolonged QT interval and their contribution to sudden death risk in coronary artery disease: the Oregon Sudden Unexpected Death Study. *Circulation* 2009;**119**:663–670.
19. Teodorescu C, Reinier K, Uy-Evanado A, Navarro J, Mariani R, Gunson K, *et al.* Prolonged QRS duration on the resting ECG is associated with sudden death risk in coronary disease, independent of prolonged ventricular repolarization. *Heart Rhythm* 2011; **8**:1562–1567.
20. Teodorescu C, Reinier K, Uy-Evanado A, Gunson K, Jui J, Chugh SS. Resting heart rate and risk of sudden cardiac death in the general population: influence of left ventricular systolic dysfunction and heart rate-modulating drugs. *Heart Rhythm* 2013;**10**:1153–1158.
21. La Rovere MT, Pinna GD, Maestri R, Mortara A, Capomolla S, Febo O, *et al.* Short-term heart rate variability strongly predicts sudden cardiac death in chronic heart failure patients. *Circulation* 2003;**107**:565–570.
22. Reinier K, Narayanan K, Uy-Evanado A, Teodorescu C, Chugh H, Mack WJ, *et al.* Electrocardiographic markers and the left ventricular ejection fraction have cumulative effects on risk of sudden cardiac death. *JACC Clin Electrophysiol* 2015;**1**:542–550.
23. Holmstrom L, Bednarski B, Chugh H, Aziz H, Pham HN, Sargsyan A, *et al.* Artificial intelligence model predicts sudden cardiac arrest manifesting with pulseless electric activity versus ventricular fibrillation. *Circ Arrhythm Electrophysiol* 2024;**17**:e012338.
24. Abdelghani SA, Rosenthal TM, Morin DP. Surface electrocardiogram predictors of sudden cardiac arrest. *Ochsner J* 2016;**16**:280–289.
25. Moffat LM, Xu D. Accuracy of machine learning models to predict in-hospital cardiac arrest: a systematic review. *Clin Nurse Spec* 2022;**36**:29–44.
26. Kwon J, Lee Y, Lee Y, Lee S, Park J. An algorithm based on deep learning for predicting in-hospital cardiac arrest. *J Am Heart Assoc* 2018;**7**:e008678.
27. Kwon J-M, Kim K-H, Jeon K-H, Lee SY, Park J, Oh B-H. Artificial intelligence algorithm for predicting cardiac arrest using electrocardiography. *Scand J Trauma Resusc Emerg Med* 2020;**28**:98.
28. Goodfellow I, Bengio Y, Courville A. *Deep Learning*. Cambridge, MA 02142, USA: The MIT Press; 2016.
29. Huang CH, Su R, Huang HC, Lin K, Foster N, Juergens N, *et al.* Subtyping Cardiac Arrest with ECG Waveforms: A Nightingale Open Science Dataset. 2021. https://doi.org/10.48815/N5WC7D. Available online: (accessed on 7 June 2023).
30. Mullainathan S, Obermeyer Z. Solving medicine's data bottleneck: Nightingale Open Science. *Nat Med* 2022;**28**:897–899.
31. Bota P, Silva R, Fred A, da Silva HP. BioSPPy: a Python toolbox for physiological signal processing, vol. 26. SoftwareX; 2024. pp. 101712. https://doi.org/10.1016/j.softx.2024.101712.
32. Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, CITRO C, *et al.* TensorFlow: large-scale machine learning on heterogeneous distributed systems. *arXiv* 2016. https://doi.org/10.48550/arXiv.1603.04467.
33. Kingma DP, Ba J. *Adam: a Method for Stochastic Optimization*: arXiv; 2017.
34. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: visual explanations from deep networks via gradient-based localization. In: *Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV)*, 2017:p.618–626.
35. Narayan SM, Wang PJ, Daubert JP. New concepts in sudden cardiac arrest to address an intractable epidemic: JACC state-of-the-art review. *J Am Coll Cardiol* 2019;**73**: 70–88.
36. Maruyama M, Ohira T, Imano H, Kitamura A, Kiyama M, Okada T, *et al.* Trends in sudden cardiac death and its risk factors in Japan from 1981 to 2005: the Circulatory Risk in Communities Study (CIRCS). *BMJ Open* 2012;**2**:e00057.
37. Albert CM, McGovern BA, Newell JB, Ruskin JN. Sex differences in cardiac arrest survivors. *Circulation* 1996;**93**:1170–1176.
38. Chugh SS, Uy-Evanado A, Teodorescu C, Reinier K, Mariani R, Gunson K, *et al.* Women have a lower prevalence of structural heart disease as a precursor to sudden cardiac arrest: the Ore-SUDS (Oregon Sudden Unexpected Death Study). *J Am Coll Cardiol* 2009;**54**:2006–2011.
39. Haukilahti MAE, Holmström L, Vähätalo J, Kenttä T, Tikkanen J, Pakanen L, *et al.* Sudden cardiac death in women. *Circulation* 2019;**139**:1012–1021.
40. Moss AJ, Greenberg H, Case RB, Zareba W, Hall WJ, Brown MW, *et al.* Long-term clinical course of patients after termination of ventricular tachyarrhythmia by an implanted defibrillator. *Circulation* 2004;**110**:3760–3765.
41. Benjamin EJ, Virani SS, Callaway CW, Chamberlain AM, Chang AR, Cheng S, *et al.* Heart disease and stroke statistics—2018 update: a report from the American Heart Association. *Circulation* 2018;**137**:e67–e492.
42. Chocron R, Jobe J, Guan S, Kim M, Shigemura M, Fahrenbruch C, *et al.* Bystander cardiopulmonary resuscitation quality: potential for improvements in cardiac arrest resuscitation. *J Am Heart Assoc* 2021;**10**:e017930.