


RESEARCH

Open Access



Machine learning based potentiating impacts of 12-lead ECG for classifying paroxysmal versus non-paroxysmal atrial fibrillation

Sungsoo Kim^{1,2}, Sohee Kwon^{2,3,4}, Mia K. Markey^{5,6}, Alan C. Bovik¹, Sung-Hwi Hong², JunYong Kim², Hye Jin Hwang⁷, Boyoung Joung², Hui-Nam Pak², Moon-Hyeong Lee² and Junbeom Park^{8*} 

Abstract

Background: Conventional modality requires several days observation by Holter monitor to differentiate atrial fibrillation (AF) between Paroxysmal atrial fibrillation (PAF) and Non-paroxysmal atrial fibrillation (Non-PAF). Rapid and practical differentiating approach is needed.

Objective: To develop a machine learning model that observes 10-s of standard 12-lead electrocardiograph (ECG) for real-time classification of AF between PAF versus Non-PAF.

Methods: In this multicenter, retrospective cohort study, the model training and cross-validation was performed on a dataset consisting of 741 patients enrolled from Severance Hospital, South Korea. For cross-institutional validation, the trained model was applied to an independent data set of 600 patients enrolled from Ewha University Hospital, South Korea. Lasso regression was applied to develop the model.

Results: In the primary analysis, the Area Under the Receiver Operating Characteristic Curve (AUC) on the test set for the model that predicted AF subtype only using ECG was 0.72 (95% CI 0.65–0.80). In the secondary analysis, AUC only using baseline characteristics was 0.53 (95% CI 0.45–0.61), while the model that employed both baseline characteristics and ECG parameters was 0.72 (95% CI 0.65–0.80). Moreover, the model that incorporated baseline characteristics, ECG, and Echocardiographic parameters achieved an AUC of 0.76 (95% CI 0.678–0.855) on the test set.

Conclusions: Our machine learning model using ECG has potential for automatic differentiation of AF between PAF versus Non-PAF achieving high accuracy. The inclusion of Echocardiographic parameters further increases model performance. Further studies are needed to clarify the next steps towards clinical translation of the proposed algorithm.

Keywords: Machine learning, 12 leads surface electrocardiogram, Atrial fibrillation

Introduction

Atrial fibrillation (AF) is the most common cardiac arrhythmia and is associated with increased risk of stroke, heart failure, and mortality [1–3]. With an aging population, AF is estimated to effect over 17.9 million patients in Europe by 2060 and 6–12 million in the USA by 2050 [1]. This increased incidence of AF results in soaring health care costs [1]. AF is classified

*Correspondence: parkjb@ewha.ac.kr

⁸ Department of Cardiology, College of Medicine, Ewha Womans University, 1071, Annyangcheon-ro, Yangcheon-gu, Seoul, Republic of Korea

Full list of author information is available at the end of the article



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

into Paroxysmal (PAF) and Non-paroxysmal atrial fibrillation (Non-PAF) subtypes based on the duration of episodes [4]. Non-PAF is further classified into persistent and long-standing atrial fibrillation. The classification of subtypes of AF is critical to determine the proper management of patients with AF. For example, according to the 2017 HRS/EHRA/ECAS/APHRS report [4], catheter ablation in symptomatic PAF is strongly supported by the evidence (i.e., class I indication), but catheter ablation is of questionable value for Non-PAF patients (i.e., class IIa/IIb indication). The outcome of catheter ablation is also significantly superior to patients with PAF than to those with non-PAF. Therefore, earlier diagnosis of symptomatic PAF could enable more prompt identification of those patients most likely to benefit from catheter ablation. Furthermore, the incidence rate of stroke varies significantly by AF subtype [4]. Patients with Non-PAF have a higher incidence of stroke and increased mortality compared to those with PAF. Thus, patients with different AF subtypes require different plans for managing their complications. Current guidelines state that distinguishing between PAF and Non-PAF requires continuous 24 h Holter ECG monitoring for 7 days [4]. However, in practice it is not plausible to perform cardiac rhythm monitoring over such a long time period for most patients. Therefore, the development of a rapid and practical classification method may provide benefits in making decisions regarding different management plans according to the patient's subtype of AF. Recent advances in machine learning for ECG analysis suggest that an algorithm might be able to automatically identify AF subtypes without prolonged monitoring. An AI model was developed to analyze ECG recorded by single-lead monitoring that appeared to have better performance than cardiologists when classifying 15 types of arrhythmia including AF [5]. A promising algorithm further predicted patients who currently do not exhibit AF on ECG and will later develop AF [6]. However, it should be noted that most prior work on machine learning for ECG analysis has relied on techniques such as neural networks that do not provide enough interpretability to elucidate pathophysiologic correlations between clinical features and the disease. Furthermore, the studies using a machine learning model to classify AF subtypes into PAF and Non-PAF are limited.

Therefore, we proposed a machine learning model that uses 12-lead surface ECG for real-time classification of subtypes with AF into PAF versus Non-PAF. We hypothesized that pathophysiologic differences between PAF and Non-PAF can be captured by different patterns of fibrillatory waves within 10 s of 12-lead surface ECG, which can enable more rapid differentiation of PAF from

Non-PAF. We applied interpretable computational algorithms to identify pathophysiologic characteristics of subtypes of atrial fibrillation in ECG data.

Methods

Participants

We retrospectively selected a total 1341 patients with AF from two University Hospitals: 741 from Severance Hospital, Seoul, South Korea and 600 from Ewha University Hospital, Seoul, South Korea from January 2008 to December 2017 (eFigure 1, Supplement). This study was approved by the Institutional Review Boards of Severance Hospital (IRB number: 2017–2301-002) and Ewha University (IRB number: 2017–10-010–002). At the enrollment stage, all patient demonstrated atrial fibrillation on ECG.

Training/validation cohort

We included 741 patients with AF at Severance Hospital, Seoul, South Korea. We collected 10 seconds of surface 12-lead ECGs, which was recorded digitally before any treatments of AF including cardioversion or catheter ablation. Patients had no history of anti-arrhythmic medication within 14 days before ECG. The exclusion criteria include patients with past medical history of (1) valve disease or valve surgery, (2) coronary bypass surgery, or (3) structural heart disease. No technical exclusion criteria were included in order to guarantee stability of our algorithm.

Test cohort

We validated the predictive value of our model with an independent test cohort. The test cohort consisted of 600 patients with AF at Ewha University school of Medicine, Seoul, South Korea. The ECG acquisition process and exclusion criteria were identical to those of the training/validation cohort.

Definition of Atrial Fibrillation classification

AF was diagnosed if AF was detected on ECG obtained from outpatient's clinic or Holter reports according to guidelines [4]. Then, AF is classified into Paroxysmal, persistent, or long-standing atrial fibrillation by the ACC/AHA/ESC guideline [4]. Persistent or long-standing atrial fibrillation is defined as Non-PAF.

Baseline characteristics and echocardiographic parameters

Baseline characteristics and echocardiographic parameters were extracted electronically from each patients' EMR (electrical medical record). The baseline characteristics were: CHA₂DS₂-VASc Score [4], age, sex assigned at birth, and history of congestive heart failure, hypertension, diabetes mellitus, stroke or TIA, and/or vascular

disease. At the point of AF diagnosis, each patient was assessed by echocardiogram to assess the heart anatomy including Left atrium (LA) volume, Left Ventricle anterior–posterior (AP) diameter, and Left Ventricle Ejection fraction.

Electrocardiogram (ECG)

In contrast to previous research using 24 hours Holter ECG [7–9], our analysis relies on only 10 seconds of routine 12-lead surface ECG. Temporal analysis [7, 8, 10, 11], frequency analysis [8, 9, 12, 13] or both [14–20] were considered to quantify AF global organization of ECG (eMtehd, Supplement). It was analyzed using conventional ECG parameters including Fibrillatory Wave Amplitude (FWA) [7, 14, 18, 20], Sample Entropy [8, 10, 11, 14, 17, 19, 21], Dominant Frequency [8, 9, 14, 18, 20], Spectral Entropy [16, 17, 19, 20] and Organization Index [13, 20]. Here, we propose the *Spectral Power Ratio* (SPR), the ratio of the power distribution in a lower frequency range versus a higher frequency range, calculated as

$$SPR_i = \frac{\int_{f_i}^{\infty} p(f)df}{\int_0^{f_i} p(f)df}$$

where f_i is the cut-off value to divide the power into a lower frequency range versus a higher frequency range for the i -th lead (eFigure 2, Supplement). An initial value of f_i was set as 10 Hz and updated during the training process.

Obtaining raw ECG data

Surface 12-Lead ECGs were recorded digitally with a 250 or 500 Hz sampling frequency using an electrophysiology recording system (GE Healthcare, Marquette, MAC5500, Waukesha, WI). ECG recording was composed of 10 s ECG data. The data were exported from the recording system to XML format and converted into CSV data file through a custom Python program.

ECG preprocessing

ECG recordings were preprocessed to reduce noise and interference for analysis of fibrillatory wave. All signal was upsampled to 1000 Hz for enhancing time alignment accuracy for later QRST complex subtraction [22]. Pan Tompkins algorithm was applied for automatic QRS detection using Butterworth bandpass filter (order: 3) [23, 24]. To avoid baseline wandering [25], high frequency noise [26] and possible powerline interference [27], ECG signal was filtered by band-pass filter between 1 and 50 Hz (6 order Chebyshev, type 2, 20-dB stop-band attenuation). Ventricular signal was cancelled by *adaptive singular value QRST cancellation*

[28, 29]. This method forms a matrix having multiple columns composed of QRST signal, and applies *Singular Value Decomposition* [30] to extract an eigen-vector of the matrix for QRST templates. Then, the template is multiplied by an adaptive coefficient, and subtracted. After QRST cancellation, the signal was filtered by an additional 3 Hz high-pass filter to suppress interference caused by possible residual T wave [31].

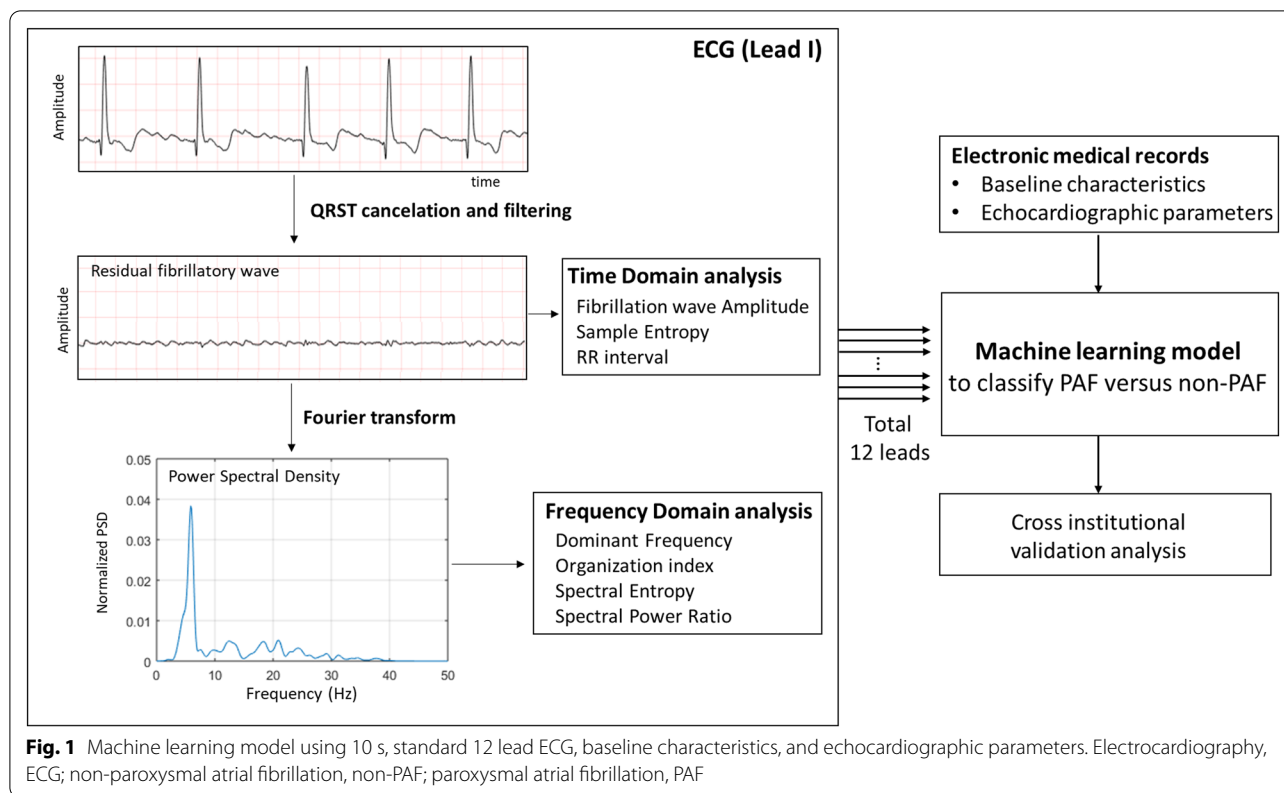
Statistical analysis and prediction model based on machine learning

Continuous variables were reported as the mean \pm the standard deviation, whereas categorical variables were reported as frequencies (percentages). Pearson Chi-square tests were applied for categorical variables, while Wilcoxon tests were used for continuous variables.

The machine learning algorithm was trained on the training/validation cohort. As the primary analysis, the ML model was trained on only ECG parameters from 12 lead. As the secondary analysis, the ML model was trained on (1) ECG parameters from 12 leads, (2) baseline characteristics, and (3) echocardiographic parameters.

Least absolute shrinkage and selection operator (LASSO) regression was applied to fit the β coefficients of the predictive models [32]. Ten-fold cross-validation was performed on the training/validation cohort. The maximum number of nonzero coefficients of the lasso coefficients was 25; the maximum number of iterations was 1000; and the convergence threshold of the coordinate descent algorithm was 0.0001. Each numeric variable was standardized with zero mean and unit standard deviation.

The predictive model was evaluated in terms of the Area Under Curve (AUC) of the Receiver Operating Characteristic (ROC) curve using an independent test cohort. In addition, the calibration curve and its c-index were evaluated. For univariate and multivariate analysis, logistic regression was performed by using a limited number of variables selected by the Lasso regression for the predictive model. For univariate and multivariate analysis, all continuous variables were dichomatized into binary values (low vs. high) using the median as the cut-off value. All hypothesis tests were 2-sided, and a 2-sided $p < 0.05$ indicated statistical significance. Calculations were performed using MATLAB 2019 (Mathworks, CA) and SAS version 9.3. Figure 1 provides an overview of our machine learning model that was built using ECG parameters from 12-lead surface ECG, baseline characteristics, and echocardiographic parameters.



Results

Table 1 provides the baseline characteristics and echocardiologic parameters for patients in the training/validation cohort and the test cohort. The training/

validation cohort consisted of 437 patients with Non-PAF and 334 patients with PAF. There was no statistically significant difference between the baseline characteristics of the two groups, except for history

Table 1 Baseline characteristics for patients with PAF and non-PAF in a training/validation cohort and a test cohort

Feature	Training and validation set (Hospital A)			Test set (Hospital B)		
	PAF (n = 334)	Non-PAF (n = 437)	p value	PAF (n = 103)	Non-PAF (n = 497)	p value
<i>Baseline characteristics</i>						
Age, year	58 (12.1)	58 (10.8)	0.75	69 (10.1)	68 (11.4)	0.42
Female gender, n (%)	91 (27%)	95 (21%)	0.09	51 (49%)	241 (48%)	0.91
Chronic Heart Failure, n (%)	21 (6%)	70 (16%)	<0.001	20 (20%)	97 (19%)	0.99
Hypertension, n (%)	164 (49%)	205 (46%)	0.56	59 (57%)	281 (56%)	0.91
Diabetes, n (%)	49 (15%)	79 (18%)	0.2	24 (23%)	121 (24%)	0.89
Cerebrovascular accident (incl. TIA), n (%)	39 (6%)	54 (6%)	0.71	21 (20%)	112 (22%)	0.69
Vascular disease, n (%)	41 (12%)	55 (13%)	0.91	5 (4%)	10 (2%)	0.15
CHA2DS2Vasc Score	1.7 (1.5)	1.73 (1.6)	0.94	2.74 (1.5)	2.75 (1.5)	0.91
<i>Echocardiographic factors</i>						
Left Atrium size, mm	40.8 (5.8)	44.4 (5.8)	<0.001	45.7 (8.1)	53.4 (9.5)	<0.001
Left Atrium volume, mm ³	35.1 (12.6)	43.3 (13.9)	<0.001	42.8 (20.2)	62.6 (37.9)	<0.001
Left Ventricular Ejection Fraction, %	65.8 (39.6)	60.8 (8.8)	0.01	56.8 (13.6)	58.8 (24.6)	0.41

Non-paroxysmal atrial fibrillation, non-PAF; paroxysmal atrial fibrillation, PAF

of congestive heart failure. 70 (16%) of Patients with Non-PAF had a history of congestive heart failure while only 21 (6%) of patients with PAF did ($p < 0.001$). For the echocardiographic parameters, the mean [SD] of the left atrium anterior–posterior diameter was larger in patients with Non-PAF (44.4 [5.8] mm) than in those with PAF (40.8 [5.8] mm) ($p < 0.001$). Similarly, left atrium volume was larger in patients with Non-PAF (43.3 [13.9] mm³) than in those with PAF (35.1 [12.6] mm³; $p < 0.001$). The left ventricle ejection fraction was lower in patients with Non-PAF (60.8 [8.8] %) than in those with PAF (65.8 [39.6] %; $p = 0.01$). The test cohort consisted of 103 patients with PAF and 497 patients with Non-PAF. None of the baseline characteristics were statistically significantly different between PAF and Non-PAF. Analysis of the echocardiographic parameters demonstrated that patients with Non-PAF had larger left atrium diameter ($p < 0.001$) and larger left atrium volume ($p < 0.001$) than did patients with PAF, which is consistent with what we observed for the training/validation cohort. However, no significant

difference was observed for left ventricular ejection fraction.

Table 2 demonstrates a mean (SD) of ECG parameters from Lead I, II, and III among 12-leads. In the training/validation cohort, a mean [\pm SD] of SPR of the lead I was higher in patients with Non-PAF (2.78 [0.50] mm) than those with PAF (2.63 [0.45]) ($p < 0.001$). FWA of the lead II was lower in patients with Non-PAF (26.4 [12.9] mm) than those with PAF (35.2 [15.8]) mm ($p < 0.001$). DF of the lead III was higher in patients with Non-PAF (6.4 [2.1] Hz) than those with PAF (6.0 [1.8] Hz; $p < 0.001$). OI of the lead I was lower in patients with Non-PAF (0.43 [0.09]) than those with PAF (0.46 [0.10]; $p < 0.001$). Spectral entropy of the lead I was higher in patients with Non-PAF (5.45 [6.0]) than those with PAF (5.28 [0.60]; $p < 0.001$). RR interval of the lead I was longer in patients with Non-PAF (865.5 [203.6] micro-sec) than those with PAF (766.2 [190.5] msec; $p < 0.001$). In the test cohort, all selected ECG parameters demonstrated similar trends from the training/validation cohort.dd

In the primary analysis, the performance of ML model using only ECG to differentiate PAF from non-PAF was

Table 2 ECG parameters for patients with PAF or non-PAF in a training/validation cohort or a test cohort

ECG parameters	Lead	Training and validation set (Hospital A)			Test set (Hospital B)		
		PAF (n = 334)	Non-PAF (n = 437)	p value	PAF (n = 103)	Non-PAF (n = 497)	p value
Fibrillatory wave amplitude, uV	I	24.6 (11.6)	19.2 (9.2)	<.001	22.7 (12.5)	18.3 (11.1)	<.001
Sample entropy	I	-.01 (1.00)	-.02 (.96)	.850	.02 (1.00)	-.02 (1.02)	.775
Dominant frequency, Hz	I	5.61 (1.62)	6.05 (3.20)	.024	5.34 (1.32)	5.83 (1.84)	.011
Organization index	I	.46 (.10)	.43 (.09)	<.001	.45 (.08)	.42 (.09)	.027
Spectral entropy	I	5.28 (.60)	5.45 (.60)	<.001	5.35 (.50)	5.49 (.54)	.020
RR interval, mm	I	766.0 (190.7)	866.4 (203.8)	<.001	647.6 (184.8)	750.9 (188.8)	<.001
Spectral power ratio	I	2.63 (.45)	2.78 (.50)	<.001	2.61 (.34)	2.73 (.46)	.008
Fibrillatory wave amplitude, uV	II	35.2 (15.8)	26.4 (12.9)	<.001	36.2 (21.9)	25.2 (16.2)	<.001
Sample entropy	II	.04 (1.07)	.04 (.95)	.912	-.01 (.94)	.06 (1.00)	.546
Dominant frequency, Hz	II	5.75 (1.19)	5.95 (1.54)	.052	5.50 (1.42)	5.94 (1.16)	<.001
Organization index	II	.49 (.09)	.47 (.09)	.001	.48 (.09)	.45 (.08)	.007
Spectral entropy	II	4.91 (.62)	4.98 (.58)	.098	5.04 (.63)	5.20 (.56)	.010
RR interval, mm	II	766.2 (190.7)	866.0 (203.3)	<.001	647.3 (185.1)	750.6 (189.5)	<.001
Spectral power ratio	II	2.58 (.47)	2.66 (.48)	.033	2.59 (.41)	2.75 (.43)	<.001
Fibrillatory wave amplitude, uV	III	40.3 (19.6)	30.1 (16.3)	<.001	40.4 (29.6)	30.4 (26.1)	<.001
Sample entropy	III	.03 (.98)	-.01 (1.03)	.600	-.08 (.97)	.06 (1.01)	.199
Dominant frequency, Hz	III	5.98 (1.84)	6.37 (2.08)	.006	5.82 (1.76)	6.17 (1.34)	.021
Organization index	III	.45 (.10)	.44 (.09)	.119	.45 (.09)	.42 (.08)	.005
Spectral entropy	III	5.17 (.68)	5.21 (.62)	.476	5.23 (.59)	5.35 (.59)	.046
RR interval, mm	III	766.1 (190.4)	866.2 (203.2)	<.001	644.5 (185.4)	750.9 (192.5)	<.001
Spectral power ratio	III	3.17 (.64)	3.30 (.62)	.006	3.04 (.50)	3.21 (.51)	.002

A mean, a standard deviation, and p value were demonstrated. For clarity, only Lead I, II, and III are presented out of 12-leads. p value is tested by t-test Non-paroxysmal atrial fibrillation, non-PAF; paroxysmal atrial fibrillation, PAF

analyzed (ECG-only model). The model was trained using by ECG data only from the training/validation cohort. In the performance evaluation using the secondary cohort (Test cohort), the model achieved AUC of 0.72 (95% CI, 0.65–0.80).

As the secondary analysis, Fig. 2 illustrates the cross-institutional discrimination properties of three

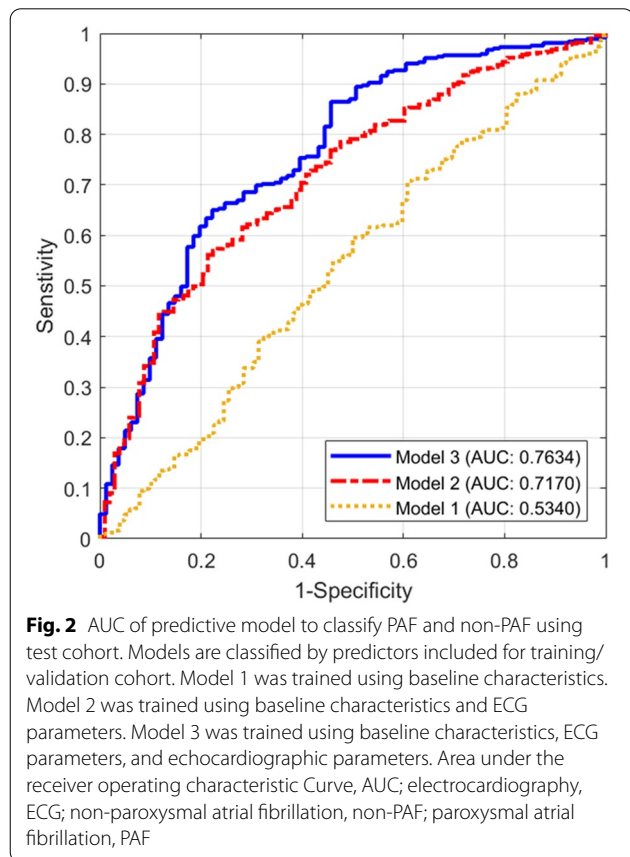


Fig. 2 AUC of predictive model to classify PAF and non-PAF using test cohort. Models are classified by predictors included for training/validation cohort. Model 1 was trained using baseline characteristics. Model 2 was trained using baseline characteristics and ECG parameters. Model 3 was trained using baseline characteristics, ECG parameters, and echocardiographic parameters. Area under the receiver operating characteristic Curve, AUC; electrocardiography, ECG; non-paroxysmal atrial fibrillation, non-PAF; paroxysmal atrial fibrillation, PAF

different prediction models. Only baseline characteristics were applied to build the Model 1. Both baseline characteristics and ECG parameters were applied to build Model 2. The baseline characteristics, ECG parameters, and echocardiographic parameters were applied to build Model 3. The performance of ROC curves increased in the order of Model 1 (0.53; 95% CI, 0.45–0.61; $p < 0.001$) Model 2 (0.72; 95% CI, 0.65–0.80; $p = 0.01$), and Model 3 (AUC = 0.76; 95% CI, 0.678–0.855 reference). The Model 3 achieves 0.756 of sensitivity and 0.603 of specificity (eTable 1, Supplement) in the cross-institutional validation. A generalized logistic regression with lasso regularization was used to select the highly predictive variables out of the 95 variables and to build a prediction model on the training/validation cohort (eFigure 3, Supplement). The prediction model was built with the 8 selected variables by lasso. The 8 selected variables were composed of left atrium AP diameter, left ventricle volume, FWA of Lead I, III, and aVR, a RR interval of V4, and SPR of Lead 1 and V2 (eTable 2, Supplement). The beta coefficients of left atrium AP diameter, left ventricle volume, a RR interval of V4, and SPR of Lead 1 and V2 had a positive amplitude in the prediction model. The beta coefficient for FWA of Lead I, III, and aVR had a negative amplitude. Baseline characteristics were not selected for the prediction model. The prediction model achieved ROC of 0.763 (95% CI 0.678–0.855) with a sensitivity of 0.756 and a specificity of 0.603 on the test cohort. A calibration curve of the prediction model demonstrated the c-index of 0.7622 on the test cohort (eFigure 4, Supplement). In this analysis, the baseline characteristic does not have any discriminating power to differentiate PAF from non-PAF: there is no statistically difference ($p > 0.05$) of AUC between “ECG-only model” vs. Model 2 (baseline characteristic + ECG).

Table 3 Beta-coefficients, OR (95% CI), and p values of multivariate logistic regression for non-PAF compared to PAF. Eight variables listed above were selected by Lasso regression

Variables*	Univariate analysis**			Multivariate analysis**		
	β -coefficient	OR (95% CI)	p value	β -coefficient	OR (95% CI)	p value
Left Atrium Anterior–Posterior diameter	1.20	3.31 (2.45–4.49)	<.001	0.74	2.11 (1.46–3.05)	<.001
Left Atrium volume	1.29	3.63 (2.68–4.90)	<.001	0.83	2.28 (1.59–3.29)	<.001
Fibrillatory wave amplitude (Lead II)	–1.11	0.33 (0.24–0.44)	<.001	–0.33	0.72 (0.48–1.08)	.110
Fibrillatory wave amplitude (Lead III)	–1.42	0.24 (0.18–0.33)	<.001	–0.94	0.39 (0.27–0.58)	<.001
Fibrillatory wave amplitude (aVR)	–0.95	0.39 (0.29–0.52)	<.001	–0.37	0.69 (0.48–0.98)	.040
RR interval (V4)	0.79	2.20 (1.64–2.94)	<.001	0.27	1.31 (0.92–1.86)	.131
Spectral Power Ratio (Lead 1)	0.55	1.73 (1.30–2.31)	<.001	0.33	1.39 (1.00–1.93)	.054
Spectral Power Ratio (V2)	0.51	1.66 (1.24–2.21)	<.001	0.40	1.50 (1.08–2.08)	.016

Area under the receiver operating characteristic Curve, AUC; confident interval, CI; electrocardiography, ECG; non-paroxysmal atrial fibrillation, non-PAF; paroxysmal atrial fibrillation, PAF

Table 3 demonstrates the estimated beta coefficients, ORs (95% CI), and p value by univariate and multivariate logistic regression analysis for the prediction model with 8 selected variables out of all parameters including baseline characteristics, ECG, and echocardiographic parameters. The multivariate OR for Non-PAF was 2.11 (95% CI 1.46–3.05) for LA diameter, 2.28 (95% CI 1.59–3.29) for LA volume, 1.50 (95% CI 1.08–2.08) for SPR of V2, 0.39 (95% CI 0.27–0.58) for FWA of Lead III, 0.69 (95% CI 0.48–0.98) for FWA of aVR, and 0.72 (95% CI 0.48–1.08) for FWA of Lead II.

Discussion

ECG is a key clinical modality for managing patients with atrial fibrillation [4]. This study is the first to develop a predictive model based on machine learning using 10 s of standard 12 lead ECG to classify subtypes of AF into PAF and Non-PAF. A fast, low cost method for classifying subtypes of atrial fibrillation would enable clinicians to decide among different treatment plans to manage patients according to their atrial fibrillation subtype. The successful cross-institutional validation of our predictive model supports the general relevance of the model for differentiating between PAF and Non-PAF.

This study showed that Non-PAF has a lower FWA, a higher DF, a lower OI, higher Spectral entropy, and longer RR interval as compared to PAF. These differences between Non-PAF and PAF can be interpreted in terms of their pathophysiologic implications. Atrial fibrillation leads to atrial remodeling and fibrosis [33, 34]. This abnormality can result in a higher number of waves and breakthroughs [35]. This structural heterogeneity may be associated with increased cancellation of electric signals causing low FWA with a complex AF and prolonged RR interval [35]. In addition, electrophysiological changes of myocytes are accompanied by structural remodeling [36]. This electrophysiological change results in shortening of atrial refractory periods causing increased left atrial rate [37], circle reentry [38], and multiple atrial sites generating electrical activities [39]. These changes may result in a lower frequency of fibrillatory wave, higher DF [18], lower OI [31], and lower Spectral Entropy [16].

In our study, SPR was the sole frequency domain parameter selected by Lasso regression in the predictive model. Among the frequency domain parameters, DF only considers a single frequency having maximum power in Power Spectral Density (PSD). To overcome this limitation of DF, OI considers several peaks in PSD with its neighbor frequency [20]. In contrast to DF or OI, SPR extracts a ratio of the power in the high frequency range to the power in the low frequency range in order to differentiate AF subtypes. From our

analysis, in Non-PAF power tends to shift from the lower frequency range to the higher frequency range, as compared to PAF. This power shift can be explained by the structural and electrophysiologic remodeling of atriums in Non-PAF compared to PAF [37, 38]. Using those ECG parameters, the ML model only using ECG achieved a fair performance to differentiate PAF from non-PAF.

In this study, we further observed that the performance of predictive models on ECG parameters improved with the addition of echocardiographic parameters. Although it has been reported that echocardiographic parameters are different between patients with Non-PAF and those with PAF^{50,51}, prior work has not investigated the independent discriminating power of echocardiographic parameter for classifying Non-PAF vs. PAF when combined with ECG parameters. Noticeably, our predictive model selected ECG parameters from different surface ECG leads. This suggests the importance of incorporating a diversity of leads to improve the performance of predictive models.

There are strengths to be noted. First, we developed a machine learning model for rapid and practical differentiating between PAF and Non-PAF. Notably, our machine learning model needs only 10 s monitoring by electrocardiograms to differentiate AF subtypes, whereas the traditional modality requires several days monitoring by Holter monitor. Second, our machine learning model is interpretable, our model therefore enables us to explain the structural and electrophysiologic remodeling of atriums in AF patients. Third, the predictive model incorporated information from both ECG and Echocardiography, the two most common cardiac assessment modalities, to build a single prediction model that demonstrated a synergic improvement in performance. Lastly, the cross-institutional validation supports the reliability of the predictive model.

The limitation of our study was known differences between the training/validation cohort and the test cohort. The baseline characteristics and echocardiographic parameters were statistically different ($p < 0.05$). However, noticeably, ECG parameters demonstrated similar statistical characteristics between the training/validation cohort and the test cohort. 49 (58%) of ECG parameters out of 84 demonstrated no significant difference (> 0.05) between patients with PAF in a training/validation cohort and patients with PAF in a test cohort. This guarantees reliability of ECG parameters to classify subtypes of atrial fibrillation into PAF and Non-PAF. In addition, our cohort is only consisted of Asian people. Additional validation with other racial groups is needed to determine if the model performance generalizes. In addition, even though different

levels of burden or different subtype of atrial fibrillation has been studied in prior research [4], this study does not demonstrate that the AF subtype which classification by ML algorithm is associated with different AF burden such as a risk of stroke. Furthermore, given the possible progressive nature of AF within paroxysmal, persistent, and long-standing, it may be difficult to estimate the clinical burden or risk based on the subtype of AF classified by an ML-based algorithm. Therefore, a long-term follow-up study using a prospective cohort is warranted.

Conclusion

The reported predictive model based on machine learning using 12 lead surface ECG can effectively classify subtypes of atrial fibrillation into Non-PAF and PAF. Furthermore, the predictive model achieved the highest performance when the available clinically relevant information including ECG, echocardiogram, and baseline characteristics, were incorporated. This study suggests the potential for predictive models based on machine learning to combine different clinical modalities, including ECG and echocardiogram. Furthermore, the predictive model enables interpretation in terms of pathophysiological differences between PAF and Non-PAF. These results may have important implications for the management of patients with atrial fibrillation according to their subtypes of atrial fibrillation.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s42444-022-00061-3>.

Additional file 1. Supplement contents includes (1) eFigure 1. Study overview, eMethod. electrocardiography analysis; (2) eFigure 2. Power spectral density of ECG signal after subtraction of QRST signal and normalized spectral power ratio. (3) eTable 1. AUC, Sensitivity, and Specificity of the predictive model to classify PAF and non-PAF. (4) eFigure 3. Lasso based training process according to Lambda; Trace plot of coefficients fit by Lasso (Left); cross-validation deviance of Lasso fit (Right). (5) eTable 2. Eight parameters selected by Lasso regression and their beta coefficient. (6) eFigure 4. Calibration curves and c-indexes for predictive model with a training/validation cohort (Left) and a test cohort (right).

Acknowledgements

Not applicable.

Authors' contributions

Sungsoo Kim; first author, data analysis, and making of machine learning algorithm. Sohee Kwon: presentation of additional idea during making machine learning algorithm. Mia K Markey: technical consultation related machine learning algorithm. Sung-Hwi Hong: data collection. JunYong Kim: data collection. Hye Jin Hwan: provides basic research-related concepts. Boyoung Joung: data collection. Hui-Nam Pak: data collection. Moon-Hyong Lee: data collection. Junbeom Park: correspondence author, leading entire study, data collection, and provides basic research-related concepts. All authors read and approved the final manuscript.

Funding

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (NRF-2017R1E1A1A01078382), by a grant from the Korean Heart Rhythm Society (KHRS2017-5), and by the Korea Medical Device Development Fund grant funded by the Korea government (the Ministry of Science and ICT, the Ministry of Trade, Industry and Energy, the Ministry of Health & Welfare, the Ministry of Food and Drug Safety) (Project Number: 9991006899).

Availability of data and materials

Not applicable.

Declarations

Ethical approval and consent to participate

All authors consent to participation in this study.

Consent for publication

We agree.

Competing interests

There is no competing interest.

Author details

¹Electrical and Computer Engineering, The University of Texas at Austin, Austin, USA. ²Department of Internal Medicine, College of Medicine, Yonsei University, Seoul, South Korea. ³Department of Epidemiology, Harvard T. H. Chan School of Public Health, Boston, MA, USA. ⁴Harvard University, Cambridge, USA. ⁵Department of Biomedical Engineering, The University of Texas at Austin, Austin, USA. ⁶Department of Imaging Physics, The University of Texas MD Anderson Cancer Center, Houston, USA. ⁷Cardiovascular Division, Department of Medicine, Beth Israel Deaconess Medical Center, Boston, MA, USA. ⁸Department of Cardiology, College of Medicine, Ewha Womans University, 1071, Anyangcheon-ro, Yangcheon-gu, Seoul, Republic of Korea.

Received: 27 October 2021 Accepted: 19 January 2022

Published online: 02 May 2022

References

- 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(9989):154–62.
- Miyasaka Y, Barnes ME, Gersh BJ, et al. Secular trends in incidence of atrial fibrillation in Olmsted County, Minnesota, 1980 to 2000, and implications on the projections for future prevalence. *Circulation*. 2006;114(2):119–25.
- Krijthe BP, Kunst A, Benjamin EJ, et al. Projections on the number of individuals with atrial fibrillation in the European Union, from 2000 to 2060. *Eur Heart J*. 2013;34(35):2746–51.
- Calkins H, Hindricks G, Cappato R, et al. 2017 HRS/EHRA/ECAS/APHS/SOLAECE expert consensus statement on catheter and surgical ablation of atrial fibrillation. *Ep Europace*. 2018;20(1):e1–160.
- Hannun AY, Rajpurkar P, Haghpanahi M, et al. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nat Med*. 2019;25(1):65.
- Attia ZI, Noseworthy PA, Lopez-Jimenez F, 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(10201):861–7.
- Alcaraz R, Rieta JJ. The application of nonlinear metrics to assess organization differences in short recordings of paroxysmal and persistent atrial fibrillation. *Physiol Meas*. 2009;31(1):115.
- Alcaraz R, Sandberg F, Sörnmo L, Rieta JJ. Classification of paroxysmal and persistent atrial fibrillation in ambulatory ECG recordings. *IEEE Trans Biomed Eng*. 2011;58(5):1441–9.
- Chiarugi F, Varanini M, Cantini F, Conforti F, Vrochos G. Noninvasive ECG as a tool for predicting termination of paroxysmal atrial fibrillation. *IEEE Trans Biomed Eng*. 2007;54(8):1399–406.

10. Nault I, Lellouche N, Matsuo S, et al. Clinical value of fibrillatory wave amplitude on surface ECG in patients with persistent atrial fibrillation. *J Interv Card Electrophysiol*. 2009;26(1):11–9.
11. Petersson R, Sandberg F, Platonov PG, Holmqvist F. Noninvasive estimation of organization in atrial fibrillation as a predictor of sinus rhythm maintenance. *J Electrocardiol*. 2011;44(2):171–5.
12. Bollmann A, Tveit A, Husser D, et al. Fibrillatory rate response to candesartan in persistent atrial fibrillation. *Europace*. 2008;10(10):1138–44.
13. Everett TH, Moorman JR, Kok L-C, Akar JG, Haines DE. Assessment of global atrial fibrillation organization to optimize timing of atrial defibrillation. *Circulation*. 2001;103(23):2857–61.
14. Alcaraz R, Hornero F, Rieta JJ. Noninvasive time and frequency predictors of long-standing atrial fibrillation early recurrence after electrical cardioversion. *Pacing Clin Electrophysiol*. 2011;34(10):1241–50.
15. Meo M, Zarzoso V, Meste O, Latcu DG, Saoudi N. Non-invasive prediction of catheter ablation outcome in persistent atrial fibrillation by exploiting the spatial diversity of surface ECG. Paper presented at: Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE2011.
16. Uldry L, Van Zaen J, Prudat Y, Kappenberger L, Vesin J-M. Measures of spatiotemporal organization differentiate persistent from long-standing atrial fibrillation. *Europace*. 2012;14(8):1125–31.
17. Nilsson F, Stridh M, Bollmann A, Sörnmo L. Predicting spontaneous termination of atrial fibrillation using the surface ECG. *Med Eng Phys*. 2006;28(8):802–8.
18. Xi Q, Sahakian AV, Frohlich TG, Ng J, Swiryn S. Relationship between pattern of occurrence of atrial fibrillation and surface electrocardiographic fibrillatory wave characteristics. *Heart Rhythm*. 2004;1(6):656–63.
19. Rezek I, Roberts SJ. Stochastic complexity measures for physiological signal analysis. *IEEE Trans Biomed Eng*. 1998;45(9):1186–91.
20. Lankveld T, Zeemering S, Scherr D, et al. Atrial fibrillation complexity parameters derived from surface ECGs predict procedural outcome and long-term follow-up of stepwise catheter ablation for atrial fibrillation. *Circul Arrhyth Electrophysiol*. 2016;9(2):e003354.
21. Richman JS, Moorman JR. Physiological time-series analysis using approximate entropy and sample entropy. *Am J Physiol Heart Circ Physiol*. 2000;278(6):H2039–2049.
22. Bollmann A, Husser D, Mainardi L, et al. Analysis of surface electrocardiograms in atrial fibrillation: techniques, research, and clinical applications. *Europace*. 2006;8(11):911–26.
23. Pan J, Tompkins WJ. A real-time QRS detection algorithm. *IEEE Trans Biomed Eng*. 1985;32(3):230–6.
24. Mitra SK, Kuo Y. *Digital signal processing: a computer-based approach*, vol. 2. New York: McGraw-Hill; 2006.
25. Boucheham B, Ferdi Y, Batouche MC. Piecewise linear correction of ECG baseline wander: a curve simplification approach. *Comput Methods Programs Biomed*. 2005;78(1):1–10.
26. Hamilton PS, Curley M, Aimi R. Effect of adaptive motion-artifact reduction on QRS detection. *Biomed Instrum Technol*. 2000;34(3):197–202.
27. Ferdjallah M, Barr RE. Adaptive digital notch filter design on the unit circle for the removal of powerline noise from biomedical signals. *IEEE Trans Biomed Eng*. 1994;41(6):529–36.
28. Alcaraz R, Rieta JJ. Adaptive singular value cancelation of ventricular activity in single-lead atrial fibrillation electrocardiograms. *Physiol Meas*. 2008;29(12):1351–69.
29. Alcaraz R, Rieta JJ. A non-invasive method to predict electrical cardioversion outcome of persistent atrial fibrillation. *Med Biol Eng Comput*. 2008;46(7):625–35.
30. Van Loan CF. *Matrix computations* (Johns Hopkins studies in mathematical sciences). Baltimore: The Johns Hopkins University Press; 1996.
31. Lankveld T, de Vos CB, Limantoro I, et al. Systematic analysis of ECG predictors of sinus rhythm maintenance after electrical cardioversion for persistent atrial fibrillation. *Heart Rhythm*. 2016;13(5):1020–7.
32. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw*. 2010;33(1):1.
33. Frustaci A, Chimenti C, Bellocci F, Morgante E, Russo MA, Maseri A. Histological substrate of atrial biopsies in patients with lone atrial fibrillation. *Circulation*. 1997;96(4):1180–4.
34. Mihm MJ, Yu F, Carnes CA, et al. Impaired myofibrillar energetics and oxidative injury during human atrial fibrillation. *Circulation*. 2001;104(2):174–80.
35. Allesie MA, de Groot NM, Houben RP, et al. Electropathological substrate of long-standing persistent atrial fibrillation in patients with structural heart disease: longitudinal dissociation. *Circul Arrhyth Electrophysiol*. 2010;3(6):606–15.
36. Kim K-B, Rodefeld MD, Schuessler RB, Cox JL, Boineau JP. Relationship between local atrial fibrillation interval and refractory period in the isolated canine atrium. *Circulation*. 1996;94(11):2961–7.
37. Sih HJ, Zipes DP, Barbari EJ, Adams DE, Olgin JE. Differences in organization between acute and chronic atrial fibrillation in dogs. *J Am Coll Cardiol*. 2000;36(3):924–31.
38. Zrenner B, Ndrepepa G, Karch MR, et al. Electrophysiologic characteristics of paroxysmal and chronic atrial fibrillation in human right atrium. *J Am Coll Cardiol*. 2001;38(4):1143–9.
39. Kamel H, Okin PM, Elkind MS, Iadecola C. Atrial fibrillation and mechanisms of stroke: time for a new model. *Stroke*. 2016;47(3):895–900.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

