ESC European Society of Cardiology

European Heart Journal - Digital Health (2024) **5**, 384–388
https://doi.org/10.1093/ehjdh/ztae004

SHORT REPORT

Artificial intelligence for ventricular arrhythmia capability using ambulatory electrocardiograms

Joseph Barker [1,2,3,4,5,*†], Xin Li [1,6,†], Ahmed Kotb [1,2], Akash Mavilakandy[2], Ibrahim Antoun[1,3], Chokanan Thaitirarot[1], Ivelin Koev[1,2], Sharon Man[1,2], Fernando S. Schlindwein [4,6], Harshil Dhutia[1,2], Shui Hao Chin[1,2], Ivan Tyukin [7], William B. Nicolson[1,2], and G. Andre Ng [1,2,4]

[1]Department of Cardiovascular Sciences, University of Leicester, Glenfield Hospital, Groby Road, Leicester LE3 9QP, UK; [2]Cardiology Department, Glenfield Hospital, University Hospitals of Leicester, Groby Road, Leicester LE3 9QP, UK; [3]Cardiology Department, Kettering General Hospital, University Hospitals of Northamptonshire NHS Group, Rothwell Rd, Kettering NN16 8UZ, UK; [4]National Institute for Health Research, Leicester Biomedical Research Centre, Groby Road, Glenfield Hospital, Leicester LE3 9QP, UK; [5]National Heart and Lung Institute, Imperial College London, Du Cane Road, Hammersmith Hospital, London W12 0HS, UK; [6]School of Engineering, University of Leicester, University Road, Leicester LE1 7RH, UK; and [7]Department of Mathematics, King's College London, Strand Campus, Strand, London WC2R 2LS, UK

Received 16 October 2023; revised 11 January 2024; accepted 15 January 2024; online publish-ahead-of-print 30 January 2024

Aims European and American clinical guidelines for implantable cardioverter defibrillators are insufficiently accurate for ventricular arrhythmia (VA) risk stratification, leading to significant morbidity and mortality. Artificial intelligence offers a novel risk stratification lens through which VA capability can be determined from the electrocardiogram (ECG) in normal cardiac rhythm. The aim of this study was to develop and test a deep neural network for VA risk stratification using routinely collected ambulatory ECGs.

Methods and results A multicentre case–control study was undertaken to assess VA-ResNet-50, our open source ResNet-50-based deep neural network. VA-ResNet-50 was designed to read pyramid samples of three-lead 24 h ambulatory ECGs to decide whether a heart is capable of VA based on the ECG alone. Consecutive adults with VA from East Midlands, UK, who had ambulatory ECGs as part of their NHS care between 2014 and 2022 were recruited and compared with all comer ambulatory electrograms without VA. Of 270 patients, 159 heterogeneous patients had a composite VA outcome. The mean time difference between the ECG and VA was 1.6 years (⅓ ambulatory ECG before VA). The deep neural network was able to classify ECGs for VA capability with an accuracy of 0.76 (95% confidence interval 0.66–0.87), F1 score of 0.79 (0.67–0.90), area under the receiver operator curve of 0.8 (0.67–0.91), and relative risk of 2.87 (1.41–5.81).

Conclusion Ambulatory ECGs confer risk signals for VA risk stratification when analysed using VA-ResNet-50. *Pyramid sampling* from the ambulatory ECGs is hypothesized to capture autonomic activity. We encourage groups to build on this open-source model.

Question Can artificial intelligence (AI) be used to predict whether a person is at risk of a lethal heart rhythm, based solely on an electrocardiogram (an electrical heart tracing)?

Findings In a study of 270 adults (of which 159 had lethal arrhythmias), the AI was correct in 4 out of every 5 cases. If the AI said a person was at risk, the risk of lethal event was three times higher than normal adults.

* Corresponding author. Email: joseph.barker@nhs.net
†Joint first authors.

Meaning In this study, the AI performed better than current medical guidelines. The AI was able to accurately determine the risk of lethal arrhythmia from standard heart tracings for 80% of cases over a year away—a conceptual shift in what an AI model can see and predict. This method shows promise in better allocating implantable shock box pacemakers (implantable cardioverter defibrillators) that save lives.

Keywords Ventricular arrhythmia • Deep learning • Risk stratification • Artificial intelligence • Implantable cardioverter defibrillator • Neural network

Introduction

Ventricular arrhythmias (VAs) can be lethal, with survival determined by access to effective defibrillation. Delays in defibrillation are associated with functional disability, care dependency, and death.[1] External defibrillators are not accessed in 90% of cases and current implantable cardioverter defibrillators (ICDs) guidelines are insufficiently accurate, including fewer than 20% of all VAs.[2,3] Novel artificial intelligence (AI) prediction methodologies may improve current guidelines to determine VA capability and therefore assign ICDs more accurately.[4] The aim of this study is to develop a deep neural network for VA risk stratification using routinely collected ambulatory electrocardiograms (ECGs) in order to predict VA capability from a patient's normal cardiac rhythm.

Methods

A multicentre retrospective deep-learning-based case–control study was undertaken at University Hospitals of Leicester (UHL: Leicester General Hospital, Glenfield Hospital & Leicester Royal Infirmary) and University Hospitals of Northamptonshire NHS Group (UHN: Kettering Hospital), UK. The eligibility criterion for the VA cohort was consecutive adults with International Classification of Diseases 10 (ICD10) diagnoses of I47.2 ventricular tachycardia (VT) and I49.0 ventricular fibrillation/flutter (VF) with ambulatory ECGs during the period between 2014 and 2022. For patients with multiple ambulatory ECGs, only the earliest recorded one was taken. There is no established power factor for AI studies, and therefore, the largest possible cohort was sought. The eligibility criterion for the comparator cohort was consecutive adults with ambulatory ECGs over a 5-day period in November 2022.

Signal processing

Three-lead ambulatory 24 h ECGs (Spacelabs Lifecard CF Holter monitors 128 Hz) were exported from Spacelabs in ISHNE (.ecg) format. Electrocardiograms were pre-processed using a second-order Bessel filter, passband between 0.1 and 50 Hz, and a notch filter at 50 Hz. R peaks in the ECG data were detected using the Pan–Tompkins algorithm. Smoothed RR intervals on the ECG were obtained by a moving average filter with a factor of 40 samples before converting to heart rate. *Pyramid sampling* was undertaken by categorizing computed heart rates into 100 distinct levels ranging from slow to fast rates; these levels were defined by uniformly distributed percentiles regardless of the time of the day. One 10 s segment was selected randomly from each level and formed one input sample. This selection process was repeated 100 times for each patient and for each of the three ECG leads. As a result, for each patient, 100 × 3D tensors were generated in the format of 100X1280X3 (*Figure 1*).

Data flow

A patient-wise partitioning strategy was employed, allocating 80% for training and 20% for testing. Within the training data, we instituted a five-fold cross-validation procedure partitioned by patients. Consequently, this yielded five distinct models from which an ensemble approach converged the predictions to deliver the final prediction model on the unseen test data (*Figure 1*).

Deep learning architecture

A transfer-learned, customized ResNet-50-based convolutional neural network was employed with a modified input layer that we name VA-ResNet-50. The architecture was organized into blocks of convolutional and identity layers with 53 convolutional layers using various filter sizes, 53 batch normalization layers, 49 ReLU layers, and 5 pooling layers. In addition to these, we incorporated three fully connected layers, two dropout layers, L2 regularization (a regularization factor of 0.01), a softmax layer, and a classification layer (*Figure 1*). We employed a learning rate of 1e−4, a decay factor of 0.1 every 5 epochs, for 20 epochs max with early stopping at 2 consecutive epochs, and adam optimization. The batch size was 64, and the dropout layer rate was 0.6. The open-source model is available on GitHub.[5] The model served to take 100 × 3D ECG tensors per patient that are individually classified by the model to generate the averaged class probability score for the per-patient prediction, with a threshold arbitrarily set to 0.5.

Statistical analysis

The performance of the neural network was evaluated in the test set using metrics such as overall accuracy, F1 score, sensitivity, specificity, area under the receiver operator curve (AUC), and area under the precision-recall curve. Confidence intervals for performance metrics were calculated using bootstrapping with 1000 iterations. All analyses were completed in MATLAB (Mathworks, USA).

Permissions and reporting

Study permissions were granted from the respective institutional review committees; registration numbers; UHN: REF8882 and UHL: REF11434. This paper was reported according to STARD2015 reporting guidelines.

Results

The study comprised 270 patients—178 UHN and 92 UHL patients. Ventricular arrhythmia occurred in 159 patients [mean (95% confidence interval, CI) age 61 years (57–65); 78 females (49%)] compared with 111 patients without VA [age 58 years (54–62); 56 females (51%)]. The VA-positive and -negative cohorts differed in terms of proportions of cardiovascular risk factors, electrocardiographically distinct diagnoses, and cardiomyopathies, but all diagnoses were represented within both cohorts. For the composite outcome of VA, VT was 88% ($n = 140$), while 24% ($n = 39$) represented VF and 6% represented both VT and VF diagnoses (*Table 1*). The mean time difference between the ECG and VA was 1.6 years, with 27% of the cohort having an ambulatory ECG before VA.

Deep learning performance

For the testing dataset, from a normal cardiac rhythm, the model was able to classify by VA capability with an F1 score of 0.79 (0.67–0.90). *Figure 1* displays the confusion matrix. The accuracy was 0.76 (95% CI 0.66–0.87). *Figure 1* shows the receiver operator characteristic curve for the test dataset; the AUC was 0.8 (0.67–0.91). The RR was 2.87 (1.41–5.81) (*Table 1*).

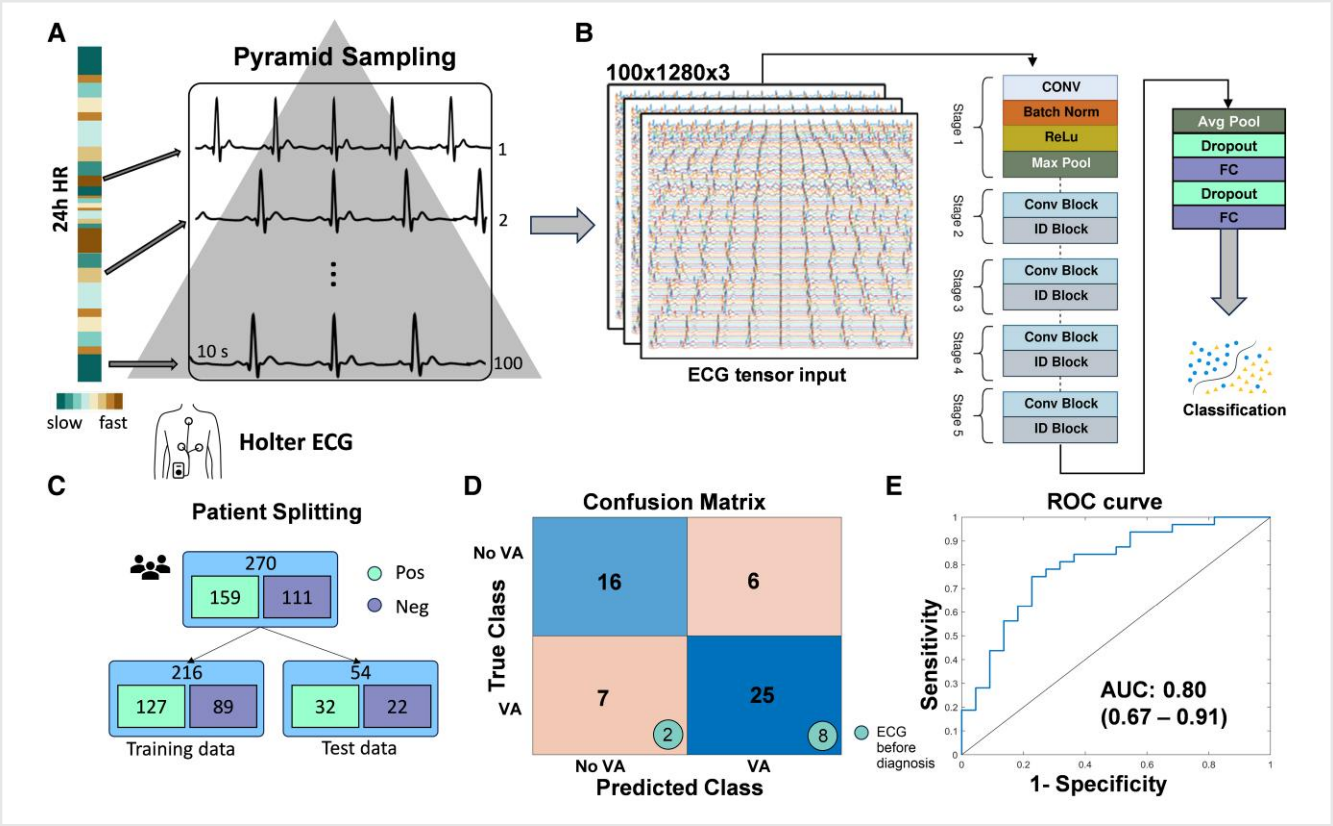

Figure 1 (*A*) A pyramid sampling schematic demonstrating 100 samples at various heart rates over a 24 h period. (*B*) VA-ResNet-50 architecture. (*C*) Patient flow. (*D*) Confusion matrix including participants before and after electrocardiogram. (*E*) Receiver operator characteristic curve. ROC, receiver operator characteristic; VA, ventricular arrhythmia; ECG, electrocardiogram; avg pool, average pooling; batch norm, batch normalization; ID block, identity block; conv block, convolutional block; FC, fully connected; ReLu, Rectified Linear Unit.

Table 1 Patient characteristics and performance metrics

Patient characteristics	Ventricular arrhythmia (*N* = 159)	No ventricular arrhythmia (*N* = 111)
Demographics		
Age in years, *n* (95% CI)	61 (57–65)	58 (54–62)
Female, *n* (%)	78 (49%)	56 (51%)
Cardiomyopathies		
Ischaemic heart disease, *n* (%)	92 (59%)	19 (17%)
Inherited cardiomyopathy, *n* (%)	24 (15%)	3 (3%)
Hypertrophic cardiomyopathy, *n* (%)	7 (4%)	1 (1%)
Heart failure (any), *n* (%)	47 (30%)	13 (12%)
Dilated cardiomyopathy, *n* (%)	10 (6%)	1 (<1%)
Myocarditis, *n* (%)	2 (1%)	0 (0%)
Valvular heart disease, *n* (%)	69 (43%)	11 (10%)
Cardiovascular risk factors		
Hypertension, *n* (%)	98 (61%)	34 (31%)
Type 2 diabetes, *n* (%)	16 (10%)	5 (5%)
Chronic obstructive pulmonary disease, *n* (%)	22 (14%)	6 (5%)
Chronic kidney disease, *n* (%)	1 (<1%)	1 (<1%)
Dyslipidaemia, *n* (%)	51 (32%)	16 (14%)
Syncope, *n* (%)	22 (14%)	7 (6%)

Continued

Table 1 *Continued*

Patient characteristics	Ventricular arrhythmia ($N = 159$)	No ventricular arrhythmia ($N = 111$)
Electrocardiographically distinct diagnoses		
Atrial fibrillation or flutter, *n* (%)	56 (35%)	14 (13%)
Left bundle branch block, *n* (%)	18 (11%)	3 (3%)
Complete heart block, *n* (%)	3 (2%)	0 (0%)
Ventricular arrhythmia outcomes		
Ventricular tachycardia, *n* (%)	131 (82%)	0 (0%)
Ventricular fibrillation/flutter, *n* (%)	36 (23%)	0 (0%)
Both ventricular tachycardia and fibrillation/ flutter	8 (5%)	0 (0%)
Presence of implantable cardioverter defibrillator	23 (15%)	0 (0%)
Time between ambulatory electrocardiogram and ventricular arrhythmia		
Mean time difference	1.6 years ($n = 159$)	NA
ECG collected before VA	0.4 years ($n = 43$, 27%)	NA
ECG collected after VA	2.1 years ($N = 117$, 73%)	NA
VA-ResNet-50 performance		
Performance metric	**Train set—80%—five-fold cross-validation results**	**Test set—20%—ensemble per patient results**
Accuracy (95% CI)	0.71 (0.6–0.82)	0.76 (0.66–0.87)
AUC (95% CI)	0.76 (0.64–0.88)	0.80 (0.67–0.91)
AUCPR (95% CI)	0.77 (0.65–0.89)	0.81 (0.64–0.91)
F1 score (95% CI)	0.76 (0.66–0.86)	0.79 (0.67–0.90)
Balanced accuracy (95% CI)	0.70 (0.53–0.83)	0.76 (0.64–0.87)
PPV (95% CI)	0.75 (0.65–0.85)	0.81 (0.67–0.95)
NPV (95% CI)	0.67 (0.53–0.81)	0.70 (0.51–0.89)
RR (95% CI)	2.32 (1.36–3.28)	2.87 (1.41–5.81)
Sensitivity (95% CI)	0.62 (0.48–0.76)	0.78 (0.64–0.92)
Specificity (95% CI)	0.78 (0.67–0.89)	0.73 (0.52–0.91)

PPV, positive predictive value; NPV, negative predictive value; RR, relative risk.

Discussion

VA-ResNet-50, a deep neural network classifier for ambulatory ECGs, demonstrates that signals exist for VA capability from the normal intrinsic cardiac rhythm. This re-look at ubiquitously available, non-invasive, cheap cardiovascular patient data holds promise to assign ICDs with greater precision and more economically than current practice.[6] The AUC of 0.8 is consistent with a growing body of evidence for VA prediction with AI.[4,7] Specifically for ambulatory ECGs, we report a lower AUC than the unpublished results of Fiorina *et al.*[8] (AUC = 0.91) to predict incident sustained VT. This difference is likely because of time disparities between the ECG and VA between cohorts, 2 weeks vs. 1.6 years. This work improves on work published by Sammani *et al.*[9] who describe an AUC of 0.67, although with an explainable auto-encoder model within a cohort of dilated cardiomyopathy. This difference might be explained by their refined 'life-threatening' VA outcome—granularity of outcome not available to ICD10. Our newly termed *pyramid sampling* could explain our good performance, as autonomic nervous activity is a prognostic marker for VAs that manifests over a range of heart rates.[10]

Limitations

A retrospective recruitment of VA meant that ambulatory ECGs were frequently available after VA, which introduces survivor bias. The traumatic nature of VA may induce electrophysiological changes to be detected. The heterogeneity of the cohort is representative of real-world VA but precludes the convention of mechanistic understanding required of evidence-based medicine. The comparator cohort is similarly comorbid but healthier than the original intended comparator cohort, meaning that the model is at risk of classifying the heart health index as opposed to a VA risk specifically. The intended comparator cohort - those with ICDs and no VA - were not available because they do not undergo ambulatory ECGs due device electrogram (EGM) availability. The outcomes were derived only from secondary care billing data, which can result in misclassification. This *ad hoc* data collection strategy likely confounds external validity, although this recruitment strategy is consistent with the field due to the suddenness and unexpected nature of VA.[4] Our current analysis does not include a systematic assessment of model stability to random and adversarial perturbations to data and model structure.[7] This, however, will be the topic of our future work.

Conclusions

Ambulatory ECGs recorded during a patient's normal activities of daily living contain signals for VA risk stratification when analysed using our open-source VA-ResNet-50 deep neural network. *Pyramid sampling* from the ambulatory ECGs is hypothesized to capture autonomic

activity. Importantly VA-ResNet-50 is agnostic of cardiomyopathy. The retrospective, *ad hoc* recruitment strategy limits generalizability. Prospective validation within other cohorts is planned, and we encourage other groups to build on this open-source model.

Acknowledgements

We acknowledge the guidance and expertise of Ms Michelle Newton, Mr Alan Bruce, Mr Graham Williams, Mr Richard Weatherhead, Ms Louise Waller, Mr Adam Brown, Ms Sally Berrill, Ms Kathy Clarke, and Ms Mei-Mei Chueng for their assistance in navigating the permissions and respective data infrastructures of the involved institutions. We acknowledge the cardiac physiologist expertise of Ms Jen Hambleton, Mr Vasco Rodrigues, and Ms Vanisha Jonas in data acquisition and analyses.

Funding

The project was funded by the John and Lucille van Geest Foundation (501100020410), reference number VG-044. J.B. is supported by the National Institute for Health and Care Research (NIHR) through the Integrated Academic Clinical Training Pathway; NIHR Academic Clinical Fellowship. I.T. is supported by the United Kingdom Research Institute UKRI Turing AI Fellowship EP/V025295/2. G.A.N. is supported by a British Heart Foundation Programme Grant (RG/17/3/32774) and NIHR Leicester Biomedical Research Centre (501100020013). X.L., W.B.N., F.S.S., and G.A.N. are supported by a Medical Research Council (501100000265) Biomedical Catalyst Developmental Pathway Funding Scheme (MR/S037306/1) and NIHR i4i grant (NIHR204553).

Conflict of interest: None declared.

Data availability

Data can be made available upon reasonable request, subject to respective institutional review board approvals.

References

1. Hayashi M, Shimizu W, Albert CM. The spectrum of epidemiology underlying sudden cardiac death. *Circ Res* 2015;**116**:1887–1906.
2. British Heart Foundation. Defibrillators used in just 10 per cent of out of hospital cardiac arrests study shows. British Heart Foundation; 2023. https://www.bhf.org.uk/what-we-do/news-from-the-bhf/news-archive/2023/june/defibrillators-used-in-just-10-per-cent-of-out-of-hospital-cardiac-arrests-study-shows (1 September 2023).
3. Wellens HJJ, Schwartz PJ, Lindemans FW, Buxton AE, Goldberger JJ, Hohnloser SH, *et al.* Risk stratification for sudden cardiac death: current status and challenges for the future. *Eur Heart J* 2014;**35**:1642–1651.
4. Barker J, Li X, Khavandi S, Koeckerling D, Mavilakandy A, Pepper C, *et al.* Machine learning in sudden cardiac death risk prediction: a systematic review. *Europace* 2022;**24**:1777–1787.
5. Li X, Barker J. (2023) VA-ResNet-50, GitHub. https://github.com/lixin0121/VA-resnet-50 (14 September 2023).
6. Dhutia H, Malhotra A, Finocchiaro G, Parpia S, Bhatia R, D'Silva A, *et al.* Diagnostic yield and financial implications of a nationwide electrocardiographic screening programme to detect cardiac disease in the young. *Europace* 2021;**23**:1295–1301. Erratum in: *Europace* 2021;23:1301.
7. Tyukin IY, Higham DJ, Gorban AN. On adversarial examples and stealth attacks in artificial intelligence systems. In: *2020 International Joint Conference on Neural Networks (IJCNN)*. Glasgow: IEEE; 2020. p1–6.
8. Fiorina L, Carbonati T, Narayanan K, Li J, Henry C, Singh J, *et al.* LB-456090-4 near-term prediction of life-threatening ventricular arrhythmias using artificial intelligence-enabled single lead ambulatory ECG. *Heart Rhythm* 2023;**20**:1084.
9. Sammani A, van de Leur RR, Henkens MTHM, Meine M, Loh P, Hassink RJ, *et al.* Life-threatening ventricular arrhythmia prediction in patients with dilated cardiomyopathy using explainable electrocardiogram-based deep neural networks. *Europace* 2022;**24**:1645–1654.
10. Ng GA, Brack KE, Patel VH, Coote JH. Autonomic modulation of electrical restitution, alternans and ventricular fibrillation initiation in the isolated heart. *Cardiovasc Res* 2007;**73**:750–760.