Development of AI for age prediction using ECG and the search for validity of AI-predicted ECG-age as a biomarker of atrial fibrillation risk

Sujeong Eom

The Graduate School
Yonsei University
Department of Biomedical Systems Informatics

Development of AI for age prediction using ECG and the search for validity of AI-predicted ECG-age as a biomarker of atrial fibrillation risk

A Master's Thesis Submitted
to the Department of Biomedical Systems Informatics
and the Graduate School of Yonsei University
in partial fulfillment of the
requirements for the degree of
Master of Science

Sujeong Eom

June 2024

This certifies that the Master's Thesis
of Sujeong Eom is approved.

Thesis Supervisor Seng Chan You

Thesis Committee Member Dukyong Yoon

Thesis Committee Member Hee Tae Yu

The Graduate School
Yonsei University
June 2024

Acknowledgements

I would like to express my deepest gratitude to my thesis advisor, Professor Seng Chan You, for his unwavering support, guidance, and invaluable mentorship throughout the entire process of writing this thesis. His expertise, encouragement, and constructive feedback have been instrumental in shaping this work.

I am also grateful to Professor Hee Tae Yu and Dukyong Yoon for their valuable insights and feedback during the review process. I would like to extend my gratitude to all the professors and colleagues, especially Professor Boyoung Joung and Seunghoon Cho, for their support and collaboration throughout this study.

I would like to thank my family for their love, encouragement, and understanding throughout this journey. Their supports have been my source of strength.

Special thanks to my colleagues and friends for their encouragement, discussions, and assistance throughout my research.

Finally, I am grateful to all those who directly or indirectly contributed to the completion of this thesis.

Thank you.

*This thesis will be submitted for consideration to a journal for potential publication.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Development of AI for age prediction using ECG and the search for validity of AI-predicted ECG-age as a biomarker of atrial fibrillation risk

Background: Applying artificial intelligence (AI) algorithms to 12-lead electrocardiograms (ECGs) to predict age is emerging as a promising method. In this study, I investigated whether the difference between predicted age from ECG waveform data and actual age, referred to as ECG-age gap or ECG-aging, is associated with the risk of incident atrial fibrillation (AF), a representative cardiac condition related to aging.

Methods: I developed an ECG-age prediction model using a large dataset (1,533,042 ECGs from 689,639 individuals) and validated it using five independent multinational datasets (637,177 ECGs from 230,838 individuals). ECG-age gap was calculated in three cohorts from Korea and the UK, with each cohort followed for 4.14±4.27, 6.08±3.81, and 2.99±1.56 years, comprising 111,483, 37,517, and 40,973 participants, respectively. Participants were classified into two groups based on ECG-age gap: Normal group (ECG-age gap <+7 years) and ECG-aging group (ECG-age gap ≥+7 years). The ability of ECG-aging to predict the risk of new-onset AF was evaluated using Cox proportional hazards models.

Results: The mean ECG-age and ECG-age gap in the three cohorts were 51.9±16.2 (0.0±6.8), 47.4±12.5 (-0.1±6.0), and 68.4±7.8 (4.7±8.7) years, respectively. The ECG-aging group had an increased risk of new-onset atrial fibrillation compared to the Normal group in each cohort, with hazard ratios of 2.50 (95% confidence interval [CI], 2.24–2.78), 1.89 (1.46–2.43), and 1.90 (1.55–2.33), respectively. The risk of incident atrial fibrillation increased with increasing ECG-age gap.

Conclusion: AI-derived ECG-aging was associated with an increased risk of incident

atrial fibrillation, indicating its potential as a risk biomarker for AF in primary prevention.

Keywords: Artificial Intelligence, Electrocardiogram, Aging, Atrial fibrillation

1. Introduction

1.1. Research background

Electrocardiography (ECG) is a simple and noninvasive diagnostic tool widely used for CVDs including AF. ECG provides a graphical representation of an individual's cardiac functioning.[1] Cardiac functioning is known to reflect body's features, such as normal aging, disease status, and heterogenous characteristics.[1]

Recently, much research has focused on applying artificial intelligence (AI) to standard 12-lead ECG data, enabling automated analysis of complex ECG features, and producing significant outcomes that contribute to medical decision-making processes. One mainstream in AI-ECG research involves the automatic diagnosis of CVDs, such as arrhythmias or coronary artery diseases, using neural networks and extensive ECG datasets.[2,3] Another direction aims at predicting future events, including mortality or CVDs. For instance, one study used resting 12-lead ECGs to develop a deep convolutional neural network (CNN) for predicting long-term cardiovascular mortality and disease risk.[4] Another study developed a CNN model using 12-lead ECGs to predict the 5-year incident AF risk.[5] These studies demonstrated that when combined with existing risk assessment tools, CNN model outcomes improved cardiovascular risk stratification.[4,5]

An important point to note is that neural networks can detect minute changes that may have been undetectable due to inconsistencies in manifestation or the small magnitude of the changes.[6]

Biological aging occurs at different rates, meaning that people of the same age can experience heterogeneous health conditions.[1] Biological aging represents a decline in functional ability, whereas chronological aging simply measures the amount of time since birth.[1] Effective methods for assessing biological aging have been proposed, including blood, genome, DNA, and physiological markers.[7] As ECG is known to convey subtle functional status, several studies have explored the possibility of AI-predicted ECG-age as a biomarker for biological aging, leading to the latest stream of AI-ECG research. Some studies have demonstrated that AI-predicted ECG-age can serve as a biomarker for

predicting aging-related CVD or mortality risks. A larger difference between AI-predicted ECG-age and chronological age was associated with all-cause mortality or CVDs.[8,9]

Atrial fibrillation (AF) is the most encountered arrhythmia and is associated with an increased risk of stroke, heart failure and mortality.[10] The global burden of AF in 2019 is estimated at 59.7 million (95% confidence interval: 45.7 to 75.3 million), which represents a doubling of the estimated cases in 1990.[11] Risk factors for AF include high systolic blood pressure, high body-mass index, and alcohol use.[12] Among all the risk factors, chronological age is the most important factor for AF.[13] Due to an aging population, the prevalence of AF continues to rise. AF is strongly connected to other cardiovascular diseases (CVDs) associated with mortality, and given the increasing prevalence, many efforts are underway to prevent the disease in its early stages.

ECGs are easy to obtain and accessible; therefore, using ECG for aging-related disease risk stratification could save a significant amount of time, effort, and budget. However, the validity of AI-predicted ECG-age as a biomarker of AF, a well-known aging-related disease, has not been sufficiently explored in diverse ethnicities, including Asians.

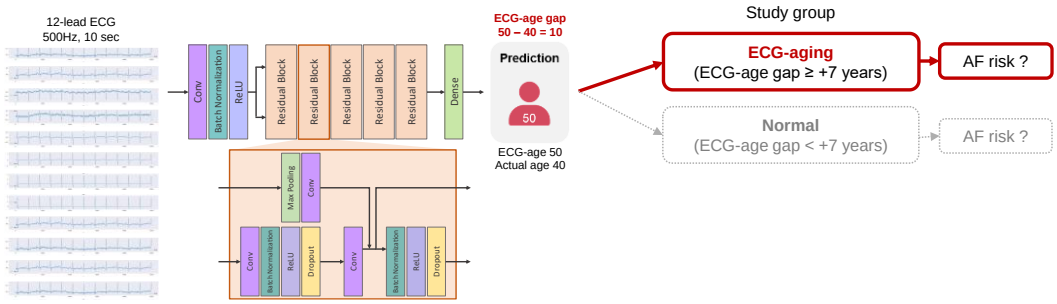
1.2. Objective

AF, a common arrhythmia, increases the risks of stroke, heart failure, and mortality.[11] Aging stands out as the single most significant risk factor for AF. With rapid aging in progress, early detection and prevention of AF are crucial for socioeconomics and quality of life. ECG, a convenient, universal, and frequently used diagnostic tool, contains comprehensive patient information, including biological aging status. Numerous studies have demonstrated the potential of AI-predicted ECG-age as a biomarker for predicting future events.

This study assumes that ECG provides multifactorial information related to an individual's biological aging. Given that aging contributes to AF, the risk of future AF through aging markers detected in ECG data can be inferred. Therefore, the objective of this study is to use standard 12-lead ECGs and deep neural network (DNN) to predict heart

age and assess the predicted age's validity as a biomarker of AF risk by conducting a multinational cohort study. The overall study design is presented in Figure 1.

Figure 1. Overall pipeline of the study



This study has two big phases. First, develop and validate ECG-age prediction DNN. The model takes ECG as an input and outputs age. Second, the association between predicted ECG-age and AF risk is estimated.

Abbreviations: AF, atrial fibrillation; Conv, convolution; DNN, deep neural network; ECG, electrocardiogram; ECG-aging, electrocardiographic aging.

2. Literature Review

2.1. ECG and deep learning

ECG is the most widespread diagnostic tool for CVDs. It captures electrical activity of the heart from different angles. There are P-wave, QRS complex, and T-wave in a single normal heartbeat. Each part is the result of repolarization or depolarization of the atrium or ventricle. Various ECG features (duration of waves or amplitudes etc.) are considered by physicians to figure out the status of the patient. Standard ECG consists of 12-leads; I, II, III, aVL, aVR, avF, V1-V6. As ECG is easy to acquire and cost-effective, it is routinely taken as the primary monitoring exam and frequently tested during any medical procedures. Therefore, it has ended up with a huge amount of ECG data stacked in each hospital. With the powerful growth of computational power and deep learning techniques, many studies have developed end-to-end automatic interpretation or diagnostic systems of ECG. Hannun et al. developed a DNN that classifies 12 rhythm classes with 91,232 single-lead ambulatory ECGs.[14] The network was superior to cardiologists. Ribeiro et al. proposed a residual CNN architecture trained with over 2 million ECGs to classify 6 different abnormalities (1st degree AV block, right bundle branch block, left bundle branch block, sinus bradycardia, AF, sinus tachycardia) in 12-lead ECGs.[2] The model outperformed cardiology medical doctors. In addition to diagnosis, Raghunath et al. developed an end-to-end DNN to predict 1-year mortality.[15] The model showed a robust performance in predicting death in a subgroup of 'normal' ECGs. Similarly, Hughes et al. developed SEER (Stanford Estimator of Electrocardiogram Risk) with CNN and 12-lead ECGs to predict 5-year cardiovascular mortality.[4] When combined with the Pooled Cohort Equations, SEER identified patients with a higher risk of atherosclerotic disease, who would not have been considered otherwise. Recently Zvuloni et al. compared feature engineering and DNN in multiclass-multilabel/binary classification and regression tasks using 12-lead ECGs.[16] The results showed that DNN didn't outperform feature engineering at classification tasks but did at regression task. Further, feature engineering and DNN combined did not bring any advantage over DNN alone.

2.2. ECG derived biomarker of aging

Human body ages at different levels.[17] Especially, there is a huge heterogeneity in health status as one gets older.[18] Therefore, accurately estimating biological aging is important and several biomarkers to express biological aging status have been developed. Biological aging biomarkers should provide additive information about certain condition risks over chronological age. Biomarkers include telomere length, DNA methylation age, transcriptomic predictors, and proteomic predictors.[17] Belsky et al. found that different biomarkers of biological aging were not correlated, suggesting the differences in the aspects that the biomarkers capture.[19]

Age is known to affect ECG.[20] QRS axis, PR interval, and QTc show linear correlation with age.[21] In this context, there are many efforts to estimate age from ECGs. ECG-age prediction studies formulate with the focus on the gap between predicted age and actual age.

Several studies were conducted with classical statistical methods (including machine learning models) to estimate age from ECG. These studies mostly included feature engineering of ECG parameters in the initial step of the main analysis. Starc et al. utilized 5 highly age-correlated ECG parameters from a healthy population and predicted functional heart age by linear regression analysis.[22] In 2014, Ball et al. introduced a Bayesian statistical model for heart age estimation from ECG.[23] Healthy non-athlete participants' heart age was in line with chronological age. Higher heart age than chronological age was revealed in subjects with risk factors or cardiac conditions. Using the same database, Lindow et al. reported that the heart age gap estimated by 10-second 12-lead ECG features and regression models well reflected cardiovascular risk and diseases.[24] With 438 ECG parameters, Hirota et al. estimated biological age with principal component analysis and Klemera and Doubal's method.[25] The predictive capability of estimated biological age for all-cause and cardiovascular death was analyzed. Principal component analysis was superior in the prediction of all-cause death but was inferior to the predictive capability of chronological age. These studies suggested that age inferred from ECG reflects cardiovascular status. The study results are limited to small populations with limited racial or ethnicity diversity (which is related to generalizability issues) and a burdensome of feature engineering process.

With the development of deep learning techniques, three main research teams have conducted consecutive ECG-age studies using DNNs. Raw ECG waveforms were used as the input for the network commonly.

Researchers from the Mayo Clinic (United States) first utilized 1-dimensional (1D) CNN to estimate age and sex with standard 10-second 12-lead ECGs.[26] The study demonstrated evidence for ECG as a biomarker and its reflection of health status. In a further study, with 25,144 participants, the associations between CNN-predicted ECG-age and chronological age gap and total or cardiovascular mortality were revealed.[1] Using the same architecture, Benavente et al. examined the relationship between ECG-age gap and CVD risk factors or age-related CVD markers in the Russian population.[27] Whether the population difference in CVD mortality is reflected in the CNN-predicted age was explored in a subsequent study.[28] The mean difference between predicted age and chronological age was higher in the Russian population who showed seven times higher CVD mortality than the Norway population. Another research team from Brazil developed a large ECG-age prediction DNN using approximately 1.3 million participants' data (namely the CODE study).[8] Remarkably, this study developed a CNN model with residual units, which are said to be efficient in time-series data. The ECG-age gap was associated with mortality even when analyzed using normal ECGs. Further, the difference between ECG-age and chronological age was associated with death and cardiovascular outcomes including atrial fibrillation in a long-term community-based population.[9] Lastly, a Taiwan research team found that excessive ECG-age compared to chronological age was associated with higher risk of all-cause mortality, cardiovascular mortality, heart failure, diabetes mellitus, chronic kidney disease, acute myocardial infarction, stroke, coronary artery disease, AF, and hypertension in a relatively healthy population (ECG12Net).[29] The study was externally validated with Brazil data only for all-cause mortality. The latter two research teams incorporated methods for model explainability in ECG-age prediction. With subsequent research using these developed ECG-age prediction algorithms, the studies provided evidence that 1) ECG-age conveys information on CVD related risk factors and 2) ECG-age is associated with future CVD risk in diverse populations. There was a single center study in Korea that examined the association of ECG-age with the risk of mortality or CVDs.[30]

2.3. Atrial fibrillation and risk prediction

AF is the most common arrhythmia with severe outcomes. Various AF risk prediction methods have been proposed. A well-known AF risk score is the CHARGE-AF (Cohorts for Heart and Aging Research in Genomic Epidemiology-AF) score published in 2013.[31] Variables including age, race, height, weight, blood pressure, current smoking, use of antihypertensive medication, diabetes, and a history of myocardial infarction and heart failure are used for the calculation. The score was validated across diverse populations in the United States and Europe. Several studies have tried to predict AF risk with 12-lead ECG and AI.[5,32-34] Christopoulos et al. and Khurshid et al. compared the performance of AF risk estimation between AI-ECG detected AF risk and the CHARGE-AF score.[5,33] They showed comparable performance while AI-ECG doesn't require any data abstraction procedure. The combination of multi-modal digital biomarkers (including demographics, clinical information and ECG features — heart rate variability, morphology, deep learning representation features) showed a robust performance in AF risk prediction.[35] This study revealed that structural changes in ECG are associated with AF.

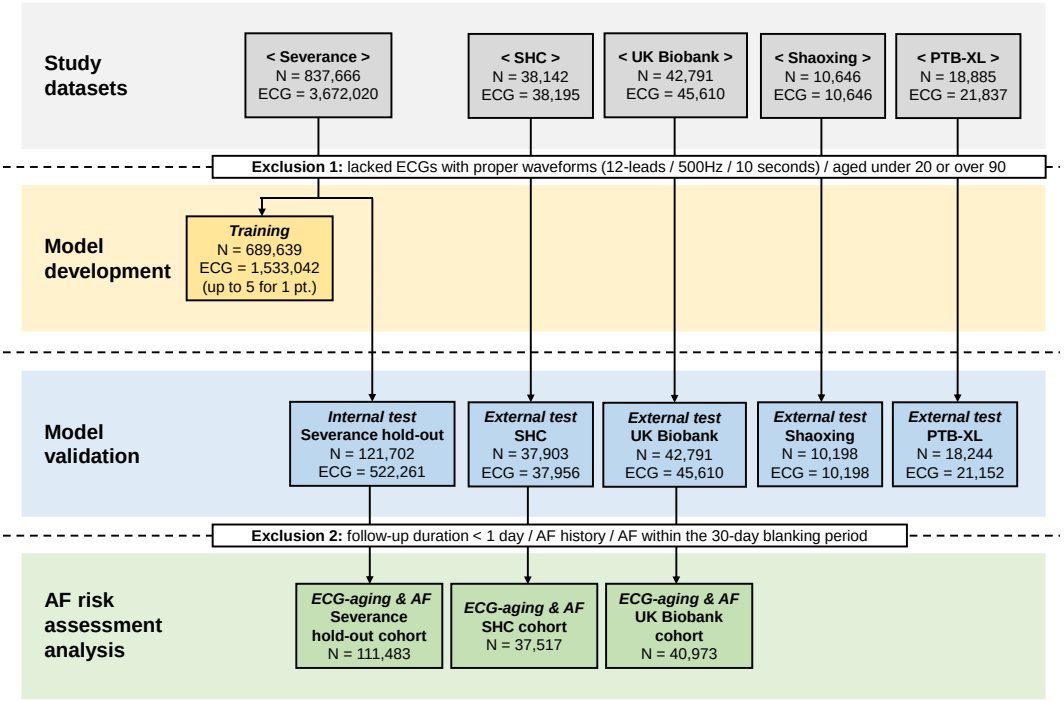
3. Materials and Methods

3.1. Data sources

Informed consent was waived for the use of deidentified data in the Severance and SHC datasets. The UK Biobank received ethical approval from the North West Multi-Centre Research Ethics Committee (11/NW/0382). Analysis of the UK Biobank dataset was conducted under application number 77793. All study analyses were approved by the Institutional Review Board of the Yonsei University Health System (4-2022-0731).

In this study, total 5 datasets from 4 different countries were utilized. 5 datasets are Severance (South Korea), Severance Health Check-up (South Korea), UK Biobank (United Kingdom), Shaoxing (China), and PTB-XL (Germany). 3,672,020 ECGs from 837,666 participants were identified in the Severance dataset. This dataset contains medical records, exams, and other health-related data from all patients who visited Severance Hospital from January 2006 to September 2021. Severance Hospital is a large tertiary referral center in South Korea. Participants who lacked ECGs with proper waveforms (12-leads / 500Hz / 10 seconds), under 20 or over 90 years old, or without age information were excluded, resulting in 3,494,908 ECGs from 811,341 participants. This was separated into training (1,443,298 ECGs from 649,072 participants), validation while training (89,744 ECGs from 40,567 participants), and internal validation dataset named Severance hold-out (522,261 ECGs from 121,702 participants). One internal validation dataset (Severance hold-out) and four external validation datasets were employed for the validation of ECG-age prediction: Severance Health Check-up (SHC; 37,956 ECGs from 37,903 participants), UK Biobank (45,610 ECGs from 42,791 participants), Shaoxing (10,198 ECGs from 10,198 participants), and PTB-XL (21,152 ECGs from 18,244 participants).[36,37] Severance hold-out, SHC, and UK Biobank datasets were further utilized as cohorts for AF risk analysis. Details are described in *Study cohorts and study design* section of the *MATERIALS AND METHODS*. A schematic flowchart of data sources with exclusion criteria is available in Figure 2 and detailed version is present in Figure 3.

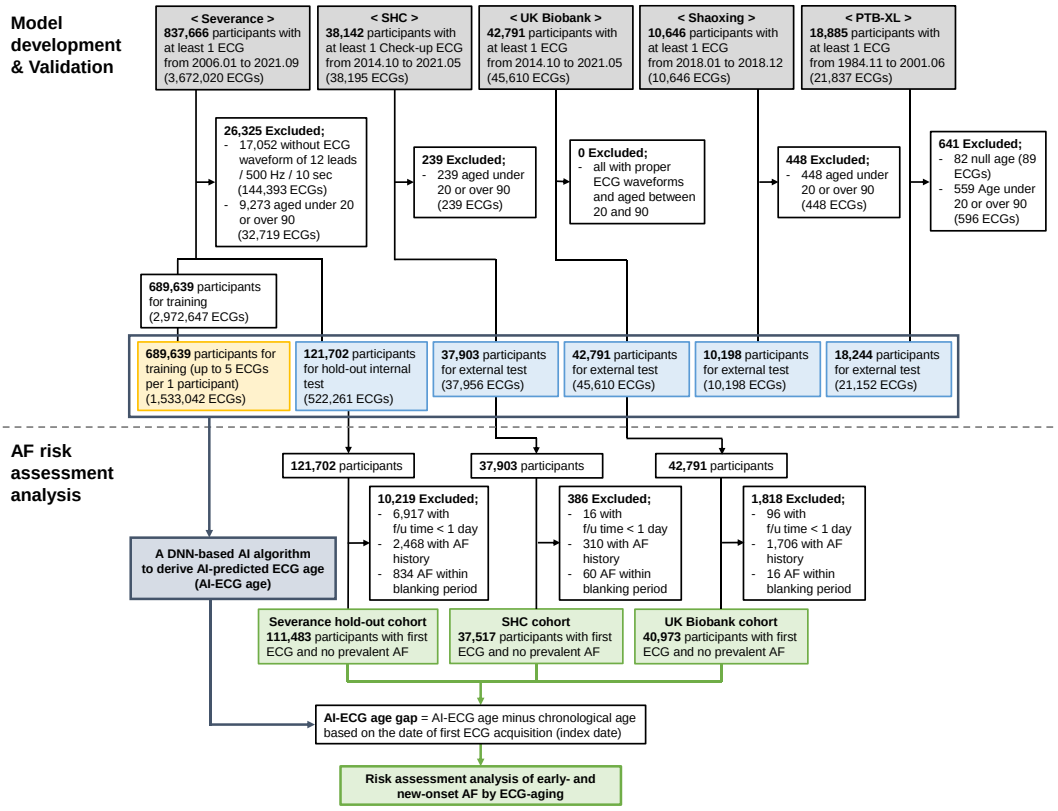
Figure 2. Schematic flowchart of datasets



[Model development] Up to five ECGs for a single participant was used for the ECG-age prediction model development. The participants for training include those for validation during the training process. [Model validation] The AI model was then validated in one hold-out test set and four external test sets. Participants without proper ECG waveforms (12-leads / 500Hz / 10 seconds), aged below 20 or above 90, or without age information were excluded from the model development and validation phase (Exclusion 1). [AF risk assessment analysis] For AF risk analysis, participants with a history of AF, a follow-up duration less than a day, or an AF occurrence during the 30-day blanking period were excluded (Exclusion 2).

Abbreviations: AF, atrial fibrillation; AI, artificial intelligence; ECG, electrocardiogram; ECG-aging, electrocardiographic aging; N, numbers of participants; pt., participants; SHC, Severance health check-up.

Figure 3. Detailed flowchart of datasets



The flow-chart includes study inclusion and exclusion criteria, and analysis flow, which is largely divided into two sections: model development & validation and AF risk assessment analysis. The AI-ECG age prediction model was trained on up to five ECGs per each participant in the training split of the Severance dataset to predict their chronological age. The participants for training include those designated for validation during the training process. The AI model was then validated in one hold-out internal test set and four external test sets.

Abbreviations: AF, atrial fibrillation; AI, artificial intelligence; DNN, deep neural network; ECG, electrocardiogram; ECG-aging, electrocardiographic aging; f/u, follow-up; SHC, Severance health check-up.

Detailed information about the five study datasets is as follows:

1. Severance

Severance Hospital, a tertiary acute-phase hospital in Seoul, South Korea, health data system contains medical records, medications, procedures, exams (encompassing a vast 12-lead ECG database), and other health-related information of all visiting patients. The Severance dataset comprised of participants who had more than one electrocardiogram (ECG) recorded between January 2006 and September 2021. The standard 12-lead ECG data was extracted from GE Healthcare's MUSE™ Cardiology Information System, with patient's information anonymized to a research ID. The ECG database includes multiple components, such as raw waveforms, measurement data (heart rate, atrial rate, ventricular rate, PR interval, QRS duration, QT interval, etc.), personal information (age, sex, etc.), and automated ECG diagnosis statements generated by the built-in software. The raw waveforms have a sampling rate of either 250Hz or 500Hz and a length of 10 seconds. Out of 837,666 participants with more than one ECG, 3,672,020 ECGs were identified, and only ECGs with a sampling rate of 500Hz were included in this study. Participants with ages below 20 or over 90 were excluded, as they are relatively rare in taking ECG tests (accounting for about 1% of the whole dataset). After exclusion, 3,494,908 ECGs from 811,341 participants were included, with a mean age of 59.8 (standard deviation [SD], 15.8) and 46.4% female. The same initial exclusion criteria set for the model development and validation phase were applied to all datasets. The Severance dataset was divided randomly into a model development (85%) and a hold-out (15%) split based on participants. Baseline data, including demographics, anthropometric measurements, and comorbidities, and the study outcomes were extracted from the Severance Clinical Research Analysis Portal (SCRAP), which is a medical big data platform integrating all in-hospital data and providing anonymized data extraction with various conditions. ECG data were linked to the SCRAP by research ID.

2. Severance Health Check-up

The Severance Health Check-up (SHC) dataset is from SHC center, an institution specializing in health check-ups. The dataset includes standard 12-lead ECGs that were extracted from GE Healthcare's MUSE™ Cardiology Information System between October 2014 to May 2021. Only participants who underwent at least one check-up ECG at the SHC center were included in this dataset. The SHC dataset was also sourced from

the SCRAP and comprises 38,142 participants with 38,195 ECGs. Data integrity was ensured through the de-identification of personal details. The same exclusion criteria as the Severance dataset were applied, resulting in 37,903 participants with 37,956 ECGs. The mean age of the dataset was 47.6 (SD, 12.2), with 47.2% female. Baseline data were extracted in the same way as the Severance dataset. There was no participant overlap between the two datasets.

3. UK Biobank

The UK Biobank is one of the largest population-scale prospective cohort studies, including >500,000 participants aged 40–69 years across the United Kingdom from 2006 to 2010, and the details of the data have been described and published elsewhere.[38,39] Approximately 9.2 million individuals living within 25 miles of the 22 assessment centers in England, Wales, and Scotland were invited and 5.4% participated in the baseline assessment. The UK Biobank has collected and continues to collect extensive phenotypic and genotypic data from questionnaires, physical measures, sample assays, accelerometry, multimodal imaging, genome-wide genotyping, and longitudinal follow-up for a wide range of health-related outcomes.[38]

For the ECG data, the UK Biobank collected it as 12-lead resting ECGs during the assessment center visits. The ECG data in the UK Biobank typically contains the raw ECG waveforms along with automated interpretations of those waveforms, providing a lot of information to analyze various ECG parameters and characteristics. The ECG data is commonly provided in the "XML" file format. XML (eXtensible Markup Language) is a markup language used for storing and transporting data. Each XML file represents the ECG data for an individual participant and is linked via the participant's unique ID. No participants were excluded according to the initial exclusion criteria. The UK Biobank dataset comprises 42,791 participants with 45,610 ECGs, a mean age of 63.9 (SD, 7.8), and 51.5% females. ECG data was obtained between October 2014 to May 2021.

4. Shaoxing

The Shaoxing dataset is a publicly available dataset of 12-lead ECG signals that was created by Chapman University and Shaoxing People's Hospital (Shaoxing Hospital Zhejiang University School of Medicine).[36] The dataset consists of 10,646 participants with

12-lead ECGs, sampled at a rate of 500Hz for 10 seconds, that feature 11 common rhythms and 67 additional cardiovascular conditions labeled by experts. After excluding 448 participants based on age limits, the final ECG dataset included 10,198 participants with 10,198 ECGs, with a mean age of 60.2 (SD, 15.9) and 44.0% female. The dataset was solely used for the validation of age prediction model.

5. PTB-XL

The PTB-XL ECG dataset is a publicly available dataset comprising 21,837 clinical 12-lead ECGs from 18,885 participants, each 10 seconds in length.[37] The raw signal data was recorded by Schiller AG devices between November 1984 and June 2001 and was curated by the Physikalisch Technische Bundesanstalt (PTB). Two cardiologists assigned potentially multiple ECG statements to each record, with a total of 71 different ECG statements conforming to the SCP-ECG standard and covering diagnostic, form, and rhythm statements. After applying exclusion criteria, 18,244 participants with 21,152 ECGs, with a mean age of 60.6 (SD, 15.8) and 47.3% female were included. PTB-XL dataset was also solely used for the validation of age prediction model.

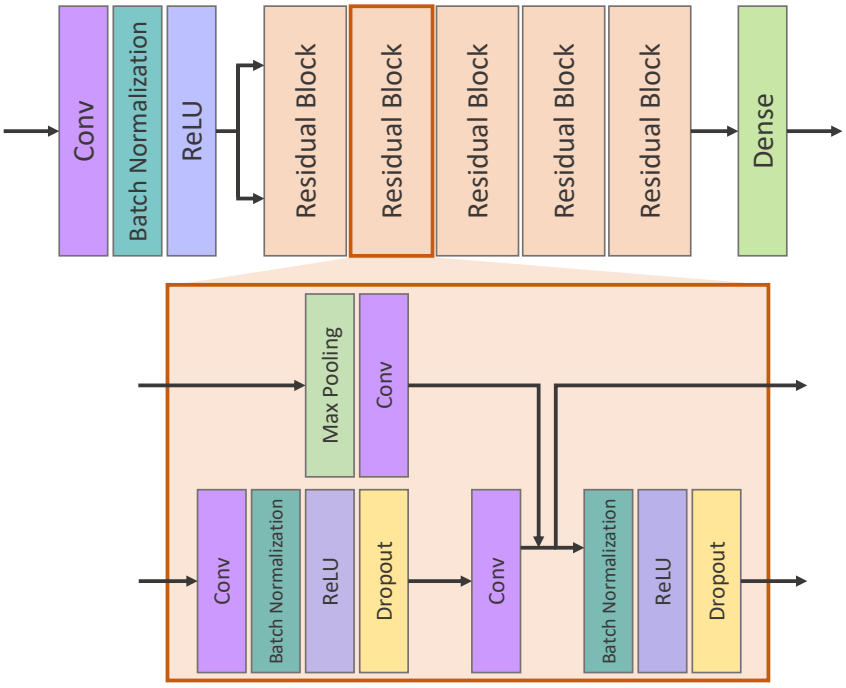
3.2. Development of the ECG-age prediction model

For model training, I used ECG data from Severance dataset. 80% of the participants (1,443,298 ECGs from 649,072 participants) were used to train the model and 5% (89,744 ECGs from 40,567 participants) were used as the validation set during training for model optimization. In case of training and validation data, maximum ECGs per participant were limited to five to prevent overfitting as some people take relatively many ECGs.

Convolutional neural network (CNN) was trained to predict age with ECG. The input of the model is an ECG waveform data, and the output is age. The unidimensional CNN with residual blocks has been shown to be effective in previous studies.[2,8] The network receives signals composed of 5,000 data points and is structured with a convolutional layer followed by five residual blocks, each containing two convolutional layers. Each residual block includes a skip connection with max pooling and a convolutional layer, enabling efficient training of sequential data.[40] Furthermore, batch normalization, ReLU activation

function, and a dropout rate of 0.5 are applied after each convolutional layer.[41,42] Figure 4 depicts the architecture of ECG-age prediction model. To address the uneven distribution of ages in the data used for model development, I applied weight correction during model training by assigning inversed weights to the input data samples based on the frequency of the given age. The model was optimized to lower the mean squared error (MSE) with weights. The Adam optimizer was used with a learning rate scheduler, with a default learning rate of 0.001 that decreased by a factor of 10 when the validation loss did not improve for 7 epochs.[43] Additionally, the training session consisted of a total of 100 epochs and was stopped if the validation loss did not improve for 30 consecutive epochs.

Figure 4. Architecture of ECG-age prediction model



The unidimensional convolutional neural network is structured with a convolutional layer followed by five residual blocks, each containing two convolutional layers. Each residual block includes a skip connection with max pooling and a convolutional layer, enabling efficient training of sequential data. Furthermore, batch normalization, ReLU

activation function, and a dropout rate of 0.5 are applied after each convolutional layer. The network receives signals composed of 5,000 data points with 8 channels.

Abbreviations: Conv, convolution; ECG, electrocardiogram.

To assess the model performance across five datasets, MSE and mean absolute error (MAE), along with their corresponding standard deviations (SD) were utilized. Further, Pearson's correlation coefficient (r) was used to assess the correlation between chronological age and the ECG-age.

I conducted an additional analysis to compare the ECG-age prediction performance with Lima's model, another known ECG-age prediction model (provided by Lima et al. in their 2021 publication).[8] Lima's model, a state-of-the-art ECG-age prediction model, was chosen for comparison due to its public availability and its extensive use of a large ECG dataset for training. This model, developed using a CNN on 12-lead ECGs, highlighted an association between the ECG-age gap (the difference between the ECG-age and chronological age) and mortality rates. Trained utilizing 1.5 million ECGs predominantly from Brazil, it achieved an MAE of 8.38 ± 7.00 in ECG-age prediction. While the model architectures of this study and Lima's model were similar, they differed in ECG sampling rates: 500Hz in my model and 400Hz in Lima's model. To compare the performance between the two models, I down sampled validation ECG datasets to 400Hz using the Fourier method.

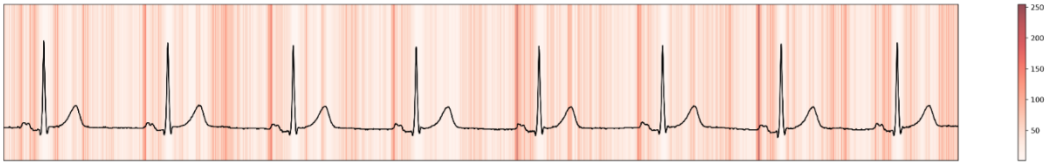
3.3. Saliency map and segmentation of ECG

Saliency maps were generated to visualize the regions of the ECG that are of relative importance for ECG-age prediction, using Severance hold-out and SHC sets. Following the ECG-age prediction, I computed the derivative of the predicted ECG-age with respect to each ECG point using a single back-propagation step.[8,44] Then a composite saliency map was created by averaging the saliency maps from each lead. The saliency values were visualized using a red colormap, where deeper shades of red indicate greater influence on ECG-age prediction. Furthermore, to identify which ECG segment received the most attention during age prediction, each ECG was divided into four segments (PR, QRS, ST,

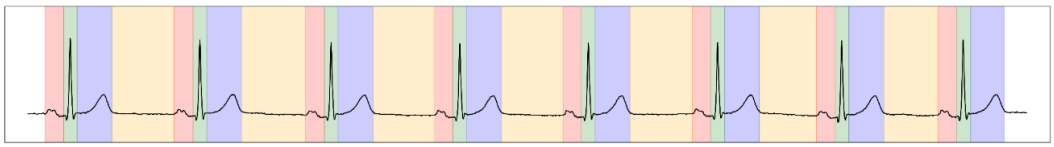
and TP) using the NeuroKit2 Python library.[45] Subsequently, the average saliency values for each segment were calculated across all ECGs. Cases where segments were not detected were excluded from the analysis. An illustrative example of a saliency map with segmentation superimposed on an ECG lead II is shown in Figure 5.

Figure 5. Illustrative example of saliency map and ECG segmentation

(A) Saliency map



(B) ECG segmentation



(A) Saliency map superimposed on ECG lead II. The gradient bar of dark red shades represents the saliency values, which indicate the degree of influence on age prediction. Deeper and darker red shades indicate more focused ECG regions with higher saliency values. (B) ECG segmentation superimposed on ECG lead II. PR, QRS, ST, and TP segments are represented by pink, green, blue, and yellow colors, respectively. Average saliency values for each segment were calculated.

Abbreviations: ECG, electrocardiogram.

3.4. Study cohorts and study design

For a longitudinal analysis of AF risk, three (Severance hold-out, SHC, and UK Biobank) datasets were included as the main study cohorts. These datasets were included

as they provided longitudinal follow-up data with participant information including demographics, comorbidities, clinical diagnoses, anthropometric measurements, and health behaviors. While up to five or all ECGs were used for model development or validation phase, only the first single ECG per participant was used for this phase.

Participants with AF history, less than a day of follow-up, or outcome occurrences during the 30-day blanking period were excluded. After the exclusion, 111,483, 37,517, and 40,973 individuals in the three cohorts were included, respectively. The blanking period was applied to exclude potential abnormal AF diagnoses around the time of the first ECG acquisition (index date).

3.5. Study variables and outcome assessment

The ECG-age was estimated by the ECG-age prediction model using the first acquired ECG in three cohorts. Then, the ECG-age gap (ECG-age - Chronological age) was calculated. Participants were classified into two groups by the threshold of ECG-age gap +7 years: ECG-aging group (age gap ≥+7 years) and Normal group (age gap <+7 years, including cases with the ECG-age younger than chronological age).[26,29,46] The cutoff +7 approximates the mean of MAEs (6.816) in validation sets.

Demographics, anthropometric measurements, comorbidities, smoking, and drinking status were extracted for baseline comparisons and risk assessment based on the index date. The study outcomes of interest were new-onset AF. New-onset AF was defined as the first diagnosis of AF after the index date. Diagnoses of comorbidities and outcomes were extracted using the International Classification of Diseases, 10th Revision (ICD-10) codes, based on more than one inpatient or two outpatient records to ensure accuracy. In the UK Biobank, self-reported non-cancer illness codes were also used, and the ICD-10 codes included converted 3-character ICD-10 codes from ICD-9-based diagnoses. Definitions and ICD-10 codes used for diagnoses are in Table 1.

Table 1. Definitions and ICD-10 codes for comorbidities and study outcomes

	Severance hold-out & SHC		UK Biobank	
Comorbidities	**Definitions**	**ICD-10 codes or conditions**	**Definitions**	**ICD-10 codes or conditions**
Atrial fibrillation	Defined from diagnosis*	ICD-10: I48	Defined from UK Biobank self-report or diagnosis*	Self-reported non-cancer illness code: 1471, 1483 ICD-10: I48
Hypertension	Defined from diagnosis*	ICD-10: I10, I11, I12, I13, I15	Defined from UK Biobank self-report or diagnosis*	Self-reported non-cancer illness code: 1065, 1072 ICD-10: I10, I11, I12, I13, I15
Diabetes mellitus	Defined from diagnosis*	ICD-10: E10, E11, E12, E13, E14	Defined from UK Biobank self-report or diagnosis*	Self-reported non-cancer illness code: 1220, 1222, 1223, 1521 ICD-10: E10, E11, E12, E13, E14

Dyslipidemia	Defined from diagnosis*	ICD-10: E78	Defined from UK Biobank self-report or diagnosis*	Self-reported non-cancer illness code: 1473 ICD-10: E78
Chronic kidney disease	Defined from eGFR (if laboratory value was not available, diagnosis code was used)	eGFR <60 mL/min per 1.73 m^2 ICD-10: N18, N19	Defined from eGFR (if laboratory value was not available, diagnosis code was used)	eGFR <60 mL/min per 1.73 m^2 Self-reported non-cancer illness code:1192, 1194 ICD-10: N18, N19
Previous myocardial infarction	Defined from diagnosis*	ICD-10: I21, I22, I25.2	Defined from UK Biobank self-report or diagnosis*	Self-reported non-cancer illness code: 1075 ICD-10: I21, I22, I25.2
Heart failure	Defined from diagnosis*	ICD-10: I11.0, I50, I97.1	Defined from UK Biobank self-report or diagnosis*	Self-reported non-cancer illness code: 1076 ICD-10: I11.0, I50, I97.1

Peripheral arterial disease	Defined from diagnosis*	ICD-10: I70, I71	Defined from UK Biobank self-report or diagnosis*	Self-reported non-cancer illness code: 1067, 1087 ICD-10: I70, I71
Ischemic stroke	Defined from diagnosis*	ICD-10: I63, I64	Defined from UK Biobank self-report or diagnosis*	Self-reported non-cancer illness code: 1583 ICD-10: I63, I64
Hemorrhagic stroke	Defined from diagnosis*	ICD-10: I60, I61, I62	Defined from UK Biobank self-report or diagnosis*	Self-reported non-cancer illness code: 1086, 1491 ICD-10: I60, I61, I62

Study outcome	**Definition**	**ICD-10 code or condition**	**Definition**	**ICD-10 code or condition**
New-onset atrial fibrillation	Defined from diagnosis* or related death without previous insurance claim for AF	ICD-10: I48	Defined from diagnosis* or related death without previous history of AF	ICD-10: I48

In the UK Biobank, the ICD-10 codes included converted 3-character ICD-10 codes from ICD-9-based diagnoses, using data-coding 1836.

* Diagnoses were established based on more than one hospital-inpatient or two outpatient records (equivalent to primary care records in the UK) corresponding to ICD-10 codes in the database to ensure accuracy.

Abbreviations: AF, atrial fibrillation; eGFR, estimated glomerular filtration rate; ICD-10, international classification of diseases 10th revision; SHC, severance health check-up.

Participants were followed until their first AF diagnosis, death, or the last follow-up date whichever happened first from the index date. For the Severance hold-out and SHC, follow-up was made until December 31, 2022 and the UK Biobank was followed until December 31, 2020.

3.6. Statistical analyses

The clinical characteristics were summarized by representing continuous variables as mean with SD, while categorical variables were expressed as numbers with corresponding percentages. For comparing categorical variables, Pearson's chi-squared test or Fischer's exact test was used, and for continuous variables, the one-way analysis of variance or Kruskal–Wallis test was used. To impute the body mass index (BMI), smoking status, and drinking status missing data, the "multivariate imputation by chained equations" R package was used to minimize bias.[47]

During the AF risk assessment phase, the relationship between the ECG-age and the risk of new-onset AF was investigated, stratified by the ECG-age groups and increasing ECG-age gap. Adjusted event rates were reported as events per 1,000 person-years and adjusted cumulative incidences of new-onset AF were graphically presented using the age- and sex-adjusted Cox regression models. Hazard ratios (HR) for new-onset AF were estimated with Cox proportional hazards models with 95% confidence intervals (CI).

To account for confounding factors, two levels of Cox proportional hazards model adjustment were conducted: [Model 1] adjusted for chronological age and sex and [Model 2] adjusted for age, sex, BMI, comorbidities (hypertension, diabetes, dyslipidemia, chronic

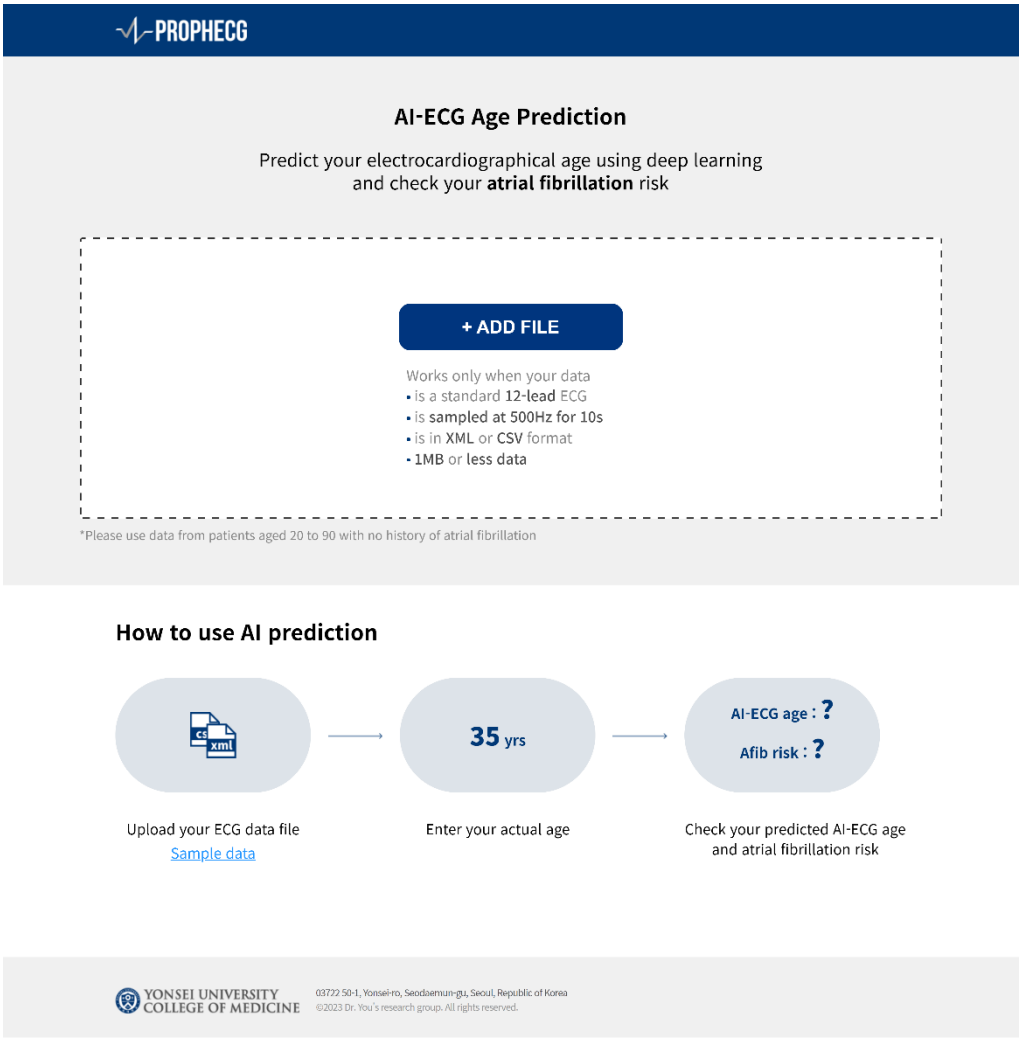
kidney disease, prior myocardial infarction, heart failure, peripheral arterial disease, and prior stroke), smoking, and drinking status.

In addition, to ascertain if the ECG-de-aging group had a reduced incidence and risk of AF compared to the Normal group, I re-categorized participants into three electrocardiographic aging (ECG-aging) groups: ECG-de-aging group (age gap <-7 years, representing decelerated aging), Normal group (age gap between -7 and +7 years), and ECG-aging group (age gap ≥+7 years, representing accelerated aging). Then the cumulative incidence and risk of new-onset AF in the three study cohorts were re-calculated.

All analyses were performed using Python version 3.9.13 (Python Software Foundation, http://www.python.org) and R version 4.2.3 (The R Foundation, www.R-project.org), and $P < 0.05$ was considered statistically significant.

The deep learning algorithm used in this study including the trained weights is accessible at the GitHub repository: https://github.com/dr-you-group/PROPHECG-Age. ECG-age prediction service using the ECG-age prediction model is openly available at the website: https://www.prophecg.com (Figure 6).

Figure 6. Screenshot of an online ECG-age prediction service



The ECG-Age prediction service provides age predictions from XML or CSV formatted raw ECG waveforms.

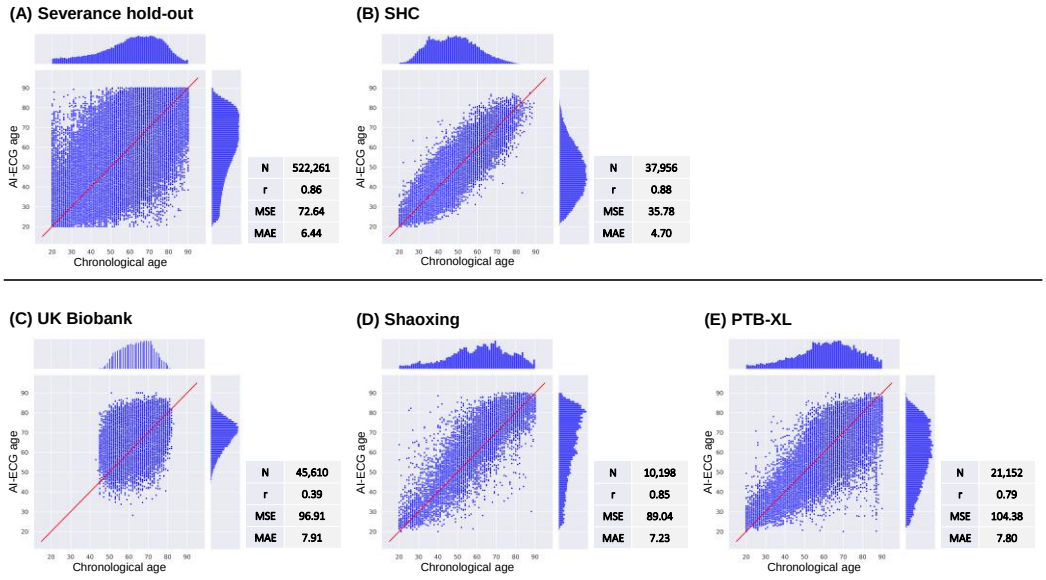
Abbreviations: AI, artificial intelligence; ECG, electrocardiogram.

4. Results

4.1. ECG-age prediction model performances and saliency analyses

For the ECG-age prediction task, the model exhibited MAEs of 6.44 (SD, 5.58), 4.70 (3.70), 7.91 (5.86), 7.23 (6.06), and 7.80 (6.59) for the Severance hold-out, SHC, UK Biobank, Shaoxing, and PTB-XL datasets, respectively. These values indicate the average deviation of the ECG-age from actual age within each dataset. Scatter plots in Figure 7 visually depict the correlation between ECG-age and chronological age, accompanied by MSE/MAE/Pearson's correlation coefficient results for each validation dataset. In contrast, Lima's model, a state-of-the-art ECG-age prediction model, yielded MAEs of 10.44 (SD, 8.55), 8.26 (6.45), 15.23 (8.34), 10.31 (8.52), and 9.11 (7.20) for the same datasets (Table 2).[12] My model consistently outperformed Lima's across all datasets.

Figure 7. Scatter plots presenting the relationship between the ECG-age and chronological age along with the model performance of ECG-age prediction



The red line represents the identity line. The lateral histograms show the distributions of predicted AI-ECG age and chronological age among participants of each cohort.

Abbreviations: AI, artificial intelligence; ECG, electrocardiogram; MAE, mean absolute error; MSE, mean squared error; N, numbers; SHC, Severance health check-up.

Table 2. Age prediction performance comparison between the developed model and SOTA model

Validation dataset	N	r		MSE		MAE	
		Lima*	Developed	Lima*	Developed	Lima*	Developed
Severance hold-out	522261	0.70	0.86	181.99	72.64	10.44	6.44
SHC	37956	0.70	0.88	109.87	35.78	8.26	4.70
UK Biobank	45610	0.17	0.39	301.43	96.91	15.23	7.91
Shaoxing	10198	0.70	0.85	178.86	89.04	10.31	7.23
PTB-XL	21152	0.75	0.79	134.98	104.38	9.11	7.80

N represents the number of ECGs of participants in the model validation phase (described in Figure 2).

* Lima's model is an age-prediction convolutional neural network based on raw 12-lead ECGs introduced by Lima et al.[8]

Abbreviations: SOTA, state-of-the-art; ECG, electrocardiogram; MAE, mean absolute error; MSE, mean squared error; N, numbers of ECGs; SHC, Severance health check-up.

The results of per-segment saliency calculations are summarized in Table 3, presenting means with SD, medians with interquartile range, and minimum-maximum values. The PR segment emerged as the most salient regions on the ECG, with the highest saliency values observed for both the Severance hold-out and SHC datasets, with mean values of 58.13 (SD, 13.73) and 59.32 (13.30), respectively.

Table 3. Results of focused ECG area in age prediction using saliency map and ECG segmentation

ECG Segment	Severance hold-out (N=512083)	SHC (N=37848)	UK Biobank (N=45563)	Shaoxing (N=10276)	PTB-XL (N=21566)
PR					
Mean (SD)	58.13 (13.73)	59.32 (13.30)	72.42 (13.43)	61.25 (13.98)	63.08 (14.46)
Median (IQR)	57.45 (48.45–67.18)	58.18 (49.07–68.37)	72.51 (63.40–81.56)	60.87 (51.54–70.50)	62.67 (52.99–73.00)
Minimum–Maximum	3.84–141.29	10.06–126.05	5.62–128.28	14.55–128.23	13.92–124.29
QRS					
Mean (SD)	26.31 (10.02)	26.24 (10.85)	49.50 (13.05)	30.29 (10.31)	29.82 (10.63)
Median (IQR)	25.14 (19.14–32.15)	24.40 (18.32–32.44)	48.69 (40.05–58.16)	29.18 (22.82–36.44)	28.86 (22.16–36.42)
Minimum–Maximum	1.78–161.94	3.21–103.15	7.97–108.06	3.55–80.22	3.98–133.82
ST					
Mean (SD)	36.04 (11.54)	34.04 (11.45)	49.92 (12.37)	38.06 (12.01)	35.64 (11.47)
Median (IQR)	35.36 (27.74–43.56)	32.93 (25.35–41.81)	49.34 (41.13–58.16)	37.23 (29.22–46.04)	34.94 (27.31–43.25)
Minimum–Maximum	1.83–137.42	5.58–90.30	9.22–123.11	5.19–123.77	3.91–103.68
TP					
Mean (SD)	51.70 (13.53)	50.57 (14.79)	44.35 (10.63)	49.11 (13.17)	47.08 (13.04)
Median (IQR)	51.71 (42.07–61.21)	49.84 (39.09–61.38)	43.77 (36.76–51.41)	48.96 (39.55–58.56)	46.65 (37.47–56.05)
Minimum–Maximum	5.19–143.37	11.87–103.47	5.72–88.07	12.80–103.17	10.34–102.44

N represents the number of ECGs of participants in the model validation phase (described in Figure 2). Cases with undetected ECG segments (10,178 in Severance hold-out dataset and 108 in the SHC dataset) were excluded from the analysis. The mean and

median values of each ECG segment, along with their corresponding SD and IQR are displayed.

Abbreviations: ECG, electrocardiogram; N, numbers of ECGs; SD, standard deviation; IQR, interquartile range; SHC, severance health check-up.

4.2. Clinical characteristics and comparison of participants

The clinical characteristics of participants across the three study cohorts are outlined in Table 4. The mean chronological age was 51.9 (16.4), 47.5 (12.2), and 63.7 (7.8) years in the Severance hold-out, SHC, and UK Biobank, respectively, with women comprising 54.1%, 47.5%, and 47.9% of the cohorts. Notably, the UK Biobank cohort, aged 43–83, exhibited a higher mean ECG-age of 68.4 (7.8) and a larger mean ECG-age gap of 4.7 (8.7) compared to the other two cohorts with age range 20-90. Beyond racial differences, variations in all clinical characteristics, including BMI, blood pressure, comorbidities, and health behaviors, were observed among the cohorts.

Table 4. Clinical characteristics of study cohorts

	Severance hold-out (N = 111483)	SHC (N = 37517)	UK Biobank (N = 40973)
Chronological age (years), mean ± SD	51.9 ± 16.4	47.5 ± 12.2	63.7 ± 7.8
ECG-age (years), mean ± SD	51.9 ± 16.2	47.4 ± 12.5	68.4 ± 7.8
ECG-age gap (years), mean ± SD	0.0 ± 6.8	-0.1 ± 6.0	4.7 ± 8.7
Sex			
Male, N (%)	51213 (45.9)	19715 (52.5)	21364 (52.1)
Female, N (%)	60270 (54.1)	17802 (47.5)	19609 (47.9)
Race†, N	111483	37517	40861
White, N (%)	0 (0.0)	0 (0.0)	39538 (96.8)
Asian, N (%)	111483 (100.0)	37517 (100.0)	467 (1.1)
Black, N (%)	0 (0.0)	0 (0.0)	303 (0.7)
Mixed, N (%)	0 (0.0)	0 (0.0)	202 (0.5)
Others, N (%)	0 (0.0)	0 (0.0)	351 (0.9)
BMI, mean ± SD	23.6 ± 3.6	23.6 ± 3.4	26.6 ± 4.2
Systolic BP, mean ± SD	125.5 ± 17.8	122.5 ± 15.3	135.2 ± 17.8
Diastolic BP, mean ± SD	76.3 ± 11.9	76.0 ± 10.8	81.5 ± 9.9
Comorbidities			
Hypertension, N (%)	12898 (11.6)	2872 (7.7)	8059 (19.7)
Diabetes, N (%)	8696 (7.8)	1450 (3.9)	1066 (2.6)
Dyslipidemia, N (%)	5197 (4.7)	2258 (6.0)	4143 (10.1)
Chronic kidney disease, N (%)	1737 (1.6)	224 (0.6)	230 (0.6)
Previous myocardial infarction, N (%)	1066 (1.0)	137 (0.4)	641 (1.6)
Heart failure, N (%)	1191 (1.1)	107 (0.3)	792 (1.9)
Peripheral arterial disease, N (%)	934 (0.8)	217 (0.6)	82 (0.2)
Previous stroke, N (%)	2910 (2.6)	277 (0.7)	220 (0.5)
Smoking status‡, N	12678	2468	40867

Never smoked, N (%)	8498 (67.0)	1839 (74.5)	24770 (60.6)
Ex-smoking, N (%)	2357 (18.6)	364 (14.7)	13476 (33.0)
Current smoking, N (%)	1823 (14.4)	265 (10.7)	2621 (6.4)
Drinking status‡, N	12682	2467	40946
Lifetime abstinence, N (%)	7307 (57.6)	1422 (57.6)	1052 (2.6)
Former drinking, N (%)	2002 (15.8)	168 (6.8)	871 (2.1)
Currently drinking, N (%)	3373 (26.6)	877 (35.5)	39023 (95.3)
Follow-up time (years), mean ± SD	4.14 ± 4.27	6.08 ± 3.81	2.99 ± 1.56

Continuous variables are presented as mean ± standard deviation and categorical variables are presented as number (percentage). The participant counts in each study cohort reflect numbers after the dataset filtering and exclusion process for observational AF risk analysis, as outlined by Exclusion 1 and 2 in Figure 2.

† N for racial data indicates the number of participants with confirmed racial information, excluding those marked as "Unknown or Prefer not to answer" in each cohort.

‡ Smoking and drinking status denote numbers and percentages within the extent of available information that participants responded to in that questionnaire (N represents the number of participants who answered the questions in the questionnaire).

Abbreviations: AF, atrial fibrillation; BMI, body mass index; BP, blood pressure; ECG, electrocardiogram; N, numbers of participants; SD, standard deviation; SHC, Severance health check-up.

Tables 5, 6, and 7 provide comparisons of baseline characteristics between the Normal and ECG-aging groups within each cohort, as well as among those excluded due to a history of AF. Across all four cohorts, the ECG-aging group tended to be younger, with a higher ECG-age and a greater proportion of men, in contrast to the Normal group. Furthermore, excluded participants with a history of AF tended to be older, predominantly male, and presented with a higher burden of comorbidities compared to study-included participants.

Table 5. Comparison of clinical characteristics between the study-included and the excluded participants in the Severance hold-out cohort

	Study-included participants		Excluded participants	P-value*
	Normal (age gap <+7)	ECG-aging (age gap ≥+7)		
No. of participants	96418	15065	2468	
Chronological age (years), mean ± SD	52.4 ± 16.5	48.4 ± 15.7	63.3 ± 12.3	<0.001
ECG-age (years), mean ± SD	50.7 ± 16.0	59.3 ± 16.0	65.0 ± 12.8	<0.001
ECG-age gap (years), mean ± SD	-1.7 ± 5.5	10.9 ± 4.0	1.6 ± 8.5	<0.001
Sex				<0.001
Male, N (%)	43780 (45.4)	7433 (49.3)	1540 (62.4)	
Female, N (%)	52638 (54.6)	7632 (50.7)	928 (37.6)	
Race†, N	96418	15065	2468	1.00
Asian, N (%)	96418 (100.0)	15065 (100.0)	2468 (100.0)	
BMI, mean ± SD	23.5 ± 3.5	23.8 ± 3.9	24.2 ± 3.6	<0.001
Systolic BP, mean ± SD	125.5 ± 17.6	126.0 ± 19.1	123.9 ± 20.0	0.002
Diastolic BP, mean ± SD	76.2 ± 11.7	77.3 ± 12.8	74.7 ± 12.5	<0.001
Comorbidities				
Hypertension, N (%)	10812 (11.2)	2086 (13.8)	794 (32.2)	<0.001
Diabetes, N (%)	7329 (7.6)	1367 (9.1)	343 (13.9)	<0.001
Dyslipidemia, N (%)	4440 (4.6)	757 (5.0)	207 (8.4)	<0.001
Chronic kidney disease, N (%)	1294 (1.3)	443 (2.9)	52 (2.1)	<0.001
Previous myocardial infarction, N (%)	817 (0.8)	249 (1.7)	59 (2.4)	<0.001
Heart failure, N (%)	945 (1.0)	246 (1.6)	329 (13.3)	<0.001
Peripheral arterial disease, N (%)	782 (0.8)	152 (1.0)	30 (1.2)	0.006
Previous stroke, N (%)	2363 (2.5)	547 (3.6)	215 (8.7)	<0.001
Smoking status‡, N	10691	1987	202	<0.001
Never smoked, N (%)	7321 (68.5)	1177 (59.2)	138 (68.3)	
Ex-smoking, N (%)	1927 (18.0)	430 (21.6)	43 (21.3)	
Currently smoking, N (%)	1443 (13.5)	380 (19.1)	21 (10.4)	
Drinking status‡, N	10695	1987	202	<0.001
Lifetime abstinence, N	6264 (58.6)	1043 (52.5)	122 (60.4)	

(%)			
Former drinking, N (%)	1622 (15.2)	380 (19.1)	34 (16.8)
Currently drinking, N (%)	2809 (26.3)	564 (28.4)	46 (22.8)

Continuous variables are presented as mean ± standard deviation and categorical variables are presented as number (percentage).

* P-values for comparing three groups were derived from Pearson's chi-squared test or Fisher's exact test for categorical variables, and the one-way analysis of variance or Kruskal–Wallis test for continuous variables.

† N for racial data indicates the number of participants with confirmed racial information, excluding those marked as "Unknown or Prefer not to answer".

‡ Smoking and drinking status denote numbers and percentages within the extent of available information that participants responded to in that questionnaire (N represents the number of participants who answered the questions in the questionnaire).

Abbreviations: BMI, body mass index; BP, blood pressure; ECG, electrocardiogram; N, numbers of participants; SD, standard deviation.

Table 6. Comparison of clinical characteristics between the study-included and the excluded participants in the SHC cohort

	Study-included participants		Excluded participants	P-value*
	Normal (age gap <+7)	ECG-aging (age gap ≥+7)		
No. of participants	33254	4263	310	
Chronological age (years), mean ± SD	47.8 ± 12.2	44.9 ± 11.8	60.8 ± 11.0	<0.001
ECG-age (years), mean ± SD	46.5 ± 12.2	55.0 ± 12.1	61.4 ± 11.9	<0.001
ECG-age gap (years), mean ± SD	-1.4 ± 4.9	10.1 ± 3.0	0.6 ± 7.6	<0.001
Sex				<0.001
Male, N (%)	17316 (52.1)	2399 (56.3)	233 (75.2)	
Female, N (%)	15938 (47.9)	1864 (43.7)	77 (24.8)	
Race†, N	33254	4263	310	1.00
Asian, N (%)	33254 (100.0)	4263 (100.0)	310 (100.0)	
BMI, mean ± SD	23.5 ± 3.3	24.3 ± 3.7	24.4 ± 3.2	<0.001
Systolic BP, mean ± SD	122.3 ± 15.2	123.7 ± 15.8	122.4 ± 14.9	0.028
Diastolic BP, mean ± SD	75.8 ± 10.8	77.1 ± 11.2	76.5 ± 10.7	0.004
Comorbidities				
Hypertension, N (%)	2503 (7.5)	369 (8.7)	157 (50.6)	<0.001
Diabetes, N (%)	1258 (3.8)	192 (4.5)	73 (23.5)	<0.001
Dyslipidemia, N (%)	1998 (6.0)	260 (6.1)	102 (32.9)	<0.001
Chronic kidney disease, N (%)	184 (0.6)	40 (0.9)	17 (5.5)	<0.001
Previous myocardial infarction, N (%)	114 (0.3)	23 (0.5)	17 (5.5)	<0.001
Heart failure, N (%)	92 (0.3)	15 (0.4)	44 (14.2)	<0.001
Peripheral arterial disease, N (%)	194 (0.6)	23 (0.5)	8 (2.6)	<0.001
Previous stroke, N (%)	243 (0.7)	34 (0.8)	22 (7.1)	<0.001
Smoking status‡, N	2204	264	98	0.233
Never smoked, N (%)	1654 (75.0)	185 (70.1)	73 (74.5)	
Ex-smoking, N (%)	321 (14.6)	43 (16.3)	18 (18.4)	
Currently smoking, N (%)	229 (10.4)	36 (13.6)	7 (7.1)	
Drinking status‡, N	2202	265	99	0.011

Lifetime abstinence, N (%)	1287 (58.4)	135 (50.9)	49 (49.5%)
Former drinking, N (%)	140 (6.4)	28 (10.6)	11 (11.1)
Currently drinking, N (%)	775 (35.2)	102 (38.5)	39 (39.4)

Continuous variables are presented as mean ± standard deviation and categorical variables are presented as number (percentage).

* P-values for comparing three groups were derived from Pearson's chi-squared test or Fisher's exact test for categorical variables, and the one-way analysis of variance or Kruskal–Wallis test for continuous variables.

† N for racial data indicates the number of participants with confirmed racial information, excluding those marked as "Unknown or Prefer not to answer".

‡ Smoking and drinking status denote numbers and percentages within the extent of available information that participants responded to in that questionnaire (N represents the number of participants who answered the questions in the questionnaire).

Abbreviations: BMI, body mass index; BP, blood pressure; ECG, electrocardiogram; N, numbers of participants; SD, standard deviation; SHC, severance health check-up.

Table 7. Comparison of clinical characteristics between the study-included and the excluded participants in the UK Biobank cohort

	Study-included participants		Excluded participants	P-value*
	Normal (age gap <+7)	ECG-aging (age gap ≥+7)		
No. of participants	24673	16300	1706	
Chronological age (years), mean ± SD	66.8 ± 7.1	59.0 ± 6.4	67.7 ± 7.3	<0.001
ECG-age (years), mean ± SD	66.0 ± 8.0	72.1 ± 5.9	72.0 ± 7.5	<0.001
ECG-age gap (years), mean ± SD	-0.8 ± 5.9	13.1 ± 4.6	4.2 ± 8.1	<0.001
Sex				<0.001
Male, N (%)	11370 (46.1)	8239 (50.5)	1062 (62.3)	
Female, N (%)	13303 (53.9)	8061 (49.5)	644 (37.7)	
Race†, N	24600	16261	1701	<0.001
White, N (%)	23924 (97.3)	15614 (96.0)	1673 (98.4)	
Asian, N (%)	248 (1.0)	219 (1.3)	12 (0.7)	
Black, N (%)	138 (0.6)	165 (1.0)	4 (0.2)	
Mixed, N (%)	95 (0.4)	107 (0.7)	2 (0.1)	
Others, N (%)	195 (0.8)	156 (1.0)	10 (0.6)	
BMI, mean ± SD	26.2 ± 4.0	27.1 ± 4.5	27.7 ± 4.7	<0.001
Systolic BP, mean ± SD	136.5 ± 18.1	133.3 ± 17.3	138.5 ± 19.1	<0.001
Diastolic BP, mean ± SD	81.1 ± 9.8	82.0 ± 10.1	82.3 ± 10.0	<0.001
Comorbidities				
Hypertension, N (%)	5076 (20.6)	2983 (18.3)	572 (33.5)	<0.001
Diabetes, N (%)	666 (2.7)	400 (2.5)	74 (4.3)	<0.001
Dyslipidemia, N (%)	2864 (11.6)	1279 (7.8)	332 (19.5)	<0.001
Chronic kidney disease, N (%)	157 (0.6)	73 (0.4)	31 (1.8)	<0.001
Previous myocardial infarction, N (%)	418 (1.7)	223 (1.4)	90 (5.3)	<0.001
Heart failure, N (%)	440 (1.8)	352 (2.2)	100 (5.9)	<0.001
Peripheral arterial disease, N (%)	50 (0.2)	32 (0.2)	8 (0.5)	0.060
Previous stroke, N (%)	143 (0.6)	77 (0.5)	41 (2.4)	<0.001
Smoking status‡, N	24606	16261	1704	<0.001
Never smoked, N (%)	14710 (59.8)	10060 (61.9)	926 (54.3)	

Ex-smoking, N (%)	8550 (34.7)	4926 (30.3)	686 (40.3)	
Currently smoking, N (%)	1346 (5.5)	1275 (7.8)	92 (5.4)	
Drinking status‡, N	24656	16290	1706	0.473
Lifetime abstinence, N (%)	640 (2.6)	412 (2.5)	34 (2.0)	
Former drinking, N (%)	527 (2.1)	344 (2.1)	43 (2.5)	
Currently drinking, N (%)	23489 (95.3)	15534 (95.4)	1629 (95.5)	

Continuous variables are presented as mean ± standard deviation and categorical variables are presented as number (percentage).

* P-values for comparing three groups were derived from Pearson's chi-squared test or Fisher's exact test for categorical variables, and the one-way analysis of variance or Kruskal–Wallis test for continuous variables.

† N for racial data indicates the number of participants with confirmed racial information, excluding those marked as "Unknown or Prefer not to answer".

‡ Smoking and drinking status denote numbers and percentages within the extent of available information that participants responded to in that questionnaire (N represents the number of participants who answered the questions in the questionnaire).

Abbreviations: BMI, body mass index; BP, blood pressure; ECG, electrocardiogram; N, numbers of participants; SD, standard deviation.

4.3. Association between ECG-aging and the risk of new-onset AF

During the mean follow-up periods of 4.14 (SD, 4.27), 6.08 (3.81), and 2.99 (1.56) years in the Severance hold-out, SHC, and UK Biobank cohorts, respectively, a total of 2,023, 456, and 538 new-onset AF events were recorded. My investigation aimed to assess the predictive utility of ECG-age as an indicator for AF prevention across diverse populations from different countries.

Table 8 illustrates a clear association between ECG-aging and the risk of new-onset AF. The event rates of new-onset AF varied between the Normal and ECG-aging groups (4.12 vs. 10.25 per 1,000 person-years in the Severance hold-out, 1.83 vs. 3.82 in the SHC, and 4.15 vs. 8.83 in the UK Biobank). The ECG-aging group consistently exhibited a higher risk of new-onset AF across all four cohorts, as evidenced by multivariable Cox regression analysis. Specifically, HRs for new-onset AF were 2.50 (95% CI, 2.24–2.78) in the Severance hold-out, 1.89 (1.46–2.43) in the SHC, and 1.90 (1.55–2.33) in the UK Biobank. Moreover, the ECG-age gap as a continuous variable was associated with an increased risk of new-onset AF. For each one ECG-age gap increase, the risk of new-onset AF increased by 6% in the Severance hold-out, 4% in the SHC, and 4% in the UK Biobank. Importantly, the significance of ECG-aging for the risk of AF was persistent even after further adjustment for potential confounders, in addition to age and sex in model 1. Figure 8 depicts cumulative incidence curves of AF, consistently indicating that the ECG-aging group had a higher risk of new-onset AF compared to the Normal group across all three cohorts.

Table 8. The incidence and risk of new-onset AF stratified by ECG-aging groups and the increasing ECG-age gap

Subtype group	No. of events / total No.	Event rates* (95% CI)	Adjusted HR (95% CI), Model 1†	P-value	Adjusted HR (95% CI), Model 2‡	P-value
Severance hold-out						
Normal (age gap <+7)	1567 / 96418	4.12 (3.67–4.62)	1 [ref]	[ref]	1 [ref]	[ref]
ECG-aging (age gap ≥+7)	456 / 15065	10.25 (9.75–10.77)	2.50 (2.24–2.78)	< 0.001	2.30 (2.06–2.56)	< 0.001
Per 1-increase in ECG-age gap			1.06 (1.05–1.07)	< 0.001	1.05 (1.05–1.06)	< 0.001
SHC						
Normal (age gap <+7)	383 / 33254	1.83 (1.45–2.31)	1 [ref]	[ref]	1 [ref]	[ref]
ECG-aging (age gap ≥+7)	73 / 4263	3.82 (3.44–4.24)	1.89 (1.46–2.43)	< 0.001	1.96 (1.52–2.54)	< 0.001
Per 1-increase in ECG-age gap			1.04 (1.03–1.06)	< 0.001	1.04 (1.03–1.06)	< 0.001
UK Biobank						
Normal	329 / 24673	4.15	1 [ref]	[ref]	1 [ref]	[ref]

(age gap <+7)		(3.47–4.97)				
ECG-aging (age gap ≥+7)	209 / 16300	8.83 (7.92–9.85)	1.90 (1.55–2.33)	< 0.001	1.76 (1.43–2.16)	< 0.001
Per 1-increase in ECG-age gap			1.04 (1.03–1.06)	< 0.001	1.04 (1.02–1.05)	< 0.001

* The event rates were adjusted for chronological age and sex and presented per 1,000 person-years.

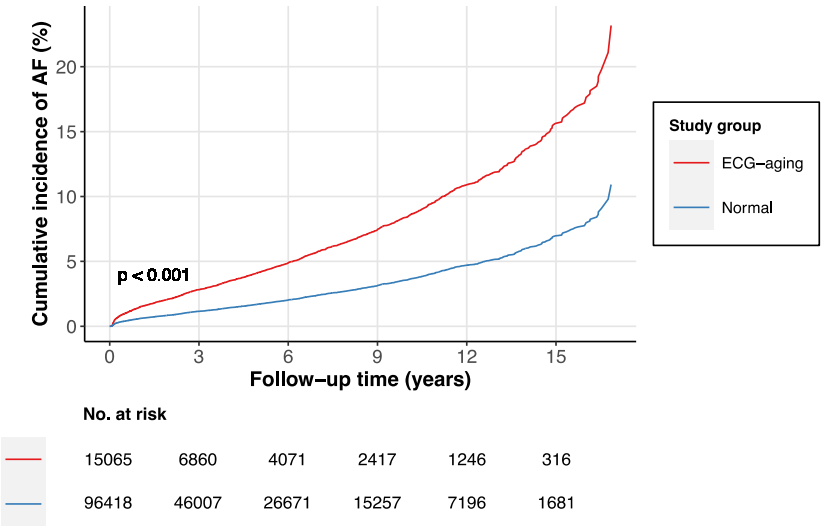
† Model 1 was adjusted for chronological age and sex.

‡ Model 2 was adjusted for chronological age, sex, BMI, comorbidities (hypertension, diabetes, dyslipidemia, chronic kidney disease, prior myocardial infarction, heart failure, peripheral arterial disease, and prior stroke), smoking, and drinking status.

Abbreviations: AF, atrial fibrillation; CI, confidence interval; ECG, electrocardiogram; ECG-aging, electrocardiographic aging; HR, hazard ratio; SHC, Severance health check-up.

Figure 8. Adjusted cumulative incidence curves of new-onset AF stratified by the ECG-aging group

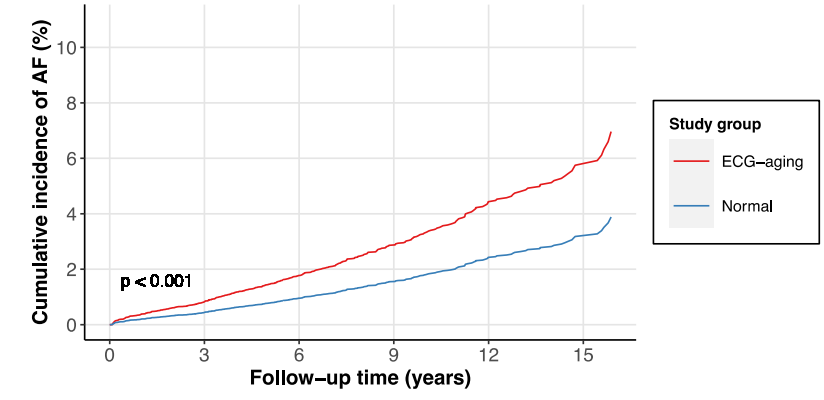
(A) Severance hold-out



No. at risk

ECG−aging	15065	6860	4071	2417	1246	316
Normal	96418	46007	26671	15257	7196	1681

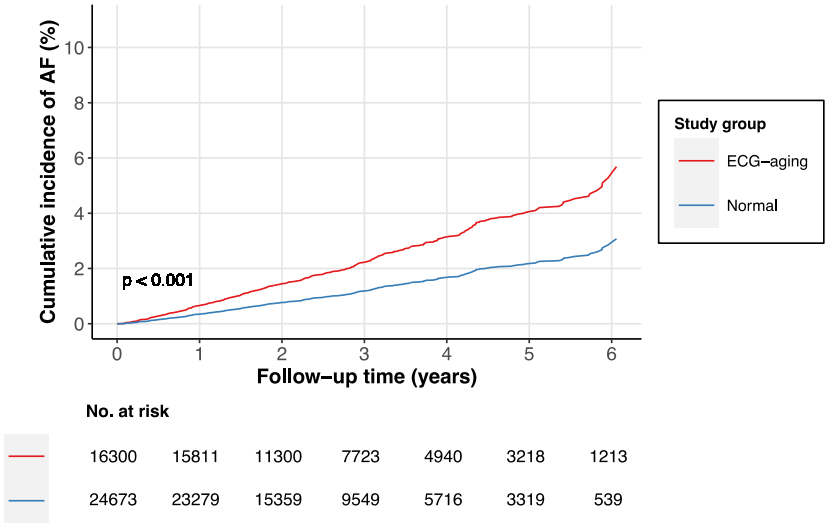
(B) SHC



No. at risk

ECG−aging	4263	3284	2000	973	405	111
Normal	33254	25528	14889	7159	2875	827

(C) UK Biobank



The incidence curves and corresponding p-values for comparing the study groups were derived from Cox regression models adjusted for chronological age and sex, and the number-at-risk table depicting the population at risk and cumulative risk over time is presented below each plot.

Abbreviations: AF, atrial fibrillation; ECG-aging, electrocardiographic aging; SHC, Severance health check-up.

To investigate the association of ECG-de-aging and AF, participants were regrouped into three ECG-aging subtypes, including the ECG-de-aging group (age gap <-7 years, indicating decelerated aging). The ECG-de-aging group showed lower event rates and risks of new-onset AF compared to the Normal-aging group except in the UK Biobank (Table 9 and Figure 9).

Table 9. The incidence and risk of new-onset AF stratified by the three subtype groups of ECG-aging

Subtype group	No. of events / total No.	Event rates* (95% CI)	Adjusted HR (95% CI), Model 1†	P-value	Adjusted HR (95% CI), Model 2‡	P-value
Severance hold-out						
ECG-de-aging (age gap <-7)	306 / 15486	3.79 (3.59–4.01)	0.85 (0.75–0.96)	0.011	0.88 (0.77–0.99)	0.041
Normal-aging (-7≤ age gap <+7)	1261 / 80932	4.35 (3.86–4.91)	1 [ref]	[ref]	1 [ref]	[ref]
ECG-aging (age gap ≥+7)	456 / 15065	10.25 (9.16–11.47)	2.42 (2.17–2.70)	< 0.001	2.24 (2.01–2.50)	< 0.001
SHC						
ECG-de-aging (age gap <-7)	58 / 4415	1.58 (1.41–1.77)	0.72 (0.54–0.95)	0.022	0.76 (0.57–1.02)	0.06
Normal-aging (-7≤ age gap <+7)	325 / 28839	1.95 (1.53–2.47)	1 [ref]	[ref]	1 [ref]	[ref]
ECG-aging (age gap ≥+7)	73 / 4263	3.82 (2.94–4.96)	1.79 (1.39–2.31)	< 0.001	1.88 (1.45–2.44)	< 0.001
UK Biobank						
ECG-de-aging (age gap <-7)	49 / 3622	3.83 (3.40–4.33)	0.82 (0.60–1.11)	0.20	0.87 (0.64–1.18)	0.36

Normal-aging (-7≤ age gap <+7)	280 / 21051	4.26 (3.55–5.11)	1 [ref]	[ref]	1 [ref]	[ref]
ECG-aging (age gap ≥+7)	209 / 16300	8.84 (6.65–11.74)	1.85 (1.51–2.28)	< 0.001	1.73 (1.41–2.13)	< 0.001

* The event rates were adjusted for chronological age and sex and presented per 1,000 person-years.

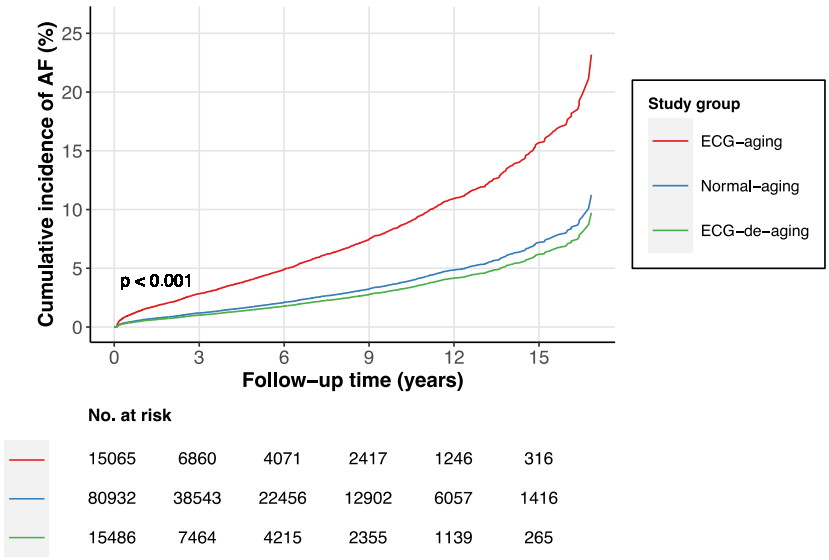
† Model 1 was adjusted for chronological age and sex.

‡ Model 2 was adjusted for chronological age, sex, BMI, comorbidities (hypertension, diabetes, dyslipidemia, chronic kidney disease, prior myocardial infarction, heart failure, peripheral arterial disease, and prior stroke), smoking, and drinking status.

Abbreviations: AF, atrial fibrillation; BMI, body mass index; CI, confidence interval; ECG, electrocardiogram; ECG-aging, electrocardiographic aging; HR, hazard ratio; Severance health check-up.

Figure 9. Adjusted cumulative incidence curves of new-onset AF stratified by the three subtype groups of ECG-aging

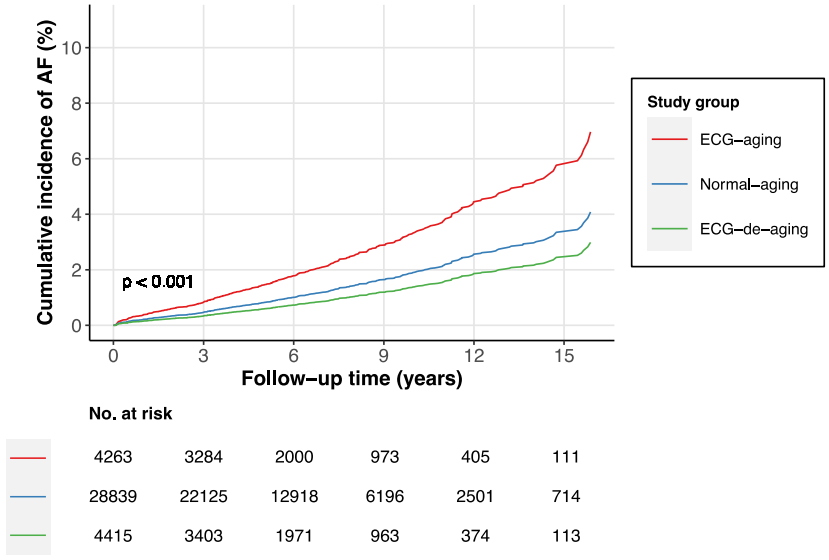
(A) Severance hold-out



No. at risk

ECG−aging	15065	6860	4071	2417	1246	316
Normal−aging	80932	38543	22456	12902	6057	1416
ECG−de−aging	15486	7464	4215	2355	1139	265

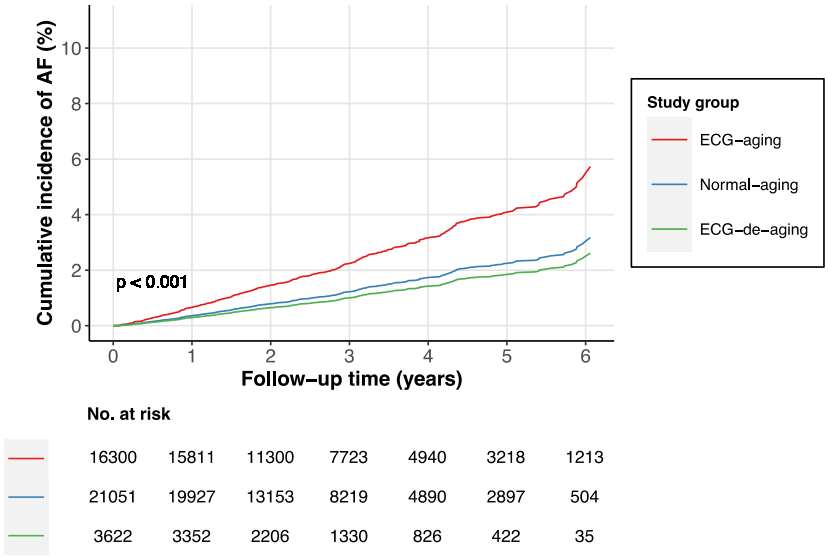
(B) SHC



No. at risk

ECG−aging	4263	3284	2000	973	405	111
Normal−aging	28839	22125	12918	6196	2501	714
ECG−de−aging	4415	3403	1971	963	374	113

(C) UK Biobank



The incidence curves and corresponding p-values for comparing the study groups were derived from Cox regression models adjusted for chronological age and sex, and the number-at-risk table depicting the population at risk and cumulative risk over time is presented below each plot.

Abbreviations: AF, atrial fibrillation; ECG-aging, electrocardiographic aging; SHC, Severance health check-up.

5. Discussion

In this large-scale, multi-national cohort study, I constructed an AI model to predict age using over 1.5 million 12-lead ECGs, achieving accuracy comparable to, and further superior to that of previous studies. The findings in this study demonstrated that individuals with “electrophysiologically older” hearts, defined as having a higher AI-predicted ECG age than their chronological age, have an increased risk of new-onset AF events, even after adjusting for known confounding covariates. Importantly, this model functioned effectively across racially and ethnically diverse populations. Further, additional analysis confirmed a decreased risk of new-onset AF in ECG-de-aging group compared.

Aging induces electrophysiological and electroanatomic alterations in the heart, which manifest in the ECG and are correlated with cardiac pathologies.[46] These age-related changes, identified by AI-ECG, likely signify unfavorable "atrial remodeling."[48] In this context, AF might be considered the heart's wrinkle, serving as a marker of the aging trajectory. Utilizing AI-ECG facilitates the early detection of this aging-related process, potentially prior to the onset of complications such as irreversible atrial myopathy, heart failure, or stroke. Nevertheless, further investigation is warranted to validate this hypothesis.

Previous research aimed to estimate heart age from ECGs by utilizing predefined variables based on P-wave, QRS complex, RR interval, and T-wave characteristics extracted from standard ECGs, albeit with limited efficacy.[22,49] AI analysis, on the other hand, leverages the entire ECG signal, potentially including unnamed signal segments. Moreover, employing nonlinear functions within neural networks enables the reflection of multiple, simultaneous, and sequential nonlinear changes in the signal, which may stem from the complex physiology of aging. Given the nonlinear and incompletely characterized nature of many biological processes, this could elucidate the superior performance of DNNs in estimating heart age compared to human-selected features. The heart age predicted by DNNs may offer more comprehensive insights into the heart's aging status than chronological age alone, facilitating the estimation and prediction of disease risk and burden.[1,29,46] Notably, an elevation in ECG-age beyond chronological age, denoted by an increased ECG-age gap, heightens the risk of cardiac diseases and cardiovascular mortality.

[8,29,46] In this study, I utilized AI to interpret standard ECGs and endeavored to estimate heart age, demonstrating that when AI-derived heart age surpasses chronological age, it furnishes valuable predictive information for future risk assessment and early stratification of AF risk. These findings underscore the notion that ECG-aging, analyzed using AI algorithms, reflects cardiac aging, given AF's association with aging, and provides more profound prognostic insights into latent cardiovascular factors than chronological age alone, particularly in forecasting future AF risk.

This research sought to elucidate the role of ECG-aging as a surrogate marker of cardiac aging for cardiovascular risk stratification, rather than merely a measure of overall aging. My overarching objective is to develop a practical and user-friendly heart age prediction model, capable of estimating an individual's cardiac age from ECGs. By identifying ECG signals associated with the aging of the cardiovascular system, the aim is to enhance cardiovascular risk prediction. This study represents an initial step towards the practical application of ECGs in the prognostic assessment of heart age, paving the way for improved clinical decision-making and patient care.

Obtaining ECGs has become increasingly convenient and cost-effective.[50] As ECGs can now be readily acquired, integrating heart age estimation into routine clinical examinations holds tremendous potential. By leveraging AI algorithms for heart age prediction, ECGs could emerge as a potent tool for assessing cardiovascular health. This innovative approach has the potential to motivate individuals to embrace healthier lifestyles and improve their cardiovascular well-being. Consequently, it could contribute to reducing healthcare expenditures and yield positive impacts on public health outcomes.

The ECG-age gap holds promise in identifying individuals who are experiencing accelerated aging of the heart beyond what is typical for their chronological age. When combined with their medical history and other clinical data, this information can facilitate the proactive identification of patients at risk for preclinical or undiagnosed cardiac conditions, thereby creating opportunities for preventive interventions. Moreover, the integration of the ECG-age gap into the ECG report can serve as a valuable reference index

for monitoring the cardiac health status of high-risk patients, enabling timely interventions and management strategies to maintain optimal cardiovascular health.

Studies have demonstrated that the ECG-age gap can improve with the resolution of illness and with lifestyle modifications such as weight loss, particularly in early observations. This suggests that it may serve as a continuous marker of relative health, offering motivation for individuals to make positive lifestyle changes.[51] The ECG-age is not a static indicator of a person's actual age at the time of an ECG. Instead, it can be regarded as a dynamic measure that reflects the physiological aging process and can vary over time. Exploring the impact of the variability of the ECG-age on long-term cardiac outcomes through follow-up ECGs obtained over time could represent a valuable starting point for future research endeavors.

This study has some limitations. First, the representativeness of the three cohorts in this study may be limited for the nationwide general populations of each cohort.[52] Second, survival bias may affect ECG-age prediction, as individuals with greater longevity are more likely to have a younger biological age, and accordingly a younger ECG-age. Nevertheless, the findings are encouraging given that the results were adjusted for chronological age. Third, some unmeasured potential confounders may still act as significant risk factors for AF. Fourth, the UK Biobank had a shorter follow-up period for AF events, and the performance of the AI model within the UK Biobank was found somewhat diminished compared to the other cohorts. Finally, although the ECG-age is considered a valuable indicator of physiological age, the nature of the relationship between physiological age and this AI-predicted index remains uncertain as there is no definitive gold standard test for physiological age.[53]

6. Conclusion

I demonstrated that ECG-aging is associated with the risk of new-onset AF and suggested its potential as a novel biomarker for AF risk prevention across diverse populations. Further studies are needed in more diverse cohorts to explore the relationship between ECG-age gap and the risk of other age-related CVDs. Still, this study suggests that the AI-predicted ECG age prediction model could be a cost-effective, non-invasive, and targeted screening tool for estimating heart age and risk stratification for primary prevention.

REFERENCES

1. Ladejobi, A. O., Medina-Inojosa, J. R., Shelly Cohen, M., Attia, Z. I., Scott, C. G., LeBrasseur, N. K., Gersh, B. J., Noseworthy, P. A., Friedman, P. A., Kapa, S., & Lopez-Jimenez, F. (2021). The 12-lead electrocardiogram as a biomarker of biological age. *European heart journal. Digital health*, *2*(3), 379–389. https://doi.org/10.1093/ehjdh/ztab043

2. Ribeiro, A. H., Ribeiro, M. H., Paixão, G. M. M., Oliveira, D. M., Gomes, P. R., Canazart, J. A., Ferreira, M. P. S., Andersson, C. R., Macfarlane, P. W., Meira, W., Jr, Schön, T. B., & Ribeiro, A. L. P. (2020). Automatic diagnosis of the 12-lead ECG using a deep neural network. *Nature communications*, *11*(1), 1760. https://doi.org/10.1038/s41467-020-15432-4

3. Han, C., Kang, K. W., Kim, T. Y., Uhm, J. S., Park, J. W., Jung, I. H., Kim, M., Bae, S., Lim, H. S., & Yoon, D. (2022). Artificial Intelligence-Enabled ECG Algorithm for the Prediction of Coronary Artery Calcification. *Frontiers in cardiovascular medicine*, *9*, 849223. https://doi.org/10.3389/fcvm.2022.849223

4. Hughes, J. W., Tooley, J., Torres Soto, J., Ostropolets, A., Poterucha, T., Christensen, M. K., Yuan, N., Ehlert, B., Kaur, D., Kang, G., Rogers, A., Narayan, S., Elias, P., Ouyang, D., Ashley, E., Zou, J., & Perez, M. V. (2023). A deep learning-based electrocardiogram risk score for long term cardiovascular death and disease. *NPJ digital medicine*, *6*(1), 169. https://doi.org/10.1038/s41746-023-00916-6

5. Khurshid, S., Friedman, S., Reeder, C., Di Achille, P., Diamant, N., Singh, P., Harrington, L. X., Wang, X., Al-Alusi, M. A., Sarma, G., Foulkes, A. S., Ellinor, P. T., Anderson, C. D., Ho, J. E., Philippakis, A. A., Batra, P., & Lubitz, S. A. (2022). ECG-Based Deep Learning and Clinical Risk Factors to Predict Atrial Fibrillation. *Circulation*, *145*(2), 122–133. https://doi.org/10.1161/CIRCULATIONAHA.121.057480

6. Attia, Z. I., Harmon, D. M., Behr, E. R., & Friedman, P. A. (2021). Application of artificial intelligence to the electrocardiogram. *European heart journal*, *42*(46), 4717–4730. https://doi.org/10.1093/eurheartj/ehab649

7. Sebastiani, P., Thyagarajan, B., Sun, F., Schupf, N., Newman, A. B., Montano, M., & Perls, T. T. (2017). Biomarker signatures of aging. *Aging cell*, *16*(2), 329–338. https://doi.org/10.1111/acel.12557

8. Lima, E. M., Ribeiro, A. H., Paixão, G. M. M., Ribeiro, M. H., Pinto-Filho, M. M., Gomes, P. R., Oliveira, D. M., Sabino, E. C., Duncan, B. B., Giatti, L., Barreto, S. M., Meira, W., Jr, Schön, T. B., & Ribeiro, A. L. P. (2021). Deep neural network-estimated electrocardiographic age as a mortality predictor. *Nature communications*, *12*(1), 5117. https://doi.org/10.1038/s41467-021-25351-7

9. Brant, L. C. C., Ribeiro, A. H., Pinto-Filho, M. M., Kornej, J., Preis, S. R., Fetterman, J. L., Eromosele, O. B., Magnani, J. W., Murabito, J. M., Larson, M. G., Benjamin, E. J., Ribeiro, A. L. P., & Lin, H. (2023). Association Between Electrocardiographic Age and Cardiovascular Events in Community Settings: The Framingham Heart Study. *Circulation. Cardiovascular quality and outcomes*, *16*(7), e009821. https://doi.org/10.1161/CIRCOUTCOMES.122.009821

10. Bai, Y., Wang, Y. L., Shantsila, A., & Lip, G. Y. H. (2017). The Global Burden of Atrial Fibrillation and Stroke: A Systematic Review of the Clinical Epidemiology of Atrial Fibrillation in Asia. *Chest*, *152*(4), 810–820. https://doi.org/10.1016/j.chest.2017.03.048

11. Roth, G. A., Mensah, G. A., Johnson, C. O., Addolorato, G., Ammirati, E., Baddour, L. M., Barengo, N. C., Beaton, A. Z., Benjamin, E. J., Benziger, C. P., Bonny, A., Brauer, M., Brodmann, M., Cahill, T. J., Carapetis, J., Catapano, A. L., Chugh, S. S., Cooper, L. T., Coresh, J., Criqui, M., … GBD-NHLBI-JACC Global Burden of Cardiovascular Diseases Writing Group (2020). Global Burden of Cardiovascular Diseases and Risk Factors, 1990-2019: Update From the GBD 2019 Study. *Journal of the American College of Cardiology*, *76*(25), 2982–3021. https://doi.org/10.1016/j.jacc.2020.11.010

12. Dong, X. J., Wang, B. B., Hou, F. F., Jiao, Y., Li, H. W., Lv, S. P., & Li, F. H. (2023). Global burden of atrial fibrillation/atrial flutter and its attributable risk factors from 1990 to 2019. *Europace : European pacing, arrhythmias, and cardiac electrophysiology : journal of the working groups on cardiac pacing, arrhythmias, and cardiac cellular electrophysiology of the European Society of Cardiology*, *25*(3), 793–803. https://doi.org/10.1093/europace/euac237

13. Staerk, L., Sherer, J. A., Ko, D., Benjamin, E. J., & Helm, R. H. (2017). Atrial Fibrillation: Epidemiology, Pathophysiology, and Clinical Outcomes. *Circulation research*, *120*(9), 1501–1517. https://doi.org/10.1161/CIRCRESAHA.117.309732

14. Hannun, A. Y., Rajpurkar, P., Haghpanahi, M., Tison, G. H., Bourn, C., Turakhia, M. P., & Ng, A. Y. (2019). Cardiologist-level arrhythmia detection and

classification in ambulatory electrocardiograms using a deep neural network. *Nature medicine*, *25*(1), 65–69. https://doi.org/10.1038/s41591-018-0268-3

15. Raghunath, S., Ulloa Cerna, A. E., Jing, L., vanMaanen, D. P., Stough, J., Hartzel, D. N., Leader, J. B., Kirchner, H. L., Stumpe, M. C., Hafez, A., Nemani, A., Carbonati, T., Johnson, K. W., Young, K., Good, C. W., Pfeifer, J. M., Patel, A. A., Delisle, B. P., Alsaid, A., Beer, D., … Fornwalt, B. K. (2020). Prediction of mortality from 12-lead electrocardiogram voltage data using a deep neural network. *Nature medicine*, *26*(6), 886–891. https://doi.org/10.1038/s41591-020-0870-z

16. Zvuloni, E., Read, J., Ribeiro, A. H., Ribeiro, A. L. P., & Behar, J. A. (2023). On Merging Feature Engineering and Deep Learning for Diagnosis, Risk Prediction and Age Estimation Based on the 12-Lead ECG. *IEEE transactions on bio-medical engineering*, *70*(7), 2227–2236. https://doi.org/10.1109/TBME.2023.3239527

17. Jylhävä, J., Pedersen, N. L., & Hägg, S. (2017). Biological Age Predictors. *EBioMedicine*, *21*, 29–36. https://doi.org/10.1016/j.ebiom.2017.03.046

18. Lowsky, D. J., Olshansky, S. J., Bhattacharya, J., & Goldman, D. P. (2014). Heterogeneity in healthy aging. *The journals of gerontology. Series A, Biological sciences and medical sciences*, *69*(6), 640–649. https://doi.org/10.1093/gerona/glt162

19. Belsky, D. W., Moffitt, T. E., Cohen, A. A., Corcoran, D. L., Levine, M. E., Prinz, J. A., Schaefer, J., Sugden, K., Williams, B., Poulton, R., & Caspi, A. (2018). Eleven Telomere, Epigenetic Clock, and Biomarker-Composite Quantifications of Biological Aging: Do They Measure the Same Thing?. *American journal of epidemiology*, *187*(6), 1220–1230. https://doi.org/10.1093/aje/kwx346

20. Bachman, S., Sparrow, D., & Smith, L. K. (1981). Effect of aging on the electrocardiogram. *The American journal of cardiology*, *48*(3), 513–516. https://doi.org/10.1016/0002-9149(81)90081-3

21. Giovanardi, P., Vernia, C., Tincani, E., Giberti, C., Silipo, F., & Fabbo, A. (2022). Combined Effects of Age and Comorbidities on Electrocardiographic Parameters in a Large Non-Selected Population. *Journal of clinical medicine*, *11*(13), 3737. https://doi.org/10.3390/jcm11133737

22. Starc, V., Leban, M. A., Šinigoj, P., Vrhovec, M., Potočnik, N., Fernlund, E., Liuba, P., & Schlegel, T. T. (2012, September). Can functional cardiac age be predicted from the ECG in a normal healthy population?. In *2012 Computing in Cardiology* (pp. 101-104). IEEE.

23. Ball, R. L., Feiveson, A. H., Schlegel, T. T., Starc, V., & Dabney, A. R. (2014). Predicting "heart age" using electrocardiography. *Journal of personalized medicine*, *4*(1), 65–78. https://doi.org/10.3390/jpm4010065

24. Lindow, T., Palencia-Lamela, I., Schlegel, T. T., & Ugander, M. (2022). Heart age estimated using explainable advanced electrocardiography. *Scientific reports*, *12*(1), 9840. https://doi.org/10.1038/s41598-022-13912-9

25. Hirota, N., Suzuki, S., Arita, T., Yagi, N., Otsuka, T., & Yamashita, T. (2021). Prediction of biological age and all-cause mortality by 12-lead electrocardiogram in patients without structural heart disease. *BMC geriatrics*, *21*(1), 460. https://doi.org/10.1186/s12877-021-02391-8

26. Attia, Z. I., Friedman, P. A., Noseworthy, P. A., Lopez-Jimenez, F., Ladewig, D. J., Satam, G., Pellikka, P. A., Munger, T. M., Asirvatham, S. J., Scott, C. G., Carter, R. E., & Kapa, S. (2019). Age and Sex Estimation Using Artificial Intelligence From Standard 12-Lead ECGs. *Circulation. Arrhythmia and electrophysiology*, *12*(9), e007284. https://doi.org/10.1161/CIRCEP.119.007284

27. Diez Benavente, E., Jimenez-Lopez, F., Attia, Z. I., Malyutina, S., Kudryavtsev, A., Ryabikov, A., Friedman, P. A., Kapa, S., Voevoda, M., Perel, P., Schirmer, H., Hughes, A., Clark, T. G., & Leon, D. A. (2021). Studying accelerated cardiovascular ageing in Russian adults through a novel deep-learning ECG biomarker. *Wellcome Open Research*, *6*, 12.

28. Benavente, E. D., Lopez-Jimenez, F., Iakunchykova, O., Malyutina, S., Kudryavtsev, A., Ryabikov, A., Friedman, P. A., Kapa, S., Perel, P., Wisgaard, T., Schirmer, H., Clark, T. G., Attia, Z. I., & Leon, D. A. (2021). Capturing population differences in rates of vascular aging using a deep learning electrocardiogram algorithm: a cross-sectional study. *medRxiv*, 2021-09.

29. Chang, C. H., Lin, C. S., Luo, Y. S., Lee, Y. T., & Lin, C. (2022). Electrocardiogram-Based Heart Age Estimation by a Deep Learning Model Provides More Information on the Incidence of Cardiovascular Disorders. *Frontiers in cardiovascular medicine*, *9*, 754909. https://doi.org/10.3389/fcvm.2022.754909

30. Baek, Y. S., Lee, D. H., Jo, Y., Lee, S. C., Choi, W., & Kim, D. H. (2023). Artificial intelligence-estimated biological heart age using a 12-lead electrocardiogram predicts mortality and cardiovascular outcomes. *Frontiers in cardiovascular medicine*, *10*, 1137892. https://doi.org/10.3389/fcvm.2023.1137892

31. Alonso, A., Krijthe, B. P., Aspelund, T., Stepas, K. A., Pencina, M. J., Moser, C. B., Sinner, M. F., Sotoodehnia, N., Fontes, J. D., Janssens, A. C., Kronmal, R. A., Magnani, J. W., Witteman, J. C., Chamberlain, A. M., Lubitz, S. A., Schnabel, R. B., Agarwal, S. K., McManus, D. D., Ellinor, P. T., Larson, M. G., … Benjamin, E. J. (2013). Simple risk model predicts incidence of atrial fibrillation in a racially and geographically diverse population: the CHARGE-AF consortium. *Journal of the American Heart Association*, *2*(2), e000102. https://doi.org/10.1161/JAHA.112.000102

32. Attia, Z. I., Noseworthy, P. A., Lopez-Jimenez, F., Asirvatham, S. J., Deshmukh, A. J., Gersh, B. J., Carter, R. E., Yao, X., Rabinstein, A. A., Erickson, B. J., Kapa, S., & Friedman, P. A. (2019). An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction. *Lancet (London, England)*, *394*(10201), 861–867. https://doi.org/10.1016/S0140-6736(19)31721-0

33. Christopoulos, G., Graff-Radford, J., Lopez, C. L., Yao, X., Attia, Z. I., Rabinstein, A. A., Petersen, R. C., Knopman, D. S., Mielke, M. M., Kremers, W., Vemuri, P., Siontis, K. C., Friedman, P. A., & Noseworthy, P. A. (2020). Artificial Intelligence-Electrocardiography to Predict Incident Atrial Fibrillation: A Population-Based Study. *Circulation. Arrhythmia and electrophysiology*, *13*(12), e009355. https://doi.org/10.1161/CIRCEP.120.009355

34. Raghunath, S., Pfeifer, J. M., Ulloa-Cerna, A. E., Nemani, A., Carbonati, T., Jing, L., vanMaanen, D. P., Hartzel, D. N., Ruhl, J. A., Lagerman, B. F., Rocha, D. B., Stoudt, N. J., Schneider, G., Johnson, K. W., Zimmerman, N., Leader, J. B., Kirchner, H. L., Griessenauer, C. J., Hafez, A., Good, C. W., … Haggerty, C. M. (2021). Deep Neural Networks Can Predict New-Onset Atrial Fibrillation From the 12-Lead ECG and Help Identify Those at Risk of Atrial Fibrillation-Related Stroke. *Circulation*, *143*(13), 1287–1298. https://doi.org/10.1161/CIRCULATIONAHA.120.047829

35. Biton, S., Gendelman, S., Ribeiro, A. H., Miana, G., Moreira, C., Ribeiro, A. L. P., & Behar, J. A. (2021). Atrial fibrillation risk prediction from the 12-lead

electrocardiogram using digital biomarkers and deep representation learning. *European heart journal. Digital health*, *2*(4), 576–585. https://doi.org/10.1093/ehjdh/ztab071

36. Zheng, J., Zhang, J., Danioko, S., Yao, H., Guo, H., & Rakovski, C. (2020). A 12-lead electrocardiogram database for arrhythmia research covering more than 10,000 patients. *Scientific data*, *7*(1), 48. https://doi.org/10.1038/s41597-020-0386-x

37. Wagner, P., Strodthoff, N., Bousseljot, R. D., Kreiseler, D., Lunze, F. I., Samek, W., & Schaeffter, T. (2020). PTB-XL, a large publicly available electrocardiography dataset. *Scientific data*, *7*(1), 154. https://doi.org/10.1038/s41597-020-0495-6

38. Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., Downey, P., Elliott, P., Green, J., Landray, M., Liu, B., Matthews, P., Ong, G., Pell, J., Silman, A., Young, A., Sprosen, T., Peakman, T., & Collins, R. (2015). UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS medicine*, *12*(3), e1001779. https://doi.org/10.1371/journal.pmed.1001779

39. Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L. T., Sharp, K., Motyer, A., Vukcevic, D., Delaneau, O., O'Connell, J., Cortes, A., Welsh, S., Young, A., Effingham, M., McVean, G., Leslie, S., Allen, N., Donnelly, P., & Marchini, J. (2018). The UK Biobank resource with deep phenotyping and genomic data. *Nature*, *562*(7726), 203–209. https://doi.org/10.1038/s41586-018-0579-z

40. Scherer, D., Müller, A., & Behnke, S. (2010, September). Evaluation of pooling operations in convolutional architectures for object recognition. In *International conference on artificial neural networks* (pp. 92-101). Berlin, Heidelberg: Springer Berlin Heidelberg.

41. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research*, *15*(1), 1929-1958.

42. Ioffe, S., & Szegedy, C. (2015, June). Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *International conference on machine learning* (pp. 448-456). pmlr.

43. Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

44. Simonyan, K., Vedaldi, A., & Zisserman, A. (2013). Deep inside convolutional networks: Visualising image classification models and saliency maps. *arXiv preprint arXiv:1312.6034*.

45. Makowski, D., Pham, T., Lau, Z. J., Brammer, J. C., Lespinasse, F., Pham, H., Schölzel, C., & Chen, S. H. A. (2021). NeuroKit2: A Python toolbox for neurophysiological signal processing. *Behavior research methods*, *53*(4), 1689–1696. https://doi.org/10.3758/s13428-020-01516-y

46. Hirota, N., Suzuki, S., Motogi, J., Nakai, H., Matsuzawa, W., Takayanagi, T., Umemoto, T., Hyodo, A., Satoh, K., Arita, T., Yagi, N., Otsuka, T., & Yamashita, T. (2023). Cardiovascular events and artificial intelligence-predicted age using 12-lead electrocardiograms. *International journal of cardiology. Heart & vasculature*, *44*, 101172. https://doi.org/10.1016/j.ijcha.2023.101172

47. Van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of statistical software*, *45*, 1-67.

48. Schotten, U., Verheule, S., Kirchhof, P., & Goette, A. (2011). Pathophysiological mechanisms of atrial fibrillation: a translational appraisal. *Physiological reviews*, *91*(1), 265–325. https://doi.org/10.1152/physrev.00031.2009

49. Ball, R. L., Feiveson, A. H., Schlegel, T. T., Starc, V., & Dabney, A. R. (2014). Predicting "heart age" using electrocardiography. *Journal of personalized medicine*, *4*(1), 65–78. https://doi.org/10.3390/jpm4010065

50. MacKinnon, G. E., & Brittain, E. L. (2020). Mobile Health Technologies in Cardiopulmonary Disease. *Chest*, *157*(3), 654–664. https://doi.org/10.1016/j.chest.2019.10.015

51. Harmon, D. M., Lopez-Jimenez, F., & Friedman, P. A. (2022). Introducing Artificial Intelligence into the Preventive Medicine Visit. Mayo Clinic proceedings, 97(8), 1575–1577. https://doi.org/10.1016/j.mayocp.2022.06.003

52. Schneider, C. V., Schneider, K. M., Teumer, A., Rudolph, K. L., Hartmann, D., Rader, D. J., & Strnad, P. (2022). Association of Telomere Length With Risk of Disease and Mortality. *JAMA internal medicine*, *182*(3), 291–300. https://doi.org/10.1001/jamainternmed.2021.7804

53. Toya, T., Ahmad, A., Attia, Z., Cohen-Shelly, M., Ozcan, I., Noseworthy, P. A., Lopez-Jimenez, F., Kapa, S., Lerman, L. O., Friedman, P. A., & Lerman, A. (2021). Vascular Aging Detected by Peripheral Endothelial Dysfunction Is

Associated With ECG-Derived Physiological Aging. *Journal of the American Heart Association, 10*(3), e018656. https://doi.org/10.1161/JAHA.120.018656

ABSTRACT IN KOREAN

심전도 활용 연령 예측 인공지능 개발 및 인공지능 예측 심전도 연령의 심방세동 위험 지표로서의 유효성 탐색

배경: 인공지능 알고리즘을 12 리드 심전도(ECG)에 적용하여 나이를 예측하는 방법이 유망하게 제시되고 있다. 본 연구에서는 심전도 파형 데이터로부터 예측된 나이와 실제 나이 간의 차이를 심전도-나이 차이(ECG-age gap) 또는 심전도-노화(ECG-aging)로 지칭하고, 이것이 노화 관련 대표 심장질환인 심방세동 발병 위험과 연관성이 있는지 조사하였다.

방법: 대규모 데이터셋(689,639 명의 1,533,042 심전도)를 활용하여 심전도-나이 예측 모델을 개발하고, 다섯 개의 독립적인 다국적 데이터셋(230,838 명의 637,177 심전도)을 사용하여 검증하였다. 심전도-나이 차이는 한국과 영국의 세 개 코호트에서 계산되었으며, 이들은 각각 4.14±4.27, 6.08±3.81 및 2.99±1.56 년간 추적되었으며 참여자는 각각 111,483, 37,517 및 40,973 명이었다. 심전도-나이 차이를 기반으로 참여자들을 두 그룹으로 분류하였다: 정상 그룹 (심전도-나이 차이<+7 년) 및 심전도-노화 그룹 (심전도-나이 차이≥+7 년). 콕스 비례위험 모델을 사용하여 심전도 노화와 신규 심방세동 발병 위험의 연관성을 평가하였다.

결과: 세 개 코호트에서 평균 심전도-나이와 평균 심전도-나이 차이는 각각 51.9±16.2 (0.0±6.8), 47.4±12.5 (-0.1±6.0) 및 68.4±7.8 (4.7±8.7) 년이었다. 심전도-노화 그룹은 각각의 코호트에서 정상 그룹과 비교하여 신규 심방세동 발병 위험의 증가와 관련이 있었으며, 각각의 코호트에서 위험비는 2.50 (신뢰구간 95%, 2.24–2.78), 1.89 (1.46–2.43) 및 1.90 (1.55–2.33) 이었다. 심전도-나이 차이가 증가함에 따라 신규 심방세동 발병 위험이 증가하였다.

결론: 인공지능에서 파생된 심전도-노화는 신규 심방세동 발병 위험과 관련이 있었으며, 이는 일차예방에서 심방세동에 대한 위험 바이오마커로의 잠재적 가능성을 나타낸다.

핵심되는 말: 인공지능, 심전도, 노화, 심방세동