ARIC Manuscript Proposal Form

ARIC Publication Admin Use Only

Publication Committee Review Date: **02/13/24**
ARIC Manuscript Proposal Number: **#4396**

1.a. Full Title: Short-term Repeatability of Artificial Intelligence Estimated Electrocardiographic Age

b. Abbreviated Title (Length 26 characters): **Repeatability of ECG-Age**

2. Writing Group [please provide a middle initial if available; EX: Adam L Williams]:
Writing group members: Katherine M. Conners, Varun Divi, Barrett Q. Buhler, Faisal F. Syed, Annie Green Howard, Eric A. Whitsel, Elsayed Z. Soliman, Christy L. Avery

I, the first author, confirm that all the coauthors have given their approval for this manuscript proposal. KMC **[please confirm with your initials electronically or in writing]**

First author [please provide a middle initial; EX: Adam L Williams]:
K **M** **C**
First name Middle initial Last name
Address: University of North Carolina Department of Epidemiology
123 W Franklin St.
Chapel Hill, NC 27516
Phone: 4847073886
E-mail: connersk@unc.edu

ARIC author to be contacted if there are questions about the manuscript and the first author does not respond or cannot be located (The ARIC author should be involved enough in ARIC to be able to point the lead author to appropriate ancillary study PIs and to be able to search ARIC manuscript proposals if the lead author doesn't have the access needed to do such a search).

Name: **Christy L. Avery**
Address: University of North Carolina Department of Epidemiology
123 W Franklin St.
Chapel Hill, NC 27516
Phone: 9199664312
E-mail: Christy_avery@unc.edu

3. Timeline: We propose to submit a short (<800 word) research letter with the below analysis. As our team has already constructed the artificial intelligence model needed to estimate ECG-age in these patients, we foresee the analysis and manuscript writing process to take no longer than six months after receiving approval.

4. Rationale:

Traditional methods for estimating risk of cardiovascular disease (CVD) rely primarily on vascular indicators such as blood pressure, cholesterol levels, and diabetes [1,2]. However, such estimates lack validation in individuals below the age of 40 and may inaccurately estimate risk in certain patient groups [3]. To enhance CVD risk stratification, assessments of cardiac structure and function are an attractive alternative given their independence from vascular risk factors and established ability to predict heart failure, stroke, and sudden cardiac death [4-7].

The electrocardiogram (ECG) is a fundamental diagnostic tool for CVD that yields extensive insight into one's cardiac structure and function. Artificial intelligence models can produce precise and impartial interpretations of ECG data and have recently been employed to estimate cardiac age from raw ECG waveforms [8]. AI-estimated ECG-age is correlated to, yet distinct from one's chronological age. This novel metric of cardiac age provides a window into the heart's functional capacity and is independent of traditional ECG features and existing biological age measures. When markedly higher than one's chronological age, AI-estimated ECG-age has successfully predicted myocardial infarction, atrial fibrillation, left ventricular systolic dysfunction, heart failure, and mortality [9-11].

Despite findings suggesting clinical implications for CVD risk assessment, there remains a scarcity of research addressing the consistency and repeatability of this novel measurement over time. Further, the modifiability of ECG-age through lifestyle and/or therapeutic intervention has been noted as a compelling yet unexplored area of research. Determining the short-term repeatability of this measurement, however, will be critical to ensure validity of future longitudinal analyses. Consistency of ECG-age measurements in the short-term will serve as a foundational step in bolstering the reliability and accuracy of future research endeavors on this metric.

5. Main Hypothesis/Study Aims: Aim: We will evaluate short-term repeatability of AI-estimated ECG-age measurements in 63 participants with repeat ECGs from the University of North Carolina's General Clinical Research Center (GCRC). Hypothesis: AI-estimated ECG-age will remain mostly unchanged with high repeatability between two short-term measurements (one to two weeks). Our findings will inform future longitudinal analyses of AI-estimated ECG-age trajectories and efforts integrating ECG-age into established risk prediction equations.

6. Design and analysis - please address the following aspects:

a) Inclusion/exclusion

Upon enrollment, participants were of good health and free of CVD, including the following conditions: diabetes, hypertension, prevalent heart failure, coronary heart disease, renal disease, chronic obstructive pulmonary disease, emphysema, utilization of class I or III antiarrhythmic medications, atrial fibrillation or flutter, and low-quality ECG-tracings (grade 5).

b) Study design

We propose a prospective panel study to assess the short-term repeatability of AI-estimated ECG-age measurements in the University of North Carolina's GCRC panel. Repeat ECG assessments were conducted on participants at two points, with one to two weeks between measurements. We will analyze available demographic data and AI-estimated ECG-age to assess short-term consistency of ECG-age measurements across participants representing community clinical practice.

c) Outcome and other variables of interest with specific reference to the time of their collection

Our primary outcome will be variation in AI-estimated ECG-age between two timepoints. Other variables of interest will include age, sex, race/ethnicity, and presence of comorbidities to describe the study population.

d) Summary of data analysis

We intend to estimate ECG-age for each participant by applying a previously developed deep neural network (DNN) machine learning model to raw ECG waveforms. In brief, our DNN

model was trained on roughly 17,000 12-lead ECGs from the publicly available PTB-XL electrocardiography dataset to estimate an individual's cardiac age through detection and extraction of features directly from the ECG waveform, not relying on traditional ECG features. The model utilizes a convolutional neural network adapted for one-dimensional signals to capture how aging impacts ECG waveforms, resembling a residual network used in image classification. We will apply this model to two 10-second standard 12-lead ECG recordings conducted by certified technicians using standardized ARIC protocols while participants rested in supine position with one to two weeks between measurements. ECGs were digitized and stored using the MAC PC Personal Cardiograph, sent via modem to the Epidemiological Cardiology Research (EPICARE) Center, and are currently stored in a secure Linux environment as raw XML files.

Once the DNN models are applied to each ECG to estimate ECG-age, we will estimate its participant-level variability between measurements at two timepoints as an intra-class correlation coefficient, i.e. the between-participant variance divided by the between + within-participant (total) variance. Additionally, we will estimate its minimally detectable change between the two measurements with 95% confidence, as follows: $MDC_{95} = SEM \times \sqrt{2} \times 1.96$, where SEM represents the standard error of measurement. Notably, these analyses follow structures and approaches established by prior manuscripts evaluating the repeatability of individual ECG metrics in the GCRC panel [12-14].

e) Any anticipated methodologic limitations or challenges if present

We do not foresee any methodologic challenges as members of our team have 1) worked with this specific dataset previously, and 2) built the machine learning model needed to estimate ECG-age successfully.

f) Will the author need Limited data to complete the proposed manuscript? ☐ Yes, Limited data is needed (Provide a brief (2-3 sentences) justification for requesting PHI data). ☒ No, De-identified data will be sufficient.

*Please note, Limited dataset access is strict and rarely provided. Limited data includes identifiable information such as dates (birthdays, visit dates, etc.). CMS, Genomic, Geocoded, Proteomics/Somalogic, and other -omic data all fall under the limited data category. De-identified data does not include dates. All dates are date adjusted to "Days since Visit 1".

7.a. Will the data be used for non-ARIC analysis or by a for-profit organization in this manuscript? (Non-ARIC analysis means that the authors are not regarded as ARIC

investigators and the "ARIC author" is essentially just a facilitator rather than an integral part of the writing group.) ☐ **Yes** ☒ **No**

b. If Yes, is the author aware that the current derived consent file ICTDER05 must be used to exclude persons with a value RES_OTH and/or RES_DNA = "ARIC only" and/or "Not for Profit" ? ☒ **Yes** ☐ **No**
(The file ICTDER is distributed to ARIC PIs annually, and contains the responses to consent updates related to stored sample use for research.)

8.a. Will the DNA data be used in this manuscript? ☐ **Yes** ☒ **No**

8.b. If yes, is the author aware that either DNA data distributed by the Coordinating Center must be used, or the current derived consent file ICTDER05 must be used to exclude those with value RES_DNA = "No use/storage DNA"? ☐ **Yes** ☐ **No** ☒ **N/A**

9. The lead author or the "sponsoring" ARIC author of this manuscript proposal has reviewed the list of existing ARIC Study manuscript proposals and has found no overlap between this proposal and previously approved manuscript proposals either published or still in active status. ARIC Investigators have access to the publications lists under the Study Members Area of the website at: https://aric.cscc.unc.edu/aric9/proposalsearch [ARIC Website□Publications□Proposal Search]

☒ Yes ☐ No

10. What are the most related manuscript proposals in ARIC (authors are encouraged to contact lead authors of these proposals for comments on the new proposal or collaboration)?
#3678 (AI ECG prediction of heart failure) and #3982 (AI ECG prediction of MI). Neither manuscript uses the GCRC data to evaluate repeatability.

11.a. Is this manuscript proposal associated with any ARIC ancillary studies or does it use current [or ongoing] ancillary study data (this includes ACHIEVE)? ☒ **Yes** ☐ **No → Skip to question 12**

11.b. If yes to 11.a., is the proposal

☒ **A. exclusively the result of an ancillary study**

☐ **B. primarily based on ARIC data with ancillary data playing a minor role (usually control variables)**

11.c. If yes to 11.a., list number 2002.05 - Repeatability of Short-Term Heart Rate Variability, UltraShort-Term Heart Rate Variability, and Spatial T Wave Axis

*ancillary studies are listed by number
https://aric.cscc.unc.edu/aric9/researchers/ancillary_studies/approved_ancillary_studies [ARIC Website Ancillary Studies Approved Ancillary Studies]

12a. Manuscript preparation is expected to be completed in one to three years. If a manuscript is not submitted for ARIC review at the end of the 3-years from the date of the approval, the manuscript proposal will expire.

12b. The NIH instituted a Public Access Policy in April, 2008 which ensures that the public has access to the published results of NIH funded research. It is **your responsibility to upload manuscripts to PubMed Central** whenever the journal does not and be in compliance with this policy. Four files about the public access policy from http://publicaccess.nih.gov/ are posted in https://aric.cscc.unc.edu/aric9/publications/policies_forms_and_guidelines [ARIC Website Publications Publication Policies, Forms, and Guidelines]. http://publicaccess.nih.gov/submit_process_journals.htm shows you which journals automatically upload articles to PubMed central.

References:

1. D'Agostino RB, Vasan RS, Pencina MJ, et al. General cardiovascular risk profile for use in primary care: the Framingham Heart Study. *Circulation.* 2008;117(6):743–753. doi: 10.1161/CIRCULATIONAHA.107.699579
2. Soureti A, Hurling R, Murray P, van Mechelen W, Cobain M. Evaluation of a cardiovascular disease risk assessment tool for the promotion of healthier lifestyles. *Eur J Cardiovasc Prev Rehabil.* 2010;17(5):519–523. doi: 10.1097/HJR.0b013e328337ccd3
3. Stone NJ, Smith SC, Jr., Orringer CE, et al. Managing atherosclerotic cardiovascular risk in young adults: JACC state-of-the-art review. *J Am Coll Cardiol.* 2022;79(8):819–836. doi: 10.1016/j.jacc.2021.12.016
4. Ha, E.T. *et al.* Relation of Left Ventricular Hypertrophy Subtype to Long-Term Mortality in Those With Subclinical Cardiovascular Disease (from the Multiethnic Study of Atherosclerosis [MESA]). *Am J Cardiol* **175**, 131-138 (2022).
5. Shenasa, M., Shenasa, H. & El-Sherif, N. Left ventricular hypertrophy and arrhythmogenesis. *Card Electrophysiol Clin* **7**, 207-20 (2015).
6. Agarwal, S.K. & Soliman, E.Z. ECG abnormalities and stroke incidence. *Expert Rev Cardiovasc Ther* **11**, 853-61 (2013).
7. Shenasa, M. & Shenasa, H. Hypertension, left ventricular hypertrophy, and sudden cardiac death. *Int J Cardiol* **237**, 60-63 (2017).
8. Syed FF, Gehi AK. Artificial Intelligence, Machine Learning, and Big Data Applications in Cardiac Electrophysiology and Arrhythmia Management. EP Lab Digest. Published September 2023. Accessed October 2023.

9. Chang CH, Lin CS, Luo YS, et al. 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

10. Lima EM, Ribeiro AH, Paixao GMM, 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

11. Brant LCC, Ribeiro AH, Pinto-Filho MM, 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. Schroeder EB, Whitsel EA, Evans GW, Prineas RJ, Chambless LE, Heiss G. Repeatability of heart rate variability measures. *J Electrocardiol*. 2004;37(3):163-172. doi:10.1016/j.jelectrocard.2004.04.004
13. Vaidean GD, Schroeder EB, Whitsel EA, et al. Short-term repeatability of electrocardiographic spatial T-wave axis and QT interval. *J Electrocardiol*. 2005;38(2):139-147. doi:10.1016/j.jelectrocard.2004.09.020
14. Snyder ML, Soliman EZ, Whitsel EA, Gellert KS, Heiss G. Short-term repeatability of electrocardiographic P wave indices and PR interval. *J Electrocardiol*. 2014;47(2):257-263. doi:10.1016/j.jelectrocard.2013.11.007

To view publications materials, click "Log in" at the top right of the ARIC website. Click "Forgot Password" if you are experiencing issues with logging in.