ESC European Society of Cardiology

European Heart Journal - Digital Health (2022) **3**, 228–235
https://doi.org/10.1093/ehjdh/ztac023

ORIGINAL ARTICLE

Artificial intelligence—electrocardiography to detect atrial fibrillation: trend of probability before and after the first episode

Georgios Christopoulos[1], Zachi I. Attia[1], Holly K. Van Houten[2], Xiaoxi Yao[2,3], Rickey E. Carter[4], Francisco Lopez-Jimenez [1], Suraj Kapa [1], Peter A. Noseworthy[1], and Paul A. Friedman[1,*]

[1]Department of Cardiovascular Medicine, Mayo Clinic, Rochester, MN 55905, USA; [2]Robert D. and Patricia E. Kern Center for the Science of Health Care Delivery, Rochester, MN, USA; [3]Division of Health Care Delivery Research, Mayo Clinic, Rochester, MN, USA; and [4]Department of Health Sciences Research, Mayo Clinic, Jacksonville, FL, USA

Received 18 November 2021; revised 9 March 2022; online publish-ahead-of-print 9 May 2022

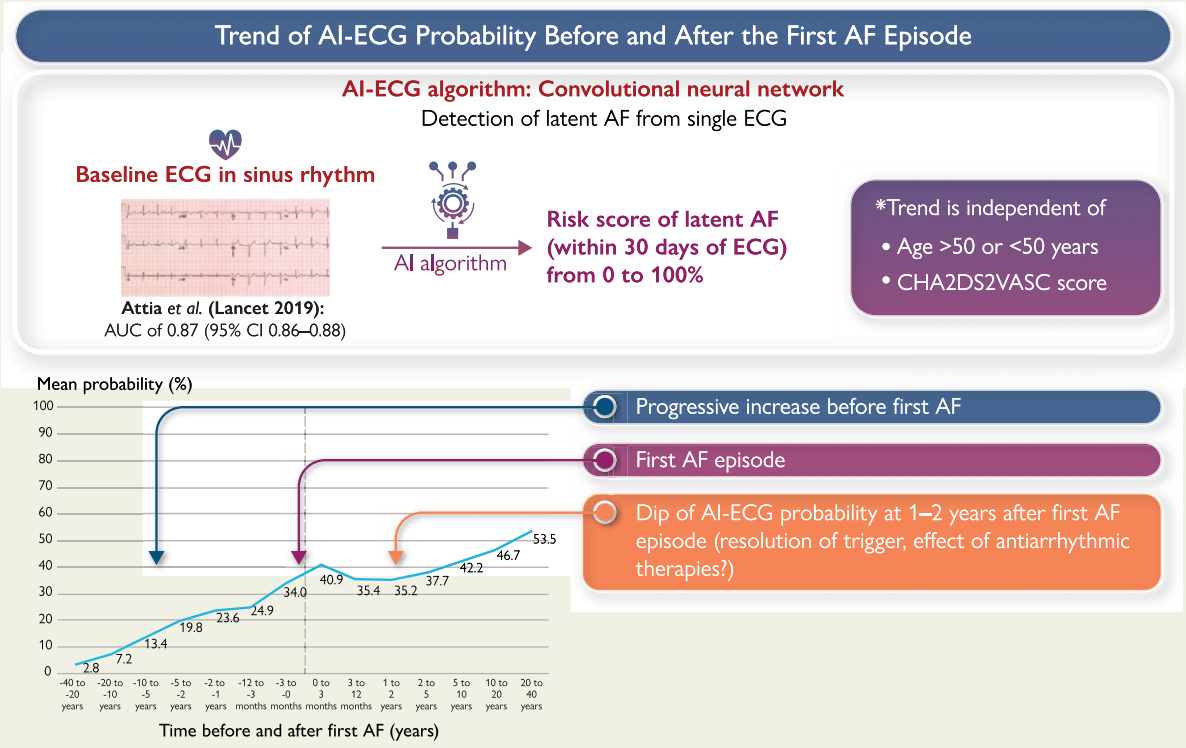
Aims Artificial intelligence (AI) enabled electrocardiography (ECG) can detect latent atrial fibrillation (AF) in patients with sinus rhythm (SR). However, the change of AI-ECG probability before and after the first AF episode is not well characterized. We sought to characterize the temporal trend of AI-ECG AF probability around the first episode of AF.

Methods and results We retrospectively studied adults who had at least one ECG in SR prior to an ECG that documented AF. An AI network calculated the AF probability from ECGs during SR (positive defined >8.7%, based on optimal sensitivity and specificity). The AI-ECG probability was reported prior to and after the first episode of AF and stratified by age and CHA_2DS_2-VASc score. Mixed effect models were used to assess the rate of change between time points. A total of 59 212 patients with 544 330 ECGs prior to AF and 413 486 ECGs after AF were included. The mean time between the first positive AI-ECG and first AF was 5.4 $\pm$ 5.7 years. The mean AI-ECG probability was 19.8% 2–5 years prior to AF, 23.6% 1–2 years prior to AF, 34.0% 0–3 months prior to AF, 40.9% 0–3 months after AF, 35.2% 1–2 years after AF, and 42.2% 2–5 years after AF ($P < 0.001$). The rate of increase prior to AF was higher for age >50 years CHA_2DS_2-VASc score $\geq$4.

Conclusion The AI-ECG probability progressively increases with time prior to the first AF episode, transiently decreases 1–2 years following AF and continues to increase thereafter.

* Corresponding author. Email: friedman.paul@mayo.edu

Graphical Abstract



Keywords Artificial intelligence • Atrial fibrillation • Machine learning • Electrocardiography

Introduction

Ambulatory electrocardiography (ECG) is routinely performed in patients suspected to have latent atrial fibrillation (AF), including patients with cerebrovascular events.[1–5] However, the diagnostic yield of a single ECG for AF detection is low because AF events can be fleeting and asymptomatic. While prolonged cardiac monitoring increases the diagnostic yield, it can be costly and burdensome for patients.[6]

Artificial intelligence (AI) is a powerful emerging modality that could help identify patients more likely to have undiagnosed AF using a 12-lead ECG obtained in sinus rhythm (SR) with a high predictive accuracy.[7] The electrophysiological basis of the AI-ECG algorithm is not completely understood; however, changes in P-wave amplitude and morphology, intra-atrial conduction block, and other subtle features undetected by the human eye may contribute to model performance.[7–10] Since progressive atrial remodelling ultimately leads to AF, the probability of AF while in SR should also progressively increase prior to AF, although this has not been studied. Determining the rate of increase of AI-ECG probability could be of clinical interest and could help predict time to first AF episode. Following an AF episode, it is unknown if the AI-ECG probability increases or decreases. However, if the AI algorithm detects anatomic substrate change, the AI-ECG probability should continue to rise following AF, whereas if it detects electrical triggers, the AI-ECG could decrease following AF after treatment of the contributing factors.

We therefore conducted a retrospective analysis to evaluate the trends of AI-ECG probability around the time of the first AF episode. The objectives of the study were (i) to evaluate AI-ECG probability over time in relation to the first AF episode, and (ii) to identify the effect of age and CHA_2DS_2-VASc on the rate of change in the AI-ECG probability in relation to the first AF episode.

Methods

Development of the neural network

The development of the convolutional neural network has been previously described.[7] After application to a single 12-lead ECG, the AI algorithm reports a probability of latent AF (0–100% scale). The AI network was trained to detect the electrocardiographic signature of AF in SR using the Keras Framework with a Tensorflow (Google, Mountain View, CA, USA) backend and Python. The input to the AI model was the ECG data contained in eight independent ECG leads (Leads I, II, and V1–6). An 8×5000 matrix was created based on this data, the long axis (5000) of which represented the temporal axis, while the short axis (8) represented the spatial axis and was only used on layer to fuse the data from all the leads. Subsequent convolutional layers were run on

the temporal axis which allowed the network to detect features of AF in SR which may be undetectable to the human eye. The AI-ECG algorithm is broadly available for use as part of the electronic medical records and can also be applied on demand to ECGs to report the AI-ECG probability on a percentage scale (0–100%).

Data sources and study population

The study was approved by the Mayo Clinic Internal Review Board. We screened consecutive patients who underwent 12-lead ECG between 26 May 1987 and 6 December 2019 at Mayo Clinic. Inclusion criteria for the study were (i) age 18 years or older, (ii) at least one ECG in AF, and (iii) at least one ECG in SR prior to detected AF. The ECGs used to train the neural network previously were not included in the present study. Patients without research authorization were excluded. Following the first AF episode, all subsequent ECGs with AF were excluded and the AI-ECG algorithm was only applied to ECGs during SR. We also excluded (i) patients with pacemakers or implantable cardioverter defibrillators placed any time prior to AF and (ii) patients on antiarrhythmic therapy (including Class I and Class III agents) at any point prior to AF, because either pacing or antiarrhythmic therapy would alter electrical conduction and therefore they would compromise the AI network's accuracy. Since antiarrhythmic therapies are commonly initiated for rhythm control of AF and device implantation is common in patients with tachycardia-bradycardia syndrome or prior to ablation of the atrioventricular node, we did not apply the same exclusion criteria after the first episode of AF. All ECGs were digital, standard 10 s, 12-lead tracings acquired in the supine position at a sampling rate of 500 Hz. A GE-Marquette ECG machine (Marquette, WI, USA) was used and the raw data were stored using the MUSE data management system. The ECGs in our laboratory are initially read by the GE-Marquette ECG system and then over-read by a physician-supervised, trained technician. Since atrial flutter is often thought along the same lines as AF from a clinical standpoint (particularly for deciding whether anticoagulation is warranted), it was also classified as AF.

AI-ECG probability and statistical analysis

The convolutional neural network was used to calculate the probability (0–100%) of future AF on ECGs during SR. Based on our previous model derivation study, we selected a probability threshold of 8.7% for a positive test, at which point the sensitivity of AI-ECG testing equals specificity.[7] The AI-ECG diagnostic accuracy using this cut-off has been internally validated in the 31-day window prior to development of AF as described by Attia *et al.*[7] We examined both AI-ECG probability as a continuous variable and rate (%) of positive or suprathreshold ECGs (probability >8.7%) as a binary variable at time intervals of 20–40 years, 10–20 years, 5–10 years, 2–5 years, 1–2 years, 3–12 months, and less than 3 months prior to the development of AF and equivalent timeframes after AF. We stratified both AI-ECG probability and rate of positive ECG screens both by age (<50 years and >50 years) and CHA_2DS_2-VASc score (0–1, 2–3, and ≥4). The CHA_2DS_2-VASc score was calculated by summing points determined by risk factors as follows: congestive heart failure (1 point), hypertension (1 point), age (1 point 65–74 years and 2 points ≥75 years), diabetes mellitus (1 point), stroke/transient ischaemic attack (2 points), vascular disease (1 point), female gender (1 point).[11] Descriptive data were presented as mean and standard deviation (SD) or median and interquartile range (IQR). Mixed effect models were used to assess the statistical significance of change between time points. The model outcome was set as the predicted risk (continuous variable), the independent variable was set as time (continuous variable), and patient identification was set as a random effect. All analyses were done using SAS Studio version 3.7 (SAS Institute Inc., Cary, NC, USA; *Tables 1 and 2*).

Table 1 Patient demographics

Age at AF (years)	
Mean (SD)	73.6 (12.2)
Median (IQR)	75 (67–82)
Age group (years)	
<50	2199 (3.7%)
>50	57 013 (96.3%)
Gender	
Female	25 896 (43.7%)
Male	33 316 (56.3%)
CHA_2DS_2-VASc	
Mean (SD)	4.0 (1.9)
Median (IQR)	4 (3–5)
CHA_2DS_2-VASc group	
0–1	5524 (9.3%)
2–3	18 402 (31.1%)
4+	35 286 (59.6%)

AF, atrial fibrillation; SD, standard deviation; IQR, interquartile range.

Results

After application of inclusion and exclusion criteria, we identified 59 212 patients with 544 330 ECGs prior to AF and 413 486 ECGs after the first AF episode (*Figure 1*). All patients contributed a mean 16.2 ± 13.5 and median 13 (IQR 7–21) ECGs. Prior to the first AF episode, patients had a mean 9.2 ± 8.2 and median 7 (IQR 4–12) ECGs, and after the first AF episode, they had 8.4 ± 9.7 and median 5 (IQR 2–11) ECGs. About 274 317 (50.3%) ECGs prior to first AF were positive (mean 5.5 ± 6.0 positive ECGs per patient). The mean and median time between the first suprathreshold ECG and first AF were 5.4 ± 5.7 and 3.7 (IQR 0.4–8.6) years, respectively.

The absolute probability and rate of positive ECGs prior and after AF are illustrated in *Figure 2*. The absolute AI probability 1–2 years prior to the first episode of AF was 23.6% (62.2% rate of positive screen) but increased to 34.0% (75.6% rate of positive screen) in the last 3 months ($P < 0.001$). Immediately after the first episode of AF, the absolute probability increased to 40.9% (86.3% rate of positive screen) ($P < 0.001$) and decreased to 35.2% (78.1% rate of positive screen) 1–2 years after index AF episode ($P < 0.001$). After this time point, the AI probability gradually increased up to 53.5% (93.4% rate of positive screen) 20–40 years post AF ($P < 0.001$).

Similar results were noted after stratifying for age and CHA_2DS_2-VASc score. Patients with age ≥50 years and CHA_2DS_2-VASc score ≥4 had both a higher AI-ECG probability and a higher rate of increase over time (*Figures 3* and *4*). The mean absolute probability at 3 months prior to first AF was 18.9% in patients <50 years but 34.6% in patients ≥50 years. Similarly, the mean absolute probability at 3 months prior to first AF was 22.8% in patients with CHA_2DS_2-VASc 0–1 but 38.1% in patients with CHA_2DS_2-VASc ≥4.

Table 2 Breakdown of data and variability of AI-ECG probability by study time period

Study timeframe	Number of ECGs	Number of patients	Number of patients with multiple ECGs	Mean (SD) change in % probability of AF[a]
20–40 years pre AF	21 990	7907	4399	0.61 (8.15)
10–20 years pre AF	108 474	26 524	19 104	3.61 (13.83)
5–10 years pre AF	118 008	33 492	23 806	4.34 (16.39)
2–5 years pre AF	107 161	35 565	23 643	3.83 (18.23)
1–2 years pre AF	45 710	23 685	8774	2.44 (18.87)
3–12 months pre AF	43 470	21 228	8272	2.63 (19.15)
0–3 months pre AF	99 517	47 001	22 706	7.15 (24.27)
0–3 months post AF	126 373	34 470	23 952	−1.59 (26.48)
3–12 months post AF	50 418	20 048	9465	1.70 (21.97)
1–2 years post AF	53 403	20 826	9841	2.73 (21.10)
2–5 years post AF	97 197	22 978	16 505	4.76 (21.23)
5–10 years post AF	66 054	12 508	9553	5.94 (22.27)
10–20 years post AF	19 603	3519	2625	7.35 (22.76)
20–40 years post AF	438	108	73	7.47 (18.83)
Total	957 816	309 859	182 718	—

AF, atrial fibrillation; ECG, electrocardiogram; SD, standard deviation.
[a]Reflects mean change of probability from earliest to latest ECG in each timeframe from the same patient.

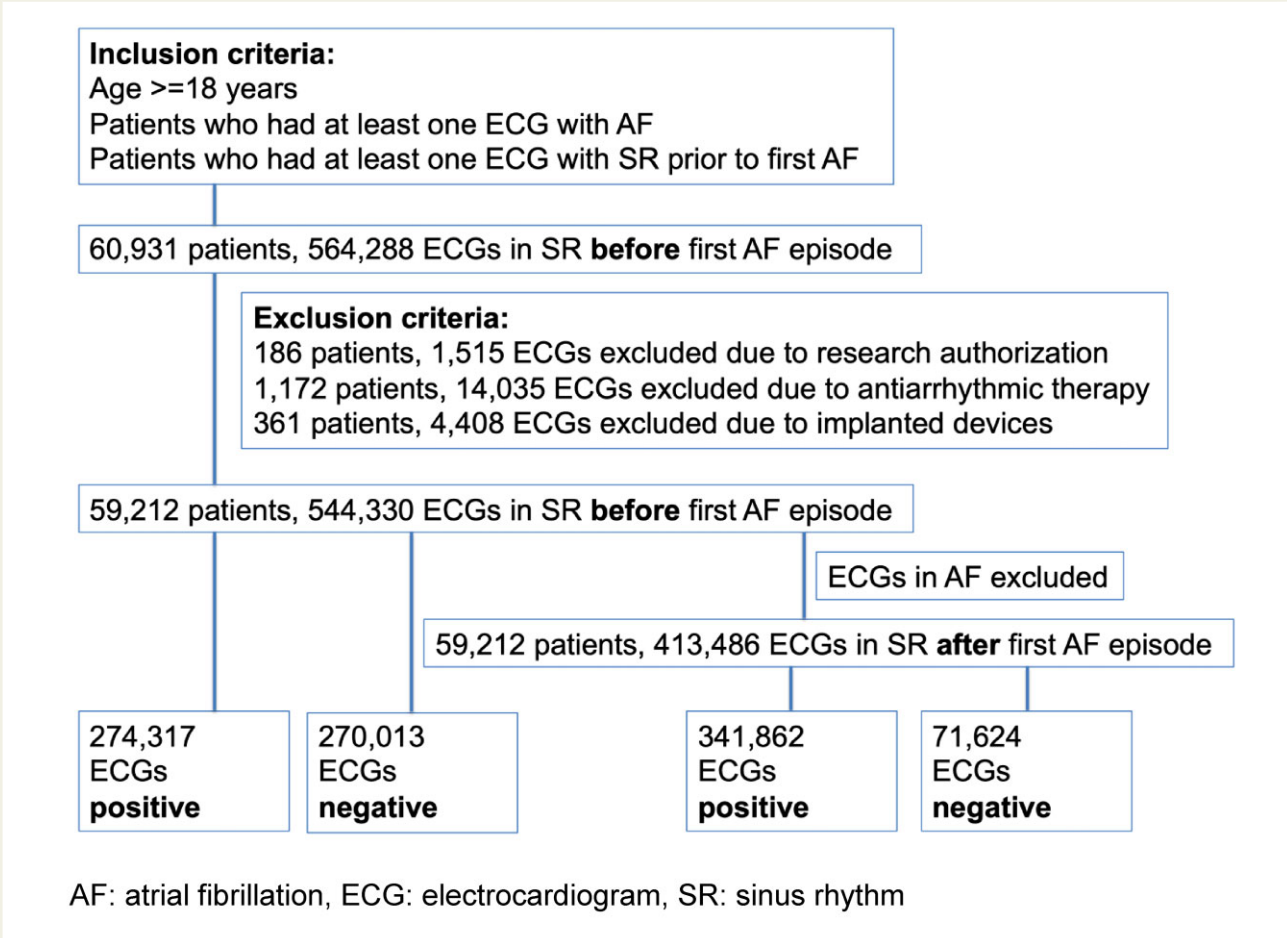


Figure 1 Study inclusion criteria and analysis plan. AF, atrial fibrillation; ECG, electrocardiogram; SR, sinus rhythm.

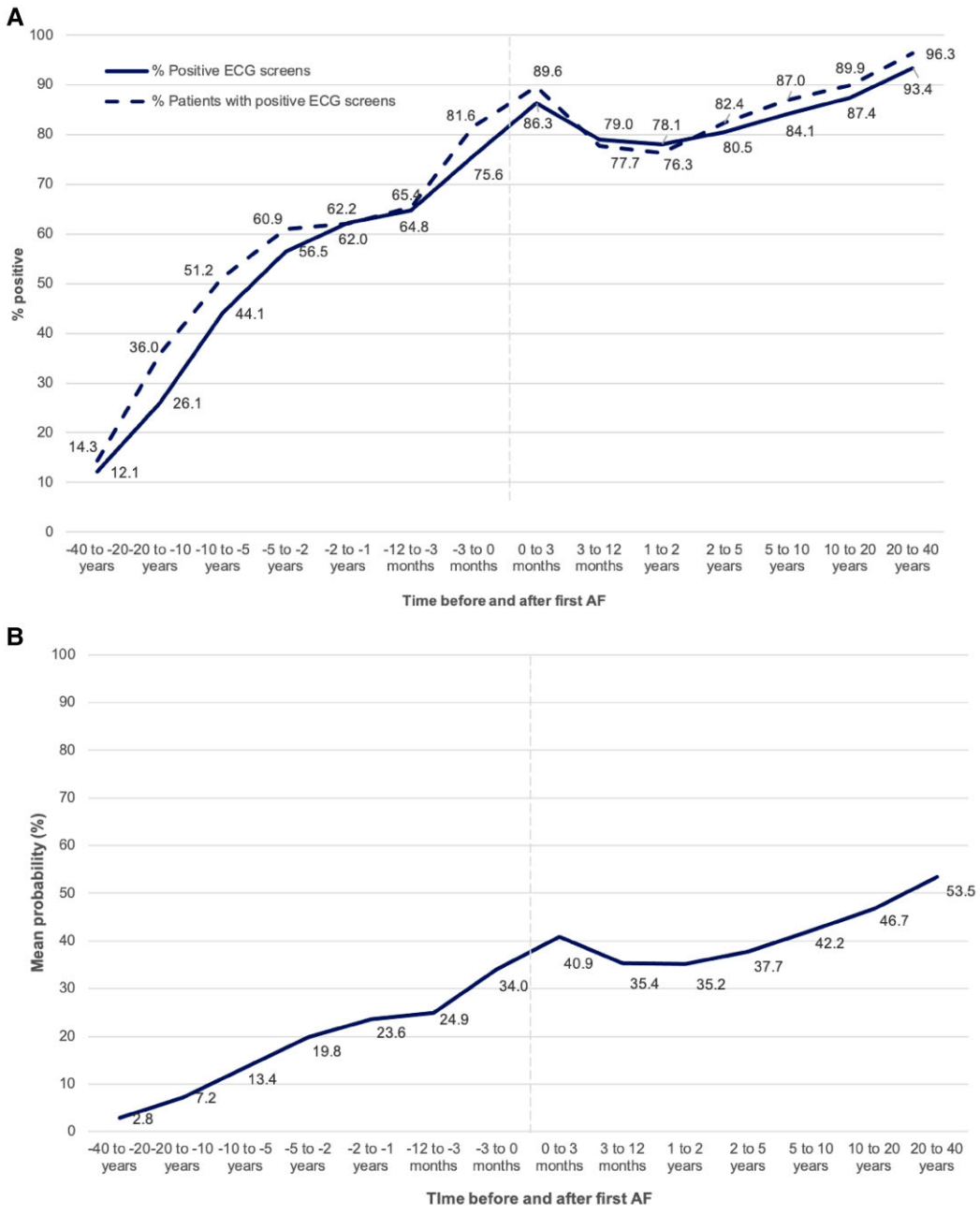


Figure 2 Positive rate of artificial intelligence electrocardiograms (*A*) and mean artificial intelligence electrocardiogram probability (*B*) as a function of time before and after first atrial fibrillation (dashed line).

Discussion

The present study demonstrates that the AI-ECG probability of AF (i) gradually increases prior to the first AF episode and (ii) continues to increase immediately after the first AF episode. Patients ≥50 years and patients with increased CHA_2DS_2-VASc score have both increased AF probability and rate of increase prior to AF. The continued increasing probability of AF on the AI-ECG prior to a documented episode suggests ongoing changes in cardiac substrate and/or electrophysiology resulting in subtle, likely non-linear electrocardiographic changes over years before the first episode of manifest, which are detectable by a neural network. It is unknown whether individuals with a rising probability of AF on the AI-ECG over time may benefit from risk factor modification to avert AF development, but our findings suggest this warrants further study.

Understanding and interpreting the AI-ECG is important for clinical use. In the study by Attia *et al*., the AI network was trained to detect AF in all ECGs in patients who never developed AF and all ECGs starting 31 days prior to the first AF in patients with at least one AF episode.[7] However, the AI-ECG's high predictive accuracy in the

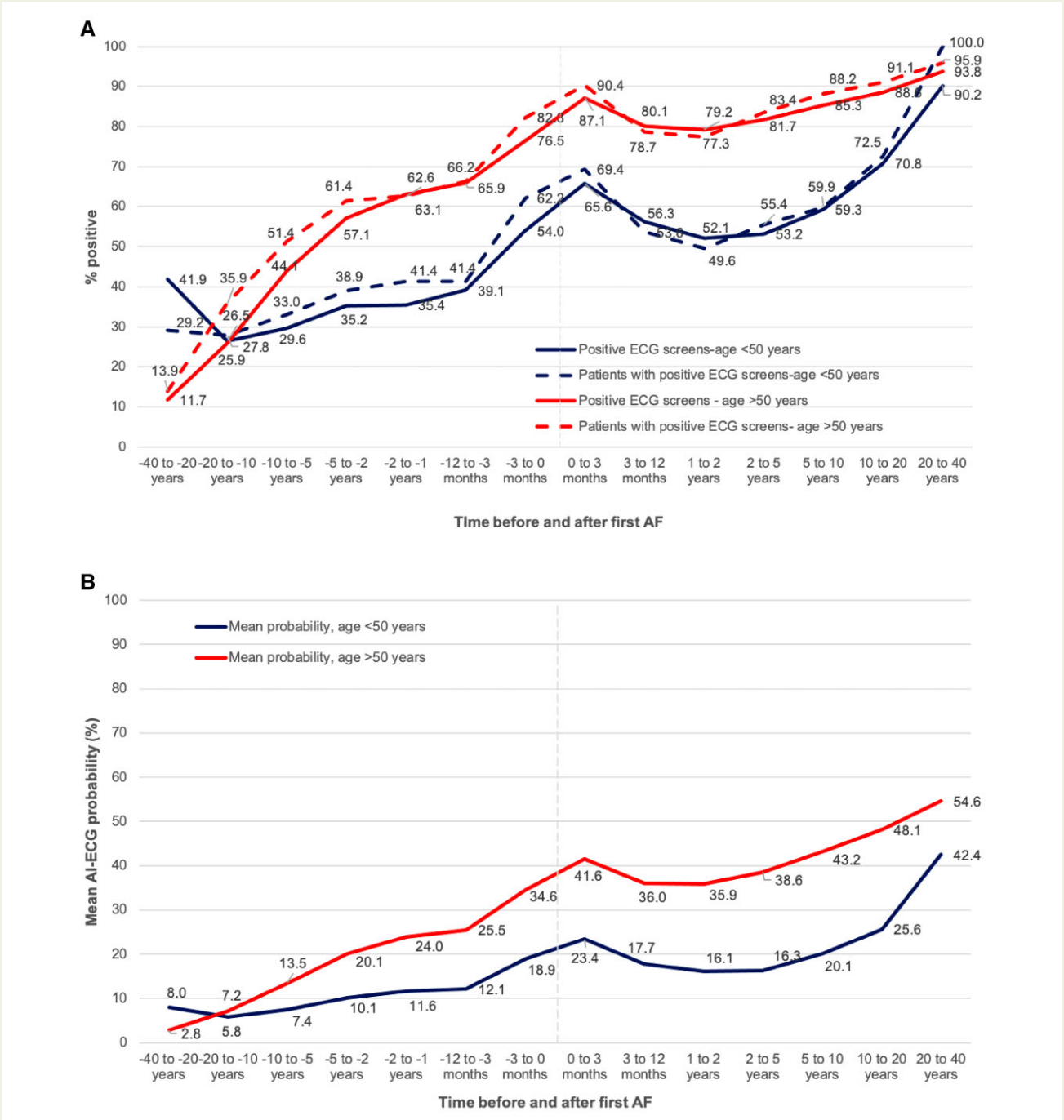


Figure 3 Impact of age on positive rate of artificial intelligence electrocardiograms (*A*) and mean artificial intelligence electrocardiogram probability (*B*) before and after first atrial fibrillation (dashed line).

interval validation data set was specifically demonstrated in the 31-day window of the study. Therefore, the AI-ECG performance was demonstrated for concurrent (within 31 days) AF. The absolute AI-ECG probability was not reported in the study but was selected to balance sensitivity and specificity (approximately 8.7% based on internal data). The performance of AI-ECG was reported with AI-ECG treated as a binary variable (i.e. suprathreshold vs. negative ECG); however, there were no implications of (i) absolute probability and (ii) future AF incidence based on this probability. In the follow-up, Mayo Clinic Study of Aging (MCSA) which showed that a probability of 50% was associated with a risk of AF of 21.5% at 2 years and 52.2% at 10 years in patients who had never had prior AF.[12] The MCSA provided a temporal component to the AI-ECG allowing for longitudinal correlation of AI-ECG absolute probability and AF incidence.

The present study adds further characterization of the AI-ECG performance by describing the trend and rate of change of AI-ECG before and after the first AF episode. Several important observations have significant clinical impact and are worthy of discussion: (i) the study clearly

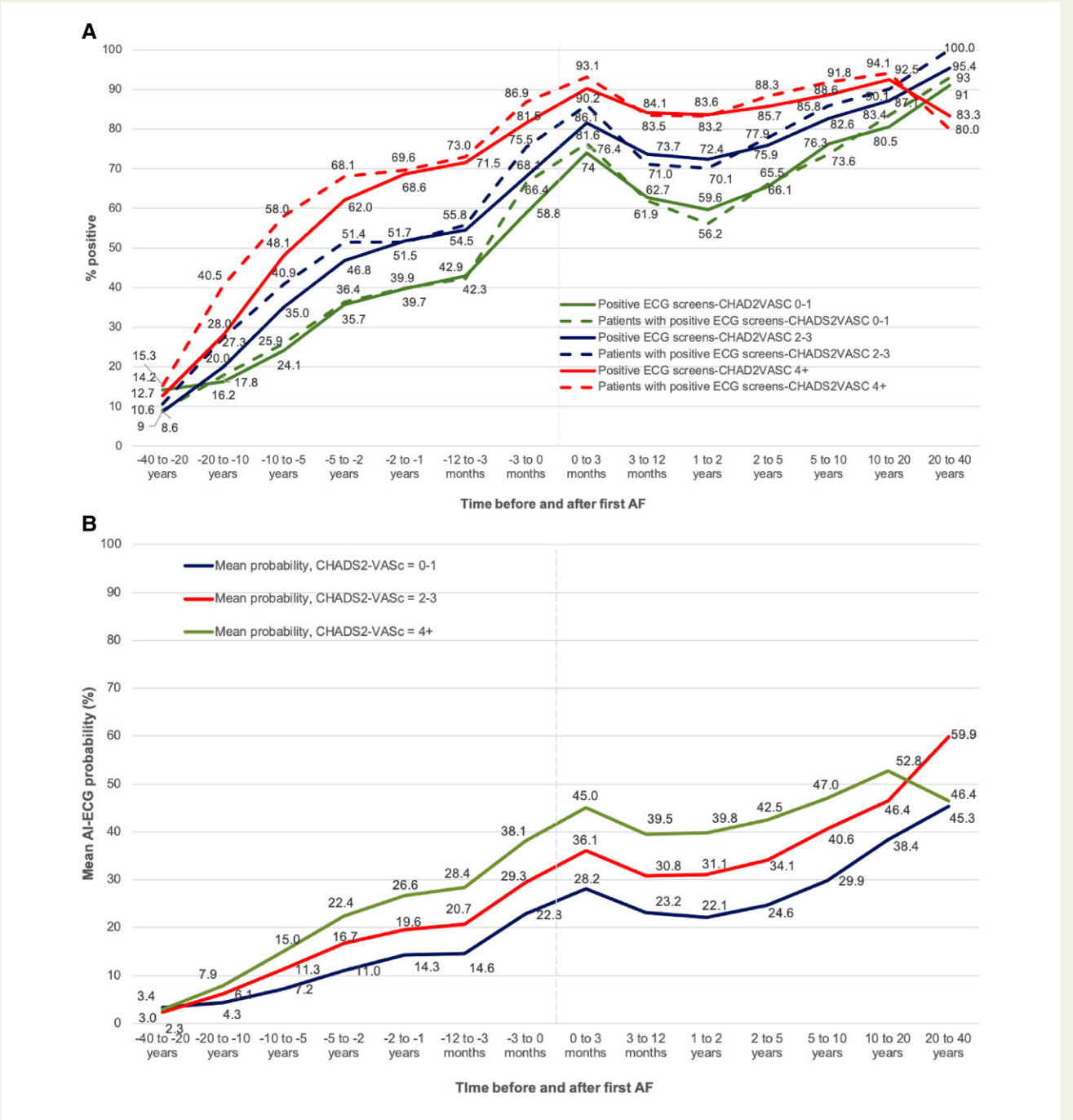


Figure 4 Impact of CHA_2DS_2-VASc score on positive rate of artificial intelligence electrocardiogram (*A*) and mean artificial intelligence electrocardiogram probability (*B*) before and after first atrial fibrillation (dashed line).

demonstrates a positive rate of increase in AI-ECG probability (and positive ECG screens) prior to AF which solidifies the predictive accuracy of AI-ECG for AF; (ii) although a single suprathreshold ECG can be treated with criticism in clinical decision-making, future ECG demonstrating higher AI-ECG probabilities with a positive trend are a concerning finding and merit clinical attention, particularly in older patients with increased risk factors for AF; and (iii) prediction of time to future AF can be made to some degrees based on sequential ECGs. For example, an increase of mean AI-ECG probability to 30–35% from 20–25% 1 year prior and confirmed by multiple ECGs could be a herald of impending AF in the next year (*Figure 2*). Such application may increase the yield of healthcare resources by utilizing monitoring on those individuals most likely to benefit. It also underscores the potential utility of serial assessment of AI-ECG screenings recognizing that it may fluctuate with variations in health and risk status.

The analysis of AI-ECG probability after the first AF episode is less important for clinical decision-making; however, it can be studied to help understand what the AI network recognizes as AF and could

potentially be a biomarker/surrogate measure for the impact of rhythm control treatments. Our analysis showed a transient increase of probability immediately following paroxysmal AF but a subsequent decrease after 1–2 years, regardless of age or CHA_2DS_2-VASc score. The transient dip in AI-ECG probability 1–2 years after the first AF episode can have different interpretations. Several triggers of AF (such as heart failure and inflammation) may only persist 0–3 months after the AF event but subsequently decrease after 1–2 years. Additionally, attempts to achieve rhythm control in AF either with antiarrhythmic agents or ablation may also decrease the AI-ECG probability in the intermediate term. The results of the present study will likely need to be individualized, since the time from first AF to second (or subsequent) AF may vary between individuals and therefore not all patients will have the same pattern of AF probability after their initial AF.

Limitations

Some important study limitations warrant discussion. The AI-ECG algorithm was applied at a single centre and generalizability of the algorithm in populations with different age, ethnicity, or prevalence of AF is uncertain. Including patients with at least one episode of AF inevitably introduces selection bias, since patients with positive AI-ECG screens may not ever develop AF. We did not analyse the AI-ECG trend for AF in patients who were never shown to have AF. As the risk of AF increases with age, it is hard to interpret whether the increase of AF probability with time is a result of impending AF episode vs. increasing age. However, increasing age alone does not explain the rapid change of AI-ECG probability in the year prior to the index AF event, nor does it explain the dip of AI-ECG 1–2 years post first AF. Additional selection bias is created by the fact that patients with symptoms (such as palpitations or shortness of breath) have higher likelihood of AF and are more likely to be screened with ECG. The true first episode of AF is challenging to determine particularly in asymptomatic patients or patients who were evaluated in different medical centres. As asymptomatic patient do not typically undergo extensive ambulatory monitoring, their first true episode of AF may significantly predate the first reported episode and therefore the rate of AI-ECG probability change may underestimate the timing of the first episode in these patients. Patients with implantable devices including pacemaker and defibrillators were excluded the study. There, no conclusions can be drawn for this patient population. Following the index AF episode, a subset of patients were treated with antiarrhythmic therapies or underwent device implantation but those patients/ECGs were not screened out as explained in the Methods section. We did not exclude patients or ECGs with subsequent AF from the analysis which can create uncertainty in the interpretation of the AI-ECG following the first AF episode. Atrial flutter was not differentiated from AF in the present study. The original AI-ECG algorithm was developed in patients with either AF or atrial flutter. However, from an electropathophysiological standpoint, atrial flutter is distinct from AF and the features that the AI-ECG algorithm detects during SR may be different in different atrial arrhythmias. Including different tachyarrhythmias in the AI model may decrease the specificity of the algorithm. Lastly, since not all AFs are associated with underlying progressive structural disease, AI-ECG may have not predicted AF in patients with triggers not anticipated by the AI network (such as non-cardiac surgery or acute metabolic disorders).

Conclusion

In conclusion, the present study demonstrates that the AI-ECG reported probability of AF gradually increases over time prior to AF with a rate that is a function of patient age and CHA_2DS_2-VASc score. The study may have implications for the prediction of AF episodes in the clinical setting.

Funding

The study was conceived, developed, and executed by the authors with no external or commercial support; funding was via internal Mayo Clinic resources.

Conflict of interest: Z.I.A., P.A.N., and P.A.F. are developers and co-inventors of the neural network. Z.I.A., R.E.C., F.L.-J., and P.A.F. are members of the Advisory Board of Anumana, the company licensing this technology.

Data availability

Study source data and analyses are available upon request from the corresponding author.

References

1. Dilaveris PE, Kennedy HL. Silent atrial fibrillation: epidemiology, diagnosis, and clinical impact. *Clin Cardiol* 2017;**40**:413–418.
2. Rosero SZ, Kutyifa V, Olshansky B, Zareba W. Ambulatory ECG monitoring in atrial fibrillation management. *Prog Cardiovasc Dis* 2013;**56**:143–152.
3. Diederichsen SZ, Haugan KJ, Brandes A, Lanng MB, Graff C, Krieger D, Kronborg C, Holst AG, Køber L, Højberg S, Svendsen JH. Natural history of subclinical atrial fibrillation detected by implanted loop recorders. *J Am Coll Cardiol* 2019;**74**:2771–2781.
4. Sposato LA, Cipriano LE, Saposnik G, Ruíz Vargas E, Riccio PM, Hachinski V. Diagnosis of atrial fibrillation after stroke and transient ischaemic attack: a systematic review and meta-analysis. *Lancet Neurol* 2015;**14**:377–387.
5. Gladstone DJ, Spring M, Dorian P, Panzov V, Thorpe KE, Hall J, Vaid H, O'Donnell M, Laupacis A, Côté R, Sharma M, Blakely JA, Shuaib A, Hachinski V, Coutts SB, Sahlas DJ, Teal P, Yip S, Spence JD, Buck B, Verreault S, Casaubon LK, Penn A, Selchen D, Jin A, Howse D, Mehdiratta M, Boyle K, Aviv R, Kapral MK, Mamdani M, for the EMBRACE Investigators and Coordinators. Atrial fibrillation in patients with cryptogenic stroke. *N Engl J Med* 2014;**370**:2467–2477.
6. Yong JHE, Thavorn K, Hoch JS, Mamdani M, Thorpe KE, Dorian P, Sharma M, Laupacis A, Gladstone DJ. Potential cost-effectiveness of ambulatory cardiac rhythm monitoring after cryptogenic stroke. *Stroke* 2016;**47**:2380–2385.
7. Attia ZI, Noseworthy PA, Lopez-Jimenez F, Asirvatham SJ, Deshmukh AJ, Gersh BJ, Carter RE, Yao X, Rabinstein AA, Erickson BJ, Kapa S, Friedman PA. 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**:861–867.
8. Schumacher K, Dagres N, Hindricks G, Husser D, Bollmann A, Kornej J. Characteristics of PR interval as predictor for atrial fibrillation: association with biomarkers and outcomes. *Clin Res Cardiol* 2017;**106**:767–775.
9. Bidstrup S, Salling Olesen M, Hastrup Svendsen J, Bille Nielsen J. Role of PR-interval in predicting the occurrence of atrial fibrillation. *J Atr Fibrillation* 2013;**6**:956.
10. German DM, Kabir MM, Dewland TA, Henrikson CA, Tereshchenko LG. Atrial fibrillation predictors: importance of the electrocardiogram. *Ann Noninvasive Electrocardiol* 2016;**21**:20–29.
11. Lip GY, Nieuwlaat R, Pisters R, Lane DA, Crijns HJ. 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**:263–272.
12. Christopoulos G, Graff-Radford J, Lopez CL, Yao X, Attia ZI, Rabinstein AA, Petersen RC, Knopman DS, Mielke MM, Kremers W, Vemuri P, Siontis KC, Friedman PA, Noseworthy PA. Artificial intelligence-electrocardiography to predict Incident atrial fibrillation: a population-based study. *Circ Arrhythm Electrophysiol* 2020; **13**:e009355.