An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction

Zachi I Attia, Peter A Noseworthy*, Francisco Lopez-Jimenez, Samuel J Asirvatham, Abhishek J Deshmukh, Bernard J Gersh, Rickey E Carter, Xiaoxi Yao, Alejandro A Rabinstein, Brad J Erickson, Suraj Kapa, Paul A Friedman*

Summary

Background Atrial fibrillation is frequently asymptomatic and thus underdetected but is associated with stroke, heart failure, and death. Existing screening methods require prolonged monitoring and are limited by cost and low yield. We aimed to develop a rapid, inexpensive, point-of-care means of identifying patients with atrial fibrillation using machine learning.

Methods We developed an artificial intelligence (AI)-enabled electrocardiograph (ECG) using a convolutional neural network to detect the electrocardiographic signature of atrial fibrillation present during normal sinus rhythm using standard 10-second, 12-lead ECGs. We included all patients aged 18 years or older with at least one digital, normal sinus rhythm, standard 10-second, 12-lead ECG acquired in the supine position at the Mayo Clinic ECG laboratory between Dec 31, 1993, and July 21, 2017, with rhythm labels validated by trained personnel under cardiologist supervision. We classified patients with at least one ECG with a rhythm of atrial fibrillation or atrial flutter as positive for atrial fibrillation. We allocated ECGs to the training, internal validation, and testing datasets in a 7:1:2 ratio. We calculated the area under the curve (AUC) of the receiver operatoring characteristic curve for the internal validation dataset to select a probability threshold, which we applied to the testing dataset. We evaluated model performance on the testing dataset by calculating the AUC and the accuracy, sensitivity, specificity, and F1 score with two-sided 95% CIs.

Findings We included 180 922 patients with 649 931 normal sinus rhythm ECGs for analysis: 454 789 ECGs recorded from 126 526 patients in the training dataset, 64 340 ECGs from 18 116 patients in the internal validation dataset, and 130 802 ECGs from 36 280 patients in the testing dataset. 3051 (8·4%) patients in the testing dataset had verified atrial fibrillation before the normal sinus rhythm ECG tested by the model. A single AI-enabled ECG identified atrial fibrillation with an AUC of 0·87 (95% CI 0·86–0·88), sensitivity of 79·0% (77·5–80·4), specificity of 79·5% (79·0–79·9), F1 score of 39·2% (38·1–40·3), and overall accuracy of 79·4% (79·0–79·9). Including all ECGs acquired during the first month of each patient's window of interest (ie, the study start date or 31 days before the first recorded atrial fibrillation ECG) increased the AUC to 0·90 (0·90–0·91), sensitivity to 82·3% (80·9–83·6), specificity to 83·4% (83·0–83·8), F1 score to 45·4% (44·2–46·5), and overall accuracy to 83·3% (83·0–83·7).

Interpretation An AI-enabled ECG acquired during normal sinus rhythm permits identification at point of care of individuals with atrial fibrillation.

Funding None.

Lancet 2019; 394: 861–67

Published Online August 1, 2019 http://dx.doi.org/10.1016/S0140-6736(19)31721-0

See Comment page 812

*Contributed equally

Department of Cardiovascular Medicine (Z I Attia MS, P A Noseworthy MD, Prof F Lopez-Jimenez MD, Prof S J Asirvatham MD, A J Deshmukh MBBS, Prof B J Gersh MB ChB, S Kapa MD, Prof P A Friedman MD), **Department of Health Sciences Research** (X Yao PhD), **Department of Neurology** (Prof A A Rabinstein MD), **Department of Radiology** (Prof B J Erickson MD), **Mayo Clinic, Rochester, MN, USA; and Department of Health Sciences Research, Mayo Clinic, Jacksonville, FL, USA** (Prof R E Carter PhD)

Correspondence to: Prof Paul Friedman, Department of Cardiovascular Medicine, Mayo Clinic, Rochester, MN 55905, USA **friedman.paul@mayo.edu**

Introduction

Atrial fibrillation is common, underdiagnosed, and associated with an increased risk of stroke, heart failure, and mortality.[1,2] Screening for atrial fibrillation can be challenging due to the low diagnostic yield of a single electrocardiograph (ECG) to detect an often fleeting arrhythmia and the cumbersome nature of prolonged monitoring. Clinical risk scores can be used to identify patients at risk but have only modest performance. Due to these limitations, major medical societies have issued inconsistent guidelines on atrial fibrillation screening.

A low-cost, widely available, and non-invasive test that facilitates identification of patients who are likely to have atrial fibrillation would have important diagnostic and therapeutic implications. For instance, up to a third of strokes have no known cause—so-called embolic stroke of undetermined source (ESUS).[3] Many of these strokes are related to atrial fibrillation, which can be underdetected due to its paroxysmal and often asymptomatic nature.[4] Patients with ESUS are at high risk of a recurrent stroke, and when atrial fibrillation is documented, anticoagulation reduces the risk of recurrent stroke and might reduce mortality.[5,6] However, empirical use of anticoagulants following ESUS, whether with warfarin or a direct oral anticoagulant, has not been shown to be beneficial[7,8] and increases risk of bleeding;[7–9] therefore,

Research in context

Evidence before this study

There is a robust literature on screening for atrial fibrillation in the general population; we did not do a formal systematic review. However, most efforts have focused on either one-time screening with a single electrocardiograph (ECG) or the use of various implantable or wearable monitors to capture infrequent atrial fibrillation episodes over time. Some studies have evaluated discrete ECG features—often P wave characteristics—as predictors of atrial fibrillation, but no individual feature has high enough predictive value to offer clinical utility using routine statistical modelling. The intensive evaluation of the ECG afforded by a convolutional neural network might be able to detect subtle, multifaceted perturbations in the ECG. We have previously shown convolution neural networks can evaluate the resting ECG for detection of antiarrhythmic drug levels, abnormal electrolytes levels, and detection of asymptomatic left ventricular dysfunction, providing proof of concept that clinically important phenomena can be detected with artificial intelligence (AI) applications to the ECG.

Added value of this study

This is the first study to our knowledge to use a convolution neural network to identify the electrocardiographic signature of atrial fibrillation present during sinus rhythm. We used an AI model to find signals in the ECG that might be invisible to the human eye but contain important information about the presence of atrial fibrillation. The AI model was trained using the standard 10-second, 12-lead ECG alone and does not require any other inputs for atrial fibrillation risk assessment. Importantly, the detection of the atrial fibrillation signal in the ECG relies on this easily obtained 10-second recording as opposed to the more invasive loop recording or cumbersome Holter monitoring. We found that an AI model can differentiate between patients with a history of (or impending) atrial fibrillation with a high degree of accuracy using a single routine ECG. Addition of multiple ECGs within an individual patient improved the model accuracy and suggests repeated measures might yield even better performance.

Implications of all the available evidence

Our study supports the hypothesis that subtle patterns on the normal sinus rhythm ECG can suggest the presence of atrial fibrillation. The ability to identify patients with potentially undetected atrial fibrillation using an inexpensive, non-invasive, widely available, point-of-care test has important practical implications for atrial fibrillation screening and potentially for the management of patients with prior stroke of unknown cause.

determination of whether atrial fibrillation is present is crucial to guide therapy.

Prolonged ambulatory cardiac rhythm monitoring is frequently used to screen for atrial fibrillation, particularly after ESUS. Approaches include insertion of implantable loop recorders and wearable patches.[1,2] These strategies are invasive or inconvenient, expensive, require a monitoring infrastructure, and have a low yield.[10]

There is growing evidence that patients who develop atrial fibrillation—even in an apparently normal heart—have structural changes in the atria that predispose towards atrial arrhythmias;[11] these changes might be important for the pathogenesis of ischaemic or embolic stroke. We have previously used machine learning in the form of deep neural networks to identify subtle patterns in the standard 12-lead ECG to identify the presence of asymptomatic ventricular dysfunction.[12]

We hypothesised that we could train a neural network to identify the subtle findings present in a standard 12-lead ECG acquired during normal sinus rhythm that are due to structural changes associated with a history of (or impending) atrial fibrillation. Such a diagnostic test could be inexpensive, widely available, and immensely useful following ESUS to guide therapy. To test this hypothesis, we trained, validated, and tested a deep neural network using a large cohort of patients from the Mayo Clinic Digital Data Vault.

Methods

Data sources and study population

We included all patients aged 18 years or older with at least one digital, normal sinus rhythm, standard 10-second, 12-lead ECG acquired in the supine position at the Mayo Clinic ECG laboratory between Dec 31, 1993, and July 21, 2017. All ECGs were acquired at a sampling rate of 500 Hz using a GE-Marquette ECG machine (Marquette, WI, USA) and the raw data were stored using the MUSE data management system. ECGs in our laboratory are intially read by the GE-Marquette ECG system and then over-read by a physician-supervised, trained technician, with corrections made to the diagnostic labels as needed. For the purposes of the present study, any ECG with a rhythm of atrial fibrillation or atrial flutter was classified as having atrial fibrillation. We chose this classification because guidelines recommend anticoagulation in the presence of either atrial fibrillation or atrial flutter and both rhythms often coexist.[13–15] The Mayo Clinic Internal Review Board approved waiver of the requirement to obtain informed consent in accordance with 45 CFR 46.116 and waiver of Health Insurance Portability and Accountability Act (HIPAA) authorisation in accordance with applicable HIPAA regulations.

Identifying study groups

The Mayo Clinic Digital Data Vault was used to extract the labels and raw data from all ECGs acquired from our

cohort of patients. We used these data to classify patients into two groups: patients positive for atrial fibrillation, who had at least one atrial fibrillation rhythm recorded on a Mayo Clinic ECG, and patients negative for atrial fibrillation, who had no ECGs with atrial fibrillation recorded and additionally had no reference to atrial fibrillation in the diagnostic codes in their electronic medical record. Patients with a diagnosis code for atrial fibrillation but no ECG documentation of atrial fibrillation were considered to have unverified atrial fibrillation and were excluded from the analysis to avoid ambiguity. ECGs with paced rhythms were also excluded.

ECG selection for patients with multiple ECGs

Many study patients had multiple ECGs recorded over the inclusion period. We defined a window of interest for each patient for the purpose of analysis (figure 1). For patients who had had at least one atrial fibrillation rhythm recorded, we defined the first recorded atrial fibrillation ECG as the index ECG and the first day of the window of interest as 31 days before the date of the index ECG. We chose this window of interest with the assumption that the structural changes associated with atrial fibrillation would be present before the first recorded atrial fibrillation episode; we chose a relatively short time interval as a conservative measure to avoid using ECGs before any structural changes developed. For patients with no ECGs with atrial fibrillation recorded, the index ECG was defined as the date of the first ECG available for that patient in the Mayo Clinic Digital Data Vault.

During training, all the ECGs in the window of interest were used to allow the network to have more samples; for the testing and validation sets, only the first normal sinus rhythm ECG within the window of interest was used to avoid repeated measurements and to mimic a real screening scenario.

Outcomes

The primary outcome of the study was the ability of the artificial intelligence (AI)-enhanced ECG to identify patients with atrial fibrillation using a standard 10-second, 12-lead ECG recorded during sinus rhythm. This performance was mathematically assessed by the area under the curve (AUC) of the receiver operating characteristic (ROC) curve, as well as the sensitivity, specificity, accuracy, and F1 score of the model.

We did a secondary analysis to determine whether use of more than one sinus rhythm ECG per patient improved the AUC of the AI-enabled ECG for the detection of a history of atrial fibrillation. We also did a secondary analysis including only the first normal sinus rhythm after the index atrial fibrillation ECG.

Overview of the AI model

We implemented a convolutional neural network (CNN) using the Keras Framework with a Tensorflow (Google; Mountain View, CA, USA) backend and Python.[16] The

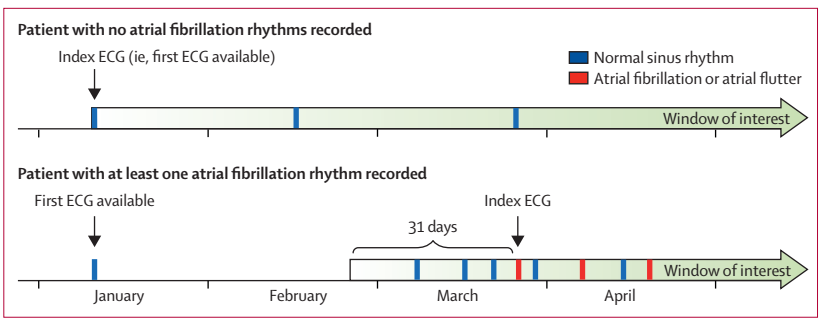

Figure 1: **ECG selection and windows of interest for patients with multiple ECGs**
The figure shows an example of ECG selection for two patients with multiple ECGs over the same year. We used all normal sinus rhythm ECGs for patients with no ECGs with atrial fibrillation recorded and the window of interest began on the date of their first ECG. For patients with at least one atrial fibrillation rhythm recorded, the first ECG recording atrial fibrillation or atrial flutter was the index ECG and the window of interest began 31 days before the index ECG. For all patients, the window of interest extended until study end. ECG=electrocardiograph.

12-lead ECG is recorded using eight physical leads and four augmented leads created as a linear function of leads I and II, which do not contain incremental information. To optimise performance, we selected only the eight independent leads (leads I, II, and V1–6) because any linear function of the leads could be learned by the models. This reduced the original 12×5000 matrix (ie, 12 leads by 10-second duration sampled at 500 Hz) to a 8×5000 matrix. The long axis (5000) represents the temporal axis and most of the convolutions were used on it to allow the model to extract morphological and temporal features, while the short axis (8) represents the lead or spatial axis and was only used on layer to fuse the data from all the leads.

The network was composed of ten residual blocks, which allow the signals to feed directly to the next layer in addition to the processing done in the current layer; this allows the network to learn even when using a very large number of layers (appendix). Each residual block was implemented using two blocks, eacah composed of a batch-normalisation layer that accounts for normalisation of the data distribution; a non-linear ReLU activation function with output zero for negative inputs and identity output for positive inputs, the non-linearity of which allows the network to create a complex non-linear representation of the ECGs for automatic feature extraction;[17] and a convolution layer. The residual blocks were completed with a shortcut link to allow gradient propagation that is implemented using a 1×1 convolution layer between the input of the residual block to its output and finally a max pooling layer.[18] The nine different residual blocks had access to a single lead and the last convolution layer fused all eight independent leads using a 1×8 convolutional layer. Following the last convolutional layer, the data were fed to a dropout layer and to the final output layer that was activated using the softmax function, which generated a probability of atrial fibrillation. The architecture of the model is available in the appendix. The model was trained on a computer with 224 GB ram and four K-80 (NVIDIA) graphics processing

See Online for appendix

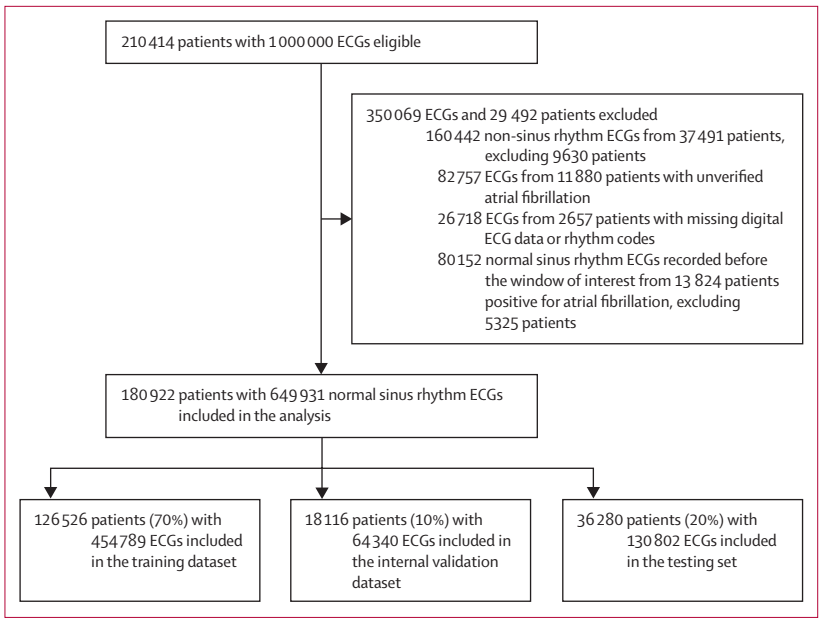

Figure 2: **Patient flow diagram**
ECG=electrocardiograph.

units (GPUs) that were used to train the model in parallel using the Keras single machine-multi GPU parallelism.

All patients and their digitally available Mayo Clinic ECGs included in the cohort were randomly assigned in a 7:1:2 ratio to one of three groups: training, internal validation, and testing datasets. The training dataset contained ECGs from 70% of the patient cohort and was used to train the network; the internal validation dataset with ECGs from 10% of the cohort was used to optimise the network and select the network hyperparameters; and the testing dataset, including ECGs from the remaining 20% of patients who were not in the training or validation datasets, was used to assess the AI-enabled ECGs' ability to detect a history of atrial fibrillation.

A ROC curve was created for the testing and validation datasets to assess the AUC of the AI-enabled ECG acquired during normal sinus rhythm to determine whether atrial fibrillation was present. Using the ROC curve for the small internal validation set, we selected a probability threshold and applied the same threshold to the testing dataset for derivation of the testing dataset accuracy, sensitivity, specificity, and F1 score.

Statistical analysis

Statistical optimisation of the CNN was done through iterative training using the Keras package. Once a final fitted model was obtained, the diagnostic performance was more formally analysed. Measures of diagnostic performance included the ROC AUC, accuracy (ie, a weighted average of sensitivity and specificity indicating the percentage of patients whose labels were predicted correctly), sensitivity, specificity, and the F1 score (ie, the harmonic mean of the sensitivity and positive predictive value). We used two-sided 95% CIs to summarise the sample variability in the estimates. We used exact (Clopper-Pearson) CIs to be conservative for accuracy, sensitivity, and specificity. The CI for the AUC was estimated using the Sun and Su optimisation of the Delong method using the pROC package[19] whereas the CI for F1 was obtained using the bootstrap method with 2000 replications. All analyses were done using R, version 3.4.2.

Role of the funding source

This study received no external funding. No entity other than the authors listed played any role in the design of the study; the collection, analysis, or interpretation of data; writing of the report; or in the decision to submit the paper for publication. ZIA and REC had full access to the data and the final decision to submit the manuscript was made by PAF.

Results

We identified 210 414 patients with 1 000 000 ECGs and, after applying exclusion criteria, included 180 922 patients with 649 931 normal sinus rhythm ECGs for analysis (figure 2). We trained the model using 454 789 ECGs recorded from 126 526 patients, with a mean of 3·6 ECGs (SD 4·8) per patient. In patients with at least one atrial fibrillation recorded in the testing dataset, 1698 (55·7%) of the 3051 first normal sinus rhythm ECGs in the window of interest were within 1 week of the index atrial fibrillation ECG (median number of days between ECGs 0, IQR –4 to 24).

Among all included patients, the mean age was 60·3 years (SD 16·5) on the date of the index ECG, 89 791 (49·6%) patients were men, and 15 419 (8·5%) had at least one recorded atrial fibrillation. In the internal validation set, there were 64 340 ECGs from 18 116 patients with a mean of 3·6 ECGs (SD 4·8) per patient. Patients had a mean age of 60·3 years (SD 16·7) at their first visit, 8983 (49·6%) were men, and 1573 (8·7%) had at least one recorded atrial fibrillation. In the testing dataset, there were 130 802 ECGs from 36 280 patients with a mean of 3·6 ECGs (4·9) per patient. Patients had a mean age of 60·1 years (16·8) at their first visit, 18 068 (49·8%) were men, and 3051 (8·4%) had at least one recorded atrial fibrillation.

When testing the model on the first sinus rhythm ECG for each patient, the ROC AUC for the detection of atrial fibrillation was 0·87 (0·86–0·88) using the internal validation set and 0·87 (0·86–0·88) using the testing dataset (table). The probability value that yielded similar sensitivity, specificity, and accuracy of 79·2% on the internal validation set was applied to the testing set and yielded an F1 score of 39·2% (95% CI 38·1–40·3), sensitivity of 79·0% (77·5–80·4), specificity of 79·5% (79·0–79·9), and an overall accuracy of 79·4% (79·0–79·9; table).We also tested the effect of using multiple sinus rhythm ECGs from the same patient, as

	AUC	Sensitivity	Specificity	F1 score	Accuracy
Main analysis	0·87 (0·86–0·88)	79·0% (77·5–80·4)	79·5% (79·0–79·9)	39·2% (38·1–40·3)	79·4% (79·0–79·9)
Secondary analysis	0·90 (0·90–0·91)	82·3% (80·9–83·6)	83·4% (83·0–83·8)	45·4% (44·2–46·5)	83·3% (83·0–83·7)

Data in parentheses are 95% CIs. In the main analysis, only the score of the first normal sinus rhythm ECG in the window of interest was used. In the secondary analysis, the highest score for all ECGs done in the first month of the window of interest was used. AUC=area under the curve. ECG=electrocardiograph.

***Table*: Model performance**

the additional data seemed likely to improve the network performance of AI-enabled ECG. Multiple ECGs provide the model with more information about each patient and might mask outliers. When testing the model on all of the sinus rhythm ECGs in the first 31 days from the study start date and selecting the average and maximum probability of atrial fibrillation scores, the AUC improved to 0·89 (0·89–0·90) using the average score on the test dataset and to 0·90 (0·90–0·91) when applying a more sensitive approach of using the score of the ECG with the highest risk (figure 3; table). Similar improvements were found when doing the same analysis on the internal validation set: the AUC improved to 0·89 (0·89–0·90) using the average score and to 0·90 (0·89–0·91) when applying a more sensitive approach of using the score of the ECG with the highest risk. In another secondary analysis on the testing dataset, we included only the first normal sinus rhythm after the onset of atrial fibrillation and the AUC of the network improved to 0·90 (0·89–0·91).

As in the primary analysis, we found the probability threshold that yielded a similar sensitivity and specificity on the internal validation set and used that to classify the patients in the testing dataset. When using the maximum score with the calculated threshold, the F1 score improved to 45·4% (95% CI 44·2–46·5), sensitivity improved to 82·3% (80·9–83·6), and specificity improved to 83·4% (83·0–83·8) with an overall accuracy of 83·3% (83·0–83·7) on the testing dataset.

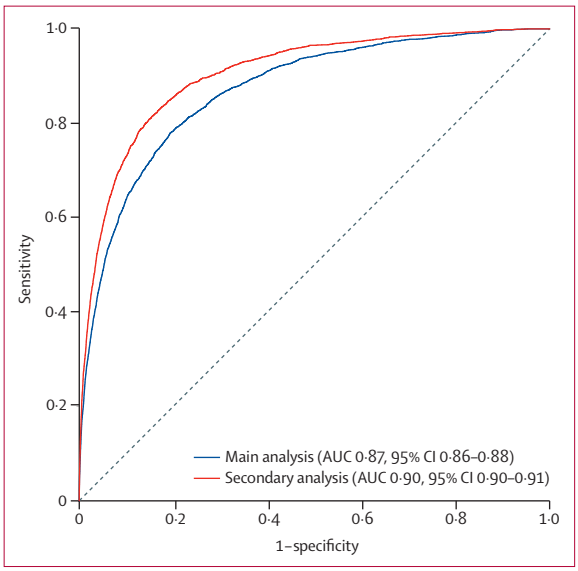

***Figure 3*: ROC curves for the convolutional neural networks on the testing dataset**

In the main analysis, only the score of the first normal sinus rhythm ECG in the window of interest was used. In the secondary analysis, the highest score for all ECGs done in the first month of the window of interest was used. ROC=receiver operating characteristic. AUC=area under the curve. ECG=electrocardiograph.

Discussion

In this study, we found that the AI-enabled ECG recorded during normal sinus rhythm performed well (AUC 0·87 for a single ECG and 0·90 for multiple) in identifying the presence of atrial fibrillation. This compares favourably with other medical screening tests such as B-type natriuretic peptide for heart failure (AUC 0·60–0·70),[20] Papanicolaou smear for cervical cancer (AUC 0·70),[21] and the CHA_2DS_2-VASc Score for stroke risk (AUC 0·57–0·72).[22] The ability to identify undetected atrial fibrillation with an inexpensive, widely available, point-of-care test—an ECG recorded during normal sinus rhythm—has important practical implications, particularly for atrial fibrillation screening efforts or for the management of patients with ESUS. This study shows the power of leveraging modern computing technology, large datasets, non-linear models, and automated features extraction using convolution layers to potentially improve diagnosis and treatment of a highly prevalent and morbid disease state.

Underdiagnosed atrial fibrillation is a major cause of ESUS. Furthermore, ESUS is associated with an increased risk of recurrent stroke, and when atrial fibrillation is found, anticoagulation significantly reduces this risk.[6] However, several large, prospective, randomised studies have shown that empirical use of anticoagulation after ESUS provides no benefit and might even cause harm, thus excluding the option of population-wide provision of anticoagulation without a clinical diagnosis of atrial fibrillation.[7,8] Although it would require further study, it is possible that this algorithm could identify a high-risk subset of patients with ESUS who could benefit from empirical anticoagulation.

Ambulatory monitoring following stroke or transient ischaemic attack for up to 48 h identifies atrial fibrillation in 2·4–13·9% of patients,[10,23] whereas prolonged monitoring with an implantable recorder detects atrial fibrillation in 30% of patients at 36 months.[1,2] Thus, short-term monitoring underdetects atrial fibrillation and long-term monitoring leaves a substantial

proportion of patients unprotected from recurrent atrial fibrillation and potential recurrent thromboembolic events, until such time as atrial fibrillation is detected. However, such monitoring is also expensive and can prove a burden to patients and clinical practices. Thus, identifying those patients who would most benefit from intensive monitoring would be valuable in patients after ESUS. Our data indicate that a simple, inexpensive, non-invasive, 10-second test—the AI-enhanced standard ECG—might permit identification of patients with underdetected atrial fibrillation. Further investigations will be necessary to confirm the diagnostic performance of AI-enabled ECG in specific populations, such as patients with ESUS or heart failure, to determine whether AI-enabled ECG could be used to refine the selection of candidates for prolonged ambulatory cardiac rhythm monitoring or to guide initiation of anticoagulation in these patients.

We also note that the threshold for a positive result could be altered for various clinical applications. The current binary cutoff was chosen to balance sensitivity and specificity, but a more sensitive cutoff point might be useful in excluding patients who do not need monitoring of atrial fibrillation after stroke or a more specific cutoff point could be used for screening of otherwise healthy people with a low pretest probability of atrial fibrillation, for instance.

The structural changes that precede atrial fibrillation, which might include myocyte hypertrophy, fibrosis, and chamber enlargement, are likely to lead to subtle ECG changes, allowing for prediction of underlying atrial fibrillation. For example, although seldom reported on ECGs, evidence of interatrial block (ie, Bayés syndrome) has been shown to correlate with both risk of incident atrial fibrillation and stroke.[24,25] Moreover, studies suggest that normal sinus rhythm on an ECG might not reflect overall atrial function. Nearly a third of patients with atrial fibrillation undergoing cardioversion have non-sinus contraction of the left atrial appendage despite apparent sinus rhythm on the surface ECG.[26] Additionally, in one study,[27] nearly a quarter of patients undergoing transoesophageal echocardiogram showed fibrillation of the left atrial appendage despite apparent sinus rhythm on the ECG. Thus, it is possible that wavelets on the ECG smaller than the readily observable P wave might reflect regional non-sinus electrical activity in these patients. A neural network trained with exposure to more than 500 000 ECGs and with sufficient depth to extract and recall subtle features not routinely appreciated or formally reported by human observers might be powerful enough to identify such features, which might account for our findings.

Once a network is trained, it can be applied to any standard digital 12-lead ECG with minimal computing power requirements; for example, a smartphone could process the signals. In the future, this might facilitate point-of-care diagnosis by allowing application of the algorithm on low-cost, widely available technologies. For instance, we have previously shown the translation of neural networks created using 12-lead ECGs to mobile, smartphone-based electrodes that typically include a single lead.[28]

The algorithm output could also be seen as a biomarker. One could draw a comparison between this algorithm and a glycated haemoglobin value in that they both provide an indicator of a disease state averaged over time. An individual ECG, which might or might not show atrial fibrillation, could be analogous to a random blood glucose test, which might or might not accurately reflect the presence of diabetes. However, just as an elevated glycated haemoglobin value can detect diabetes in a patient with an isolated normal fasting glucose, the AI-enabled ECG can detect atrial fibrillation in a patient during normal sinus rhythm.

This result is borne out by prior work from our group using neural networks to screen for the presence of asymptomatic left ventricular dysfunction using a standard 12-lead ECG, wherein the network effectively identified patients with ventricular dysfunction (AUC 0·93).[12] We found that the network predicted the future development of ventricular dysfunction, indicating that early disease affects myocytes' ability to generate electrical currents in a subtle manner before development of overt dysfunction. That finding suggests that, in a similar manner, the present network might identify structural disease before atrial fibrillation develops.

Other research teams have evaluated discrete ECG features (eg, PR interval, P wave dispersion, P wave signal averaging) as predictors of atrial fibrillation, but no individual feature has high enough predictive value to offer clinical utility. It is likely that these various features interact in some non-linear fashion that cannot be accounted for through traditional statistical methods or algorithmic approaches.

Computational approaches such as neural networks afford the ability to consider complex datasets in the context of all contained data rather than preselected discrete data elements. However, a key limitation in existing neural networks is explainability. Identifying these features could be of importance because they might offer novel findings that could provide new therapeutic targets or allow for more certainty for clinicians who are otherwise trying to understand what drives the network's interpretation. Finding ways to peer into this so-called black box is an area of active ongoing investigation.

Our work is best understood in the context of its limitations. Patients were considered negative for atrial fibrillation if they did not have any verified atrial fibrillation, but it is likely that some patients in this group had undetected atrial fibrillation and were thus labelled incorrectly. As a corollary, we can hypothesise that some of the false-positive patients identified by the network as having a history of atrial fibrillation despite

being classified as negative for atrial fibrillation might actually have had undiagnosed atrial fibrillation.

Although the prevalence of atrial fibrillation in our study is similar to other large clinical ECG datasets,[29] it might be higher than in the general population.[30] The network, therefore, has been trained for retrospective classification of clinically indicated ECGs more so than for atrial fibrillation prediction in unselected patients. We anticipate that the network would perform well in other datasets of clinically indicated ECGs but would need further prospective calibration before widespread application to screening of a broader, ostensibly healthy population is justified.

In conclusion, an AI-enabled ECG acquired during normal sinus rhythm permits point-of-care identification of individuals with a high likelihood of atrial fibrillation. This result could have important implications for atrial fibrillation screening and for the management of patients with unexplained stroke.

Contributors

ZIA, PAN, and PAF designed the study. ZIA developed the neural network. ZIA and REC did the statistical analysis. ZIA, PAN, and PAF wrote the manuscript. ZIA, PAN, FLJ, SJA, AJD, BJG, REC, XY, AAR, BJE, SK, and PAF critically reviewed the manuscript.

Declaration of interests

We declare no competing interests.

Acknowledgments

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

References

1 Sanna T, Diener HC, Passman RS, et al. Cryptogenic stroke and underlying atrial fibrillation. *N Engl J Med* 2014; **370:** 2478–86.
2 Gladstone DJ, Spring M, Dorian P, et al. Atrial fibrillation in patients with cryptogenic stroke. *N Engl J Med* 2014; **370:** 2467–77.
3 Hart RG, Diener HC, Coutts SB, et al. Embolic strokes of undetermined source: the case for a new clinical construct. *Lancet Neurol* 2014; **13:** 429–38.
4 Martin DT, Bersohn MM, Waldo AL, et al. Randomized trial of atrial arrhythmia monitoring to guide anticoagulation in patients with implanted defibrillator and cardiac resynchronization devices. *Eur Heart J* 2015; **36:** 1660–68.
5 Seiffge DJ, Werring DJ, Paciaroni M, et al. Timing of anticoagulation after recent ischaemic stroke in patients with atrial fibrillation. *Lancet Neurol* 2019; **18:** 117–26.
6 Lip GYH, Banerjee A, Boriani G, et al. Antithrombotic therapy for atrial fibrillation: CHEST guideline and expert panel report. *Chest* 2018; **154:** 1121–201.
7 Hart RG, Sharma M, Mundl H, et al. Rivaroxaban for stroke prevention after embolic stroke of undetermined source. *N Engl J Med* 2018; **378:** 2191–201.
8 Mohr JP, Thompson JL, Lazar RM, et al. A comparison of warfarin and aspirin for the prevention of recurrent ischemic stroke. *N Engl J Med* 2001; **345:** 1444–51.
9 Chimowitz MI, Lynn MJ, Howlett-Smith H, et al. Comparison of warfarin and aspirin for symptomatic intracranial arterial stenosis. *N Engl J Med* 2005; **352:** 1305–16.
10 Seet RC, Friedman PA, Rabinstein AA. Prolonged rhythm monitoring for the detection of occult paroxysmal atrial fibrillation in ischemic stroke of unknown cause. *Circulation* 2011; **124:** 477–86.
11 Kottkamp H. Human atrial fibrillation substrate: towards a specific fibrotic atrial cardiomyopathy. *Eur Heart J* 2013; **34:** 2731–38.
12 Attia ZI, Kapa S, Lopez-Jimenez F, et al. Screening for cardiac contractile dysfunction using an artificial intelligence-enabled electrocardiogram. *Nature Med* 2019; **25:** 70–74.
13 January CT, Wann LS, Alpert JS, et al. 2014 AHA/ACC/HRS guideline for the management of patients with atrial fibrillation: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and the Heart Rhythm Society. *J Am Coll Cardiol* 2014; **64:** e1–76.
14 Page RL, Joglar JA, Caldwell MA, et al. 2015 ACC/AHA/HRS guideline for the management of adult patients with supraventricular tachycardia: executive summary: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines and the Heart Rhythm Society. *J Am Coll Cardiol* 2016; **67:** 1575–623.
15 Kirchhof P, Benussi S, Kotecha D, et al. 2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS. *Eur Heart J* 2016; **37:** 2893–962.
16 van Rossum G. Python tutorial, technical report CS-R9526. Amsterdam: Stichting Mathematisch Centrum, 1995.
17 Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift. *Proc Int Conf Mach Learn* 2015; **37:** 448–56.
18 Nagi J, Ducatelle F, Di Caro G, et al. Max-pooling convolutional neural networks for vision-based hand gesture recognition. 2011 IEEE International Conference on Signal and Image Processing Applications; Kuala Lumpur; Nov 16–18, 2011; 342–47.
19 Robin X, Turck N, Hainard A, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* 2011; **12:** 77.
20 Bhalla V, Isakson S, Bhalla MA, et al. Diagnostic ability of B-type natriuretic peptide and impedance cardiography: testing to identify left ventricular dysfunction in hypertensive patients. *Am J Hypertens* 2005; **18:** 73s–81s.
21 Chen Y, Cui Z, Xiao Z, et al. PAX1 and SOX1 methylation as an initial screening method for cervical cancer: a meta-analysis of individual studies in Asians. *Ann Transl Med* 2016; **4:** 365.
22 Wu JT, Wang SL, Chu YJ, et al. CHADS2 and CHA2DS2-VASc scores predict the risk of ischemic stroke outcome in patients with interatrial block without atrial fibrillation. *J Atheroscler Thromb* 2017; **24:** 176–84.
23 Rabinstein AA. Prolonged cardiac monitoring for detection of paroxysmal atrial fibrillation after cerebral ischemia. *Stroke* 2014; **45:** 1208–14.
24 Martinez-Selles M, Masso-van Roessel A, Alvarez-Garcia J, et al. Interatrial block and atrial arrhythmias in centenarians: prevalence, associations, and clinical implications. *Heart rhythm* 2016; **13:** 645–51.
25 Arboix A, Marti L, Dorison S, Sanchez MJ. Bayes syndrome and acute cardioembolic ischemic stroke. *World J Clin Cases* 2017; **5:** 93–101.
26 Bellotti P, Spirito P, Lupi G, Vecchio C. Left atrial appendage function assessed by transesophageal echocardiography before and on the day after elective cardioversion for nonvalvular atrial fibrillation. *Am J Cardiol* 1998; **81:** 1199–202.
27 Warraich HJ, Gandhavadi M, Manning WJ. Mechanical discordance of the left atrium and appendage: a novel mechanism of stroke in paroxysmal atrial fibrillation. *Stroke* 2014; **45:** 1481–84.
28 Yasin OZ, Attia Z, Dillon JJ, et al. Noninvasive blood potassium measurement using signal-processed, single-lead ecg acquired from a handheld smartphone. *J Electrocardiol* 2017; **50:** 620–25.
29 Norberg J, Backstrom S, Jansson JH, Johansson L. Estimating the prevalence of atrial fibrillation in a general population using validated electronic health data. *Clin Epidemiol* 2013; **5:** 475–81.
30 Schnabel RB, Yin X, Gona P, et al. 50 year trends in atrial fibrillation prevalence, incidence, risk factors, and mortality in the Framingham Heart Study: a cohort study. *Lancet* 2015; **386:** 154–62.