



OPEN

## Artificial intelligence predicts clinically relevant atrial high-rate episodes in patients with cardiac implantable electronic devices

Min Kim<sup>1,14</sup>, Younghyun Kang<sup>2,14</sup>, Seng Chan You<sup>5</sup>, Hyung-Deuk Park<sup>2</sup>, Sang-Soo Lee<sup>2</sup>, Tae-Hoon Kim<sup>3</sup>, Hee Tae Yu<sup>3</sup>, Eue-Keun Choi<sup>6</sup>, Hyoung-Seob Park<sup>7</sup>, Junbeom Park<sup>8</sup>, Young Soo Lee<sup>9</sup>, Ki-Woon Kang<sup>10</sup>, Jaemin Shim<sup>11</sup>, Jung-Hoon Sung<sup>12</sup>, Il-Young Oh<sup>13</sup>, Jong Sung Park<sup>4</sup>✉ & Boyoung Joung<sup>3</sup>✉

To assess the utility of machine learning (ML) algorithms in predicting clinically relevant atrial high-rate episodes (AHREs), which can be recorded by a pacemaker. We aimed to develop ML-based models to predict clinically relevant AHREs based on the clinical parameters of patients with implanted pacemakers in comparison to logistic regression (LR). We included 721 patients without known atrial fibrillation or atrial flutter from a prospective multicenter (11 tertiary hospitals) registry comprising all geographical regions of Korea from September 2017 to July 2020. Predictive models of clinically relevant AHREs were developed using the random forest (RF) algorithm, support vector machine (SVM) algorithm, and extreme gradient boosting (XGB) algorithm. Model prediction training was conducted by seven hospitals, and model performance was evaluated using data from four hospitals. During a median follow-up of 18 months, clinically relevant AHREs were noted in 104 patients (14.4%). The three ML-based models improved the discrimination of the AHREs (area under the receiver operating characteristic curve: RF: 0.742, SVM: 0.675, and XGB: 0.745 vs. LR: 0.669). The XGB model had a greater resolution in the Brier score (RF: 0.008, SVM: 0.008, and XGB: 0.021 vs. LR: 0.013) than the other models. The use of the ML-based models in patient classification was associated with improved prediction of clinically relevant AHREs after pacemaker implantation.

Multiple clinical trials have demonstrated that longer atrial high-rate episodes (AHREs) are associated with an increased risk of atrial fibrillation (AF), ischemic stroke, and adverse cardiovascular outcomes<sup>1–5</sup>. Therefore, the current European Society of Cardiology (ESC) guidelines recommend comprehensive cardiovascular evaluation, including the stroke risk, medical comorbidities, and risk factors, in patients with AHREs detected by implanted devices. Notably, in patients with a longer duration of AHREs, more intensive monitoring can be more useful<sup>6,7</sup>. Previous clinical studies have demonstrated multiple predictors of AHREs in patients with pacemakers<sup>1,4,8–10</sup>. However, these inconsistent predictors paradoxically indicate that the etiology is ambiguous. Therefore, precisely estimating clinically relevant AHREs is an important part of optimal patient management after pacemaker implantation. Machine learning (ML) offers a computational and alternative approach to standard predictive modeling that recognizes complex characteristics within data<sup>11</sup>. To date, there have been several investigations

<sup>1</sup>Division of Cardiology, Chungbuk National University Hospital, Cheongju, Korea. <sup>2</sup>Medtronic Korea, Seoul, Korea. <sup>3</sup>Division of Cardiology, Department of Internal Medicine, Severance Cardiovascular Hospital, Yonsei University College of Medicine, 50-1 Yonsei-ro, Seodaemun-gu, Seoul 03722, Republic of Korea. <sup>4</sup>Division of Cardiology, Dong-A University Hospital, 26 Daesingongwon-ro, Seo-gu, Busan 49201, Republic of Korea. <sup>5</sup>Department of Preventive Medicine and Public Health, Yonsei University College of Medicine, Seoul, Korea. <sup>6</sup>Department of Cardiology, Seoul National University Hospital, Seoul, Korea. <sup>7</sup>Division of Cardiology, Keimyung University Hospital, Daegu, Korea. <sup>8</sup>Department of Cardiology, Ewha Womans University Hospital, Seoul, Korea. <sup>9</sup>Division of Cardiology, Daegu Catholic University Hospital, Daegu, Korea. <sup>10</sup>Division of Cardiology, Eulji University Hospital, Daejeon, Korea. <sup>11</sup>Department of Cardiology, Korea University Hospital, Seoul, Korea. <sup>12</sup>Department of Cardiology, CHA Bundang University Hospital, Seongnam, Korea. <sup>13</sup>Department of Cardiology, Seoul National University Bundang Hospital, Seongnam, Korea. <sup>14</sup>These authors contributed equally: Min Kim and Younghyun Kang. ✉email: thinkmed@dae.ac.kr; cby6908@yuhs.ac

of ML algorithms for coronary artery disease genetics<sup>12</sup>, cardiac resynchronization therapy outcomes, such as mortality and heart failure (HF)-related hospitalization<sup>13</sup>, HF with preserved ejection fraction<sup>14</sup>, and estimation of ventricular tachycardia recurrence and mortality after catheter ablation<sup>15</sup> in the cardiology era. We hypothesized that ML algorithms can produce a predictive model for clinically relevant AHREs in individual patients that can be more useful than previously reported predictors. Herein, we sought to determine which class of ML algorithms has the highest predictive accuracy using data from a prospective multicenter registry.

## Materials and methods

**Study design.** The cohort of patients in this study was derived from the evaluation of the Atrial Fibrillation occurrence in patients after Pacemaker implantation (AF-Pacemaker study), a prospective, multicenter, observational registry study performed in patients with AF aged > 18 years attending any of the 11 tertiary hospital centers comprising all geographical regions of Republic of Korea. The study enrollment period started in September 2017 and ended in July 2020.

The AF-Pacemaker study aimed to investigate the occurrence and management (including ablation therapy) of device-detected AF episodes in patients with pacemaker implants through a prospective, non-randomized, non-blinded, observational, multicenter design. The study was conducted in compliance with the ethical rules of the Declaration of Helsinki as a statement of ethical principles for medical research involving human subjects by the World Medical Association and approved by the Institutional Review Board of Yonsei University Health System (1-2017-0008). This study was registered at ClinicalTrials.gov (NCT03303872, First posted on October 6, 2017). The ethics committees of all 11 tertiary hospitals (Severance Hospital, Seoul National University Bundang Hospital, Seoul National University Hospital, Donga University Medical Center, Keimyung University Hospital, Ewha Womans University Medical Center, Daegu Catholic University Medical Center, Korea University Medical Center, Eulji University Hospital, CHA Bundang Medical Center, and Kangneung Asan Medical Center) approved this study, and all patients provided informed consent for their inclusion.

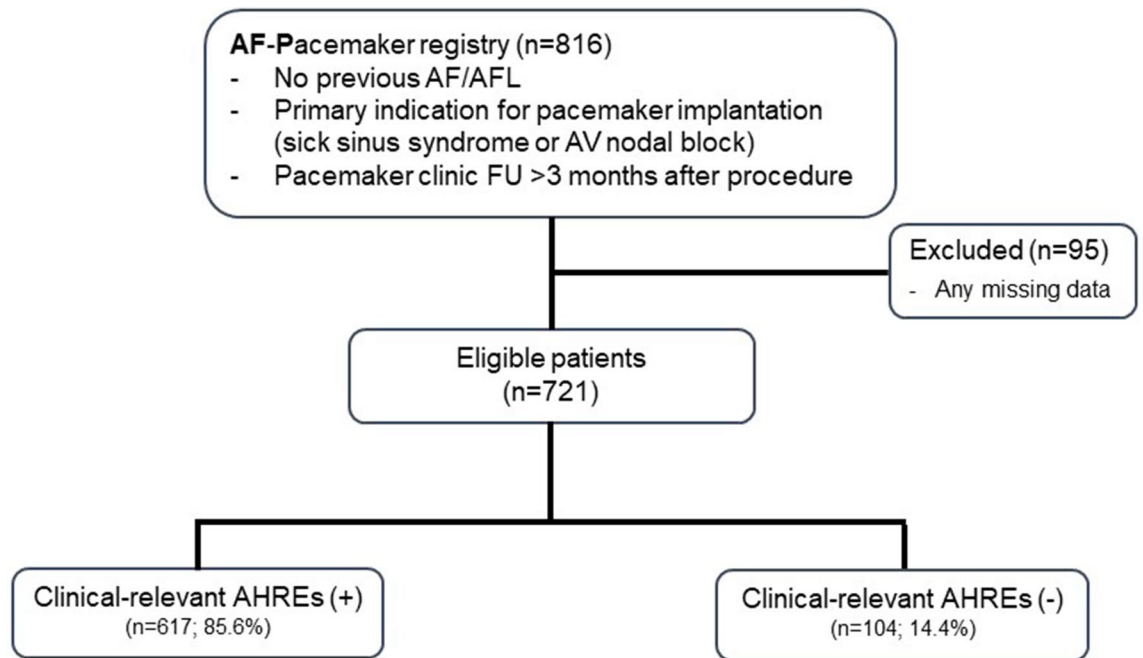
**Study population.** The study population included patients (1) eligible for permanent pacemaker implantation according to the guidelines on cardiac pacemaker implantation for sick sinus syndrome (sinus bradycardia, sinus pause of  $\geq 3$  s, tachy-brady syndrome, sinus node dysfunction, and chronotropic incompetence) or atrioventricular block (high-degree/complete atrioventricular block), (2) with an atrial sensing capability.

A total of 816 consecutive patients who were implanted with a St. Jude Medical dual-chamber rate-adaptive pacemaker (Assurity PM2240) with stored electrogram capabilities were enrolled. The pacemakers were incorporated with bipolar atrial and ventricular leads (Tendril MRI LPA1200M, Isoflex Optim 1944/1948, and Tendril ST Optim 1888TC) in all patients. The atrial and ventricular leads were placed in the right atrial appendage and right ventricular apex, respectively. We excluded patients with severe liver dysfunction (aspartate aminotransferase/alanine aminotransferase level  $\geq 3$  times the normal upper limit) or severe renal dysfunction (serum creatinine level of  $\geq 3.5$  mg/dL or creatinine clearance of  $\leq 30$  mL/min), including conditions requiring dialysis; pregnant or lactating patients; or those malignant cancer, dilated cardiomyopathy, hypertrophic cardiomyopathy, severe valvular heart disease, or life expectancy of  $\leq 12$  months from enrollment. Further, 95 patients with missing data for analysis were excluded (Fig. 1). Based on the available data from 721 patients 567 patients were used for model development and 154 for validation.

**Data collection, Pacemaker programming and AHREs detection.** For the AF-Pacemaker study, data were collected by independent clinical research coordinators via Web-based case report forms on the Internet-based Clinical Research and Trial management system (iCReaT), a data management system established by the Centers for Disease Control and Prevention, Ministry of Health and Welfare, Republic of Korea (iCReaT Study No. C170004). Each center could see its data and those of the other participating centers.

An AHRE detection rate of 220 beats/min was programmed, and storage of up to four atrial electrograms of 12-s duration on automatic detection of AHREs was activated. At every visit, the longest duration of all AHREs was ascertained. The longest AHRE (> 6 min) was defined as clinically relevant AHREs. Device interrogation information was obtained at regular clinic visits every 6 months after pacemaker implantation. An interval of up to 3 months before and after each clinic visit was allowed. The AHREs were interrogated to compare electrocardiogram (ECG) traces and pacemaker AHREs data at the end of the ambulatory monitoring period. The clinicians were blinded to the atrial diagnostic data. A detailed study protocol has been published previously<sup>16</sup>.

**Model development. Predictive feature selection.** We collected 28 variables available before the time of pacemaker implantation, including baseline patient demographic characteristics, clinical information, medications, and 12-lead ECG, laboratory examination, Holter monitoring, treadmill test, and transthoracic echocardiography findings. The predictors were age, sex, body mass index, current or former smoking status, current or former alcohol consumption, baseline heart rate, baseline systolic blood pressure, baseline diastolic blood pressure, indications for pacemaker implantation (sinus node dysfunction or atrioventricular node disease), estimated glomerular filtration rate (eGFR), left atrium (LA) diameter, left ventricular ejection fraction (LVEF), QRS duration, corrected QT (QTc) interval, presence of HF, hypertension, diabetes mellitus, prior stroke or transient ischemic attack (TIA), vascular disease, chronic kidney disease, presence of ventricular premature contraction, presence of atrial premature contraction, dyslipidemia, and medications, including renin-angiotensin-aldoosterone system blockers, beta adrenergic receptor blockers, calcium channel blockers, diuretics, and statins. Continuous variables were evaluated using point-biserial correlation coefficients<sup>17</sup> in relation to the clinically relevant AHREs, and the correlation of discrete variables was measured using Cramér's V<sup>18</sup>. Considering the two indicators, variables showing a low degree of correlation value using the knee point with the clinically rel-



\*Abbreviation: AF, atrial fibrillation; AFL, atrial flutter; AHREs, atrial high-rate episodes; AV, atrioventricular; FU, follow-up

**Figure 1.** Study population selection process.

event AHREs were excluded from the modeling process (Supplementary Table 1 and Supplementary Fig. 1). In addition, all continuous variables were normalized using Z-score normalization. Finally, the predictive models were constructed with subset data consisting of 10 variables based on model performance index for maximum prediction performance, excluding 18 variables, as shown in Supplementary Table 2.

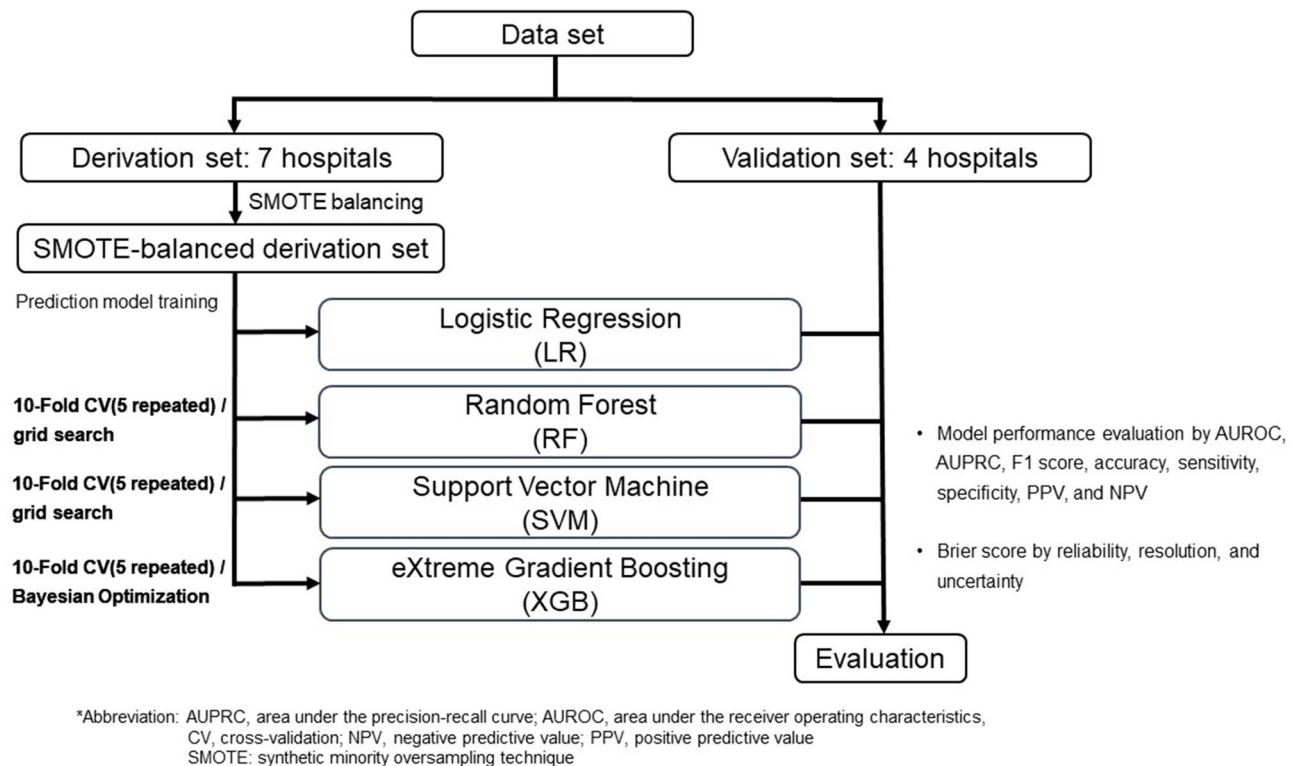
**Imbalanced data preprocessing.** When a predictive model is trained on imbalanced data, it tends to classify patterns biased into majority classes and ignores the characteristics of minority classes. Considering the imbalance problem of the study population, we applied a balancing method using the synthetic minority oversampling technique (SMOTE)<sup>19</sup>, and the minority class (clinically relevant AHREs) was oversampled. Clinically relevant AHREs were noted in 14.4% of the original dataset but in 30.8% of the SMOTE-balanced dataset after oversampling.

**Model derivation and algorithms.** To develop ML algorithms, we divided the study population into a training set, in which the prediction algorithm for clinically relevant AHREs was derived, and a test set, in which the algorithm was evaluated based on hospital units. The training set was derived from the patient data of seven hospitals and the test set from the patient data of the four remaining hospitals. We developed three ML-based predictive models: the random forest (RF)<sup>20</sup> algorithm, support vector machine (SVM)<sup>21</sup> algorithm, and extreme gradient boosting (XGB)<sup>22</sup> algorithm, in comparison to conventional logistic regression (Fig. 2).

RF algorithm is an ensemble learning method that creates multiple prediction results by multiple decision trees and determine the outcome by majority vote from decision tree results. We developed our RF model with 300 decision trees (Supplementary Fig. 2) and the average tree depth is  $16 \pm 2.1$  (mean  $\pm$  SD). SVM is an algorithm that finds a hyperplane that divides two categories of data in a multi-dimensional space of data. Lastly, XGB is an ensemble algorithm to create a strong decision tree classifier by combining weak tree classifiers built sequentially that each subsequent tree is trained from residuals of the previous tree to reduce its error.

We applied five repetitions of the tenfold cross validation (CV)<sup>23</sup> method to validate predictive models with hyperparameters tuned by a grid search algorithm for RF and SVM, and Bayesian optimization<sup>24</sup> algorithm for XGB. To determine how each variable affects the prediction of outcome, we calculated the feature importance of individual variables using the average absolute deviation by sensitivity analysis for each model<sup>25</sup>.

**Statistical analysis.** Continuous variables were presented as means  $\pm$  standard deviations for normally distributed values or medians and interquartile intervals for non-normally distributed values and categorical variables as numbers and percentages in each group. The baseline characteristics of the two groups were compared using Student's t-test or Wilcoxon test for continuous variables and Pearson's  $\chi^2$  test or Fisher's exact test for categorical variables. To evaluate the model discrimination and accuracy, we measured the area under the receiver operating characteristic (AUROC) curve, area under the precision–recall curve (AUPRC), F1-score, accuracy,



**Figure 2.** Flow diagram for the modeling process.

sensitivity, specificity, negative predictive value (NPV), and positive predictive value (PPV) and calculated the Brier score<sup>26</sup>, including reliability, resolution, and uncertainty. All analyses were performed using R (version 4.0.2; R Foundation for Statistical Computing, Vienna, Austria). The dedicated R packets and versions to validate and apply the 3 ML models has been published in GitHub (RF, R package version 4.6–14 <https://CRAN.R-project.org/package=randomForest>; SVM, R package version 0.9–29 <https://CRAN.R-project.org/package=kernlab>; XGB, R package version 1.4.1.1. <https://CRAN.R-project.org/package=xgboost>). Statistical significance was set at  $p$  values of  $<0.05$ .

## Results

Between September 2017 and July 2020, 721 patients were eligible for inclusion; their median age was 73 years, and 61.5% were women. The median follow-up duration was 18 months. Atrioventricular node disease (56.7%) was the most common indication for cardiac implantable electronic device implantation, followed by sinus node dysfunction (43.3%). A total of 104 patients (14.4%) experienced AHREs lasting for  $>6$  min, which were defined as clinically relevant AHREs in this study. The patients who experienced clinically relevant AHREs were significantly more likely to have a higher rate of sinus node dysfunction for pacemaker implantation, experience  $>1\%$  atrial premature contraction on pre-procedural Holter monitoring, use beta-adrenergic receptor blockers, and have a shorter QRS duration than those who did not experience clinically relevant AHREs (Table 1). The derivation set consisted of 567 patients from seven hospitals, and the validation set consisted of 154 patients from four hospitals. Supplementary Table 3 shows the characteristics of these two sets. The prevalence of former or current smokers and alcohol consumers and dyslipidemia, baseline systolic blood pressure, and LA diameter on echocardiography were higher in the derivation set. Further, this set of patients had lower baseline eGFR and LVEF and shorter QRS duration and QTc interval on electrocardiography.

**Model performance.** In the ML-based models, improvement in discrimination was achieved using the same data of the validation set (Table 2). The AUROCs achieved by each model was 0.742 for the RF algorithm, 0.675 for the SVM algorithm, and 0.745 for the XGB algorithm, which were numerically higher than 0.669 for logistic regression. The AUPRC and F1-score also improved, especially those for the XGB algorithm. The ROC and PRC curve plots are shown in Fig. 3. Using the Brier score, we achieved better performance in all three models based on the reliability, with lower values indicating higher agreement between observed and predicted risks, in comparison with logistic regression. The XGB algorithm had a higher resolution value, which indicated a prediction more accurate than that of the other algorithms across the spectrum risk.

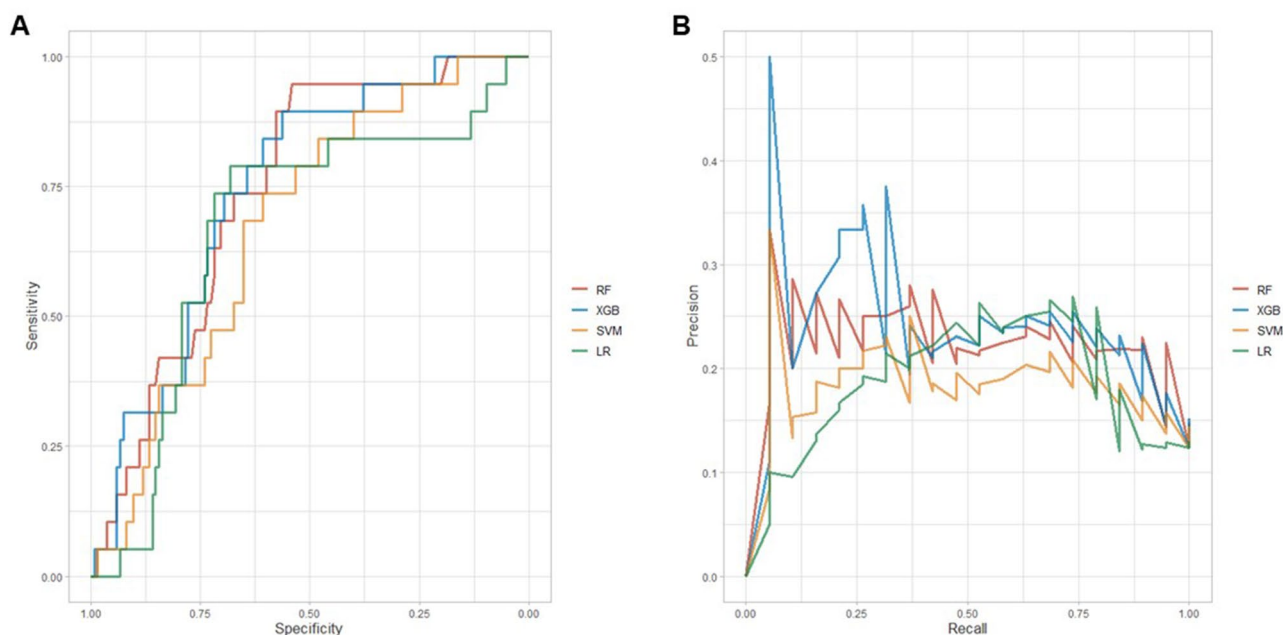
**Feature importance of the individual variables.** We observed that beta-adrenergic receptor blocker use, prior stroke or TIA, LA diameter, and  $>1\%$  atrial premature contraction on Holter monitoring before pacer

Variables	Total (n = 721)	Clinically relevant AHREs (-) (n = 617)	Clinically relevant AHREs (+) (n = 104)	p value
<b>Demographic</b>				
Age, (years)	73 (65, 79)	73 (65, 78)	74 (68, 80)	0.155
Male, n (%)	285 (39.5)	242 (39.2)	43 (41.3)	0.763
Body mass index, (kg/m <sup>2</sup> )	24.1 (22.2, 26.2)	24.2 (22.2, 26.2)	23.9 (22.1, 26.1)	0.566
Smoking, n (%)				
Former/Current	90 (12.5)	78 (12.6)	12 (11.5)	0.877
Alcohol, n (%)	93 (12.9)	80 (13.0)	13 (12.5)	1.000
Former/Current				
<b>Clinical</b>				
Heart failure, n (%)	26 (3.6)	23 (3.7)	3 (2.9)	1.000
Hypertension, n (%)	484 (67.1)	416 (67.4)	68 (65.4)	0.767
Diabetes, n (%)	196 (27.2)	173 (28.0)	23 (22.1)	0.256
Prior stroke/TIA, n (%)	80 (11.1)	64 (10.4)	16 (15.4)	0.181
Vascular disease, n (%)	71 (9.8)	58 (9.4)	13 (12.5)	0.422
Dyslipidemia, n (%)	222 (30.8)	197 (31.9)	25 (24.0)	0.134
Chronic kidney disease, n (%)	59 (8.2)	50 (8.1)	9 (8.7)	1.000
CHA <sub>2</sub> DS <sub>2</sub> VAS score*	3 (2, 4)	3 (2, 4)	3 (2, 4)	0.454
CHA <sub>2</sub> DS <sub>2</sub> VAS score, group				0.955
0, n (%)	21 (2.9)	18 (2.9)	3 (2.9)	
1, n (%)	104 (14.4)	90 (14.6)	14 (13.5)	
≥ 2, n (%)	596 (82.7)	509 (82.5)	87 (83.7)	
Pacemaker indication				<0.001
Sick sinus syndrome, n (%)	312 (43.3)	248 (40.2)	64 (61.5)	
AV block, n (%)	409 (56.7)	369 (59.8)	40 (38.5)	
Baseline systolic blood pressure, (mmHg)	135 (121, 148)	135 (122, 148)	133 (120, 145)	0.174
Baseline diastolic blood pressure, (mmHg)	71 (64, 80)	71 (64, 80)	72 (63, 80)	0.735
Baseline heart rate, (/min)	60 (50, 72)	60 (50, 72)	60 (50, 72)	0.954
Baseline eGFR, (mL/min/1.73 m <sup>2</sup> )	78.0 (62.0, 91.0)	77.0 (63.0, 92.0)	80.5 (62.0, 88.3)	0.894
<b>Electrocardiogram</b>				
QRS duration, (ms)	106 (90, 142)	108 (90, 144)	98 (88, 134)	0.038
QTc interval, (ms)	455 (422, 488)	455 (423, 488)	451 (420, 477)	0.301
<b>Echocardiography</b>				
LA diameter, (mm)	40 (36, 45)	40 (35, 45)	42 (37, 45)	0.094
LVEF, (%)	65 (60, 70)	65 (60, 70)	65 (60, 70)	0.868
<b>Holter recording</b>				
APC > 1% at pre-implantation, n (%)	53 (7.4)	38 (6.2)	15 (14.4)	0.005
VPC > 1% at pre-implantation, n (%)	49 (6.8)	39 (6.3)	10 (9.6)	0.306
<b>Medications</b>				
ARB/ACEi, n (%)	312 (43.3)	273 (44.2)	39 (37.5)	0.239
Beta adrenergic receptor blocker, n (%)	112 (15.5)	84 (13.6)	28 (26.9)	0.001
Calcium channel blocker, n (%)	224 (31.1)	194 (31.4)	30 (28.8)	0.678
Statin, n (%)	306 (42.4)	267 (43.3)	39 (37.5)	0.320
Diuretics, n (%)	162 (22.5)	139 (22.5)	23 (22.1)	1.000

**Table 1.** Baseline characteristics of the patients with and without clinically relevant AHREs. The data are presented as number (%), median [IQR]. ACEi angiotensin-converting-enzyme inhibitor, AHREs atrial high-rate episodes, APC atrial premature complex, ARB angiotensin receptor blocker, AV atrioventricular, GFR glomerular filtration rate, LA left atrium, LVEF left ventricular ejection fraction, QTc corrected QT, TIA transient ischemic attack, VPC ventricular premature complex. \*The CHA<sub>2</sub>DS<sub>2</sub>-VAS score is a measure of the risk of stroke in patients with atrial fibrillation, with scoring ranging from 0 to 9 and higher scores indicating greater risk. Congestive heart failure, hypertension, age 75 years or older (doubled), diabetes, stroke (doubled), vascular disease, age 65 to 74 years, sex category (female).

	Logistic regression	RF	SVM	XGB
<b>Model performance</b>				
AUROC (95% CI)	0.669 (0.536–0.803)	0.742 (0.637–0.835)	0.675 (0.561–0.789)	0.745 (0.631–0.847)
AUPRC (95% CI)	0.182 (0.104–0.274)	0.224 (0.119–0.397)	0.182 (0.102–0.337)	0.240 (0.125–0.424)
F1 score (95% CI)	0.853 (0.783–0.881)	0.888 (0.845–0.925)	0.865 (0.821–0.905)	0.896 (0.857–0.931)
Accuracy (95% CI)	0.753 (0.677–0.819)	0.805 (0.734–0.865)	0.773 (0.698–0.836)	0.818 (0.748–0.876)
Sensitivity (95% CI)	0.815 (0.739–0.876)	0.881 (0.815–0.931)	0.830 (0.755–0.889)	0.889 (0.823–0.936)
Specificity (95% CI)	0.316 (0.126–0.566)	0.263 (0.091–0.512)	0.368 (0.163–0.616)	0.316 (0.126–0.566)
PPV (95% CI)	0.194 (0.074–0.375)	0.238 (0.082–0.472)	0.233 (0.099–0.423)	0.286 (0.113–0.522)
NPV (95% CI)	0.894 (0.824–0.943)	0.895 (0.830–0.941)	0.903 (0.837–0.949)	0.902 (0.839–0.947)
<b>Brier score</b>				
Overall	0.181	0.138	0.158	0.141
Reliability	0.086	0.038	0.058	0.054
Resolution	0.013	0.008	0.008	0.021
Uncertainty	0.108	0.108	0.108	0.108

**Table 2.** Performance characteristics of models in the validation set for predicting clinically relevant AHREs in patients with pacemaker. *AHREs* atrial high-rate episodes, *AUPRC* area under the precision–recall curve, *AUROC* area under receiver operating characteristic, *CI* confidence interval, *NPV* negative predictive value, *PPV* positive predictive value, *RF* random forest, *SVM* support vector machine, *XGB* extreme gradient boosting.



**Figure 3.** Receiver operating characteristic curve analysis (A) and precision–recall curve analysis (B) for each model.

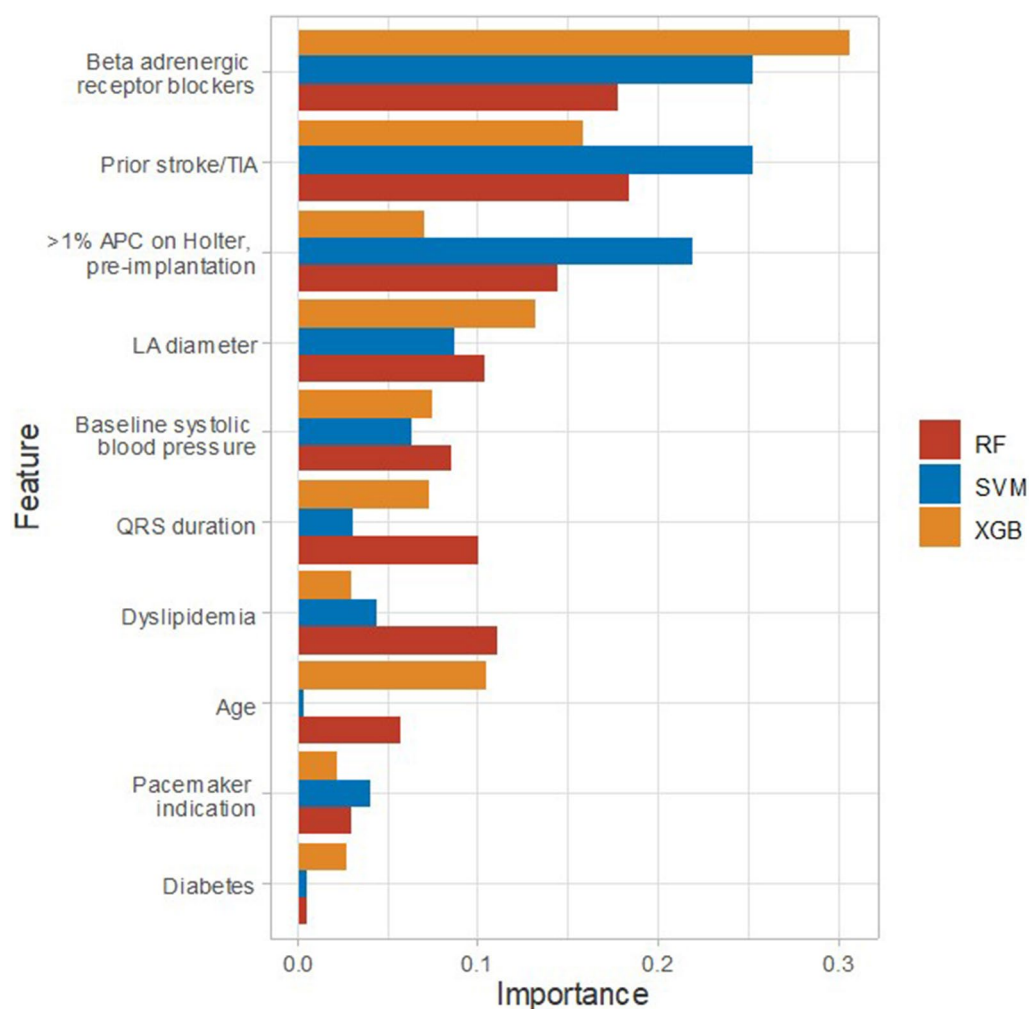
maker implantation consistently influenced the modeling process for predicting clinically relevant AHREs using the three ML-based models (Table 3 and Fig. 4).

## Discussion

In this prospective, multicenter, observational registry study, we found that the three ML algorithms (RF, SVM, and XGB) were better at identifying individuals who will develop clinically relevant AHREs among those with pacemaker implants. The XGB algorithm showed better performance and Brier scores than did the other algorithms. We also defined and calculated the index of feature importance of the variables in all three ML-based models. We found consistent feature importances for beta-adrenergic receptor blocker use, prior stroke or TIA, > 1% atrial premature contraction on Holter monitoring before pacemaker implantation, and LA diameter on echocardiography.

Rank	Random Forest		Support Vector Machine		eXtreme Gradient Boosting	
	Features	Importance*	Features	Importance*	Features	Importance*
1	Prior stroke/TIA	0.184	Beta adrenergic receptor blockers	0.254	Beta adrenergic receptor blockers	0.306
2	Beta adrenergic receptor blockers	0.178	Prior stroke/TIA	0.253	Prior stroke/TIA	0.158
3	> 1% APC on Holter, pre-implantation	0.145	> 1% APC on Holter, pre-implantation	0.220	LA diameter	0.132
4	Dyslipidemia	0.111	LA diameter	0.088	Age	0.105
5	LA diameter	0.104	Baseline systolic blood pressure	0.063	Baseline systolic blood pressure	0.075
6	QRS duration	0.101	Dyslipidemia	0.044	QRS duration	0.073
7	Baseline systolic blood pressure	0.085	Pacemaker indication	0.040	> 1% APC on Holter, pre-implantation	0.071
8	Age	0.057	QRS duration	0.030	Dyslipidemia	0.030
9	Pacemaker indication	0.030	Diabetes	0.005	Diabetes	0.028
10	Diabetes	0.006	Age	0.003	Pacemaker indication	0.023

**Table 3.** Feature importance in the three machine learning-based models. *AAD* average absolute deviation, *APC* atrial premature contraction, *CI* confidence interval, *TIA* transient ischemic attack, *LA* left atrium. \* Feature importance was measured from sensitivity analysis.



**Figure 4.** Feature importance index plot for each model.

The detection of clinically relevant AHREs with cardiac implantable electronic devices is becoming increasingly attractive for predicting the progression to AF or reducing the risk of embolic stroke. Although various predictors of AHREs<sup>4,8,27,28</sup> have been identified, they did not show highly consistent results. ML is powerful and has become ubiquitous and indispensable for solving complex medical problems. It offers an improved description and development of decision support tools to predict clinical events and encourage steps forward.

Although it has been reported that AHREs  $\geq 30$  s or  $\geq 2$  min is associated with cerebrovascular events in previous studies<sup>29,30</sup>, other studies have reported the more relevant diagnostic points for AHREs ( $> 6$  min)<sup>1,31,32</sup>. The current ESC guidelines also suggest that a minimum of 5 to 6 min of AHREs is associated with progression to AF, ischemic stroke, and major adverse cardiovascular events<sup>3,5,6</sup>.

To our knowledge, this is the first study to apply an ML algorithm to predict clinically relevant AHREs in patients with pacemaker implants and demonstrate improved prediction compared with that obtained with traditional statistical methods.

We selected the RF, SVM, and XGB algorithms and applied them to each predictive model in comparison to the logistic regression model. Because our dataset was relatively small and imbalanced, we used the five repeated tenfold CV method, instead of specifying a validation set in a training set separately, and the SMOTE to oversample the data set. In this study, we assessed the AUROC, AUPRC, F1-score, accuracy, sensitivity, specificity, NPV, and PPV and calculated the Brier score to evaluate the performance metrics of each model. The model with the XGB algorithm achieved the best performance. The predictive models also provide an opportunity to understand the features that may contribute to clinically relevant AHREs. This approach can identify the quantitation of feature importance for each variable, and we found consistent (beta-adrenergic receptor blocker use, prior stroke or TIA,  $> 1\%$  atrial premature contraction on Holter monitoring before pacemaker implantation, and LA diameter) factors in the predictive models. Some of these features (prior stroke or TIA and LA diameter) have been described previously<sup>8,16</sup>.

This study had several strengths. The predictive models were based on simple and readily available clinical characteristics. Although the models appear only as mathematical exercise, it can provide information that is indirectly helpful to clinicians. Recently, Perino et al.<sup>33</sup> reported that with increasing AHREs lasting for  $> 6$  min to  $> 24$  h, the stroke risk increased in those who did not receive anticoagulation treatment and mostly decreased in those who did. Vergara et al.<sup>34</sup> investigated the temporal association between AHREs and the risk of ventricular arrhythmias (VA). AHREs that precede VA increased the risk of VA recurrence. In these regards, the predicted probability for AHREs lasting for  $> 6$  min could be shared with the patient, and anticoagulation or antiarrhythmic therapy could serve as a critical element to prevent ischemic stroke or the recurrence of VA. The temporal association was observed. We compared the model evaluation performance metrics, and the XGB model yielded greater discrimination and accuracy than did the other models and logistic regression. Meanwhile, the SVM model exhibited relatively worse performance metrics than did the RF and XGB models because the ensemble model tends to have better performance and robustness than the single model. Our study showed that ML algorithms may play a role in precision cardiology. Although the baseline characteristics of the derivation and validation sets were slightly different, the model performance metrics showed acceptable results.

This study has some limitations. As the data sets, especially the validation set, were relatively small and contained limited features, the current analysis findings may not necessarily be representative of the predictors of clinically relevant AHREs in patients with implanted pacemakers. However, the patients were prospectively enrolled from 11 tertiary centers, which can yield some degree of generalizability. In this study, 61.5% were women. Many women were enrolled, but the ratio of men and women was comparable to similar studies in the same region<sup>35,36</sup>. Further assessment and improvement of the applicability of the predictive models are necessary for larger and different-race populations.

Finally, the models developed herein used only clinically relevant AHREs as the outcome data, not progression to clinical AF, ischemic stroke, MACEs, and death. Further improvements in multiple outcome prediction with higher accuracy should be explored.

## Conclusion

Our study illustrated the utility of ML algorithms in estimating clinically relevant AHREs in patients with implanted pacemakers using easily obtainable preimplantation features. Classification of patients using these models can support clinical decisions for anticoagulation therapy to prevent adverse outcomes in selected patients. From this perspective, models need to be built and validated individually for each diagnosis. More high-quality evidence can be obtained by applying ML algorithms; consequently, the data obtained can aid in the optimal management of these patients with shared decision-making.

## Data availability

The data that support the findings of this study are available from a web-based case report form on the Internet-based Clinical Research and Trial Management System (iCReaT), a data management system established by the Centers for Disease Control and Prevention, Ministry of Health and Welfare, Republic of Korea (iCReaT study no. C170004) but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of all investigators of the AF-Pacemaker study.

Received: 21 May 2021; Accepted: 8 November 2021

Published online: 07 January 2022

## References

1. Healey, J. S. et al. Subclinical atrial fibrillation and the risk of stroke. *N Engl. J. Med.* **366**, 120–129. <https://doi.org/10.1056/NEJMa1105575> (2012).
2. Glotzer, T. V. et al. Atrial high rate episodes detected by pacemaker diagnostics predict death and stroke: report of the Atrial Diagnostics Ancillary Study of the MDe Selection Trial (MOST). *Circulation* **107**, 1614–1619. <https://doi.org/10.1161/01.CIR.0000057981.70380.45> (2003).



3. Mahajan, R. *et al.* Subclinical device-detected atrial fibrillation and stroke risk: a systematic review and meta-analysis. *Eur. Heart J.* **39**, 1407–1415. <https://doi.org/10.1093/eurheartj/ehx731> (2018).
4. Gonzalez, M. *et al.* Newly detected atrial high rate episodes predict long-term mortality outcomes in patients with permanent pacemakers. *Heart Rhythm* **11**, 2214–2221. <https://doi.org/10.1016/j.hrthm.2014.08.019> (2014).
5. Pastori, D. *et al.* Atrial high-rate episodes and risk of major adverse cardiovascular events in patients with cardiac implantable electronic devices. *Clin. Res. Cardiol.* **109**, 96–102. <https://doi.org/10.1007/s00392-019-01493-z> (2020).
6. Hindricks, G. *et al.* 2020 ESC Guidelines for the diagnosis and management of atrial fibrillation developed in collaboration with the European Association for Cardio-Thoracic Surgery (EACTS). *Eur. Heart J.* **42**, 373–498. <https://doi.org/10.1093/eurheartj/ehaa612> (2021).
7. Pastori, D. *et al.* Incidence of bleeding in patients with atrial fibrillation and advanced liver fibrosis on treatment with vitamin K or non-vitamin K antagonist oral anticoagulants. *Int. J. Cardiol.* **264**, 58–63. <https://doi.org/10.1016/j.ijcard.2018.01.097> (2018).
8. Pedersen, K. B. *et al.* Subclinical atrial fibrillation in patients with recent transient ischemic attack. *J. Cardiovasc. Electrophysiol.* **29**, 707–714. <https://doi.org/10.1111/jce.13470> (2018).
9. Lopes, R. D. *et al.* Rationale and design of the apixaban for the reduction of thrombo-embolism in patients with device-detected sub-clinical atrial fibrillation (ARTESiA) trial. *Am. Heart J.* **189**, 137–145. <https://doi.org/10.1016/j.ahj.2017.04.008> (2017).
10. Belkin, M. N. *et al.* Incidence and clinical significance of new-onset device-detected atrial tachyarrhythmia: a meta-analysis. *Circ. Arrhythm. Electrophysiol.* **11**, e005393. <https://doi.org/10.1161/CIRCEP.117.005393> (2018).
11. Obermeyer, Z. & Emanuel, E. J. Predicting the future - big data, machine learning, and clinical medicine. *N Engl. J. Med.* **375**, 1216–1219. <https://doi.org/10.1056/NEJMp1606181> (2016).
12. Drenos, F., Grossi, E., Buscema, M. & Humphries, S. E. Networks in coronary heart disease genetics as a step towards systems epidemiology. *PLoS ONE* <https://doi.org/10.1371/journal.pone.0125876> (2015).
13. Kalscheur, M. M. *et al.* Machine learning algorithm predicts cardiac resynchronization therapy outcomes: lessons from the COM-PANION trial. *Circ. Arrhythm. Electrophysiol.* **11**, e005499. <https://doi.org/10.1161/CIRCEP.117.005499> (2018).
14. Shah, S. J. *et al.* Phenomapping for novel classification of heart failure with preserved ejection fraction. *Circulation* **131**, 269–279. <https://doi.org/10.1161/CIRCULATIONAHA.114.010637> (2015).
15. Vergara, P. *et al.* Predictive score for identifying survival and recurrence risk profiles in patients undergoing ventricular tachycardia ablation: the I-VT score. *Circ. Arrhythm. Electrophysiol.* **11**, e006730. <https://doi.org/10.1161/CIRCEP.118.006730> (2018).
16. Kim, M. *et al.* Prevalence and predictors of clinically relevant atrial high-rate episodes in patients with cardiac implantable electronic devices. *Korean Circ. J.* **51**, 235–247 (2021).
17. Tate, R. F. Correlation between a discrete and a continuous variable point-biserial correlation. *Ann. Math. Stat.* **25**, 603–607 (1954).
18. Cramér H. *Mathematical Methods of Statistics* (PMS-9). Princeton: Princeton University Press <https://doi.org/10.1515/9781400883868> (2016).
19. Chawla, N. V., Bowyer, K. W., Hall, L. O. & Kegelmeyer, W. P. SMOTE: Synthetic minority over-sampling technique. *J. Artif. Intell. Res.* **16**, 321–357. <https://doi.org/10.1613/jair.953> (2002).
20. Breiman, L. Random forests. *Mach. Learn.* **45**, 5–32. <https://doi.org/10.1023/A:1010933404324> (2001).
21. Cortes, C. & Vapnik, V. Support-vector networks. *Mach. Learn.* **20**, 273–297. <https://doi.org/10.1007/Bf00994018> (1995).
22. Chen, M. H. *et al.* XGBoost-based algorithm interpretation and application on post-fault transient stability status prediction of power system. *IEEE Access* **7**, 13149–13158. <https://doi.org/10.1109/Access.2019.2893448> (2019).
23. Stone, M. & Jensen, L. Cross-validatory choice of weights for inter-block and intra-block estimation in balanced incomplete block-designs - reply. *Biometrics* **36**, 553–553 (1980).
24. Kim, Y. & Chung, M. An approach to hyperparameter optimization for the objective function in machine learning. *Electronics-Switz* <https://doi.org/10.3390/electronics8111267> (2019).
25. Cortez, P. & Embrechts, M. J. Using sensitivity analysis and visualization techniques to open black box data mining models. *Inform. Sci.* **225**, 1–17. <https://doi.org/10.1016/j.ins.2012.10.039> (2013).
26. Brier, G. W. Verification of forecasts expressed in terms of probability. *Mon. Weather Rev.* **78**, 1. [https://doi.org/10.1175/1520-0493\(1950\)078%3c0001:Vofeit%3e2.0.Co;2](https://doi.org/10.1175/1520-0493(1950)078%3c0001:Vofeit%3e2.0.Co;2) (1950).
27. Glotzer, T. V. *et al.* The relationship between daily atrial tachyarrhythmia burden from implantable device diagnostics and stroke risk The TRENDS study. *Circ-Arrhythmia Electrophysiol.* **2**, 474–480. <https://doi.org/10.1161/Circep.109.849638> (2009).
28. Cheung, J. W. *et al.* Newly detected atrial fibrillation following dual chamber pacemaker implantation. *J. Cardiovasc. Electrophysiol.* **17**, 1323–1328. <https://doi.org/10.1111/j.1540-8167.2006.00648.x> (2006).
29. Nakano, M. *et al.* Impact of atrial high-rate episodes on the risk of future stroke. *J. Cardiol.* **74**, 144–149. <https://doi.org/10.1016/j.jjcc.2019.01.006> (2019).
30. Lu, W. D. & Chen, J. Y. The optimal cutoff of atrial high-rate episodes for neurological events in patients with dual chamber permanent pacemakers. *Clin. Cardiol.* **44**, 871–879. <https://doi.org/10.1002/clc.23626> (2021).
31. Witt, C. T. *et al.* Early detection of atrial high rate episodes predicts atrial fibrillation and thromboembolic events in patients with cardiac resynchronization therapy. *Heart Rhythm* **12**, 2368–2375. <https://doi.org/10.1016/j.hrthm.2015.07.007> (2015).
32. Turakhia, M. P. *et al.* Atrial fibrillation burden and short-term risk of stroke: case-crossover analysis of continuously recorded heart rhythm from cardiac electronic implanted devices. *Circ. Arrhythm. Electrophysiol.* **8**, 1040–1047. <https://doi.org/10.1161/CIRCEP.114.003057> (2015).
33. Perino, A. C. *et al.* Practice variation in anticoagulation prescription and outcomes after device-detected atrial fibrillation. *Circulation* **139**, 2502–2512. <https://doi.org/10.1161/CIRCULATIONAHA.118.038988> (2019).
34. Vergara, P. *et al.* Are atrial high-rate episodes associated with increased risk of ventricular arrhythmias and mortality?. *JACC Clin. Electrophysiol.* **5**, 1197–1208. <https://doi.org/10.1016/j.jacep.2019.06.018> (2019).
35. Kim, B. S. *et al.* Predictors and long-term clinical outcomes of newly developed atrial fibrillation in patients with cardiac implantable electronic devices. *Med. Baltimore* **95**, e4181. <https://doi.org/10.1097/MD.0000000000004181> (2016).
36. Park, Y. J., Kim, J. S., Park, K. M., On, Y. K. & Park, S. J. Subclinical atrial fibrillation burden and adverse clinical outcomes in patients with permanent pacemakers. *Stroke* **52**, 1299–1308. <https://doi.org/10.1161/STROKEAHA.120.031822> (2021).

## Acknowledgements

This study was supported by research grants from the Korean Healthcare Technology R&D project funded by the Ministry of Health & Welfare (HI15C1200 and HC19C0130) and Abbott Medical Korea Ltd.

## Author contributions

B.J. designed the study, and M.K., Y.K., S.C.Y. analyzed and interpreted the data. M.K. and Y.K. drafted the manuscript, and all authors did the final approval of the manuscript submission. H.D.P., S.S.L., T.H.K., H.T.Y., E.K.C. H.S.P., J.P., Y.S.L., K.W.K., J.S., J.H.S., and I.Y.O. interpreted data and contributed to acquiring patients' clinical data. J.S.P., and B.J. contributed to critical review of the manuscript.

## Funding

This research was supported by a grant of Patient-Centered Clinical Research Coordinating Center (PACEN) funded by the Ministry of Health & Welfare, Republic of Korea (grant number: HI19C0481, HC19C013, HI15C1200) and Abbott Medical Korea Ltd.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-03914-4>.

**Correspondence** and requests for materials should be addressed to J.S.P. or B.J.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**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/>.

© The Author(s) 2022