Journal of Electrocardiology 76 (2023) 61–65

Contents lists available at ScienceDirect

Journal of Electrocardiology

journal homepage: www.jecgonline.com

An ECG-based machine learning model for predicting new-onset atrial fibrillation is superior to age and clinical features in identifying patients at high stroke risk

Sushravya Raghunath, PhD [a,b,1], John M. Pfeifer, MD, MPH [a,b,1], Christopher R. Kelsey, BS [a], Arun Nemani, PhD [b], Jeffrey A. Ruhl, MS [a], Dustin N. Hartzel, BS [a], Alvaro E. Ulloa Cerna, PhD [a], Linyuan Jing, PhD [a], David P. vanMaanen, MS [a], Joseph B. Leader, BA [a,b], Gargi Schneider, MD [a], Thomas B. Morland, MD [a], Ruijun Chen, MD [a,b], Noah Zimmerman, PhD [b], Brandon K. Fornwalt, MD, PhD [a,b,2], Christopher M. Haggerty, PhD [a,b,*]

[a] Geisinger, Danville, PA, USA
[b] Tempus Labs Inc., Chicago, IL, USA

ARTICLE INFO

Keywords:
Atrial fibrillation
Ischemic stroke
Machine learning
Artificial intelligence

ABSTRACT

Background:

Several large trials have employed age or clinical features to select patients for atrial fibrillation (AF) screening to reduce strokes. We hypothesized that a machine learning (ML) model trained to predict AF risk from 12-lead electrocardiogram (ECG) would be more efficient than criteria based on clinical variables in indicating a population for AF screening to potentially prevent AF-related stroke.

Methods:

We retrospectively included all patients with clinical encounters in Geisinger without a prior history of AF. Incidence of AF within 1 year and AF-related strokes within 3 years of the encounter were identified. AF-related stroke was defined as a stroke where AF was diagnosed at the time of stroke or within a year after the stroke. The efficiency of five methods was evaluated for selecting a cohort for AF screening. The methods were selected from four clinical trials (mSToPS, GUARD-AF, SCREEN-AF and STROKESTOP) and the ECG-based ML model. We simulated patient selection for the five methods between the years 2011 and 2014 and evaluated outcomes for 1 year intervals between 2012 and 2015, resulting in a total of twenty 1-year periods. Patients were considered eligible if they met the criteria before the start of the given 1-year period or within that period. The primary outcomes were numbers needed to screen (NNS) for AF and AF-associated stroke.

Results:

The clinical trial models indicated large proportions of the population with a prior ECG for AF screening (up to 31%), coinciding with NNS ranging from 14 to 18 for AF and 249–359 for AF-associated stroke. At comparable sensitivity, the ECG ML model indicated a modest number of patients for screening (14%) and had the highest efficiency in NNS for AF (7.3; up to 60% reduction) and AF-associated stroke (223; up to 38% reduction).

Conclusions:

An ECG-based ML risk prediction model is more efficient than contemporary AF-screening criteria based on age alone or age and clinical features at indicating a population for AF screening to potentially prevent AF-related strokes.

* Corresponding author at: Geisinger, 100 North Academy Avenue, Danville, PA 17822-4400, USA.
E-mail address: chaggerty3@gatech.edu (C.M. Haggerty).
[1] Contributed equally.
[2] Contributed equally.

https://doi.org/10.1016/j.jelectrocard.2022.11.001

Available online 8 November 2022
0022-0736/© 2022 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Introduction

Atrial fibrillation (AF) is a significant risk factor for ischemic stroke, which can be substantially mitigated with oral anticoagulation therapy [1]. However, in over 25% of strokes related to AF, stroke is the presenting manifestation [2], which implies that improved detection of asymptomatic AF would help reduce the incidence of stroke. Numerous large clinical trials in recent years have thus sought to evaluate various age-based or other clinical criteria as the basis for effective and efficient screening for undiagnosed AF [3–6]. These studies have generally found that such screening improves the yield of new AF diagnoses, but the non-specific criteria employed have resulted in low-efficiency screens that are challenging to scale. In fact, the US Preventive Services Task Force (USPSTF) specifically cited this limited accuracy as part of its assertion of insufficient evidence to support systematic AF screening [7].

Advances in machine learning (ML) may provide new opportunities to shift this paradigm. Several recent studies have demonstrated potential for ML models trained on large digital electrocardiogram (ECG) datasets to identify patients with undiagnosed AF and/or at risk for developing new-onset AF [8–10]. We have previously reported [8] the model architecture and training paradigm of predictive models that were capable of identifying patients at risk for AF and AF-related stroke in a simulated retrospective deployment, with numbers needed to screen (NNS) over a five-year span of 9 and 162, respectively. Such findings suggest that an opportunistic screening approach informed by an ML model would be more efficient than the primarily age-based approaches used in recent trials.

The objective of this study was to quantify this potential improvement by directly comparing the performance characteristics and diagnostic yield of patients indicated for screening by the ECG-based ML model against the patients who would have been indicated for screening as part of four different trials: mHealth Screening to Prevent Strokes (mSToPS) [3], STROKESTOP [4], SCREEN-AF [5], and reducinG stroke by screening for UndiAgnosed atRial fibrillation in elderly inDividuals (GUARD-AF) [11]. We hypothesized that case ascertainment using ML predictions would be more efficient than the trial-based criteria based on lower NNS.

Materials and methods

The Geisinger Institutional Review Board approved this retrospective analysis with a waiver of consent.

Data and endpoints

We identified patients with an ECG study or encounters in internal medicine, family medicine, or cardiology that were not canceled or no-show at a Geisinger clinic between 2011 and 01-01 to 2015-12-31 with follow-up available until Aug 2021 from the Geisinger electronic health record (EHR; Epic Systems, Madison, WI). Further, the population was restricted to patients with any history of an encounter with cardiology or cardiac-related tests, such as ECG. Finally, the included population for this study was defined as patients over 50-years-old with no known prior AF and with sufficient follow-up of at least 1 year in the absence of a documented AF diagnosis within a year of encounter (Fig. 1). Several phenotypes and endpoints were defined with respect to each encounter:

1. Documentation of new-onset AF within 1 year after the encounter
2. Incidence of AF-related stroke within three years after the encounter
3. CHA_2DS_2-VASc score [12]
4. ML model-based prediction of AF risk [8] (for ECG encounters only)
5. Patient eligibility for mSToPS [3], GUARD-AF [11], SCREEN-AF [5], or STROKESTOP [4] based on respective inclusion and exclusion criteria.
6. Available follow-up for censoring with at least one of the following qualifying encounters: ECG, echocardiography, outpatient visit with internal medicine, family medicine or cardiology, any inpatient encounter or any surgical procedure.

Details for each of these definitions are provided below.

EHR-based definitions of atrial fibrillation, stroke, and CHA_2DS_2-VASc[12]

New-onset AF was defined based on the presence of a diagnosis code for AF (including atrial flutter) on the patient problem list, >1 inpatient or outpatient encounters, or a reported finding of AF on a 12-lead ECG. Diagnoses occurring within 30 days following cardiac surgery were excluded. As noted in prior work, this definition was found to have high positive (94%) and negative (98%) predictive values (PPV and NPV, respectively) via physician chart review in a balanced sample of 200 charts [8].

Incidence of ischemic stroke was defined based on either an internal "Get with the Guidelines" registry or a custom phenotype comprising an emergency department-to-hospital admission encounter coded as a stroke with a concurrent head computed tomography or magnetic resonance imaging study. On a blinded physician review of 200 outcome-balanced charts, this algorithm achieved 85% PPV and 100% NPV. Of note, the stroke registry captured data from the 3 main Geisinger hospitals during the inclusion and follow-up observational periods (2011-01-01 to 2018-12-31). A stroke was considered AF-related if AF was first identified at the time of the stroke or up to 1 year after the stroke.

CHA_2DS_2-VASc [12] scores were computed for every encounter based on the results of validated EHR-derived phenotypes for heart failure, hypertension, diabetes, stroke, transient ischemic attack, thromboembolism, myocardial infarction, and peripheral artery disease. Each encounter was classified as to whether the CHA_2DS_2-VASc score was ≥2.

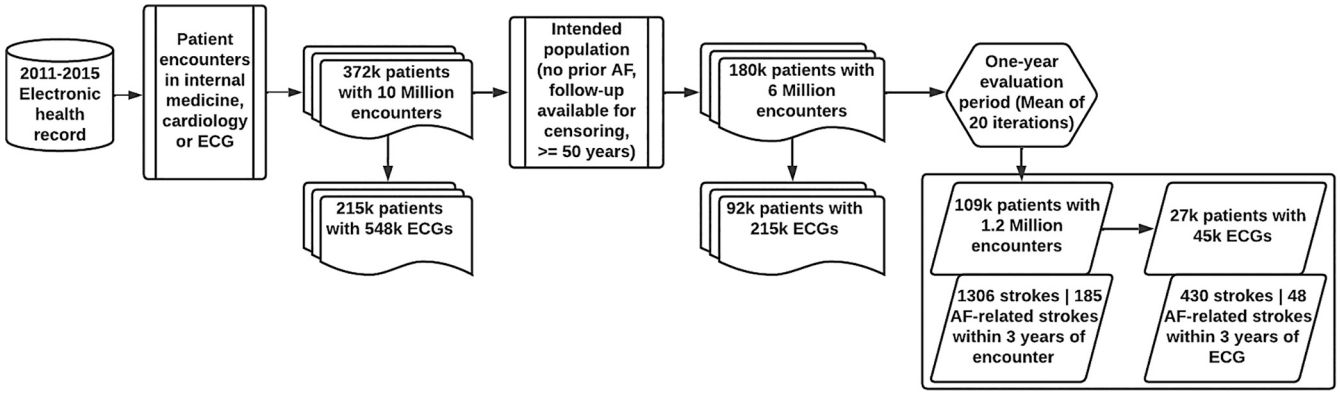

Fig. 1. Flowchart illustration of data ingestion and preparation.

Machine learning model predictions

The development of the ML model used to predict new-onset AF followed the method described previously [8]. Briefly, this study simulated a clinical deployment in which 320,000 digital 12-lead ECG studies (100,000 unique patients) from the Geisinger ECG database (GE MUSE) acquired between 1984 and 2009 were used to train a deep convolutional neural network model to predict risk of new diagnosis of AF within 1 year of ECG. Model inputs comprised solely of digital voltage data for eight non-derived leads (I, II, and V1-V6), as well as patient age and sex. Voltage traces were sampled at 500 Hz for 10 s for all leads (upsampled from 250 Hz, as needed). The full model with raw prediction scores and variable operating points is identified as the 'full ECG ML model', and the specific operating points chosen for comparison to the trial models are subsequently identified as the 'ECG ML model'. These fixed points were selected to yield expected sensitivities of 30% and 50% based on performance in the training set.

Clinical trial eligibility

Within the broad dataset of patients described above, we further implemented the inclusion/exclusion criteria from the comparison trials to identify the earliest date at which a given patient met criteria for each trial. We note that these criteria were not used to include or exclude patients from our analysis, but to classify the subset of patients who were indicated for screening with respect to each trial. A patient was considered indicated for screening by a given trial during the time interval after they first met inclusion criteria but before they first met exclusion criteria (as applicable) as follows:

a) mSToPS [3]: Included patients ≥75 or ≥ 55/65 years (male/female) with 1 or more specified comorbidities (prior cerebrovascular accident, heart failure, diagnosis of diabetes and hypertension, mitral valve disease, left ventricular hypertrophy, chronic obstructive pulmonary disease requiring home oxygen, sleep apnea, history of pulmonary embolism, history of myocardial infarction or obesity); excluded patients with atrial tachycardia, already prescribed anticoagulation, or with pacemaker or defibrillator.
b) SCREEN-AF [5]: Included patients ≥75 years taking an antihypertensive medication; excluded patients already receiving oral anticoagulation or with an implanted pacemaker, defibrillator, or loop recorder.
c) GUARD-AF [3,11]: Included patients ≥70 years; excluded patients taking oral anticoagulation, with a contraindication to oral anticoagulation, or with an implanted pacemaker, defibrillator, or loop recorder.
d) STROKESTOP [4,5]: Included patients 75 or 76 years.

Study design and analysis

We sampled 20 random, potentially overlapping 1-year intervals between 2011 and 01-01 and 2015-12-31 using a bootstrap procedure to serve as separate evaluation periods. Any patient used in the ECG ML model training was excluded. For each 1-year interval, we selected a random ECG encounter for each patient, excluding patients with no ECG in the period. Determination of AF screening indication for each model was defined as of that encounter date. The AF endpoint was considered positive if a new AF diagnosis was clinically documented within 1 year after the encounter date. The stroke endpoint was considered positive if an AF-associated stroke occurred within 3 years after the encounter date.

In a secondary analysis, we evaluated the full population (returning all patients without an ECG who had been excluded). An ECG encounter was chosen if the patient had an ECG in the interval, to serve as the required input for the ECG ML model prediction. Otherwise, we selected a random encounter for patients without a completed ECG to serve as the date for indicated AF screening. If a prior ECG was available within the year prior to that random encounter, the ML model score from that ECG was imputed via age-shifting to that encounter. If no prior ECG was available, a model score was set to 0 (no indicated screening).

The analysis was completed independently for each 1-year interval and the results were aggregated to report mean and 95% confidence intervals. The primary outcomes were NNS to detect a new case of 1-year incident AF and a 3-year incident AF-related stroke. We performed a two-tailed paired *t*-test to estimate differences at a significance level of $p = 0.01$ to account for five comparisons.

Results

Patients and clinical outcomes

The demographics and characteristics of patients included in the study are provided in Table 1. Across each interval evaluated, the mean 1-year incidence of new AF diagnoses within the ECG cohort was 4.0% and the mean number of AF-related stroke events was 48. Of note, 38% of new AF and 26% of AF-related strokes occurred in patients with an ECG in a given period. Furthermore, the majority of AF diagnoses in cases of AF-related stroke occurred in the first month after stroke (Supplementary Table 1).

Screening performance - ECG encounters

Results for each model applied to patients with an ECG encounter within each period are shown in Fig. 2 and Table 2. As shown by the curves in Fig. 2, the 'full' ECG ML model achieved higher sensitivity for AF and AF-associated stroke for a given size of the population indicated for screening (depending on the operating point selected). At an expected sensitivity of 50%, the ECG ML model indicated 14.4% of the population for screening and had the highest sensitivity and significantly lower NNS for AF of 7.3 compared with the other models ($p < 0.0001$ for all the four comparisons; up to 60% reduction). For AF-related stroke, the NNS was again lower (223.1), although the sensitivity of mSToPS and GUARD-AF was higher, with twice the number of patients indicated for screening. As expected, the alternate operating point of 30% sensitivity indicated few patients for screening (6.4%) and further reduced NNS for AF (5.3) and AF-related stroke (196.3) with the tradeoff of even lower sensitivity compared with mSToPS and GUARD-AF.

Screening performance - All encounters

Results for each model applied to patients with any encounter within a given 1-year period are shown in Supplementary Table 2. The models varied considerably in the percentage of patients indicated for screening, with the ECG ML model (@30% sensitivity) having the lowest percentage (1.9%) and mSToPS having the highest (31.0%). The

Table 1
Average demographic characteristics of the patient population from 20 iterations of a 1-year time window used for analysis.

Cohort: Age > 50 & no prior history of AF with internal medicine, cardiology or ECG encounters	All encounters	ECG encounters
Number of unique patients	109,233	27,446
Age in years (median)	64	65
Sex - female	55%	54%
Race - white	98%	97%
1-year AF events	2908 (2.6%)	1103 (4.0%)
3-year stroke events	1306 (1.2%)	430 (1.6%)
3-year AF-related stroke events	185 (0.2%)	48 (0.2%)
CHA_2DS_2-VASc score (mean | median)	1.6 | 1.0	1.8 | 2.0
CHA_2DS_2-VASc score ≥ 2	46%	48%
Heart failure	4%	4%
Diabetes Mellitus (I & II)	21%	18%

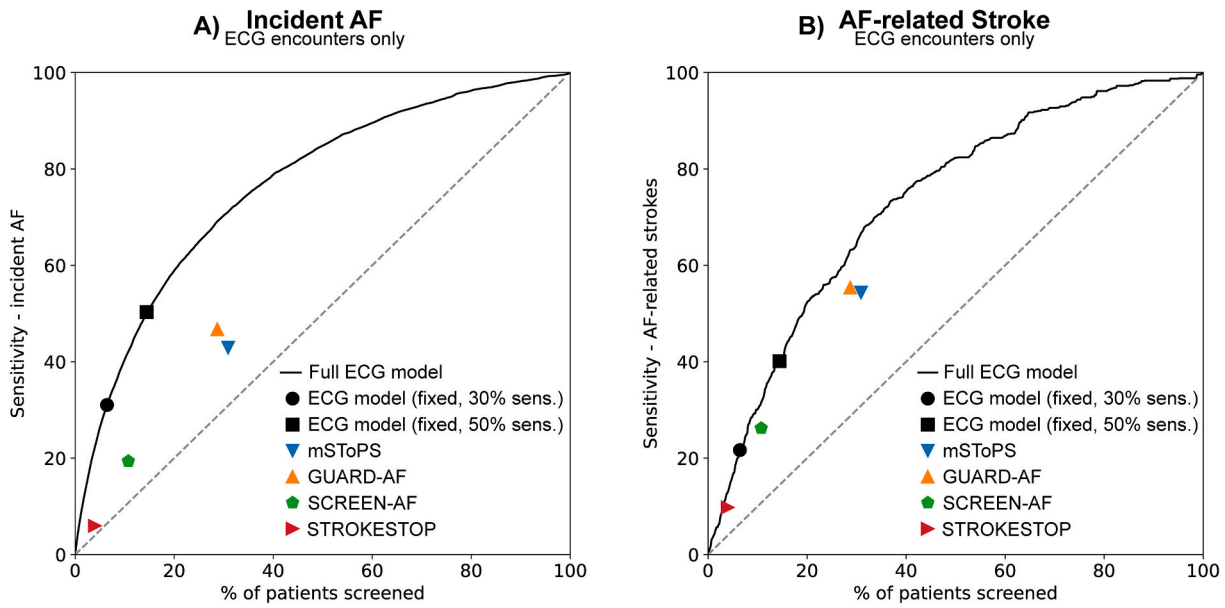

Fig. 2. Cumulative gain plots depicting proportion of stratified patients indicated for screening (x-axis) and sensitivity (y-axis) based on model-specific criteria. The left and right figures illustrate the cumulative gain curves as mean values of 20 iterations for prediction of incident AF within 1 year of the ECG encounter and incident AF-related stroke within 3-years of the ECG encounter.

Table 2
The performance statistics of varying models for indicating patients for screening based on data from ECG encounters.

		ECG ML model @ 30% sensitivity	ECG ML model @ 50% sensitivity	mSToPS	GUARD-AF	SCREEN-AF	STROKESTOP
Data	N			27,446 [25907–28,984]			
	AF incidence (%)			4.0 [3.8–4.1]			
Indicated for screening	n (%)	6.4 [6.2–6.6]	14.4 [14.1–14.7]	30.9 [30.5–31.3]	28.7 [28.5–29]	10.7 [10.8–10.9]	2.1 [2.0–2.1]
	Mean CHA_2DS_2-VASc score	2.6 [2.6–2.6]	2.5 [2.5–2.5]	2.3 [2.3–2.3]	2.7 [2.7–2.7]	3.3 [3.3–3.3]	3.1 [3.1–3.1]
	CHA_2DS_2-VASc ≥2 (%)	79.4 [78.6–80.2]	76.6 [75.8–77.3]	71.5 [71.1–71.9]	89.5 [89.4–89.7]	100.0 [100.0–100.0]	100.0 [100.0–100.0]
New-onset AF	Sensitivity (%)	31.0 [30.1–31.9]	50.3 [49.8–50.8]	42.9 [42.3–43.6]	46.7 [46.5–47]	19.4 [19.0–19.7]	3.2 [2.9–3.5]
	Specificity (%)	94.6 [94.4–94.7]	87.0 [86.8–87.3]	69.6 [69.3–69.9]	71 [71.7–72.3]	89.6 [89.5–89.8]	98.0 [97.9–98.0]
	NNS per case	5.3 [5–5.5]	7.3 [7.1–7.5]	18.3 [17.5–19.1]	15.6 [14.8–16.4]]	14.1 [13.3–14.9]	16.8 [15.4–18.1]
AF-related strokes	Total strokes	47.7 [41.8–53.6]					
	Sensitivity (%)	21.7 [19.6–23.8]	40.1 [37.3–42.9]	54.4 [52.1–56.6]	55.4 [54.1–56.7]	26.2 [24.2–28.2]	5.5 [4.7–6.3]
	NNS per case	196.3 [156.5–236]	223.1 [200.4–245.9]	359.4 [300.3–418.5]	317.1 [284.7–349.6]	259.4 [223.3–295.7]	248.8 [218.6–279.1]

sensitivity for capturing 1-year new-onset AF within the screening population similarly varied from 6.4% (STROKESTOP) to 47.9% (GUARD-AF), corresponding with an NNS-AF of 22.4 and 21.1, respectively. As expected, the ECG ML model maintained the lowest NNS-AF, but had poor sensitivity within the group of patients with no recent ECG available for analysis. Similar trends were observed with respect to AF-related strokes.

Discussion

AF remains an underdiagnosed disease, as evidenced by the high rates of patients who suffer a stroke as their initial presentation of AF. Improved recognition of AF in these minimally symptomatic or asymptomatic individuals therefore has the potential to significantly impact patient outcomes through stroke reduction. However, recommendations for AF screening remain a point of debate, with numerous trials employing different screening criteria and international guidelines (Europe [13] vs. US[7]) reaching diverging conclusions. While part of this debate justifiably centers around the efficacy of stroke prevention in screen-detected AF, consideration for the accuracy and efficiency of the screening approach also remains a concern because the burden on the healthcare system could become unreasonable if too many patients are indicated for test-based screening.

The main finding of the present work was that an ECG-based ML model trained to predict risk for new AF within 1 year was more efficient than contemporaneous screening approaches employed by large clinical trials for detecting AF after an ECG. As demonstrated by Fig. 2, across its full range of operating points, the ML model could match the fixed sensitivity of the clinical trial models for both AF and AF-related stroke while screening fewer patients. To better evaluate performance and

tradeoffs along that curve, we reported data at two operating points, corresponding with 30% and 50% sensitivities. The 50% sensitivity operating point yielded AF and AF-related stroke rates comparable to mSToPS and GUARD-AF. However, this performance still indicated over 14% of patients for screening, which may represent a considerable burden on the clinical personnel responsible for model implementation. Notably, this includes concerns of 'alert fatigue' and the implications of high false positive rates on end users. In contrast, the 30% sensitivity point indicated only 6.4% of the population, which was further enriched for disease (NNS = 5.3), and thus may be a more pragmatic option in some settings despite decreased sensitivity.

Several limitations of this work must be acknowledged. This study represents a retrospective simulation of screening performance within a single healthcare system with a predominantly white patient population, so the generalizability of the model to new centers and populations, and similarity to real-world outcomes and adherence to screening must be explicitly demonstrated through future work. However, the reported AF screening performance and the determination of AF-associated strokes were both reliant on clinically recognized AF, whereas a prospective screening effort is likely to uncover otherwise unrecognized AF. Consequently, these findings could represent an underestimate of the potential for new AF detection, although the optimal timing and duration of model-directed patch monitoring to maximize that detection remain an open question that ongoing prospective efforts should help inform (http://ClinicalTrials.gov identifier: NCT05442203). Additionally, as evidenced by results within the full population (not restricted to those with available ECGs), a large proportion of patients presenting with new AF or AF-related stroke did not have an ECG in the preceding 1–3 years, representing a fundamental limitation for this approach. Given the increasing evidence for opportunities in AI-based ECG analysis as well as wearable technology with integrated ECG capabilities, it is feasible that availability of ECG data will continue to increase and close this gap. Finally, the ECG ML model predictions are based on a convolutional neural network, and thus are not inherently explainable or interpretable. While attempts at saliency mapping [14] (Supplementary Fig. 1) often align with intuitive expectations, further work is needed to better interpret model output [15].

Conclusions

An ML ECG-based risk prediction model is more efficient (lower NNS at a similar sensitivity) than contemporary AF-screening criteria in selecting a population for additional AF and AF-related stroke screening.

Disclosures

Geisinger investigators receive funding from Tempus for ongoing development of predictive modeling technology. Tempus and Geisinger have jointly applied for predictive modeling patents. None of the Geisinger investigators have ownership interest in any of the intellectual property resulting from the partnership.

CRediT authorship contribution statement

Sushravya Raghunath: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Visualization, Project administration. **John M. Pfeifer:** Conceptualization, Methodology, Writing – review & editing, Supervision. **Christopher R. Kelsey:** Software, Validation, Resources, Data curation, Writing – review & editing. **Arun Nemani:** Methodology, Software, Writing – review & editing. **Jeffrey A. Ruhl:** Validation, Data curation, Project administration, Writing – review & editing. **Dustin N. Hartzel:** Software, Validation, Resources, Data curation, Writing – review & editing. **Alvaro E. Ulloa Cerna:** Methodology, Validation, Resources, Data curation, Writing – review & editing. **Linyuan Jing:** Methodology, Validation, Resources, Data curation, Writing – review & editing. **David P. vanMaanen:** Software, Resources, Data curation, Writing – review & editing. **Joseph B. Leader:** Validation, Resources, Data curation, Writing – review & editing. **Gargi Schneider:** Validation, Data curation, Writing – review & editing. **Thomas B. Morland:** Validation, Data curation, Writing – review & editing. **Ruijun Chen:** Validation, Data curation, Writing – review & editing. **Noah Zimmerman:** Resources, Writing – review & editing, Supervision. **Brandon K. Fornwalt:** Conceptualization, Methodology, Writing – review & editing, Supervision, Funding acquisition. **Christopher M. Haggerty:** Conceptualization, Methodology, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Visualization, Supervision, Project administration, Funding acquisition.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jelectrocard.2022.11.001.

References

[1] Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham study. Stroke 1991;22:983–8.
[2] Freedman B, Potpara TS, Lip GYH. Stroke prevention in atrial fibrillation. Lancet 2016;388:806–17.
[3] Steinhubl SR, et al. Effect of a home-based wearable continuous ECG monitoring patch on detection of undiagnosed atrial fibrillation: the mSToPS randomized clinical trial. JAMA 2018;320:146–55.
[4] Svennberg E, et al. Mass screening for untreated atrial fibrillation: the STROKESTOP study. Circulation 2015;131:2176–84.
[5] Gladstone DJ, et al. Screening for atrial fibrillation in the older population: a randomized clinical trial. JAMA Cardiol 2021;6:558–67.
[6] Ashburner JM, et al. Design and rationale of a pragmatic trial integrating routine screening for atrial fibrillation at primary care visits: the VITAL-AF trial. Am Heart J 2019;215:147–56.
[7] US Preventive Services Task Force, et al. Screening for atrial fibrillation: us preventive services task force recommendation statement. JAMA 2022;327 (360–367).
[8] Raghunath S, et al. 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 2021;143:1287–98.
[9] Khurshid S, et al. ECG-based deep learning and clinical risk factors to predict atrial fibrillation. Circulation 2022;145:122–33.
[10] Attia ZI, et al. An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction. Lancet 2019;394:861–7.
[11] Singer DE, et al. ReducinG stroke by screening for UndiAgnosed atRial fibrillation in elderly inDividuals (GUARD-AF): rationale and design of the GUARD-AF randomized trial of screening for atrial fibrillation with a 14-day patch-based continuous ECG monitor. Am Heart J 2022. https://doi.org/10.1016/j.ahj.2022.04.005.
[12] Lip GYH, Nieuwlaat R, Pisters R, Lane DA, Crijns HJGM. Refining clinical risk stratification for predicting stroke and thromboembolism in atrial fibrillation using a novel risk factor-based approach: the euro heart survey on atrial fibrillation. Chest 2010;137:263–72.
[13] Hindricks G, et al. 2020 ESC guidelines for the diagnosis and management of atrial fibrillation developed in collaboration with the European Association for Cardio-Thoracic Surgery (EACTS): the task force for the diagnosis and management of atrial fibrillation of the European Society of Cardiology (ESC) developed with the special contribution of the European heart rhythm association (EHRA) of the ESC. Eur Heart J 2021;42:373–498.
[14] Selvaraju RR, et al. Grad-CAM: visual explanations from deep networks via gradient-based localization. In: 2017 IEEE International Conference on Computer Vision (ICCV); 2017. https://doi.org/10.1109/iccv.2017.74.
[15] Ghassemi M, Oakden-Rayner L, Beam AL. The false hope of current approaches to explainable artificial intelligence in health care. The Lancet Digital Health 2021;3: e745–50.