ECG-Age for Cardiovascular Outcome Prediction: Evaluation in an Adult Congenital Heart Disease Cohort

Seyedeh Somayyeh Mousavi[1], Cheryl L. Raskind-Hood[2], Alexandra Haffner[3], Gari D. Clifford[1,4], Wendy M Book[2,3], Reza Sameni[1,4]
[1] Department of Biomedical Informatics, Emory University, Atlanta, USA
[2] Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, USA
[3] Department of Medicine, Emory University, Atlanta, USA
[4] Department of Biomedical Engineering, Georgia Institute of Technology, Atlanta, USA

Abstract

Electrocardiogram (ECG)-based estimated age has been proposed as a marker of cardiac well-being, beyond chronological age (CA). Significant gaps between CA and ECG-based age estimates (ECG-Age) have been associated with adverse cardiovascular outcomes, but their physiological interpretation and generalization across different cohorts remains unclear. We developed a machine learning model to estimate ECG-Age from interpretable ECG features and evaluated its utility for outcome prediction. The model was trained and validated on the CODE15 dataset using 1,316 ECG features. The model achieved a correlation coefficient of 0.83 between CA and the estimated ECG-Age. The trained model was then applied to an adult congenital heart disease (ACHD) dataset with different ventricular loading conditions. We show that the ECG-Age of different ACHD ventricular loading groups is significantly higher than their CA. Demographics, ECG-Age estimates, and the 1,316 ECG features were further used to predict adverse cardiac outcomes within 3–24 months after ECG acquisition using four input feature set scenarios:(1) sex and CA, (2) sex, CA, and ECG-Age, (3) sex, CA, and ECG features, and (4) sex, CA, ECG-Age, and ECG features. The results show that while ECG-Age provides incremental predictive value beyond sex and CA alone, it does not provide additional predictive value when ECG features are included in the feature set for predicting adverse cardiac outcomes in the ACHD population.

1. Introduction

Aging is a biological process that gradually affects physiological functions. Chronological age (CA) is therefore a major factor in clinical risk assessment [1]. Biological age, estimated by artificial intelligent models from physiological measurements such as electrocardiogram (ECG), has been proposed as an indicator of an individual's functional state beyond CA. Significant gaps between CA and ECG-based age estimates (ECG-Age) of, e.g., above eight years, have been associated with adverse cardiovascular outcomes [2]. However, existing ECG-Age models are dominantly based on deep learning (DL) approaches, which limit their interpretability.

Subjects with adult congenital heart disease (ACHD) can potentially benefit from the notion of ECG-Age as a measure of their lifelong cardiovascular disease progression and severity. Subjects with ACHD often require lifelong follow-ups to reduce the risk of progressive complications. Our previous studies highlighted the application of ECG to predict adverse cardiac outcomes in the ACHD population [3–5]. ACHD patients are often categorized based on ventricular loading conditions into four groups: defects primarily affecting the left or systemic side of the heart ("Left-side"), defects primarily affecting the right or pulmonic side of the heart ("Right-side"), patients with a univentricular circulation ("Single-ventricle"), and other defects that do not primarily affect cardiac loading conditions ("Neither") [5].

In this work, we develop a machine learning (ML) model to estimate ECG-Age from a comprehensive set of ECG-derived interpretable features. The model performance is compared with a state-of-the-art DL-based model [2]. We evaluate the model performance in an ACHD population, across different anatomical groups. Finally, we explore whether ECG-Age provides incremental predictive value for predicting adverse cardiac outcomes beyond CA and ECG features in the ACHD population.

2. Method

2.1. Interpretable ECG–Age Estimation

We used 12-lead ECGs from the CODE15 dataset [2] to build our interpretable ECG-Age models. ECG Records from subjects aged 18–90 years with available sex information (male or female) were selected. Records with missing leads in any of the 12 leads were excluded from analysis. The resulting subset comprises 333,711 ECGs (7–10 s sampled at 400 Hz) from 228,297 unique subjects.

The ECG recordings were preprocessed to remove base-

line wander and out-of-band noise. A 60 Hz notch filter was applied to remove powerline interference using a second-order infinite impulse response filter with a quality factor of 45, implemented with zero-phase forward-backward filtering. The baseline wander was then estimated using a moving median filter with a sliding window length of 0.6 s, followed by a moving average filter with a sliding window length of 0.3 s. The resulting baseline was subtracted from the ECG. The recordings were further filtered using a band-pass filter with a passband of 0.1–100 Hz [5].

We developed an ECG feature extraction toolset comprising a comprehensive set of clinically relevant ECG characteristics, heart rate metrics, time interval and amplitude measurements, area under waveform components, component slopes, beat-wise regularity and waveform quality measures, average beat morphology, Hjorth descriptors, and inter-beat variability metrics derived using singular value decomposition on stacked beats [5, 6]. The toolset can be applied to datasets with varying ECG lengths, arbitrary lead configurations, and different sampling frequencies. The ECG feature extraction toolset is available in the Open-Source Electrophysiological Toolbox (OSET) [7]. Using this toolset, a total of 1,316 ECG features were extracted from each 12-lead ECG record. All features were computed for each lead separately, except for the average beat morphology, which was derived exclusively from leads I and II.

Multiple regression models were trained and evaluated. A CatBoost regression model trained and cross-validated on 80% of the subjects and tested on the remaining 20% outperformed all models. The performance of the developed models was evaluated to enable comparison with other studies using various metrics, including mean error (ME), mean absolute error (MAE), standard deviation (SD), and the correlation coefficient (ρ) between CA and ECG-Age. We also used Shapley Additive Explanations (SHAP) plots to highlight the most important features contributing to accurate age estimation by the ML model.

2.2. Predicting Adverse ACHD Outcomes

For predicting adverse cardiac outcomes, we used an ECG cohort of adult male and female ACHD subjects, aged 18 to 90 years, with one or more 12-lead ECG recordings (10 s, sampled at 500 Hz) collected during an Emory Healthcare encounter. ACHD patients were identified based on CHD-related International Classification of Diseases, Ninth and Tenth Revision, Clinical Modification diagnosis codes (ICD-9-CM and ICD-10-CM). The dataset includes 62,330 ECG records from 10,974 subjects. The study was approved by the Emory University Institutional Review Board under Study ID #7815.

For outcome labeling, an ECG record was considered positive if any of the following clinical outcomes was documented in the patient chart within 3 or 24 months after ECG acquisition: (1) death, (2) cardiac hospitalization, or (3) cardiac surgery or intervention. The 3–24 months prediction horizon was selected by an expert cardiologist as a clinically actionable time window. The outcome-labeling procedure is detailed in [5]. After excluding subjects without sufficient follow-up for the selected prediction horizon, 12,775 ECG records from 5,353 subjects were included in the outcome-prediction analysis.

Table 1. Performance of developed machine learning and open-source high-performance deep learning models for estimating ECG-Ages on the CODE15 (20%) and CHD datasets. (ME: Mean Error, MAE: Mean Absolute Error, STD: Standard Deviation, ρ: Correlation Coefficient).

Model	Dataset	# Records	ME	STD	MAE	STD	ρ
Our Model	CODE15	66,517	**0.06**	**10.68**	**8.37**	**6.64**	**0.84**
DL Model [2]	CODE15	66,517	2.14	11.14	8.72	7.26	0.83
Our Model	ACHD	62,330	12.39	14.12	14.78	11.59	0.44
DL Model [2]	ACHD	62,330	16.56	16.83	18.82	14.25	0.37

For adverse cardiac outcome prediction, we trained and evaluated four CatBoost models with different feature sets as input: (1) sex and CA as a baseline demographics-only model, (2) sex, CA, and ECG-Age (pretrained on the CODE15 dataset), (3) sex, CA, and comprehensive ECG features (1,316), and (4) sex, CA, ECG- Age, and ECG features.

CatBoost classifiers were trained and evaluated with 80% training and 20% testing subject-level data splitting and applying class weighting. Model performances were evaluated using Receiver Operating Characteristic (ROC) curves and the Area Under the ROC Curve (AUROC) values. In addition, to systematically assess the incremental predictive value of individual feature sets, we performed pairwise comparisons of model AUROCs using DeLong's statistical test [8], which evaluates whether differences between two correlated ROCs obtained from models evaluated on the same dataset are statistically significant.

3. Results

3.1. ECG-Age Estimation Performance

Table 1 summarizes the performance of the CatBoost and the DL-based model [2] for age estimation using 20% of the CODE15 dataset (used as unseen test) and the ACHD datasets. Our ECG-Age model achieved an MAE of 8.37 years, an STD of 6.64 years, and a correlation coefficient of 0.84 using the CODE15 (20%), outperforming the DL model across all metrics. Fig. 1 presents the scatter plots of CA vs. ECG-Age, their histograms, and 75% and 95% contour plots on CODE15.

As shown in Table 1, both ECG-Age models report significantly higher ECG-Age than CA, with MEs of 12.39 and 16.56 years for our model and the DL model [2], respectively, highlighting that individuals in the ACHD population generally have sicker hearts, with ECGs that re-

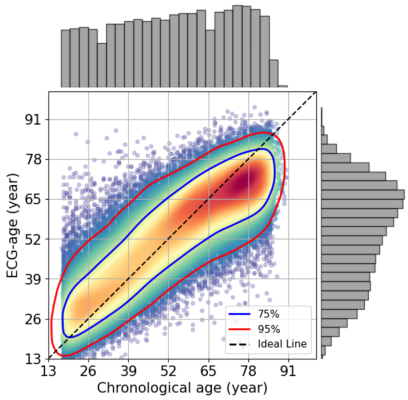

Figure 1. Scatter plot and 75% and 95% contour plots of age vs. ECG-age using the developed model for CODE15.

Table 2. Performance of the developed ECG-Age estimator on the ACHD cohort stratified by ventricular loading conditions. (ME: Mean Error, MAE: Mean Absolute Error, STD: Standard Deviation, ρ: Correlation Coefficient).

Ventricular loading conditions	# Record	ME	STD	MAE	STD	ρ
Left-side	23,665	12.38	13.45	14.45	11.21	0.52
Neither	3,046	6.38	11.32	9.92	8.39	0.50
Right-side	31,006	10.81	13.71	13.67	10.85	0.44
Single-ventricular	4,613	26.95	12.93	27.21	12.38	0.26

semble those of older age groups. Table 2 summarizes the ACHD ECG-Age estimation results stratified by ventricular loading conditions. The Left-side group had the highest correlation coefficient (0.52), while the Single-ventricle group had the lowest correlation coefficient (0.26) between CA and ECG-Age values. Fig. 2 shows the 75th percentile contour plots of the CA vs. ECG-Age of the ACHD cohort stratified by ventricular loading conditions. Fig. 3 presents the SHAP analysis of the top 15 features of the ECG-Age estimation model on the ACHD dataset, indicating that features related to morphological characteristics, amplitude, and Hjorth descriptors were among the top-ranked features. Sex is the second-most-important feature because the age of males and females had different distributions in the ACHD cohort (men were on average older than women).

3.2. ACHD Adverse Outcome Prediction

Fig. 4 presents the ROC curves for the ML models developed for predicting adverse cardiac outcomes within 3 to 24 months after ECG acquisition, using the four input feature sets detailed in Section 2.2. Scenarios 1 to 4 achieved AUROCs of 0.53, 0.59, 0.67, and 0.67, respectively. Table 3 summarizes the DeLong test results, p-values and z-scores, for the pairwise comparison between the four scenarios.

4. Discussion

i) ECG-Age is not an independent biomarker for ACHD: Given the growing interest in introducing ECG-Age as a marker of cardiovascular risk, if ECG-Age represented

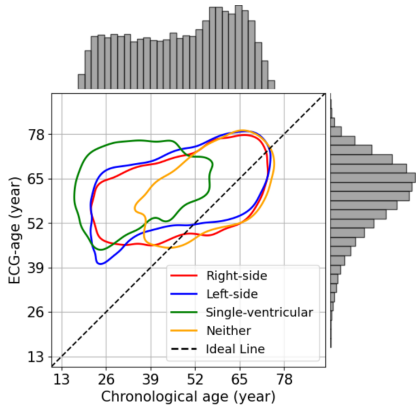

Figure 2. Age vs ECG-Age 75% contours from the proposed model across ACHD ventricular loading conditions

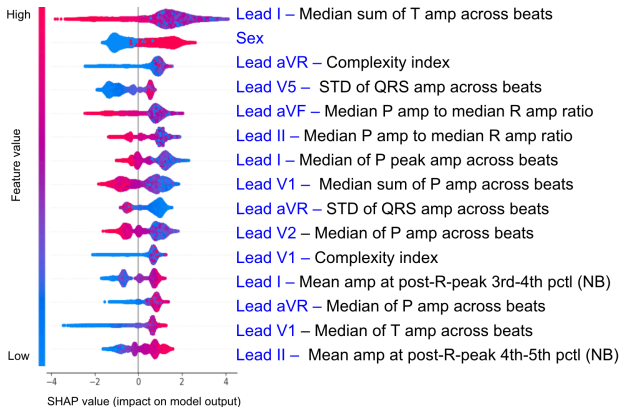

Figure 3. Top features by SHAP importance for ECG-Age estimation in the ACHD cohort (NB: normalized beat)

a stable physiological marker, independently developed models with comparable population-level performance would be expected to yield concordant estimates at the individual level. The observed variability between models (Table 1) suggests that ECG-Age is not uniquely identifiable from the ECG and instead reflects model-dependent regression of age-correlated features. Therefore, the exact deviations of ECG-Age from CA should not be interpreted as a biomarker for cardiovascular risk. Nonetheless, significant deviations of, e.g., over 8–10 years between CA and ECG-Age, may indeed reflect severe cardiac conditions, as in the ACHD cohort.

ii) ECG-Age distribution varies significantly across ACHD ventricular loading groups: According to Fig. 2, while CA is distributed across adulthood, ECG-Age exhibits a narrower distribution centered around 60–70 yrs. Across all ventricular loading groups, the model predicts ECG-Ages that are generally higher than subjects' CAs for younger individuals (<40 yrs), while the contour lines are closer to the identity line at older ages (>40 yrs). The Right-side and Left-side groups exhibit similar CA vs.

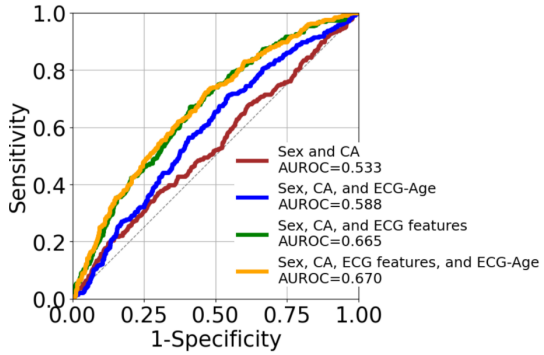

Figure 4. ROC curves of the best-performing ML models for predicting adverse cardiac outcomes within 3–24 mos. after ECG recordings in the ACHD population using sex, chronological age (CA), ECG-Age, and ECG features.

Table 3. Pairwise DeLong test between the developed ML models for predicting adverse cardiac outcomes within 3–24 months after ECG recordings in the ACHD population using demographics (sex and age), ECG-Age, and ECG features (w/wo: with/without).

Scenario	z**-score**	p**-value**
Dem. (w/wo) ECG-Age	2.67	7×10^{-3}
Dem. (w/wo) ECG-Features	5.46	$< 10^{-7}$
Dem. (w/wo) ECG-Age & ECG-Features	5.65	$< 10^{-7}$
Dem. & ECG-Age (w/wo) ECG-Features	4.17	$< 3 \times 10^{-5}$
Dem. & ECG-Age (vs.) Dem. & ECG-Features	4.44	$< 9 \times 10^{-6}$
Dem. & ECG-Features (w/wo) ECG-Age	0.67	0.50

ECG-Age distribution. In contrast, the Single-ventricle group demonstrates a distinct pattern. Individuals in this group tend to have higher ECG-Age estimates despite younger CAs, resulting in the largest ECG-Age gap among the ventricular loading groups. This observation is consistent with the lower survival observed in this population. The Neither group exhibits an intermediate pattern, with ECG-Age remaining generally higher than CA but approaching the identity line at older ages.

iii) Direct ECG features vs. ECG-Age for outcome prediction: According to Table 3 the model incorporating sex, CA, and ECG features significantly outperformed the model using only sex and CA (z-score $= 5.46$, p-value $< 10^{-7}$), confirming our previous findings that ECG features provide additional discriminatory information for ACHD outcome prediction [5]. Although adding ECG-Age to the model containing only sex and CA improved predictive performance (z-score $= 2.67$, p-value $= 0.007$), incorporating ECG-Age into the model including sex, CA, and ECG features did not result in a statistically significant improvement (z-score $= 0.67$, p-value $= 0.50$). This finding suggests that ECG features already capture most of the predictive information provided by ECG-Age.

5. Conclusion

We developed a feature-based, interpretable ECG-Age estimation model with performance comparable to that of DL-based models, and evaluated its application across different ventricular loading groups and for predicting adverse outcomes in an ACHD cohort. Although ECG-Age was significantly higher than CA across the different ACHD groups, as a feature for ML, it provided incremental predictive value beyond demographics and ECG features for outcome prediction. There is also a significant discrepancy between ECG-Ages predicted by complementary models. In conclusion, ECG-Age is model dependent, not an independent biological biomarker; application-specific ML models with engineered ECG features are more predictive than ECG-Age alone.

Acknowledgments

This study was supported by the NIH NHLBI Award R2 1HL172209, and the NIH NCATS Award UL1TR002378.

References

[1] Cooney MT, Vartiainen E, Laatikainen T, De Bacquer D, McGorrian C, Dudina A, Graham I, SCORE, investigators F. Cardiovascular risk age: concepts and practicalities. Heart 2012;98(12):941–946.

[2] Lima EM, Ribeiro AH, Paixão GMM, Ribeiro MH, Pinto-Filho MM, Gomes PR, Oliveira DM, Sabino EC, Duncan BB, Giatti L, Barreto SM, Meira Jr W, Schön TB, Ribeiro ALP. Deep neural network-estimated electrocardiographic age as a mortality predictor. Nature Communications August 2021;12(1). ISSN 2041-1723.

[3] Book WM, Hurst JW, Parks WJ, Hopkins KL. Electrocardiographic predictors of right ventricular volume measured by magnetic resonance imaging late after total repair of tetralogy of fallot. Clinical Cardiology November 1999; 22(11):740–746. ISSN 1932-8737.

[4] Mousavi SS, Raskind-Hood CL, Haffner A, Robichaux C, Ivey LC, Sameni R, Book WM. Electrocardiogram Characteristics and Outcomes in an Adult Congenital Heart Disease Population. Annals of Noninvasive Electrocardiology 2026; 31(5):e70224. E70224 5717826.

[5] Mousavi SS, Raskind-Hood CL, Haffner A, Robichaux C, Ivey LC, Sameni R, M Book W. Electrocardiogram-Based Machine Learning Model for Predicting Adverse Cardiac Outcomes in Adult Congenital Heart Disease, May 2026.

[6] Mousavi SS, Karimi S, Hassannia MS, Koscova Z, Rad AB, Albert D, Clifford GD, Sameni R. Estimating Blood Pressure from the Electrocardiogram: Findings of a Large-Scale Negative Results Study. Physiological Measurement 2025;.

[7] Mousavi S, Karimi S, Sameni R. The Open-Source Electrophysiological Toolbox (OSET) ECG Feature Subset, 2006–2025. URL https://github.com/alphanumericslab/OSET/tree/master/matlab/tools/ecg/ecg_features.

[8] DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach. Biometrics September 1988;44(3):837. ISSN 0006-341X.

Address for correspondence:
Reza Sameni, 101 Woodruff Circle, Atlanta, GA 30322, USA
rsameni@dbmi.emory.edu