



OPEN ACCESS

EDITED BY

Dexter Canoy,
Newcastle University, United Kingdom

REVIEWED BY

Eric S. Ho,
Lafayette College, United States
Vito Starc,
University of Ljubljana, Slovenia

*CORRESPONDENCE

Robert Leung
✉ robert.leung@umassmemorial.org

[†]These authors have contributed equally to this work

RECEIVED 09 January 2024

ACCEPTED 13 June 2024

PUBLISHED 28 June 2024

CITATION

Leung R, Wang B, Gottbrecht M, Doerr A, Marya N, Soni A, McManus DD and Lin H (2024) Association between deep neural network-derived electrocardiographic-age and incident stroke.
Front. Cardiovasc. Med. 11:1368094.
doi: 10.3389/fcvm.2024.1368094

COPYRIGHT

© 2024 Leung, Wang, Gottbrecht, Doerr, Marya, Soni, McManus and Lin. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Association between deep neural network-derived electrocardiographic-age and incident stroke

Robert Leung^{1*†}, Biqi Wang^{1,2†}, Matthew Gottbrecht³, Adam Doerr¹, Neil Marya^{1,4}, Apurv Soni^{1,2}, David D. McManus^{1,2,3} and Honghuang Lin^{1,2}

¹Program in Digital Medicine, Department of Medicine, UMass Chan Medical School, Worcester, MA, United States, ²Division of Health Systems Science, Department of Medicine, UMass Chan Medical School, Worcester, MA, United States, ³Division of Cardiology, Department of Medicine, UMass Chan Medical School, Worcester, MA, United States, ⁴Division of Gastroenterology, Department of Medicine, UMass Chan Medical School, Worcester, MA, United States

Background: Stroke continues to be a leading cause of death and disability worldwide despite improvements in prevention and treatment. Traditional stroke risk calculators are biased and imprecise. Novel stroke predictors need to be identified. Recently, deep neural networks (DNNs) have been used to determine age from ECGs, otherwise known as the electrocardiographic-age (ECG-age), which predicts clinical outcomes. However, the relationship between ECG-age and stroke has not been well studied. We hypothesized that ECG-age is associated with incident stroke.

Methods: In this study, UK Biobank participants with available ECGs (from 2014 or later). ECG-age was estimated using a deep neural network (DNN) applied to raw ECG waveforms. We calculated the Δ age (ECG-age minus chronological age) and classified individuals as having normal, accelerated, or decelerated aging if Δ age was within, higher, or lower than the mean absolute error of the model, respectively. Multivariable Cox proportional hazards regression models adjusted for age, sex, and clinical factors were used to assess the association between Δ age and incident stroke.

Results: The study population included 67,757 UK Biobank participants (mean age 65 ± 8 years; 48.3% male). Every 10-year increase in Δ age was associated with a 22% increase in incident stroke [HR, 1.22 (95% CI, 1.00–1.49)] in the multivariable-adjusted model. Accelerated aging was associated with a 42% increase in incident stroke [HR, 1.42 (95% CI, 1.12–1.80)] compared to normal aging. In addition, Δ age was associated with prevalent stroke [OR, 1.28 (95% CI, 1.11–1.49)].

Conclusions: DNN-estimated ECG-age was associated with incident and prevalent stroke in the UK Biobank. Further investigation is required to determine if ECG-age can be used as a reliable biomarker of stroke risk.

KEYWORDS

stroke, ECG-age, UK Biobank, AI, DNN

Introduction

Stroke is the second leading cause of death and disability worldwide despite improvements in prevention and treatment. It continues to pose a massive burden on both individual and societal levels. Stroke alone accounts for approximately 116 million global disability-adjusted life-years (DALYs) lost in 2016 and over US\$891 billion in 2019 (1, 2). Globally,

the prevalence of stroke is expected to rise, leading to further increases in stroke-related costs (3). In the US and other developed nations, the main driver for increased stroke prevalence is advanced age, as age is the most important non-modifiable risk factor. Moreover, the literature suggests that evidence-based guidelines for stroke care are less likely to be followed with older age (4). As a result, stroke risk reduction among older patients remains crucial.

Predicting which patients are at highest risk for a certain disease enable early intervention, thereby improving patient outcomes and reducing pressure on the healthcare system through primary prevention. However, at present many stroke risk calculators like the American College of Cardiology/American Heart Association (ACC/AHA) Pooled Cohort Equations (PCE) and the Framingham Stroke Risk Score (FSRS) are derived from conventional statistical methods that include a few predictors quantified by human observers. These models oversimplify complex relationships (5, 6). In contrast, a study by Dritsas and Trigka demonstrated the potential of various machine learning (ML) algorithms in predicting stroke (7), an approach that has been used to improve the prediction of other diseases like hypertension and diabetes mellitus (8, 9). Further, both the PCE and FSRS tend to underperform in certain ethnicities and socioeconomic classes. They also do not incorporate novel risk markers that some consider important for risk assessment (10). Fortunately, one of the key applications of machine learning (ML) in healthcare is risk stratification and clinical decision support. ML has the potential to circumvent these limitations and outperform current stroke risk prediction tools (5).

The ML method utilized depends on the type of data. Deep learning (DL) models based on neural networks have been used to extract features from imaging and ECGs as a basis for predicting cardiovascular disease (CVD) risk (5). One such model was a deep neural network (DNN)-based age prediction model developed on the Clinical Outcomes in Digital Electrocardiography (CODE) data set to predict an individual's age based on ECG waveform, otherwise known as the electrocardiographic age (ECG-age) (11). Existing literature has shown that DNN-estimated age is highly correlated to chronological age with the delta age (ECG-age minus chronological age) being a predictor of overall mortality (12). A recent study showed that advanced "biological aging," as predicted by ECG data, was associated with increased risk of all-cause mortality, myocardial infarction (MI), atrial fibrillation (AF), and heart failure, highlighting the vital role of ECG as a biomarker of CVD risk and its potential role in stroke prediction (11, 13).

We hypothesize that ECG changes with age relate strongly to stroke risk. This study aimed to evaluate whether DNN-estimated ECG-age can predict incident stroke in the large-scale and long-term UK Biobank.

Methods

Study population

Our study population is comprised of more than half a million volunteers aged 40–69 years who enrolled in the UK

Biobank between the years 2006 and 2010 (14). These participants lived within 25 miles of one of the 22 assessment centers located throughout England, Wales, and Scotland. On enrollment in the UK Biobank, participants provided signed consent and answered questions about sociodemographic, lifestyle, environmental, and health-related factors. 12-lead (at-rest) ECGs were performed at an imaging assessment center for UK Biobank. Exact details of the procedure can be found at <https://biobank.ctsu.ox.ac.uk/crystal/refer.cgi?id=510>. ECG acquisition was approximately 10 s per individual. Participants who had an ECG done during their clinic visit (2014 or later) or subsequent clinic visit (2019 or later) were included in our study. For participants with two ECGs, only the first valid ECG was accounted. We excluded individuals with prevalent stroke ($n = 808$), which left 67,757 individuals for the association analysis of ECG-age and incident stroke.

Clinical variables

All the lifestyle and clinical risk factors of stroke were measured at the same time as the ECG measurement. Lifestyle factors (smoking, alcohol consumption, physical activity, and diet) were self-reported and were dichotomized as previous publications presented (15). For blood pressure, two measures were taken one minute apart using a digital blood pressure monitor. Body mass index was calculated as weight in kilograms divided by height in meters squared. Particulate matter (PM) air pollution (PM_{2.5}, unit micro-g/m³) was collected from local environment air pollution data in the year 2010 (Data Field 24,006). Another air pollutant, PM₁₀ (PM with an aerodynamic diameter of <10 micro-g), was measured in 2007 (Data Field 24,019) and 2010 (Data Field 24,005). Health-related outcomes including diabetes mellitus (DM), chronic kidney diseases (CKD), and dyslipidemia were followed up based on the primary care data (READ code), hospital inpatient data (ICD-9/ICD-10 codes), death registry (ICD-10 code), and self-reported medical conditions. Their first occurrence dates were derived and provided by the UK Biobank: DM included type 1, type 2, and other diabetes (Data Fields 130,706, 130,708, 130,710, 130,712, 130,714), CKD Data Field was 132,032, and dyslipidemia was 130,814.

Study outcomes

Ischemic stroke and hemorrhagic (intracerebral and subarachnoid subtypes) stroke were the outcomes of interest. These outcomes were defined by self-reported medical conditions, ICD-9, or ICD-10 codes (16, 17). The Data Field related to the stroke outcomes were 42,006, 42,008, 42,010, and 42,012. The included participants were followed up since they took the first ECG until whichever came first: newly diagnosed stroke, or their last follow-up time in the UK Biobank, in which they were free of stroke, death, or end of August 2023.

DNN-estimated ECG-age

The ECG-age was obtained using a DNN that applies the raw ECG waveform in an end-to-end approach as opposed to classical automatic analysis methods (18). The model was trained to predict a person's age by extracting features directly from the data instead of relying on traditional ECG interpretation (18, 19). Through this learning process, the DNN captured how aging affects ECG waveform (11). For this study, the raw ECG signals used to generate the ECG-age model were derived from the CODE study, which was further refined using the UK Biobank ECG data. The CODE study is part of the Telehealth Network of Minas Gerais (TNMG) with a database containing ECGs obtained from 2010 to 2017 in Brazilian primary care settings (20). To date, the CODE data set has been recognized as the world's largest ECG database used to develop artificial intelligence (AI)-ECG applications with 1,558, 415 participants (21). The ECG-age model in the CODE study and its development have been previously described (12). The model uses a DNN to make predictions like the residual network proposed for image classification except with unidimensional signals (11). More information on the code used in the DNN-estimated ECG-age model training, evaluation, and statistical analysis can be found at <https://github.com/antonior92/ecg-age-prediction>.

Statistical analyses

To assess the DNN-estimated ECG-age as a predictor for stroke risk, we used the Δ age (ECG-age minus the chronological age) as the independent variable and thereby capturing the excess risk caused by a greater decline in cardiovascular health than expected by chronological aging. We also separated participants into three aging categories based on Δ age mean absolute error (MAE) (12). Those with an ECG-age older than the chronological age by \geq MAE years were considered accelerated aging. Those with an ECG-age younger than the chronological age by \geq MAE years were considered decelerated aging. Those with an ECG-age within the MAE were considered normal aging. We used Cox proportional hazards regression models to assess the association of Δ age and incident stroke: the first model was chronological age and sex adjusted, and the second multivariable model was additionally adjusted for SBP, BMI, particulate matter air pollution, smoking status, diet, physical activity level, alcohol consumption, diabetes, CKD, and dyslipidemia. The results were expressed as stroke risk of Δ age per 10 years.

For the secondary analysis, we assessed the association of Δ age with prevalent stroke by logistic regression models and adjusted for chronological age, sex, and additional clinical risk factors. Several sensitivity analyses included: (A) we excluded participants with prevalent atrial fibrillation at the time of ECG measurement, (B) we assessed the 5-year stroke risk by the Cox model, (C) we assessed the ability of Δ age to predict incident stroke using the C-statistics where we calculated the change in the predictive capacity with and without the Δ age in the multivariable Cox models, and the net reclassification index was also generated to

assess the prediction ability (22), (D) we explored the non-linear relationship of Δ age and the hazard of incident stroke by using the fractional polynomials and splines.

In our sensitivity analysis, we performed stratified analysis in two age groups (chronological age <60 , or ≥ 60), and in men and women separately to assess the associations of Δ age and incident stroke. We used the multivariable-adjusted Kaplan-Meier curves which were achieved by the inverse probability weighting method to illustrate the cumulative risk of stroke across decelerated, normal, and accelerated aging groups. We used *P* values less than 0.05 (two-sided) as the statistical significance level. All the analyses were performed using R software package version 4.2.1 (<https://www.r-project.org/>).

Results

Our primary analysis includes 67,757 participants (68,565 for the secondary analysis) from the UK Biobank. The mean age was 65 ± 8 years and 48.3% were men. The clinical characteristics are shown in Table 1. Those in the accelerated aging cohort were more likely male, had a higher BMI, and more current smoking, alcohol consumption, unhealthy diet, and less physical activity (Supplementary Table S1). ECG-age was generated with a DNN

TABLE 1 Baseline characteristics of participants with ECG-age in the UK Biobank.

	Overall	No stroke	Incident stroke	<i>P</i> value
<i>n</i>	67,757	67,378	379	
Chronological age [mean (SD)]	65 (8)	65 (8)	68 (7)	<0.001
Men (%)	32,706 (48.3)	32,487 (48.2)	219 (57.8)	<0.001
Δ age [mean (SD)]	-0.01 (5.88)	-0.02 (5.88)	1.09 (6.34)	<0.001
Body mass index [mean (SD)]	26.6 (4.5)	26.6 (4.5)	26.9 (4.4)	0.30
Systolic blood pressure [mean (SD)]	141 (19)	141 (19)	146 (21)	<0.001
Diastolic blood pressure [mean (SD)]	79 (10)	79 (10)	81 (12)	0.005
Current smoker (%)	2,273 (3.4)	2,259 (3.4)	14 (3.7)	0.82
Healthy drinking (%)	44,985 (66.9)	44,727 (66.9)	258 (68.8)	0.47
Physically active (%)	53,261 (87.2)	52,970 (87.2)	291 (86.4)	0.71
Healthy diet (%)	29,686 (45.0)	29,525 (45.0)	161 (43.9)	0.71
Prevalent diabetes (%)	3,846 (5.7)	3,807 (5.7)	39 (10.3)	<0.001
Prevalent chronic kidney diseases (%)	1,444 (2.1)	1,428 (2.1)	16 (4.2)	0.008
Prevalent dyslipidemia (%)	16,013 (23.6)	15,901 (23.6)	112 (29.6)	0.008
Incident ischemic stroke (%)	304 (0.4)	0 (0.0)	304 (80.2)	<0.001
Incident hemorrhagic stroke (%)	86 (0.1)	0 (0.0)	86 (22.7)	<0.001
Incident hemorrhagic intracerebral stroke (%)	65 (0.1)	0 (0.0)	65 (17.2)	<0.001
Incident hemorrhagic subarachnoid stroke (%)	32 (0.0)	0 (0.0)	32 (8.4)	<0.001
Follow-up years [mean (SD)]	3.94 (2.55)	3.95 (2.56)	2.75 (1.88)	<0.001

Δ age, ECG predicted age—chronological age; *P* values were from the *t*-test (continuous variables) or Chi-square test (categorical variables).

(based on a previous model from the CODE study) using raw ECG waveform data. On the other hand, participants within the decelerated age group had less prevalent DM, CKD, and dyslipidemia (Supplementary Table S1). A total of 808 participants had a stroke before or at the time an ECG was performed. During an average of 3.94 ± 2.55 years of follow-up, an additional 379 participants were diagnosed with stroke.

Association of Δ age with incident stroke

Table 2 shows the association of Δ age with incident stroke in the chronological age- and sex-adjusted, and multivariable-adjusted models. Every 10-year increase in Δ age was associated with a 22% increase in the risk of incident stroke [HR = 1.22 (95% CI, 1.00–1.49)] after adjusting for multiple risk factors. The association remained significant if we restricted to 5 years risk of incident stroke in both age- and sex- adjusted and

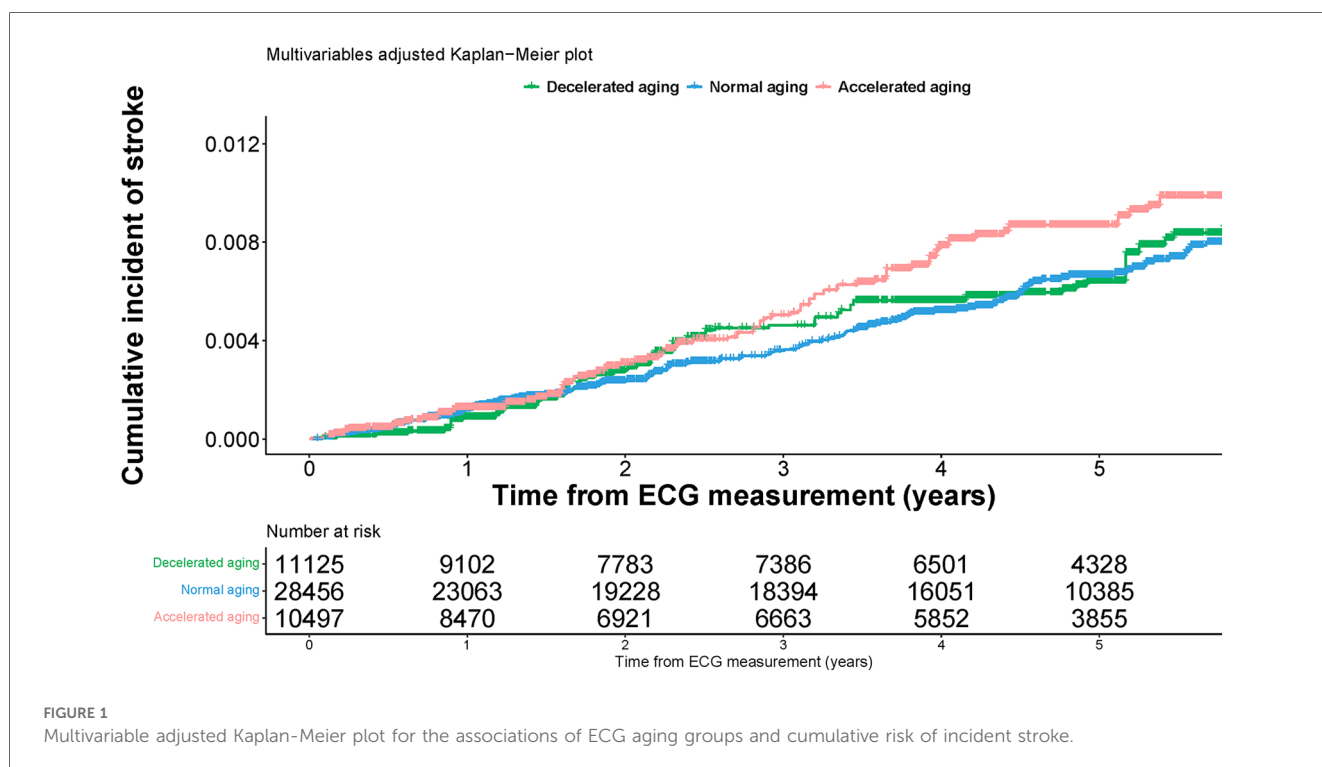
multivariable-adjusted models (Supplementary Table S2). The association of Δ age with incident stroke in the multivariable-adjusted model was attenuated when we excluded prevalent atrial fibrillation cases (Supplementary Table S3). However, the inclusion of Δ age did not significantly improve the prediction of incident stroke with a change of C-statistics of 0.003 [95% CI, -0.003 – 0.009] in the multivariable model. There was also no significant change in terms of net reclassification improvement (NRI) ($P > 0.05$).

Figure 1 shows the multivariable-adjusted cumulative risk of incident stroke across three aging groups. Participants in the accelerated aging group had a higher probability of developing stroke compared to the normal and decelerated aging groups (log-rank test P value = 0.02). We also fit a smoothing spline fit for the relative hazard of incident stroke as a function of Δ age (Supplementary Figure S1). The graph did not support a nonlinear relationship between Δ age and hazard of incident stroke ($P = 0.07$ for the nonlinear term).

TABLE 2 Associations of Δ age (per 10 years) and incident stroke.

	Chronological age and sex adjusted model					Multivariable adjusted model				
	Total	Incident cases	HR (95% CI)	<i>P</i> value	<i>P</i> interact	Total	Incident cases	HR (95% CI)	<i>P</i> value	<i>P</i> interact
All	67,757	379	1.38 (1.17, 1.64)	<0.001	–	50,100	282	1.22 (1.00, 1.49)	0.05	–
Baseline chronological age <60	18,266	60	1.06 (0.68, 1.64)	0.81	0.18	14,009	46	0.82 (0.49, 1.38)	0.45	0.12
Baseline chronological age \geq 60	49,491	319	1.45 (1.20, 1.74)	<0.001		36,091	236	1.31 (1.05, 1.63)	0.02	
Men	32,706	219	1.49 (1.19, 1.87)	<0.001	0.28	24,501	170	1.34 (1.03, 1.74)	0.03	0.25
Women	35,051	160	1.24 (0.96, 1.61)	0.11		25,599	112	1.07 (0.78, 1.47)	0.66	

Multivariable model adjusted for chronological age, sex, BMI, current smoking, drinking, physical activity, diet, systolic blood pressure, prevalent diabetes, prevalent CKD, and prevalent dyslipidemia.



Two additional analyses on comparisons of the stroke risk models based on PCE with or without ECG Δ age and revised FSRS (R-FSRS) with or without ECG Δ age were conducted. Although there was subtle improvement in the C statistics when Δ age was added to both models, these findings were not statistically significant (Supplementary Table S8).

Sensitivity analysis

In the age and sex-stratified analysis, Δ age was positively associated with stroke incidence in individuals older than 60 years and in men (HR = 1.31 [95% CI, 1.05–1.63] and 1.34 [95% CI, 1.03–1.74] respectively). However, the difference did not reach the pre-defined significance threshold (both with interaction *P* values greater than 0.05). Similar association results were seen in the incidence of ischemic stroke (Supplementary Table S4). Because we had very limited incident hemorrhagic stroke cases, we did not find any association of Δ age and hemorrhagic stroke (Supplementary Table S1). Moreover, in Supplementary Table S5, accelerated aging was associated with a 42% increase in incident stroke [HR = 1.42 (95% CI, 1.12–1.80)] in the age- and sex-adjusted model. Decelerated aging was associated with a 10% decrease in incident stroke, but this finding was not statistically significant.

Association of Δ age with prevalent stroke

As shown in Supplementary Table S6, participants with prevalent stroke were more likely to be male and have higher Δ age, BMI, and systolic blood pressure. Like their incident stroke counterparts, they were more likely to have prevalent DM, CKD, and dyslipidemia. Supplementary Table S7 shows the association between Δ age and prevalent stroke in the age- and sex-adjusted, and multivariable-adjusted models. There was a significant association between prevalent stroke and Δ age [OR, 1.28 (95% CI, 1.11–1.49)], suggesting that those with a history of stroke were 28% more likely to exhibit accelerated aging.

Discussion

In this study, we found that Δ age (the difference between DNN-estimated ECG-age and chronological age) was associated with incident stroke in UK Biobank. ECG-age was also associated with prevalent stroke even after adjusting for known risk factors. These results indicate that ECG-age may reflect an accelerated compromise in cardiac electrical function (11). Accelerated aging was able to predict incident stroke in the age- and sex-adjusted model. All in all, our findings suggest that ECG-age is connected to stroke in the community.

DNNs and other AI-related technologies are being increasingly implemented in healthcare. These technologies not only have the potential to transform various aspects of patient care, but also some studies are already suggesting that AI can perform as well

as if not better than humans at certain tasks (23). As mentioned previously, commonly used stroke risk stratification tools (i.e., PCE, FSRS) fail to incorporate newer risk markers including pro-BNP and left ventricular hypertrophy, which are associated with increased risk of stroke (24, 25). Moreover, traditional risk calculators like the PCE tend to under- or overestimate according to ethnicity, socioeconomic class, and presence of inflammatory states. Our results demonstrated that the addition of ECG Δ age in risk models like the PCE and R-FSRS could improve their accuracy in stroke risk assessment.

Current literature is rife with examples demonstrating significant relationships between specific ECG features, age, and CVD risk. A study by van der Wall et al. found that the most important features for the prediction of physiologic age were T wave morphology indices in leads V4 and V5, and *P* wave amplitude in leads aVR and II (26). Recent findings in population-based studies suggest that abnormal *P* wave terminal force in lead I, a marker of left atrial abnormality, is strongly associated with incident stroke (27). Moreover, lower heart variability has been shown to be associated with increased risk of CV events and mortality (28).

Similarly, previous studies have linked AI-generated ECG-age to cardiovascular risk and outcomes. A study by Raghunath et al. incorporated a DNN-based model using ECG waveforms, which predicted all-cause mortality (29). The ECG-age model developed by the CODE study was associated with all-cause mortality in multiple validation cohorts (12). One of these is the ELSA-Brazil cohort study (Brazilian Longitudinal Study of Adult Health), wherein accelerated aging was able to predict 1-year overall mortality (12). Others have examined the connection between Δ age and clinical outcomes. Chang et al. found that—compared to chronological age—older DNN-estimated ECG-age was correlated to all-cause mortality and CVD (30). Our study adds to these findings by demonstrating that ECG-age is also able to predict incident stroke in the community.

Unfortunately, since the DNN remains partly enigmatic in terms of interpretation, the ways in which ECG-age can explain cardiovascular risk may be complicated (11). Lima et al. found no significant differences between common ECG features (heart rate, *P* duration, QRS axis and duration, RR interval, and QTc interval) among subjects with accelerating, normal, or decelerating aging (12). To better understand this finding, the same study performed an analysis restricted to normal ECGs (per conventional standards) that reported a significant relationship between ECG-age and death. It appears that ECG-age prediction does not solely rely on traditional ECG abnormalities. This hypothesis is further corroborated by studies merging traditional and deep learning methods, suggesting that traditional ECG features alone do not fully explain the age prediction (19). Interestingly, DNN-estimated ECG-age appears to be a proxy for biological aging stemming from a single input, possibly capturing the residual risk from traditional and unknown risk factors (11). However, this lack of explainability could potentially be mitigated by incorporating aspects of the Bayesian approach. This approach has been used to infer 'heart age' from a patient's chronological age and sex by assessing and

quantifying an individual's expected vs. actual ECG findings (31). Various ECG parameters were used as modifiers of the Bayesian-predicted heart age including QRS-T angle, a strong prognostic factor for all-cause mortality and CV events (31, 32).

There are a few limitations to address in this study. First, use of deep learning does not allow for feature extraction or explainability, leading to uncertainty about how ECG-age detects stroke risk. Future research should concentrate on understanding the components associated with DNN-estimated ECG-age and investigating if ECG-age quantifies modifiable excess risk amenable to early intervention (11). Moreover, our study population was predominantly healthy, better educated, and of white European descent. On the other hand, the DNN model used to predict ECG-age in our study was trained by extracting features from the CODE dataset, which is derived from a population that is generally sicker than the UK Biobank cohort. Also, compared to the hold-out split or CODE-15% cohort, mean age and the prevalence of CV risk factors were higher in both validation cohorts used in the CODE study. As such, the generalizability of our results to more diverse populations is unknown. Lastly, we cannot infer the causality of ECG-age to stroke due to the observational nature of our study and potential residual confounding factors.

In conclusion, DNN-estimated ECG-age was associated with incident and prevalent stroke in a large population-based study. Further analysis is needed to determine if ECG-age can be used as a practical biomarker of stroke risk.

Data availability statement

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

Ethics statement

Ethical approval was not required for the study involving humans in accordance with the local legislation and institutional requirements. Written informed consent to participate in this study was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and the institutional requirements.

References

- Saini V, Guada L, Yavagal DR. Global epidemiology of stroke and access to acute ischemic stroke interventions. *Neurology*. (2021) 97(20 Suppl 2):S6–S16. doi: 10.1212/WNL.0000000000012781
- Feigin VL, Brainin M, Norrving B, Martins S, Sacco RL, Hacke W, et al. World stroke organization (WSO): global stroke fact sheet 2022. *Int J Stroke*. (2022) 17(1):18–29. doi: 10.1177/17474930211065917 Erratum in: *Int J Stroke*. 2022 Apr;17(4):478.
- Mukherjee D, Patil CG. Epidemiology and the global burden of stroke. *World Neurosurg*. (2011) 76(6 Suppl):S85–90. doi: 10.1016/j.wneu.2011.07.023
- Sharrief A, Grotta JC. Stroke in the elderly. *Handb Clin Neurol*. (2019) 167:393–418. doi: 10.1016/B978-0-12-804766-8.00021-2
- Chahine Y, Magoon MJ, Maidu B, Del Álamo JC, Boyle PM, Akoum N. Machine learning and the conundrum of stroke risk prediction. *Arrhythm Electrophysiol Rev*. (2023) 12:e07. doi: 10.15420/aer.2022.34
- Brindle PM, McConnachie A, Upton MN, Hart CL, Davey Smith G, Watt GC. The accuracy of the framingham risk-score in different socioeconomic groups: a prospective study. *Br J Gen Pract*. (2005) 55(520):838–45. PMID: 16281999

Author contributions

RL: Conceptualization, Writing – original draft, Writing – review & editing. BW: Data curation, Formal Analysis, Investigation, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing. MG: Writing – review & editing. AD: Writing – review & editing. NM: Writing – review & editing. AS: Writing – review & editing. DM: Conceptualization, Writing – review & editing. HL: Conceptualization, Methodology, Supervision, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article.

This study is supported by the American Heart Association Grant (20SFRN35360180), and NIH grants U01AG068221 and R01AG080670. DDM's time is supported by R01HL155343, U01HL146382, U54HL143541, and R33HL158541.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcvm.2024.1368094/full#supplementary-material>

7. Dritsas E, Trigka M. Stroke risk prediction with machine learning techniques. *Sensors (Basel)*. (2022) 22(13):4670. doi: 10.3390/s22134670
8. Dritsas E, Fazakis N, Kocsis O, Fakotakis N, Moustakas K. Long-Term hypertension risk prediction with ML techniques in ELSA database. In: Simos DE, Pardalos PM, Kotsireas IS, editors. *Learning and Intelligent Optimization. LION 2021. Lecture Notes in Computer Science*. Berlin/Heidelberg, Germany: Springer (2021). p. 113–20.
9. Fazakis N, Kocsis O, Dritsas E, Alexiou S, Fakotakis N, Moustakas K. Machine learning tools for long-term type 2 diabetes risk prediction. *IEEE Access*. (2021) 9:103737–57. doi: 10.1109/ACCESS.2021.3098691
10. Lloyd-Jones D. Strengths and limitations of the ASCVD risk score and what should go in the risk discussion. *American College of Cardiology*. (2014). Available online at: <https://www.acc.org/Latest-in-Cardiology/Articles/2014/08/25/14/48/Strengths-and-Limitations-of-the-ASCVD-Risk-Score-and-What-Should-Go-in-the-Risk-Discussion>
11. Brant LCC, Ribeiro AH, Pinto-Filho MM, Kornej J, Preis SR, Fetterman JL, et al. Association between electrocardiographic age and cardiovascular events in community settings: the framingham heart study. *Circ Cardiovasc Qual Outcomes*. (2023) 16(7):e009821. doi: 10.1161/CIRCOUTCOMES.122.009821
12. Lima EM, Ribeiro AH, Paixão GMM, Ribeiro MH, Pinto-Filho MM, Gomes PR, et al. Deep neural network-estimated electrocardiographic age as a mortality predictor. *Nat Commun*. (2021) 12(1):5117. doi: 10.1038/s41467-021-25351-7
13. Ott G, Schaubelt Y, Alcaraz JM, Haverkamp W, Strodtzoff N. Uncovering ECG changes during health aging using explainable AI. *arXiv*. (2023) 19(4):2310.07463v1. doi: 10.48550/arXiv.2310.07463
14. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med*. (2015) 12(3):e1001779. doi: 10.1371/journal.pmed.1001779
15. Zhang C, Miao X, Wang B, Thomas RJ, Ribeiro AH, Brant LCC, et al. Association of lifestyle with deep learning predicted electrocardiographic age. *Front Cardiovasc Med*. (2023) 10:1160091. doi: 10.3389/fcvm.2023.1160091
16. Woodfield R, Grant I, UK Biobank Stroke Outcomes Group, UK Biobank Follow-Up and Outcomes Working Group, Sudlow CL. Accuracy of electronic health record data for identifying stroke cases in large-scale epidemiological studies: a systematic review from the UK Biobank stroke outcomes group. *PLoS One*. (2015) 10(10):e0140533. doi: 10.1371/journal.pone.0140533
17. Woodfield R, UK Biobank Stroke Outcomes Group, UK Biobank Follow-up and Outcomes Working Group, Sudlow CLM. Accuracy of patient self-report of stroke: a systematic review from the UK Biobank stroke outcomes group. *PLoS One*. (2015) 10(9):e0137538. doi: 10.1371/journal.pone.0137538
18. Ribeiro AH, Ribeiro MH, Paixão GMM, Oliveira DM, Gomes PR, Canazart JA, et al. Automatic diagnosis of the 12-lead ECG using a deep neural network. *Nat Commun*. (2020) 11(1):1760. doi: 10.1038/s41467-020-15432-4 Erratum in: *Nat Commun*. 2020 May 1;11(1):2227.
19. Zvuloni E, Read J, Ribeiro AH, Ribeiro ALP, Behar JA. On merging feature engineering and deep learning for diagnosis, risk prediction and age estimation based on the 12-lead ECG. *IEEE Trans Biomed Eng*. (2023) 70(7):2227–36. doi: 10.1109/TBME.2023.3239527
20. Ribeiro ALP, Paixão GMM, Gomes PR, Ribeiro MH, Ribeiro AH, Canazart JA, et al. Tele-electrocardiography and bigdata: the CODE (clinical outcomes in digital electrocardiography) study. *J Electrocardiol*. (2019) 57S:S75–8. doi: 10.1016/j.jelectrocard.2019.09.008
21. Siontis KC, Noseworthy PA, Attia ZI, Friedman PA. Artificial intelligence-enhanced electrocardiography in cardiovascular disease management. *Nat Rev Cardiol*. (2021) 18(7):465–78. doi: 10.1038/s41569-020-00503-2
22. Inoue E. NRI for Risk Prediction Models with Time to Event and Binary Response Data. (2018). Available online at: <https://cran.r-project.org/web/packages/nricens/nricens.pdf> (Accessed October 27, 2023).
23. Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. *Future Healthc J*. (2019) 6(2):94–8. doi: 10.7861/futurehosp.6-2-94
24. Lind L, Siegbahn A, Lindahl B, Stenemo M, Sundström J, Ärnlöv J. Discovery of new risk markers for ischemic stroke using a novel targeted proteomics chip. *Stroke*. (2015) 46(12):3340–7. doi: 10.1161/STROKEAHA.115.010829
25. Bots ML, Nikitin Y, Salonen JT, Elwood PC, Malyutina S, de Concalves AF, et al. Left ventricular hypertrophy and risk of fatal and non-fatal stroke. EUROSTROKE: a collaborative study among research centres in Europe. *J Epidemiol Community Health*. (2002) 56(Suppl 1):i8–13. doi: 10.1136/jech.56.suppl_1.i8
26. van der Wall HEC, Hassing GJ, Doll RJ, van Westen GJP, Cohen AF, Selder JL, et al. Cardiac age detected by machine learning applied to the surface ECG of healthy subjects: creation of a benchmark. *J Electrocardiol*. (2022) 72:49–55. doi: 10.1016/j.jelectrocard.2022.03.001
27. Okin PM, Kamel H, Kjeldsen SE, Devereux RB. Electrocardiographic left atrial abnormalities and risk of incident stroke in hypertensive patients with electrocardiographic left ventricular hypertrophy. *J Hypertens*. (2016) 34(9):1831–7. doi: 10.1097/HJH.0000000000000989
28. Fang SC, Wu YL, Tsai PS. Heart rate variability and risk of all-cause death and cardiovascular events in patients with cardiovascular disease: a meta-analysis of cohort studies. *Biol Res Nurs*. (2020) 22(1):45–56. doi: 10.1177/1099800419877442 Erratum in: *Biol Res Nurs*. 2020;22(3):423–425.
29. Raghunath S, Ulloa Cerna AE, Jing L, van Maanen DP, Stough J, Hartzel DN, et al. Prediction of mortality from 12-lead electrocardiogram voltage data using a deep neural network. *Nat Med*. (2020) 26(6):886–91. doi: 10.1038/s41591-020-0870-z
30. Chang CH, Lin CS, Luo YS, Lee YT, Lin C. Electrocardiogram-based heart age estimation by a deep learning model provides more information on the incidence of cardiovascular disorders. *Front Cardiovasc Med*. (2022) 9:754909. doi: 10.3389/fcvm.2022.754909
31. Lindow T, Palencia-Lamela I, Schlegel TT, Ugander M. Heart age estimated using explainable advanced electrocardiography. *Sci Rep*. (2022) 12(1):9840. doi: 10.1038/s41598-022-13912-9
32. Aro AL, Huikuri HV, Tikkanen JT, Juhani Juntila M, Rissanen HA, Reunanen A, et al. QRS-T angle as a predictor of sudden cardiac death in a middle-aged general population. *EP Europace*. (2012) 14(6):872–6. doi: 10.1093/europace/eur393