

# Electrocardiogram-based deep learning to predict mortality in paediatric and adult congenital heart disease

Joshua Mayourian<sup>1,2</sup>, Amr El-Bokl<sup>1,2</sup>, Platon Lukyanenko<sup>3</sup>, William G. La Cava<sup>3</sup>, Tal Geva<sup>1,2</sup>, Anne Marie Valente<sup>1,2</sup>, John K. Triedman<sup>1,2</sup>, and Sunil J. Ghelani <sup>1,2,\*</sup>

<sup>1</sup>Department of Cardiology, Boston Children's Hospital, Boston, MA, USA; <sup>2</sup>Department of Pediatrics, Harvard Medical School, Boston, MA, USA; and <sup>3</sup>Department of Pediatrics, Computational Health Informatics Program, Boston Children's Hospital, Harvard Medical School, Boston, MA, USA

Received 29 March 2024; revised 28 June 2024; accepted 13 September 2024

## Abstract

**Background and Aims** Robust and convenient risk stratification of patients with paediatric and adult congenital heart disease (CHD) is lacking. This study aims to address this gap with an artificial intelligence-enhanced electrocardiogram (ECG) tool across the lifespan of a large, diverse cohort with CHD.

**Methods** A convolutional neural network was trained (50%) and tested (50%) on ECGs obtained in cardiology clinic at the Boston Children's Hospital to detect 5-year mortality. Temporal validation on a contemporary cohort was performed. Model performance was evaluated using the area under the receiver operating characteristic and precision-recall curves.

**Results** The training and test cohorts composed of 112 804 ECGs (39 784 patients; ECG age range 0–85 years; 4.9% 5-year mortality) and 112 575 ECGs (39 784 patients; ECG age range 0–92 years; 4.6% 5-year mortality from ECG), respectively. Model performance (area under the receiver operating characteristic curve 0.79, 95% confidence interval 0.77–0.81; area under the precision-recall curve 0.17, 95% confidence interval 0.15–0.19) outperformed age at ECG, QRS duration, and left ventricular ejection fraction and was similar during temporal validation. In subgroup analysis, artificial intelligence-enhanced ECG outperformed left ventricular ejection fraction across a wide range of CHD lesions. Kaplan–Meier analysis demonstrates predictive value for longer-term mortality in the overall cohort and for lesion subgroups. In the overall cohort, precordial lead QRS complexes were most salient with high-risk features including wide and low-amplitude QRS complexes. Lesion-specific high-risk features such as QRS fragmentation in tetralogy of Fallot were identified.

**Conclusions** This temporally validated model shows promise to inexpensively risk-stratify individuals with CHD across the lifespan, which may inform the timing of imaging/interventions and facilitate improved access to care.

\* Corresponding author. Tel: +1 617 355 4278, Email: [sunil.ghelani@cardio.chboston.org](mailto:sunil.ghelani@cardio.chboston.org)

© The Author(s) 2024. Published by Oxford University Press on behalf of the European Society of Cardiology. All rights reserved. For commercial re-use, please contact [reprints@oup.com](mailto:reprints@oup.com) for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact [journals.permissions@oup.com](mailto:journals.permissions@oup.com).

## Structured Graphical Abstract

### Key Question

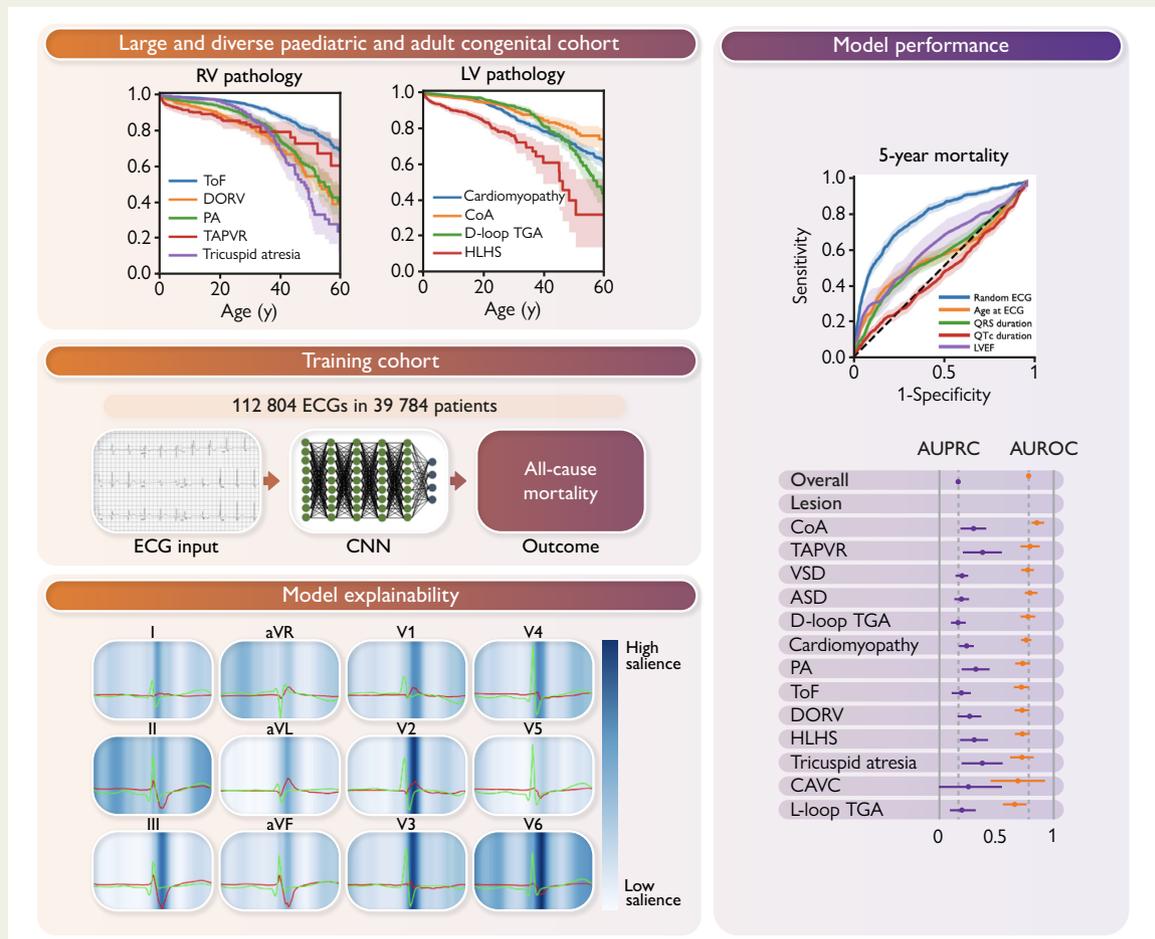
Can artificial intelligence-enhanced electrocardiogram (AI-ECG) predict 5-year mortality in paediatric and adult patients with congenital heart diseases?

### Key Finding

AI-ECG successfully predicted 5-year mortality and outperformed conventional markers such as age, ejection fraction, and QRS duration.

### Take Home Message

AI-ECG can provide inexpensive and convenient risk stratification in patients with congenital heart disease across lifespan, which may inform clinical decision-making and improve access to care.



A large and diverse paediatric and adult congenital heart disease cohort was used to train and test an artificial intelligence-enhanced electrocardiogram algorithm to accurately predict 5-year mortality across a range of congenital heart disease lesions. In an effort to interpret model behaviour, model explainability analysis was performed. AI-ECG, artificial intelligence-enhanced electrocardiogram; ASD, atrial septal defect; CAVC, complete atrioventricular canal defect; CNN, convoluted neural network; CoA, coarctation of the aorta; DORV, double outlet right ventricle; ECG, electrocardiogram; HLHS, hypoplastic left heart syndrome; LV, left ventricle; LVEF, left ventricular ejection fraction; PA, pulmonary atresia; RV, right ventricle; TAPVR, totally anomalous pulmonary venous return; TGA, transposition of the great arteries; ToF, tetralogy of Fallot; VSD, ventricular septal defect.

### Keywords

Congenital heart disease • Mortality • Risk stratification • Electrocardiogram • Artificial intelligence

## Introduction

With recent medical and surgical advancements, most children born with congenital heart diseases (CHDs) now survive into adulthood.<sup>1</sup> In the USA alone, it is estimated that approximately 2.4 million

individuals are living with CHD, including approximately 1 million children and 1.4 million adults.<sup>2</sup> These conditions require lifelong follow-up involving frequent clinic visits and cardiac testing with significantly higher associated costs throughout the lifespan.<sup>3-5</sup> The necessity to

facilitate the management of this growing complex population with a wide range of structural anomalies and distinct long-term consequences has motivated efforts to risk stratify CHD patients across the lifespan.<sup>6</sup> Doing so may reduce a significant portion of unnecessary and expensive testing, while simultaneously identifying individuals that might benefit from closer monitoring or intervention thereby improving clinical outcomes.

However, the development of robust risk prediction models—particularly in CHD—faces multiple challenges including a paucity of big data, limited data sources, reliance on expensive imaging modalities (with an associated need for paediatric cardiologist expertise), and sub-optimal model performance.<sup>6</sup> This premise motivated multiple recent calls for artificial intelligence applications to improve risk stratification in CHD.<sup>6,7</sup>

Major adverse outcomes in this unique population often relate to ventricular dysfunction or arrhythmias, and extraction of conventional electrocardiogram (ECG) features such as QRS duration in CHD [e.g. tetralogy of Fallot (ToF)] has shown value in risk stratification.<sup>8</sup> In addition, deep learning-based artificial intelligence-enhanced ECG (AI-ECG) algorithms show promise for diagnostic and prognostic applications in adults,<sup>9–15</sup> making it similarly conceivable that AI-ECG may aid risk stratification also in the CHD population. However, there remains a paucity of available AI-ECG applications to congenital cardiology<sup>16</sup> given (i) the absence of extensive data sets, impeding similar research; (ii) AI-ECG algorithms derived from adults with structurally normal hearts would be expected to have poor generalizability to CHD cohorts. To this end, there remain (to our knowledge) no AI-ECG algorithms to predict mortality in CHD.

In this work, we leverage an exhaustive electronic database at a large congenital heart centre to address this gap. To do so, a convolutional neural network was trained to predict 5-year mortality using >100 000 ECGs on nearly 40 000 patients and tested on an equally sized cohort as well as on a contemporary cohort. Subgroup analysis defined model performance within a range of specific CHDs. Survival analysis investigated longer-term survival for patients. Finally, saliency mapping and median waveform analysis provided lesion-specific model explainability.

## Methods

Our study adheres to the TRIPOD + AI guidelines.<sup>17</sup>

### Study population and patient assignment

Patient data were utilized from the Boston Children's Hospital between 1990 (the earliest available digitized ECG in the clinic) until June 2018 (to minimize right censoring in predicting 5-year mortality). Inclusion criteria consisted of any patient presenting to the cardiology clinic at the Boston Children's Hospital with at least one ECG performed. Temporal validation for predicting 1-year mortality was performed on any patient presenting to the cardiology clinic at the Boston Children's Hospital with at least one ECG performed from July 2018 to July 2022 (to minimize right censoring in predicting 1-year mortality). All data were retrieved from an internal database in November 2023.

Similar to prior work,<sup>13,16</sup> a group stratified design was used to partition the main cohort at the patient level such that all ECGs for a given patient were restricted to either the training or internal testing cohort. Patients were randomly partitioned 50:50 into training and internal testing cohorts. Patients from the training data set were excluded from the temporal validation cohort.

### Data retrieval

Raw ECG signals were exported from the MUSE ECG data management system (GE Healthcare, Chicago, IL, USA), which contains 1D vectors of data (sampling rate of 250 Hz for 10-s duration) for each lead (I, II, and V1–6). From these vectors, Einthoven's law<sup>18</sup> and the Goldberger equation<sup>19</sup> were implemented to obtain leads III, aVF, aVL, and aVR. Other data retrieved from an internal database at the Boston Children's Hospital include age, sex, and physician-reviewed ECGs measurements (e.g. QRS interval, QRS axis, T axis, P axis, PR interval, QT interval, QTc, and heart rate).

Patients with known CHD, cardiomyopathy, pre-excitation syndromes, and channelopathies were identified based on the institutional Fyler coding system,<sup>20</sup> which has been mapped into the International Paediatric and Congenital Cardiac Code International Classification of Diseases-11 nomenclatures.<sup>21</sup>

### Quality control and data pre-processing

An ECG was discarded if any lead was not 2500 samples long or if any lead recording had missing lead information. After quality control, a high pass filter was utilized<sup>22</sup> to account for recording errors (e.g. baseline wander and electrical interference) with cut-off frequency 0.8 Hz, rejection band 0.2 Hz, ripple in passband 0.5 dB, and attenuation in rejection band 40 dB. Finally, the ECG was then trimmed to 2048 samples (~8 s) to facilitate conveniently working with convolutional neural networks.

### Definition of outcomes

The primary outcome was 5-year mortality after an ECG. All-cause mortality was used as the primary outcome given (i) heart failure is the most common cause of death in adult CHD<sup>23</sup> and (ii) it aligns with previous AI-ECG work<sup>14,15</sup> and recent CHD work.<sup>24</sup> To obtain all-cause mortality, date of death was retrieved from an internal institutional database. Secondary outcomes include 1-, 2-, 3-, and 4-year mortality after an ECG. For temporal validation, 1-year mortality was used to minimize right censoring.

To perform a secondary Kaplan–Meier survival analysis for time to mortality, we obtained the date last known alive within the internal database (i.e. last recorded event in the institutional database for a given patient).

### Model selection, architecture, and training

Similar to our prior work,<sup>16</sup> the AI-ECG model was developed solely on the training set, which was further partitioned 95% for training and 5% for validation to perform hyperparameter tuning. 12 × 2048 ECG samples were used as inputs to a convolutional neural network that is similar to the residual network previously described<sup>25</sup> that is adapted for unidimensional signals. A diagram and details of the architecture used in this study are shown in Mayourian *et al.*<sup>16</sup> The output of the last block is fed into a fully connected layer with a sigmoid activation function as the outcomes (1-, 2-, 3-, 4-, and 5-year mortality) within this single model are not mutually exclusive.

The final hyperparameters were obtained via a grid search on the training set among the following options: kernel size [3, 9, 17], batch size [8, 32, 64], and initial learning rate [0.01, 0.001, 0.0001, 0.00001]. The average cross-entropy was minimized using the Adam optimizer. Maximum 150 epochs were used with early stopping based on validation loss. The model with the lowest validation loss during hyperparameter tuning was selected as the final model. The final hyperparameters were kernel size 17, batch size 32, and learning rate 0.001.

### Performance evaluation and statistical analyses

Given the potential for loss to follow-up, model performance for all binary outcomes (e.g. 1- and 5-year mortality) was assessed only on ECGs with documented follow-up after the outcome timeframe or mortality events within the outcome timeframe. Consistent with prior works,<sup>16,26–28</sup>

multiple ECGs per patient were allowed in the training cohort. In contrast, model performance was evaluated on one ECG per patient. The ECG selected for testing was either the first available, the last available, or a randomly selected ECG per patient.

Given the imbalanced data set (i.e. low prevalence of mortality), the area under the receiver operating characteristic curve (AUROC) and area under the precision-recall [i.e. positive predictive value (PPV) sensitivity] curve (AUPRC) were computed. To benchmark the model, age at ECG (previously used to benchmark AI-ECG predictions of mortality<sup>14</sup>), QRS duration (a conventional ECG predictor of mortality in ToF<sup>29</sup>), and QTc duration (an independent risk factor for sudden cardiac death<sup>30</sup>) were used. In addition, left ventricular ejection fraction (LVEF)—an established determinant of cardiovascular morbidity and mortality<sup>31,32</sup>—from an echo within 2 days of the paired ECG was used as a benchmark when available. Other performance metrics evaluated included PPV, negative predictive value (NPV), sensitivity, and specificity. These metrics were calculated at the classification threshold maximizing the sensitivity and specificity (i.e. the Youden index) in the training set. For all metrics, a higher value is indicative of better performance. Resampling with 1000 bootstraps was implemented to obtain performance metric median and 95% confidence intervals (CIs). Area under the receiver operating characteristic curves were compared when applicable via the DeLong test.<sup>33</sup>

## Subgroup analyses

Subgroup analyses were performed on the test set for each disease of interest. Given that conduction disturbances are also a marker of disease progression in CHD,<sup>8</sup> we also assessed model performance when stratifying by QRS duration and presence of complete right bundle branch block. Area under the receiver operating characteristic curves and AUPRCs were calculated for each subgroup.

## Survival analysis

Cox proportional hazards regression was used to evaluate factors associated with time from ECG until all-cause mortality. Patients who did not experience death were censored at the time of last known follow-up. Patients with unknown follow-up time after ECG were excluded from this secondary analysis. The Kaplan–Meier methodology was used to estimate primary outcome event rates.

## Model explainability

In an effort to interpret model behaviour, analyses of median waveforms and saliency mapping were performed.

Median waveform analysis is a technique to visually represent single beats of the highest and lowest risk ECGs.<sup>16,27</sup> Herein, a subset of test set ECGs (100 for the overall cohort, 25 for specific lesions) with the highest predicted probability for a given outcome were used to create high-risk median waveforms, and a corresponding set with the lowest predicted probabilities was used to create low-risk median waveforms. Median waveforms were generated in each lead using the NeuroKit Python toolbox<sup>34</sup> by (i) detecting QRS complexes, (ii) interpolating all ECGs to the same heart rate, (iii) computing the median voltage across beats for each patient, and (iv) computing the median voltage across patients for each time bin in the cardiac cycle.<sup>16,27</sup>

Saliency mapping aims to identify which features of the ECG input contribute to model prediction by highlighting components of the ECG where a change in input (i.e. ECG voltage) leads to a relatively large change in prediction.<sup>27</sup> Saliency maps were created using a Shapley Additive Explanations (SHAP) framework<sup>35</sup> for the high-risk ECGs. The above median waveform steps were similarly implemented on SHAP values over time. Darker regions in saliency maps correspond to greater contribution to the prediction of 5-year mortality.

## Software

Programming codes used to perform the analyses are available upon reasonable request. The convolutional neural network used the Keras

framework with a TensorFlow (Google) backend using Python 3.9.<sup>36</sup> Deep learning was executed on institutional graphics processing units. All other pre- and post-processing codes were written in Python 3.9<sup>36</sup> and R 4.0,<sup>37</sup> which was executed locally.

# Results

## Internal patient population characteristics

The training cohort comprised of 112 804 ECGs from 39 784 patients [median age at ECG 7.7 [interquartile range (IQR), 1.6–14.8; range 0–85] years; 52% male]. The first digitized ECG included in this cohort was from 1990. As shown in [Table 1](#), a wide range of CHD lesions were included, including 11% with ventricular septal defects, 4.5% with cardiomyopathy, 3.8% with coarctation of the aorta, 3.5% with ToF, 2.2% with D-loop transposition of the great arteries, 1.2% with hypoplastic left heart syndrome, 1.0% with L-loop transposition of the great arteries, 0.8% with dextrocardia, and 0.7% with tricuspid atresia. The internal testing cohort included 112 575 ECGs from 39 784 distinct patients [median age at ECG 7.9 (IQR, 1.5–14.8; range 0–92) years; 52% male] and composed of a similar breakdown of congenital heart lesions. The contemporary cohort included 42 927 ECGs from 25 537 patients. Numerous differences in baseline characteristics were noted in the contemporary cohort including higher prevalence of all disease groups (except ventricular septal defects) and older age at ECG [median age at ECG 10.5 (IQR, 3.1–16.7) years].

In the training cohort, there were 806 (2.0%) mortality events at median age 18.9 (IQR, 8.0–32.4) years. In the testing cohort, there were 870 (2.2%) mortality events at median age 19.4 (IQR, 7.5–32.2) years. Five-year mortality after ECGs in the training and test cohorts was 4.9% and 4.6%, respectively. Similar 1-year mortality events (1.0%) were noted across internal test and temporal validation cohorts, with a younger age of mortality in the temporal validation cohort [median age 13.1 (IQR, 4.7–22.6) years]. Electrocardiograms with mortality within 5 years had higher heart rates, longer QRS and QTc intervals, and lower paired echo ejection fraction (see [Supplementary data online, Table S1](#)).

Survival in the training and testing cohorts was similar [hazard ratio 1.0 (95% CI 0.9–1.0),  $P = .3$ ; log-rank  $P = .2$ ] and spanned across the lifespan ([Figure 1A](#)). As shown in [Figure 1B and C](#), a large portion of patients survive to adulthood, with survival rates varying by lesion. The lowest long-term survival was lesions that typically involve single ventricle palliation (tricuspid atresia for RV pathology, hypoplastic left heart syndrome for left ventricular pathology).

## Model performance

During testing ([Figure 2A](#)), the model achieved the following performance in 5-year mortality when using the random, last, and first ECGs: AUROCs of 0.79 (95% CI 0.77–0.81), 0.82 (95% CI 0.80–0.83), and 0.75 (95% CI 0.72–0.77), respectively, and AUPRCs of 0.17 (95% CI 0.15–0.19), 0.25 (95% CI 0.23–0.28), and 0.11 (95% CI 0.09–0.12), respectively. The randomly selected ECG performance outperformed age at ECG [AUROC 0.58 (95% CI 0.55–0.61); AUPRC 0.11 (95% CI 0.08–0.13);  $P < .001$ ], QRS duration [AUROC 0.57 (95% CI 0.54–0.60); AUPRC 0.07 (95% CI 0.06–0.08);  $P < .001$ ], QTc duration [AUROC 0.48 (95% CI 0.45–0.50); AUPRC 0.05 (95% CI 0.04–0.06);  $P < .001$ ], and paired echo LVEF [AUROC 0.62 (95% CI 0.57–0.68) [AUPRC 0.10 (95% CI 0.07–0.15);  $P < .001$ ] when predicting 5-year mortality ([Figure 2B](#)), as well as 1-year mortality (see [Supplementary](#)

**Table 1** Internal cohort baseline characteristics

Characteristic	Training	Testing	Temporal validation
<b>Demographics</b>			
Patients	39 784	39 784	25 537
Sex			
Female	19 061 (48%)	19 009 (48%)	12 496 (49%)
Male	20 718 (52%)	20 769 (52%)	13 036 (51%)
Unknown	5 (<0.1%)	6 (<0.1%)	5 (<0.1%)
Diagnosis <sup>a</sup>			
ToF	1374 (3.5%)	1352 (3.4%)	1029 (4.0%)
Cardiomyopathy	1810 (4.5%)	1773 (4.5%)	1273 (5.0%)
ASD	1498 (3.8%)	1551 (3.9%)	2638 (10%)
CAVC	248 (0.6%)	205 (0.5%)	339 (1.3%)
CoA	1500 (3.8%)	1470 (3.7%)	1298 (5.1%)
DORV	526 (1.3%)	530 (1.3%)	515 (2.0%)
D-loop TGA	884 (2.2%)	914 (2.3%)	729 (2.9%)
HLHS	487 (1.2%)	465 (1.2%)	553 (2.2%)
L-loop TGA	380 (1.0%)	389 (1.0%)	343 (1.3%)
PA	621 (1.6%)	627 (1.6%)	509 (2.0%)
TAPVR	240 (0.6%)	257 (0.6%)	262 (1.0%)
Tricuspid Atresia	271 (0.7%)	247 (0.6%)	195 (0.8%)
VSD	4388 (11%)	4196 (11%)	2674 (10%)
Dextrocardia	304 (0.8%)	326 (0.8%)	290 (1.1%)
WPW syndrome	1070 (2.7%)	1003 (2.5%)	678 (2.7%)
Channelopathy	223 (0.6%)	231 (0.6%)	283 (1.1%)
<b>ECG characteristics</b>			
ECGs	112 804	112 575	42 927
Age at ECG (years)	7.7 (1.6, 14.8)	7.9 (1.5, 14.8)	10.5 (3.1, 16.7)
CRBBB	5749 (5.1%)	5873 (5.2%)	2572 (6.0%)
<b>Outcomes</b>			
Mortalities	806 (2.0%)	870 (2.2%)	225 (0.9%)
Age of death (years)	18.9 (8.0, 32.4)	19.4 (7.5, 32.2)	13.1 (4.7, 22.6)
1-year mortality after ECG <sup>b</sup>	909 (1.2%)	781 (1.0%)	283 (1.0%)
5-year mortality after ECG <sup>b</sup>	2964 (4.9%)	2779 (4.6%)	

Age intervals denote interquartile range.

ASD, atrial septal defect; CoA, coarctation of the aorta; CAVC, complete atrioventricular canal defect; CRBBB, complete right bundle branch block; DORV, double outlet right ventricle; HLHS, hypoplastic left heart syndrome; LVEF, left ventricular ejection fraction; PA, pulmonary atresia; ToF, tetralogy of Fallot; TAPVR, total anomalous pulmonary venous return; TGA, transposition of the great arteries; VSD, ventricular septal defect; WPW, Wolff-Parkinson-White.

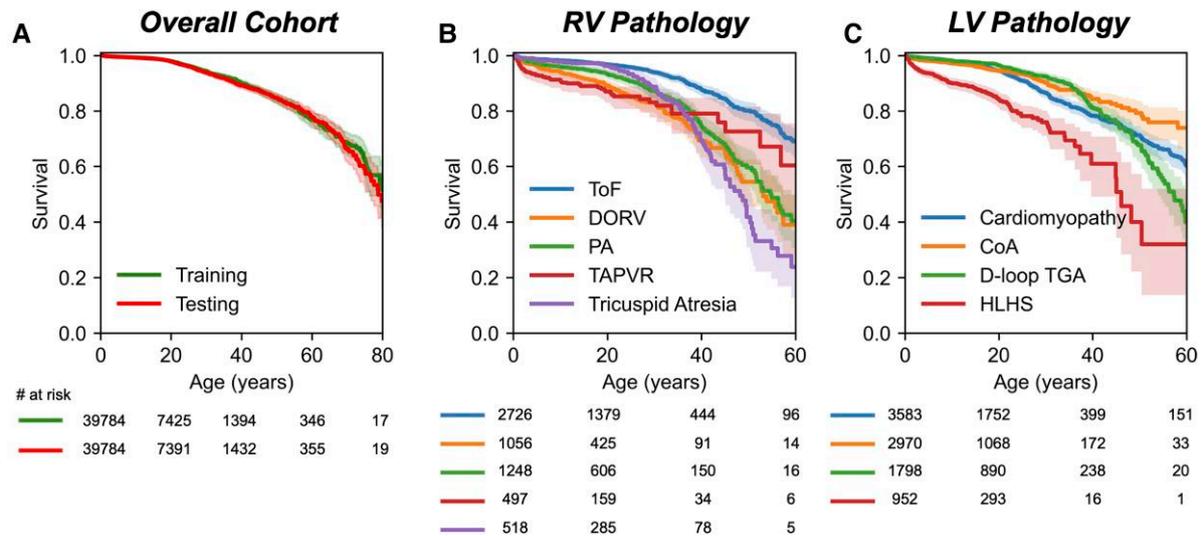
<sup>a</sup>Note select lesions included, some of which overlap.

<sup>b</sup>Denominator is equal to the total number of ECGs with either mortality events within the outcome timeframe or with documented follow-up after the outcome timeframe.

data online, Figure S1). Artificial intelligence-enhanced ECG similarly outperformed these metrics when using the last available ECG (see Supplementary data online, Figure S2).

When using a random ECG, the sensitivity, specificity, and PPV were 0.66 (95% CI 0.62–0.70), 0.78 (95% CI 0.78–0.79), and 12.4% (95% CI

11.6%–13.1%), respectively, with 76.4% (95% CI 75.6%–77.1%) predicted negative (see Supplementary data online, Table S2). A NPV of ~98% was achieved independent of ECG used (see Supplementary data online, Table S2). Model calibration is shown in Supplementary data online, Figure S3.



**Figure 1** Kaplan–Meier survival analysis of training and testing cohorts. (A) Kaplan–Meier curve survival analysis of training (green) and testing (red) cohorts demonstrates the large, diverse cohorts across the congenital heart disease lifespan. Lesion-specific survival curves for (B) left ventricular and (C) right ventricular pathology are shown when pooling training and testing cohorts. Number at risk within each group inset below. CoA, coarctation of the aorta; DORV, double outlet right ventricle; HLHS, hypoplastic left heart syndrome; LV, left ventricular; PA, pulmonary atresia; RV, right ventricular; TAPVR, total anomalous pulmonary venous return; TGA, transposition of the great arteries; ToF, tetralogy of Fallot

## Contemporary model performance

During temporal validation, AUROC of 0.79 (95% CI 0.74–0.83) and AUPRC of 0.04 (95% CI 0.03–0.06) was achieved to predict 1-year mortality (Figure 2C).

## Subgroup analysis

In a subgroup analysis (Figure 3), model performance appeared lesion and age dependent. When using a random ECG per patient, model performance in extra-cardiac pathophysiology (i.e. coarctation of the aorta and total anomalous pulmonary venous return) was as follows: AUROCs of 0.86 (95% CI 0.81–0.91) and 0.80 (95% CI 0.72–0.88), respectively, and AUPRCs of 0.30 (95% CI 0.19–0.41) and 0.38 (95% CI 0.20–0.56), respectively. For a wide range of congenital heart lesions with accumulating pathophysiologic myocardial burden over time (e.g. cardiomyopathy, pulmonary atresia, double outlet right ventricle, hypoplastic left heart syndrome, and tricuspid atresia), lower AUROC with higher AUPRC patterns were noted. These trends were consistent when using the last available ECG, and less so when using the first available ECG. Similarly, the dextrocardiac subgroup had a lower AUROC [0.71 (95% CI 0.61–0.81)] with higher AUPRC [0.21 (95% CI 0.10–0.32)]. Model performance was lower in L-loop transposition of the great arteries (Figure 3). Across all lesions with sufficient data available for comparison, AUROC and AUPRC trended higher for AI-ECG compared with LVEF (see Supplementary data online, Figure S4).

Given that conduction disturbances are also a marker of disease progression in CHD,<sup>8</sup> we also assessed model performance when stratifying by QRS duration and presence of complete right bundle branch block. In patients with QRS duration  $\leq 120$  ms, AUROC of 0.77 (95% CI 0.75–0.80) and AUPRC of 0.14 (95% CI 0.12–0.16) were achieved. For QRS duration  $> 120$  ms, AUROC of 0.76 (95% CI 0.72–0.80) and AUPRC of 0.20 (95% CI 0.15–0.24) were obtained. In patients with complete right bundle branch block, AUROC of 0.78 (95% CI 0.73–0.83) and AUPRC of 0.21 (95% CI 0.13–0.29) were achieved. In patients without

complete right bundle branch block, AUROC of 0.79 (95% CI 0.77–0.81) and AUPRC of 0.18 (95% CI 0.15–0.21) were obtained.

## Survival analysis

Longer-term survival was assessed when stratifying patients into low- (<threshold) or high-risk ( $\geq$ threshold) groups based on AI-ECG predictions (Figure 4). When using a random ECG, there was 15-year survival of 96% and 80% for low- vs. high-risk ECGs, respectively. When using a random ECG, high-risk patients were 4.9 times (95% CI 4.3–5.6) more likely to experience mortality ( $P < .001$ ). This phenomenon was more pronounced when using the last ECG, with 15-year survival of 95% and 72% in low- and high-risk groups, respectively, and a hazard ratio of 7.2 (95% CI 6.3–8.3) ( $P < .001$ ).

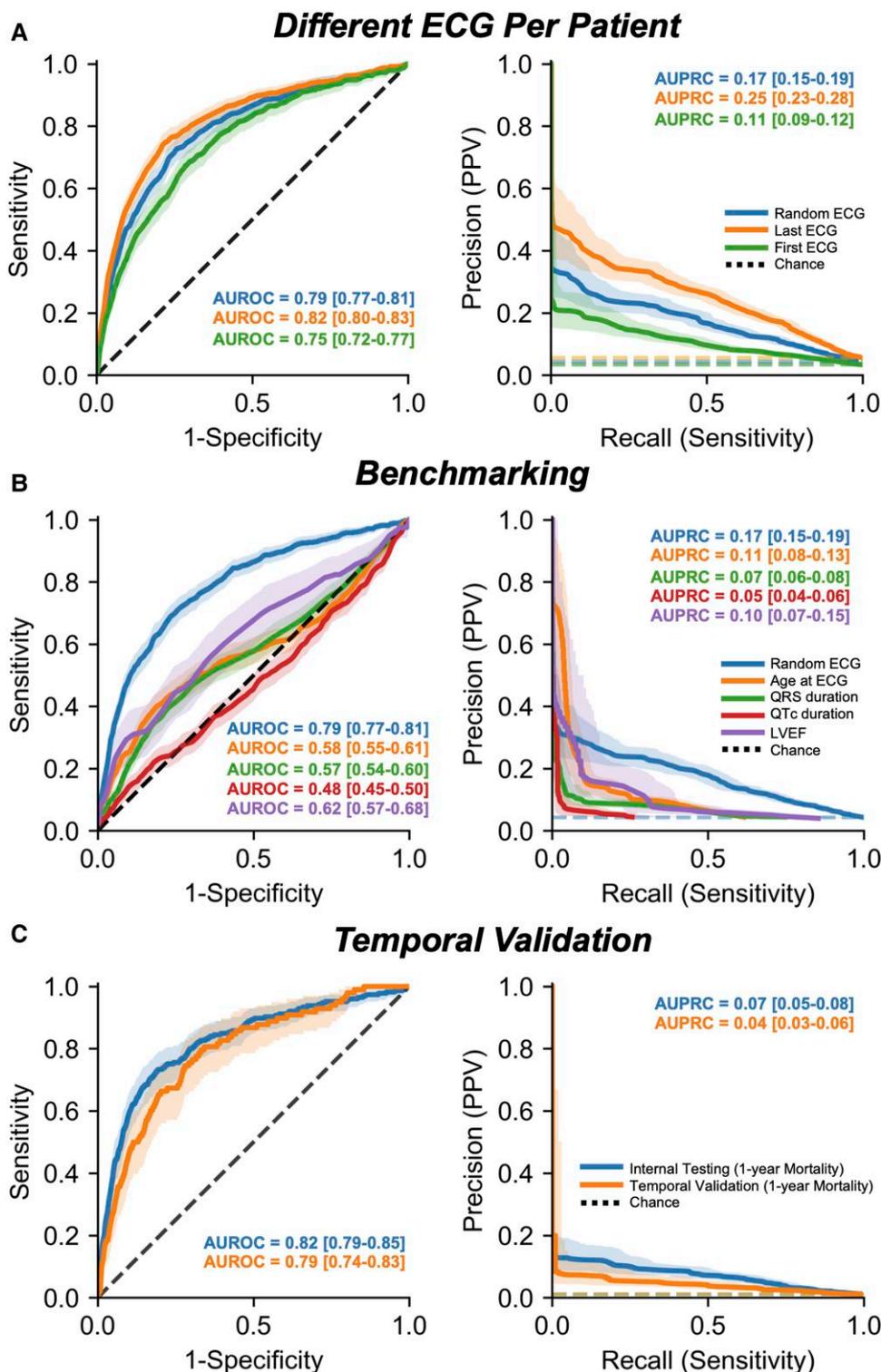
In the internal test cohort, the AI-ECG model predictions [c-index 0.74 (95% CI 0.72–0.76)] outperformed QRS duration [c-index 0.52 (95% CI 0.50–0.54)] and LVEF [c-index 0.64 (95% CI 0.60–0.68)] in Cox model survival discrimination (Table 2). The addition of age and/or LVEF as predictors to AI-ECG provided no added value in the c-index (Table 2).

Lesion-specific survival analyses demonstrated effective risk stratification in nearly all lesions (Figure 5), including dextrocardia [hazard ratio 2.4 (95% CI 1.1–5.3);  $P = .02$ ; Supplementary data online, Figure S5], but not complete atrioventricular canal defects [hazard ratio 3.5 (95% CI 0.8–15.8);  $P = .10$ ; Supplementary data online, Figure S6].

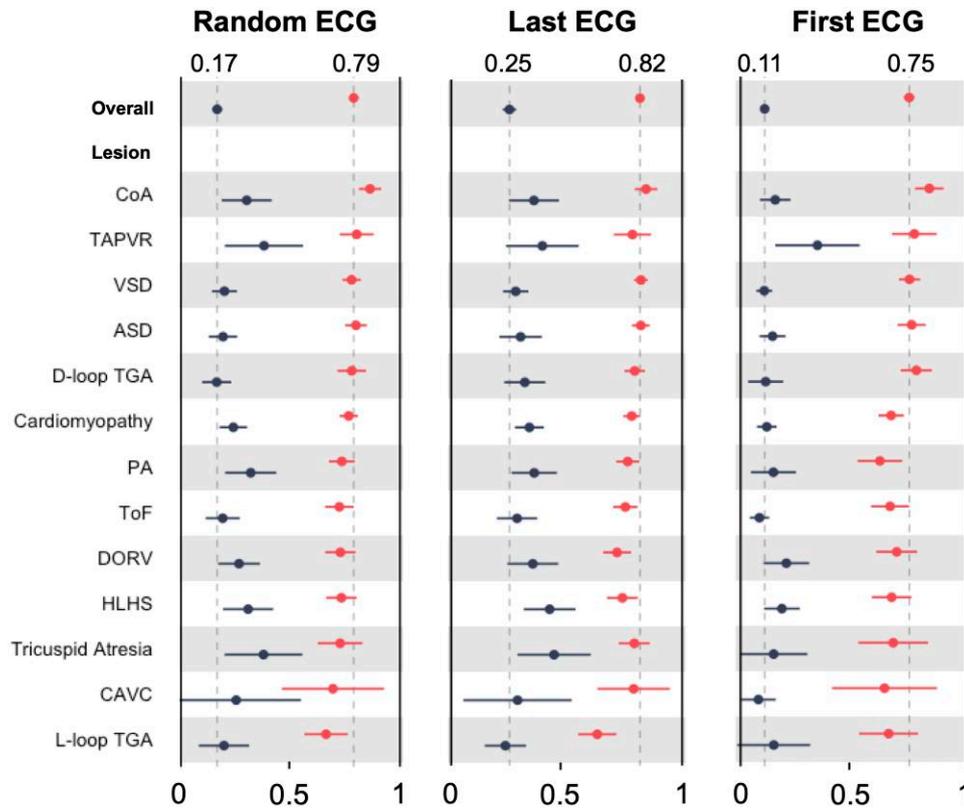
## Model explainability

The most salient features (Figure 6) of an ECG to predict 5-year mortality in the overall cohort include S waves (limb lead III and precordial leads V2, V3, and V6) and T waves (limb leads II and V6). High-risk features to predict 5-year mortality include wide QRS complexes with deep S waves and low-amplitude waveforms (Figure 6).

When performing saliency mapping and median waveform analysis within subgroups, lesion-specific high-risk signatures were identified



**Figure 2** Electrocardiogram-based deep learning model performance. (A) Artificial intelligence-enhanced electrocardiogram model performance to predict 5-year mortality evaluated using a random (blue), last (orange), and first (green) electrocardiogram per patient. (B) Performance benchmarking of the random electrocardiogram (blue) to age at electrocardiogram (orange), QRS duration (green), QTc duration (red), and left ventricular ejection fraction (purple). (C) Comparison of internal testing (blue) and temporal validation (orange) performance to predict 1-year mortality. Area under the receiver operating characteristic curve and area under the precision-recall curve metric values for each model and outcome are inset. Dotted line represents chance. 95% confidence intervals are shown using bootstrapping. AUROC, area under the receiver operating characteristic curve; AUPRC, area under the precision-recall curve; ECG, electrocardiogram; LVEF, left ventricular ejection fraction; PPV, positive predictive value



**Figure 3** Model performance in congenital heart disease subgroups. Forest plot showing artificial intelligence-enhanced electrocardiogram area under the area under the receiver operating characteristic curve (red) and area under the precision-recall curve (black) performance when stratifying by lesion when using a random (left), last (middle), and first (right) electrocardiogram. Area under the receiver operating characteristic curve and area under the precision-recall curve metric values for each model and outcome are inset. 95% confidence intervals are shown using bootstrapping. ASD, atrial septal defect; CoA, coarctation of the aorta; CAVC, complete atrioventricular canal defect; DORV, double outlet right ventricle; ECG, electrocardiogram; HLHS, hypoplastic left heart syndrome; PA, pulmonary atresia; ToF, tetralogy of Fallot; TAPVC, total anomalous pulmonary venous connection; TGA, transposition of the great arteries; VSD, ventricular septal defect

(Figure 6). In cardiomyopathy, V2 and V3 were less salient than the overall cohort; however, high-risk features appeared similar to the overall cohort. In right-sided myopathy such as ToF, V6 was less salient than the overall cohort; in addition, while low-risk features included a right bundle branch block, high-risk features included QRS fragmentation. In left-sided pathology such as hypoplastic left heart syndrome, saliency maps were more focused on the QRS complex compared with the overall cohort. High-risk features included tall and wide R waves in V1–3 with a deep S wave in II–III and lateral precordial leads.

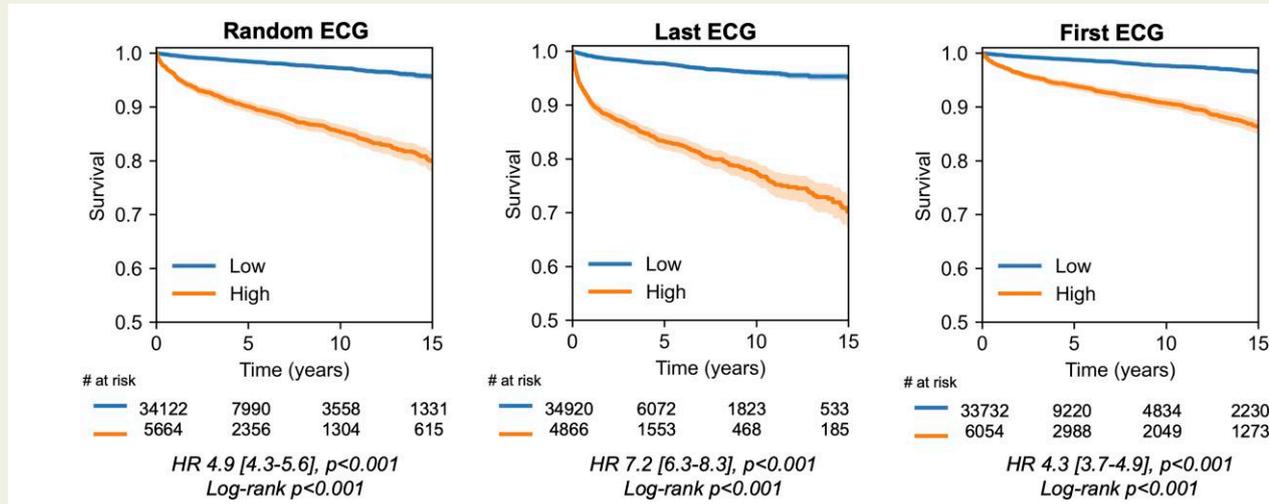
## Discussion

For several decades, risk stratification has been of great interest in the CHD field using conventional and artificial intelligence approaches.<sup>6,7,24</sup> The ongoing challenge in developing robust risk prediction models in CHD has led to multiple recent calls for AI applications to improve risk stratification in CHD.<sup>6,7</sup> In this work, we address this gap by developing and validating the first (to our knowledge) ECG-based deep learning algorithm to predict mortality in children and adults with CHD. We demonstrate that AI-ECG can successfully predict 5-year mortality, outperforming conventional markers such as age, LVEF, and QRS duration. The encouraging performance when using the first available, the last available, or a random ECG per patient demonstrates

the promise of our model to predict mortality during initial and follow-up assessments to aid in lifelong risk stratification. Model explainability analysis provides transparency and interpretability for clinicians and may help generate hypotheses for underlying myopathy ECG signatures predictive of mortality. Altogether, our findings demonstrate the prognostic value of AI-ECG in CHD across the lifespan, which may (i) enhance current risk stratification strategies, (ii) prioritize patients for diagnostic studies and/or interventions, and (iii) facilitate improved access to care (*Structured Graphical Abstract*).

## Conventional electrocardiogram predictors of mortality in congenital heart disease

There is a paucity of conventional ECG analyses to predict morbidity and mortality in patients with CHD. The majority of applications are in ToF: several studies have recognized severe QRS prolongation as a risk factor for mortality in this patient population.<sup>38</sup> However, the sensitivity of QRS duration > 180 ms to predict mortality was <50% in recent studies.<sup>39</sup> Subsequent work by Bokma et al.<sup>29</sup> demonstrated QRS fragmentation—related to myocardial fibrosis and dysfunction—is superior to QRS duration in predicting mortality. QRS fragmentation has also been associated with mortality in cardiomyopathy.<sup>40</sup> In addition, a



**Figure 4** Kaplan–Meier survival analysis based on artificial intelligence-enhanced electrocardiogram risk stratification. Kaplan–Meier curve survival analysis when stratifying patients as low- (blue) or high-risk (orange) based on artificial intelligence-enhanced electrocardiogram predictions using random (left), last (middle), or first (right) electrocardiogram per patient. Number at risk within each group inset below. Hazard ratio inset below (high- vs. low-risk group based on artificial intelligence-enhanced electrocardiogram predictions) with 95% confidence interval using Cox regression analysis. P-value statistic below based on log-rank testing. ECG, electrocardiogram; HR, hazard ratio

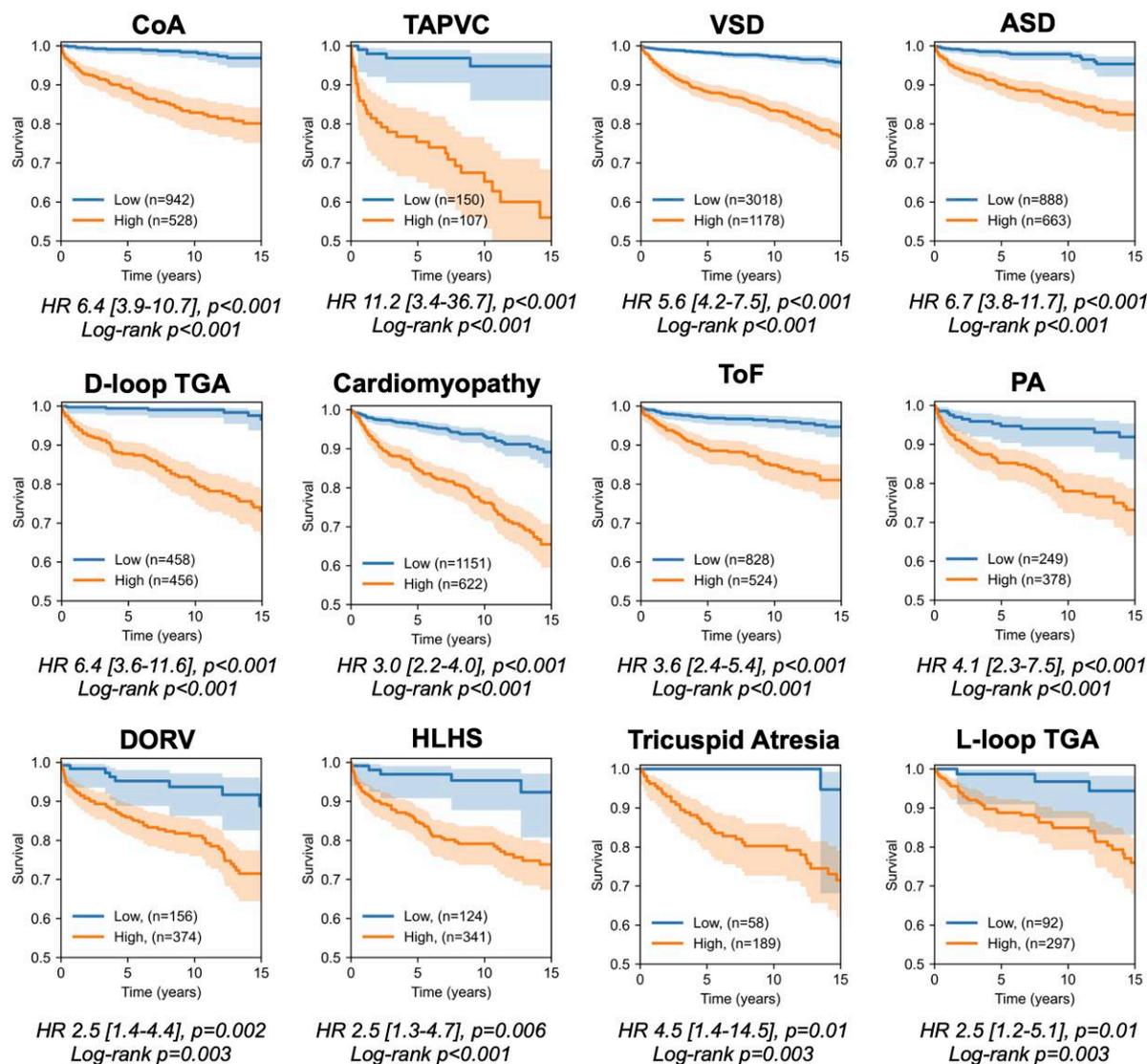
**Table 2** Cox models on the internal test cohort incorporating artificial intelligence-enhanced electrocardiogram predictions

Cox model	Variable z-score	P-value	c-index (95% CI)
Age only	8.4	<2e-16	0.58 (0.56–0.60)
QRS duration only	5.8	5e-9	0.52 (0.50–0.54)
LVEF only	-11.5	<2e-16	0.64 (0.60–0.68)
AI-ECG only	15.2	<2e-16	0.74 (0.72–0.76)
Age ± LVEF			0.65 (0.61–0.69)
Age	7.0	3e-12	
LVEF	-10.3	<2e-16	
Age ± AI-ECG			0.71 (0.69–0.73)
Age	6.7	2e-11	
AI-ECG	14.0	<2e-16	
LVEF ± AI-ECG			0.72 (0.70–0.74)
LVEF	-6.4	2e-10	
AI-ECG	10.9	<2e-16	
Age ± LVEF ± AI-ECG			0.72 (0.70–0.74)
Age	6.0	2e-9	
LVEF	-5.7	1e-8	
AI-ECG	10.2	<2e-16	

AI-ECG, artificial intelligence-enhanced electrocardiogram; CI, confidence interval; LVEF, left ventricular ejection fraction.

longer QRS has been previously predictive of mortality in hypoplastic left heart syndrome,<sup>41</sup> D-loop transposition of the great arteries,<sup>42</sup> and Fontan circulation,<sup>43</sup> all with limited performance.

This has motivated more recent efforts to incorporate imaging modality data into risk prediction algorithms.<sup>44</sup> However, even such algorithms continue to have suboptimal performance and are reliant on expensive



**Figure 5** Lesion-specific Kaplan–Meier survival analysis based on artificial intelligence-enhanced electrocardiogram risk stratification. Lesion-specific Kaplan–Meier curve survival analysis when stratifying patients as low- (blue) or high-risk (orange) based on artificial intelligence-enhanced electrocardiogram predictions. Hazard ratio inset below (high- vs. low-risk group based on artificial intelligence-enhanced electrocardiogram predictions) with 95% confidence interval using Cox regression analysis. P-value statistic below based on log-rank testing. Initial sample size in each cohort inset. ASD, atrial septal defect; CoA, coarctation of the aorta; DORV, double outlet right ventricle; HR, hazard ratio; HLHS, hypoplastic left heart syndrome; PA, pulmonary atresia; ToF, tetralogy of Fallot; TAPVC, total anomalous pulmonary venous connection; TGA, transposition of the great arteries; VSD, ventricular septal defect

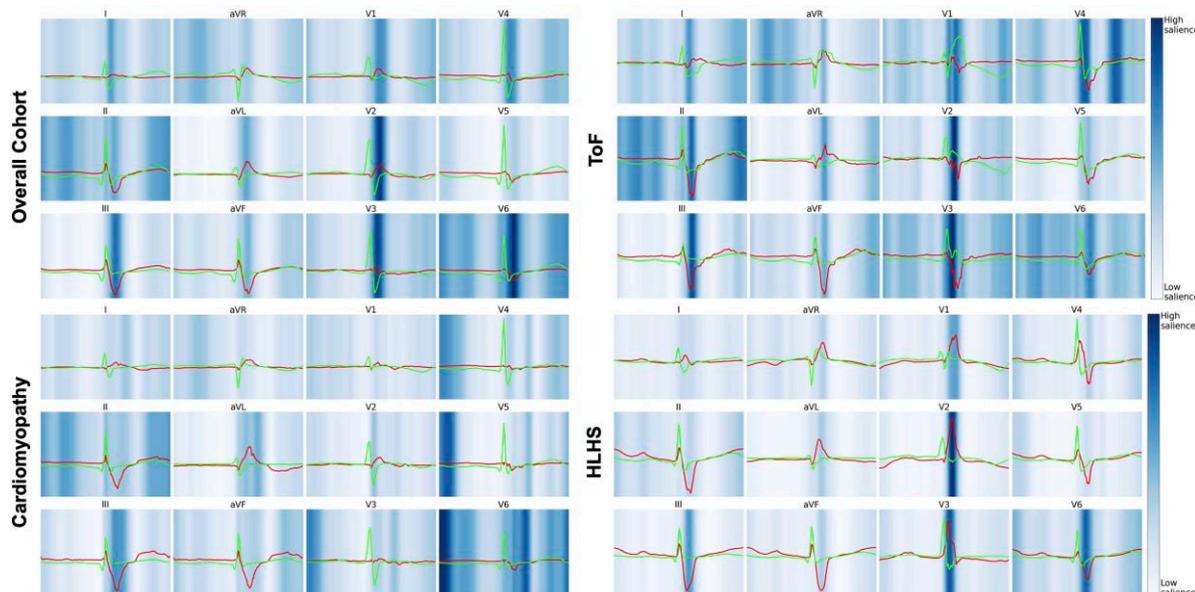
modalities (e.g. cardiac magnetic resonance imaging) that require subspecialized expertise. As shown herein, AI-ECG provides an inexpensive, ubiquitous alternative that is predictive across a wide range of lesions.

## Clinical significance and implications

Imaging modalities (e.g. cardiac magnetic resonance and echocardiography) conventionally used to aid in risk stratification in CHD have practice limitations (time-intensive, resource-consuming, and need for subspecialist expertise) that hinder its widespread use. In contrast, ECGs is rapid and cost-effective and can be conveniently acquired at every cardiology visit, which facilitates easier and more frequent use that may guide clinical

decision-making. We envision this AI-ECG algorithm may serve as a screening or surveillance tool and potentially improve access to care.

Given the objective to risk stratify patients from a preventative lens, we opted to train and test using ECGs from cardiology clinic. In this setting, AI-ECGs could be of significant screening and surveillance value. As shown in [Supplementary data online, Table S2](#), a NPV of 98% was achieved to predict 5-year mortality, with 15-year survival of 96% for low-risk ECGs. To this end, as a screening tool, low-risk AI-ECG predictions may help reduce follow-up frequency, the need for non-invasive imaging, diagnostic catheterizations, or implantable cardioverter-defibrillators. On the other hand, given the 5-year mortality PPV of 12%, it may help identify high-risk patients requiring closer monitoring.



**Figure 6** Explainability of artificial intelligence-enhanced electrocardiogram predictions. Visualization of median waveforms generated in each lead using electrocardiograms from the highest (red) and lowest (green) artificial intelligence-enhanced electrocardiogram predictions of the overall cohort, as well as cardiomyopathy, tetralogy of Fallot, and hypoplastic left heart syndrome subgroups. Saliency mapping demarcates regions of the electrocardiogram waveform having greatest (dark blue) and least (light blue) influence on each outcome. Saliency was averaged over the highest predicted electrocardiograms for each outcome. HLHS, hypoplastic left heart syndrome; ToF, tetralogy of Fallot

As a surveillance tool, a congenital cardiologist could conceivably monitor AI-ECG predictions at each cardiology visit. Monitoring AI-ECG predictions over time may provide insight into responsiveness to interventions (e.g. pulmonary valve replacement in ToF) or inter-stage monitoring of single ventricle patients.

Finally, for low-resource settings with limited access to advanced modalities, this algorithm may help improve access to care. Notably, despite the majority of CHD patients being adults, approximately half remain without regional CHD services.<sup>45</sup> This technology may therefore contribute to the democratization of specialty expertise and circumvent the requirement for specialized cardiac magnetic resonance knowledge to risk stratify certain lesions (e.g. ToF).

### Significance of temporal validation

From 1990 (the first ECG available in this cohort) through today, CHD has rapidly evolved. For example, surgical and medical advances have contributed to the nearly 40% decrease in CHD mortality from 1999 to 2017,<sup>46</sup> with >97% of children with CHD expected to reach adulthood.<sup>47</sup> From an institutional standpoint, the Boston Children's Hospital volume and complexity have significantly increased (Table 1). Given the range of confounding factors that could affect mortality, we opted to perform temporal validation with reassuringly similar performance.

### Artificial intelligence-enhanced electrocardiogram model insights gained

The overall CHD and cardiomyopathy cohorts had similar high-risk features, suggesting common final signatures on an ECG indicative of high-risk mortality. Indeed, both CHD and cardiomyopathy have non-trivial overlap, and both largely contribute to paediatric heart failure.<sup>48</sup> The common features noted herein include wide QRS complexes with

deep S waves and low-amplitude waveforms, which may correspond to heterogeneous slow activation of the myocardium due to scar and/or myocardial stress. In addition, select high-risk patterns were noted in each disease process. For example, QRS fragmentation was noted in high-risk ECGs for ToF (Figure 6). Future work is required to investigate the relation of these features with progressive cardiomyopathy and paediatric heart failure.

Finally, we note variation in model performance by disease subtype. Interestingly, effective risk stratification was achieved in a majority of diseases (including dextrocardia), but not complete atrioventricular canal defects. The reasons for poor performance in this group are not readily apparent but may involve the high heterogeneity of this group given its association with heterotaxy syndrome and situs abnormalities, trisomy 21, and inherently abnormal QRS axis (superior axis deviation) compared with the rest of the cohort.

### Limitations and future directions

There are several limitations of this work. First, although heart failure is the most common cause of death in adult CHD,<sup>23</sup> all-cause mortality rather than cardiac mortality was used for the primary outcome in this study. In addition, while all-cause mortality is routinely coded at our institution, it is conceivable that positive outcomes are undocumented. We attempted to mitigate this limitation in our binary outcome analysis by including only ECGs with mortality events within the outcome timeframe or documented follow-up after the outcome timeframe. Similarly, in our survival analysis, we censored at time of the last known follow-up. Nevertheless, future use of state or national death index database should be considered. Second, while several recent AI-ECG works also used all-cause mortality as the primary endpoint,<sup>14,15</sup> similar clinically meaningful outcomes such as heart transplant are also of

interest. Third, while temporal validation was achieved, it is of great interest to obtain external validation for each CHD lesion. Fourth, only one example of thresholding was used to evaluate model performance, as further consideration is required to weigh the impact of resultant false negatives (which may lead to clinical consequences of missed pathology) and false positives (which may lead to unnecessary extraneous testing), as well as optimally set thresholds across institutions. To this end, multicentre external validation to further refine thresholds for clinical implementation is warranted. Similarly, multicentre collaboration via federated learning<sup>49</sup> may help improve training/testing sample sizes, which may further improve performance. Given our objective to develop an inexpensive and convenient risk stratification tool, only ECG inputs were utilized; nevertheless, multimodal inputs may lead to improved model performance (especially for complex lesions) requiring further investigation.<sup>50</sup> The limitations of saliency mapping must be noted.<sup>51</sup> Lastly, diagnostic categories in this study are quite heterogeneous. For example, the cardiomyopathy category includes dilated, hypertrophic, and restrictive cardiomyopathy, among others. Similarly, multiple CHD diagnoses can occur simultaneously, such that patients can be assigned into multiple categories. In addition, the cohort includes patients with and without repairs for CHD lesions.

Future work therefore includes model refinement and external validation for each lesion of interest, multicentre collaboration, consideration of multimodal inputs, and prospective trials (to determine how to properly implement such tools to support clinical decision-making). Finally, we note that a recent randomized clinical trial implementing AI-ECG alerts led to decreased all-cause mortality in the general adult population.<sup>52</sup> It is similarly of great interest to implement a similar study design in the distinct CHD population.

## Conclusions

In conclusion, these findings demonstrate the promise of AI-ECG to inexpensively and conveniently risk stratify individuals with CHD across the lifespan. This tool may facilitate the prioritization of patients for future interventions/studies, provide meaningful insight into novel ECG waveforms suggestive of mortality, and potentially reduce disparities by improving access to care. Future multicentre collaboration and prospective trials are warranted.

## Acknowledgements

The authors would like to acknowledge Boston Children's Hospital's High-Performance Computing Resources Clusters Enkefalos 2 (E2) made available for conducting the research reported in this publication.

## Supplementary data

Supplementary data are available at *European Heart Journal* online.

## Declarations

## Disclosure of Interest

All authors declare no disclosure of interest for this contribution.

## Data Availability

Requests for the Boston Children's Hospital data and related materials will be internally reviewed to clarify if the request is subject to intellectual property or confidentiality constraints. Shareable data and materials will be released under a material transfer agreement for non-commercial research purposes. The use of the Boston Children's Hospital data was approved by its Institutional Review Boards.

## Funding

Funding support received from the Thrasher Research Fund Early Career Award (J.M.), Boston Children's Hospital Electrophysiology Research Education Fund (J.M. and J.K.T.) and National Institutes of Health grant R00-LM012926 from the National Library of Medicine (W.G.L.).

## Ethical Approval

Institutional Review Board or Ethics Committee approval was obtained.

## Pre-registered Clinical Trial Number

None supplied.

## References

- Baumgartner H, De Backer J, Babu-Narayan SV, Budts W, Chessa M, Diller GP, et al. 2020 ESC guidelines for the management of adult congenital heart disease. *Eur Heart J* 2021;**42**:563–645. <https://doi.org/10.1093/eurheartj/ehaa554>
- Gilboa SM, Devine OJ, Kucik JE, Oster ME, Riehle-Colarusso T, Nembhard WN, et al. Congenital heart defects in the United States: estimating the magnitude of the affected population in 2010. *Circulation* 2016;**134**:101–9. <https://doi.org/10.1161/CIRCULATIONAHA.115.019307>
- Arth AC, Tinker SC, Simeone RM, Ailes EC, Cragan JD, Grosse SD. Inpatient hospitalization costs associated with birth defects among persons of all ages—United States, 2013. *MMWR Morb Mortal Wkly Rep* 2017;**66**:41–6. <https://doi.org/10.15585/mmwr.mm6602a1>
- Nandi D, Rossano JW. Epidemiology and cost of heart failure in children. *Cardiol Young* 2015;**25**:1460–8. <https://doi.org/10.1017/S1047951115002280>
- Chowdhury D, Johnson JN, Baker-Smith CM, Jaquiss RDB, Mahendran AK, Curren V, et al. Health care policy and congenital heart disease: 2020 focus on our 2030 future. *J Am Heart Assoc* 2021;**10**:e020605. <https://doi.org/10.1161/JAHA.120.020605>
- Opotowsky AR, Khairy P, Diller G, Kasparian NA, Brophy J, Jenkins K, et al. Clinical risk assessment and prediction in congenital heart disease across the lifespan: JACC scientific statement. *J Am Coll Cardiol* 2024;**83**:2092–111. <https://doi.org/10.1016/j.jacc.2024.02.055>
- Bruce C, Gatzoulis MA, Brida M. Digital technology and artificial intelligence for improving congenital heart disease care: alea iacta est. *Eur Heart J* 2024;**45**:1386–9. <https://doi.org/10.1093/eurheartj/ehad898>
- Gatzoulis MA, Till JA, Somerville J, Redington AN. Mechano-electrical interaction in tetralogy of Fallot. QRS prolongation relates to right ventricular size and predicts malignant ventricular arrhythmias and sudden death. *Circulation* 1995;**92**:231–7. <https://doi.org/10.1161/01.CIR.92.2.231>
- 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;**25**:70–4. <https://doi.org/10.1038/s41591-018-0240-2>
- Attia ZI, Kapa S, Yao X, Lopez-Jimenez F, Mohan TL, Pelliikka PA, et al. Prospective validation of a deep learning electrocardiogram algorithm for the detection of left ventricular systolic dysfunction. *J Cardiovasc Electrophysiol* 2019;**30**:668–74. <https://doi.org/10.1111/jce.13889>
- Yao X, Rushlow DR, Inselman JW, McCoy RG, Thacher TD, Behnken EM, et al. Artificial intelligence-enabled electrocardiograms for identification of patients with low ejection fraction: a pragmatic, randomized clinical trial. *Nat Med* 2021;**27**:815–9. <https://doi.org/10.1038/s41591-021-01335-4>
- Adedinsawo D, Carter RE, Attia Z, Johnson P, Kashou AH, Dugan JL, et al. Artificial intelligence-enabled ECG algorithm to identify patients with left ventricular systolic dysfunction presenting to the emergency department with dyspnea. *Circ Arrhythm Electrophysiol* 2020;**13**:e008437. <https://doi.org/10.1161/CIRCEP.120.008437>
- Vaid A, Johnson KW, Badgeley MA, Somani SS, Bicak M, Landi I, et al. Using deep-learning algorithms to simultaneously identify right and left ventricular dysfunction

- from the electrocardiogram. *JACC Cardiovasc Imaging* 2022;**15**:395–410. <https://doi.org/10.1016/j.jcmg.2021.08.004>
14. Raghunath S, Cerna U, Jing AE, vanMaanen L, Stough DP, Hartzel J, et al. Prediction of mortality from 12-lead electrocardiogram voltage data using a deep neural network. *Nat Med* 2020;**26**:886–91. <https://doi.org/10.1038/s41591-020-0870-z>
  15. Lima EM, Ribeiro AH, Paixao GMM, Ribeiro MH, Pinto-Filho MM, Gomes PR, et al. Deep neural network-estimated electrocardiographic age as a mortality predictor. *Nat Commun* 2021;**12**:5117. <https://doi.org/10.1038/s41467-021-25351-7>
  16. Mayourian J, La Cava WG, Vaid A, Nadkarni GN, Ghelani SJ, Mannix R, et al. Pediatric ECG-based deep learning to predict left ventricular dysfunction and remodeling. *Circulation* 2023;**149**:917–31. <https://doi.org/10.1161/CIRCULATIONAHA.123.067750>
  17. Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Van Calster B, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ* 2024;**385**:e078378. <https://doi.org/10.1136/bmj-2023-078378>
  18. Einthoven W. Weiteres über das elektrokardiogramm. *Archiv für die gesamte Physiologie des Menschen und der Tiere* 1908;**122**:517–84. <https://doi.org/10.1007/BF01677829>
  19. Goldberger E. A simple, indifferent, electrocardiographic electrode of zero potential and a technique of obtaining augmented, unipolar, extremity leads. *Am Heart J* 1942;**23**:483–92. [https://doi.org/10.1016/S0002-8703\(42\)90293-X](https://doi.org/10.1016/S0002-8703(42)90293-X)
  20. Colan SD. Early database initiatives: the Fyler codes. In: Barach PR, Jacobs JP, Lipshultz SE, Laussen PC (eds.), *Pediatric and Congenital Cardiac Care*. Vol 1. Outcomes Analysis. London: Springer London, 2015, 163–9.
  21. Jacobs JP, Franklin RCG, Beland MJ, Spicer DE, Colan SD, Walters HL III, et al. Nomenclature for pediatric and congenital cardiac care: unification of clinical and administrative nomenclature—the 2021 International Paediatric and Congenital Cardiac Code (IPCCC) and the Eleventh Revision of the International Classification of Diseases (ICD-11). *World J Pediatr Congenit Heart Surg* 2021;**12**:E1–E18. <https://doi.org/10.1177/21501351211032919>
  22. Gustafsson S, Gedon D, Lampa E, Ribeiro AH, Holzmann MJ, Schon TB, et al. Development and validation of deep learning ECG-based prediction of myocardial infarction in emergency department patients. *Sci Rep* 2022;**12**:19615. <https://doi.org/10.1038/s41598-022-24254-x>
  23. Diller G-P, Kempny A, Alonso-Gonzalez R, Swan L, Uebing A, Li W, et al. Survival prospects and circumstances of death in contemporary adult congenital heart disease patients under follow-up at a large tertiary centre. *Circulation* 2015;**132**:2118–25. <https://doi.org/10.1161/CIRCULATIONAHA.115.017202>
  24. Diller GP, Kempny A, Babu-Narayan SV, Henriks M, Brida M, Uebing A, et al. Machine learning algorithms estimating prognosis and guiding therapy in adult congenital heart disease: data from a single tertiary centre including 10 019 patients. *Eur Heart J* 2019;**40**:1069–77. <https://doi.org/10.1093/eurheartj/ehy915>
  25. Ribeiro AH, Ribeiro MH, Paixao GMM, Oliveira DM, Gomes PR, Canazart JA, et al. Automatic diagnosis of the 12-lead ECG using a deep neural network. *Nat Commun* 2020;**11**:1760. <https://doi.org/10.1038/s41467-020-15432-4>
  26. Ulloa-Cerna AE, Jing L, Pfeifer JM, Raghunath S, Ruhl JA, Rocha DB, et al. rECHOmmed: an ECG-based machine learning approach for identifying patients at increased risk of undiagnosed structural heart disease detectable by echocardiography. *Circulation* 2022;**146**:36–47. <https://doi.org/10.1161/CIRCULATIONAHA.121.057869>
  27. Khurshid S, Friedman S, Reeder C, Di Achille P, Diamant N, Singh P, et al. ECG-based deep learning and clinical risk factors to predict atrial fibrillation. *Circulation* 2022;**145**:122–33. <https://doi.org/10.1161/CIRCULATIONAHA.121.057480>
  28. Sangha V, Nargesi AA, Dhingra LS, Khunte A, Mortazavi BJ, Ribeiro AH, et al. Detection of left ventricular systolic dysfunction from electrocardiographic images. *Circulation* 2023;**148**:765–77. <https://doi.org/10.1161/CIRCULATIONAHA.122.062646>
  29. Bokma JP, Winter MM, Vehmeijer JT, Vliegen HW, van Dijk AP, van Melle JP, et al. QRS fragmentation is superior to QRS duration in predicting mortality in adults with tetralogy of fallot. *Heart* 2017;**103**:666–71. <https://doi.org/10.1136/heartjnl-2016-310068>
  30. Straus SM, Kors JA, De Bruin ML, van der Hooft CS, Hofman A, Heeringa J, et al. Prolonged QTc interval and risk of sudden cardiac death in a population of older adults. *J Am Coll Cardiol* 2006;**47**:362–7. <https://doi.org/10.1016/j.jacc.2005.08.067>
  31. Solomon SD, Anavekar N, Skali H, McMurray JJ, Swedberg K, Yusuf S, et al. Influence of ejection fraction on cardiovascular outcomes in a broad spectrum of heart failure patients. *Circulation* 2005;**112**:3738–44. <https://doi.org/10.1161/CIRCULATIONAHA.105.561423>
  32. Klem I, Klein M, Khan M, Yang EY, Nabi F, Ivanov A, et al. Relationship of LVEF and myocardial scar to long-term mortality risk and mode of death in patients with nonischemic cardiomyopathy. *Circulation* 2021;**143**:1343–58. <https://doi.org/10.1161/CIRCULATIONAHA.120.048477>
  33. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;**44**:837–45. <https://doi.org/10.2307/2531595>
  34. Makowski D, Pham T, Lau ZJ, Brammer JC, Lespinasse F, Pham H, et al. NeuroKit2: a Python toolbox for neurophysiological signal processing. *Behav Res Methods* 2021;**53**:1689–96. <https://doi.org/10.3758/s13428-020-01516-y>
  35. Lundberg SM, Erion G, Chen H, DeGrave A, Prutkin JM, Nair B, et al. From local explanations to global understanding with explainable AI for trees. *Nat Mach Intell* 2020;**2**:56–67. <https://doi.org/10.1038/s42256-019-0138-9>
  36. Team PC. *Python: A Dynamic, Open Source Programming Language*. Wilmington, DE: Python Software Foundation; 2015. <https://www.python.org/>
  37. Team RC. *R: A Language and Environment for Statistical Computing*. Indianapolis, Indiana, United States: R Foundation for Statistical Computing; 2019. <https://www.r-project.org/>
  38. Gatzoulis MA, Balaji S, Webber SA, Siu SC, Hokanson JS, Poile C, et al. Risk factors for arrhythmia and sudden cardiac death late after repair of tetralogy of Fallot: a multicentre study. *Lancet* 2000;**356**:975–81. [https://doi.org/10.1016/S0140-6736\(00\)02714-8](https://doi.org/10.1016/S0140-6736(00)02714-8)
  39. Valente AM, Gauvreau K, Assenza GE, Babu-Narayan SV, Schreier J, Gatzoulis MA, et al. Contemporary predictors of death and sustained ventricular tachycardia in patients with repaired tetralogy of Fallot enrolled in the INDICATOR cohort. *Heart* 2014;**100**:247–53. <https://doi.org/10.1136/heartjnl-2013-304958>
  40. Sha J, Zhang S, Tang M, Chen K, Zhao X, Wang F. Fragmented QRS is associated with all-cause mortality and ventricular arrhythmias in patient with idiopathic dilated cardiomyopathy. *Ann Noninvasive Electrocardiol* 2011;**16**:270–5. <https://doi.org/10.1111/j.1542-474X.2011.00442.x>
  41. Karikari Y, Abdulkarim M, Li Y, Looma RS, Zimmerman F, Husayni T. The progress and significance of QRS duration by electrocardiography in hypoplastic left heart syndrome. *Pediatr Cardiol* 2020;**41**:141–8. <https://doi.org/10.1007/s00246-019-02237-6>
  42. Broberg CS, van Dissel A, Minnier J, Aboulhossn J, Kauling RM, Ginde S, et al. Long-term outcomes after atrial switch operation for transposition of the great arteries. *J Am Coll Cardiol* 2022;**80**:951–63. <https://doi.org/10.1016/j.jacc.2022.06.020>
  43. Gearhart A, Bassi S, Rathod RH, Beroukhim RS, Lipsitz S, Gold MP, et al. Ventricular dyssynchrony late after the Fontan operation is associated with decreased survival. *J Cardiovasc Magn Reson* 2023;**25**:66. <https://doi.org/10.1186/s12968-023-00984-3>
  44. Mayourian J, Sleeper LA, Lee JH, Lu M, Geva A, Mulder B, et al. Development and validation of a mortality risk score for repaired tetralogy of Fallot. *J Am Heart Assoc* 2024;**13**:e034871. <https://doi.org/10.1161/JAHA.123.034871>
  45. Saliccioli KB, Oluyomi A, Lupo PJ, Ermi PR, Lopez KN. A model for geographic and sociodemographic access to care disparities for adults with congenital heart disease. *Congenit Heart Dis* 2019;**14**:752–9. <https://doi.org/10.1111/chd.12819>
  46. Lopez KN, Morris SA, Sexson Tejtel SK, Espallat A, Salemi JL. US mortality attributable to congenital heart disease across the lifespan from 1999 through 2017 exposes persistent racial/ethnic disparities. *Circulation* 2020;**142**:1132–47. <https://doi.org/10.1161/CIRCULATIONAHA.120.046822>
  47. Mandalenakis Z, Giang KW, Eriksson P, Liden H, Synnergren M, Wahlander H, et al. Survival in children with congenital heart disease: have we reached a peak at 97%? *J Am Heart Assoc* 2020;**9**:e017704. <https://doi.org/10.1161/JAHA.120.017704>
  48. Hinton RB, Ware SM. Heart failure in pediatric patients with congenital heart disease. *Circ Res* 2017;**120**:978–94. <https://doi.org/10.1161/CIRCRESAHA.116.308996>
  49. Goto S, Solanki D, John JE, Yagi R, Homilius M, Ichihara G, et al. Multinational federated learning approach to train ECG and echocardiogram models for hypertrophic cardiomyopathy detection. *Circulation* 2022;**146**:755–69. <https://doi.org/10.1161/CIRCULATIONAHA.121.058696>
  50. Jain S, Elias P, Poterucha T, Randazzo M, Lopez Jimenez F, Khera R, et al. Artificial intelligence in cardiovascular care—part 2: applications: JACC review topic of the week. *J Am Coll Cardiol* 2024;**83**:2487–96. <https://doi.org/10.1016/j.jacc.2024.03.401>
  51. Ghassemi M, Oakden-Rayner L, Beam AL. The false hope of current approaches to explainable artificial intelligence in health care. *Lancet Digit Health* 2021;**3**:e745–50. [https://doi.org/10.1016/S2589-7500\(21\)00208-9](https://doi.org/10.1016/S2589-7500(21)00208-9)
  52. Lin CS, Liu WT, Tsai DJ, Lou YS, Chang CH, Lee CC, et al. AI-enabled electrocardiography alert intervention and all-cause mortality: a pragmatic randomized clinical trial. *Nat Med* 2024;**30**:1461–70. <https://doi.org/10.1038/s41591-024-02961-4>