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

Seyedeh Somayyeh Mousavi[1], Wendy M Book[2], Reza Sameni[1,3]
[1] Department of Biomedical Informatics, Emory University, USA
[2] Department of Medicine, Emory University, USA
[3] Department of Biomedical Engineering, Georgia Institute of Technology, USA

Biological age, inferred from physiological measurements, has been proposed as an indicator of functional state beyond chronological age (CA). However, it lacks a standard, clinically accepted definition. In cardiology, differences between CA and electrocardiogram (ECG)–based age estimates (ECG-age) have been associated with adverse cardiovascular outcomes, but their physiological interpretation remains unclear—whether they reflect intrinsic aging processes or serve as a surrogate for aggregate cardiovascular risk.

We developed a machine learning model to estimate ECG-age from ECG–derived interpretable features and evaluated its utility for outcome prediction. The model was trained and validated on the CODE-15 12-lead dataset (345,779 ECGs, 233,770 subjects) using over 1,300 explainable features of waveform morphology and rhythm.

Model performance was benchmarked against state-of-the-art deep learning–based ECG-age models. The model achieved a correlation coefficient of 0.83 with CA, comparable to models on the same dataset. Agreement between the two approaches showed a correlation of 0.88 with a mean absolute error of 8.4 years, indicating substantial differences in individual-level estimates. The trained model was then applied to an adult congenital heart disease (ACHD) cohort (63,979 ECGs, 11,028 subjects). The ECG-age estimates were used to predict adverse cardiac outcomes (cardiac hospitalization, intervention, surgery, or death) within 3–90 months following ECG acquisition. Outcome prediction using CA alone, ECG-age alone, and their combination resulted in AUROC of 0.603, 0.587, and 0.627, respectively.

These results show that ECG-age provides incremental predictive value beyond CA in ACHD outcome prediction. However, if ECG-age represented 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 suggests that ECG-age is not uniquely identifiable from the ECG and instead reflects model-dependent regression of age-correlated features. Therefore, deviations of ECG-age from CA should not be interpreted as a biomarker for cardiovascular risk.