Reviews in Cardiovascular Medicine

Rev. Cardiovasc. Med. **2022**; 23(7): 231
https://doi.org/10.31083/j.rcm2307231

Original Research

A Territory-Wide Study of Arrhythmogenic Right Ventricular Cardiomyopathy Patients from Hong Kong

Ishan Lakhani[1,†], Jiandong Zhou[2,†], Sharen Lee[1], Ka Hou Christien Li[3], Keith Sai Kit Leung[4], Jeremy Man Ho Hui[1], Yan Hiu Athena Lee[1], Guoliang Li[5], Tong Liu[6], Wing Tak Wong[7], Ian Chi Kei Wong[8,9], Ngai Shing Mok[10], Chloe Miu Mak[11], Qingpeng Zhang[2,*], Gary Tse[6,*]

[1]Cardiovascular Analytics Group, Laboratory of Cardiovascular Physiology, Hong Kong, China
[2]School of Data Science, City University of Hong Kong, Hong Kong, China
[3]Faculty of Medicine, Newcastle University, NE1 7RU Newcastle upon Tyne, UK
[4]Aston Medical School, Aston University, B4 7ET Birmingham, UK
[5]Arrhythmia Unit, Department of Cardiovascular Medicine, First Affiliated Hospital of Xi'an Jiaotong University, 710061 Xi'an, Shaanxi, China
[6]Tianjin Key Laboratory of Ionic-Molecular Function of Cardiovascular Disease, Department of Cardiology, Tianjin Institute of Cardiology, Second Hospital of Tianjin Medical University, 300211 Tianjin, China
[7]State Key Laboratory of Agrobiotechnology (CUHK), School of Life Sciences, Chinese University of Hong Kong, Hong Kong, China
[8]Centre for Safe Medication Practice and Research, Department of Pharmacology and Pharmacy, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, China
[9]School of Pharmacy, University College London, WC1E 6BT London, UK
[10]Department of Medicine and Geriatrics, Princess Margaret Hospital, Hong Kong Hospital Authority, Hong Kong, China
[11]Department of Pathology, Hong Kong Children's Hospital, Hospital Authority, Hong Kong, China
*Correspondence: garytse86@gmail.com (Gary Tse); qingpeng.zhang@cityu.edu.hk (Qingpeng Zhang)
[†]These authors contributed equally.

Academic Editor: Stefan Peters
Submitted: 17 November 2021 Revised: 12 May 2022 Accepted: 16 May 2022 Published: 24 June 2022

Abstract

Background: Arrhythmogenic right ventricular cardiomyopathy/dysplasia (ARVC/D) is a hereditary disease characterized by fibrofatty infiltration of the right ventricular myocardium that predisposes affected patients to malignant ventricular arrhythmias, dual-chamber cardiac failure and sudden cardiac death (SCD). The present study aims to investigate the risk of detrimental cardiovascular events in an Asian population of ARVC/D patients, including the incidence of malignant ventricular arrhythmias, new-onset heart failure with reduced ejection fraction (HFrEF), as well as long-term mortality. **Methods and Results**: This was a territory-wide retrospective cohort study of patients diagnosed with ARVC/D between 1997 and 2019 in Hong Kong. This study consisted of 109 ARVC/D patients (median age: 61 [46–71] years; 58% male). Of these, 51 and 24 patients developed incident VT/VF and new-onset HFrEF, respectively. Five patients underwent cardiac transplantation, and 14 died during follow-up. Multivariate Cox regression identified prolonged QRS duration as a predictor of VT/VF ($p < 0.05$). Female gender, prolonged QTc duration, the presence of epsilon waves and T-wave inversion (TWI) in any lead except aVR/V1 predicted new-onset HFrEF ($p < 0.05$). The presence of epsilon waves, in addition to the parameters of prolonged QRS duration and worsening ejection fraction predicted all-cause mortality ($p < 0.05$). Clinical scores were developed to predict incident VT/VF, new-onset HFrEF and all-cause mortality, and all were significantly improved by machine learning techniques. **Conclusions**: Clinical and electrocardiographic parameters are important for assessing prognosis in ARVC/D patients and should in turn be used in tandem to aid risk stratification in the hospital setting.

Keywords: arrhythmogenic right ventricular cardiomyopathy/dysplasia; heart failure; ventricular arrhythmias; mortality

1. Introduction

Arrhythmogenic right ventricular cardiomyopathy/dysplasia (ARVC/D) is a rare hereditary condition presenting at an incidence of 1 in 2500 to 1 in 5000 in the general population, with notable geographical variations in disease prevalence [1]. ARVC/D is characterized by genetic mutations in desmosomal genes [2,3] and accompanying aberrations in cardiomyocyte cell-cell adhesion, leading to early cardiac regional anatomical abnormalities, typically confined to the right ventricular inflow tract, outflow tract, and apex, which together constitute the "triangle of dysplasia" [4]. Disease progression is in turn dominated by diffuse thinning of the right ventricular wall with cardiomyocyte loss and corresponding fibrofatty replacement of the myocardium [5]. These pathological alterations not only disturb the native electrical conduction system, thereby predisposing affected patients to malignant ventricular arrhythmias and sudden cardiac death (SCD) [6,7], but also potentially induce left ventricular dysfunction and subsequent dual-chamber cardiac failure [8].

The definitive diagnosis of ARVC/D is challenging owing to the absence of a single set of parameters sufficiently specific to the disease [1]. As such, the current diagnostic criterion seeks to amalgamate a series of clinical, pathological and genetic features most commonly observed in affected patients, amongst which electrocardiographic and echocardiographic parameters are the most prominent [9]. The evident heterogeneity in the phenotypic presentation and complications associated with ARVC/D poses difficulties to optimal management [10]. Current therapies are primarily geared towards the prevention of lethal ventricular arrhythmias, and implantable cardioverter-defibrillator (ICD) placement has hitherto proven to be the only effective strategy in reducing long-term mortality [1]. Such dilemmas in the management of these patients are further compounded by the apparent underreporting of prognostic markers to assist risk stratification in the clinical setting.

The present study aims to investigate the risk of detrimental cardiovascular events in an Asian population of ARVC/D patients, including the incidence of malignant ventricular arrhythmias, new-onset heart failure with reduced ejection fraction (HFrEF), as well as long-term mortality. Moreover, the prognostic importance of several clinical parameters will be examined in an attempt to identify possible markers with predictive value that could improve overall assessment and therapeutic guidance. These markers will be utilized to not only construct scoring systems, but also, for the first time, employ machine learning algorithms designed to enhance outcome prediction.

2. Methods

2.1 Diagnosis of ARVC/D

In 1994, an International Task Force (ITF) proposed an initial criterion for ARVC/D recognition, based on six major categories: (i) global and/or regional dysfunction and structural alterations of the right ventricle, (ii) tissue characterization of the right ventricular wall, (iii) repolarization abnormalities, (iv) depolarization abnormalities, (v) cardiac arrhythmias, and (vi) family history. Each category comprised of one or more major and/or minor requirements, from which several different permutations of major and minor variable combinations were considered diagnostic of ARVC/D. With time, the discovery of new associated histological, electrocardiographic and echocardiographic parameters with greater sensitivity for the detection of early stage disease led to the proposition of the revised ITF criteria in 2010. The modified criteria elaborated upon the initial guidelines in greater detail, the specifics of which can be found elsewhere [11].

2.2 Study Population and Their Baseline Characteristics

This study was approved by The Joint Chinese University of Hong Kong – New Territories East Cluster Clinical Research Ethics Committee. The current study included ARVC/D patients who presented to public hospitals managed by the Hospital Authority of Hong Kong between January 1999 to December 2019. Patient data was obtained using the Clinical Management System (CMS), an electronic health database that is connected to the territory-wide Clinical Data Analysis and Reporting System (CDARS). Both systems are integrative centralized platforms that permit the extraction of clinical data for analysis and reporting. The collaborative use of CMS and CDARS systems allowed for the retrieval of comprehensive medical records, including disease diagnoses, clinical comorbidities, electrocardiographic indices, echocardiographic parameters and operative procedures. Our teams have used these systems for studying other ion channelopathies in the territory [12,13]. In the present study, ARVC/D subjects were recruited by International Statistical Classification of Diseases (ICD) coding and with subsequent diagnostic confirmation by cardiologist review. Collected patient data included: (1) age, (2) gender, (3) age at ARVC/D diagnosis, (4) family history of ARVC/D and VF/SCD, (5) presentation of palpitations, (6) presentation of syncope and the number of episodes, (7) presentation of premature ventricular contractions (PVCs) and PVC burden, (8) pre-existing ventricular tachycardia/ventricular fibrillation (VT/VF) prior to ARVC/D diagnosis and the number of episodes, (9) incident non-sustained ventricular tachycardia (NSVT) and the number of episodes, (10) performance of electrophysiological study (EPS) and presentation of EPS-induced VT/VF, (11) performance of 24-hour ECG Holter, (12) performance of the exercise stress test, (13) ICD implantation, and (14) operative heart transplantation.

Further data collection involved using these electronic databases to obtain echocardiographic reports closest to the date of ARVC/D diagnosis in order to determine left ventricular ejection fractions (LVEF) and confirm the presence of right ventricular morphological pathologies consistent with ARVC/D diagnosis, including right ventricular dyskinesia, dilatation, aneurysms, fibrofatty replacement and systolic dysfunction. Likewise, automated electrocardiogram (ECG) recordings taken closest to the date of ARVC/D diagnosis were also extracted for the following indices: (1) ventricular rate, (2) P-wave duration, (3) PR-interval, (4) QRS duration, (5) QT and QTc interval, (6) T-wave inversion, (7) R-wave amplitude in V5, (8) S-wave amplitude in V1, (9) manifestation of epsilon waves, (10) P-wave axis: representing the net vectorial direction of atrial depolarization, (11) QRS axis and T-wave axis: representing the net vectorial depolarization and repolarization, respectively. Moreover, the primary long-term outcome assessed was incident VT/VF post-ARVC/D diagnosis. Secondary outcomes derived included: (1) new-onset HFrEF defined as LVEF $\leq$40%, and (2) all-cause mortality.

2.3 Statistical Analysis

Descriptive statistics were presented as median [interquartile range] or as count (percentage) as appropriate.

Table 1. Baseline characteristics for patients with and without incident VT/VF.

Characteristics	Patients with incident VT/VF	Patients without incident VT/VF	*p* value
Total (%)	49 (45)	60 (55)	
Demographics			
Female sex (%)	16 (33)	30 (50)	0.068
Age (years)	65.0 (45.0–71.0)	59.0 (46.5–71.5)	0.901
Clinical features			
Pre-existing VT/VF (%)	21 (43)	16 (27)	0.076
Family history of ARVC/D (%)	2 (4)	6 (10)	0.250
Family history of VF/SCD (%)	6 (12)	5 (8)	0.447
Syncope (%)	25 (51)	17 (28)	0.012
Palpitations (%)	32 (65)	29 (48)	0.4683
Left ventricular ejection fraction (%)	54 (41.7–62.0)	60.7 (54.4–65.0)	0.1119
Right ventricular pathologies			
Dilatation	37 (76)	29 (48)	0.005
Dyskinesia	27 (55)	22 (37)	0.052
Aneurysm	8 (16)	4 (7)	0.124
Fibrofatty replacement	14 (29)	22 (37)	0.805
Systolic dysfunction	29 (59)	16 (27)	0.001
ECG parameters			
QRS duration (ms)	108.0 (96.0–129.0)	94.5 (86.0–103.5)	0.002
QTc duration (ms)	448.5 (412.0–477.0)	431.0 (414.0–470.0)	0.706
PRI interval (ms)	167.0 (141.0–195.0)	163.0 (150.0–196.0)	0.680
R wave amplitude in V5 (%)	0.7 (0.5–1.3)	0.9 (0.5–1.9)	0.274
S wave amplitude in V1 (%)	0.4 (0.2–0.7)	0.4 (0.2–1.0)	0.667
Epsilon waves (%)	13 (27)	5 (8)	0.023
Premature ventricular contractions (%)	30 (61)	33 (55)	0.432
T-wave inversion in any lead except aVR/V1 (%)	34 (69)	32 (53)	0.312
T-wave inversion in 2 of 3 inferior leads (%)	20 (41)	17 (28)	0.347

*Abbreviations: VT/VF, ventricular tachycardia/ventricular fibrillation; VF/SCD, ventricular fibrillation/sudden cardiac death.

The study population was stratified according to the presence or absence of incident VT/VF. The Mann-Whitney U test was used to compare continuous variables. Chi-squared test with Yates' correction was used for 2×2 contingency data, and Pearson's Chi-squared test was used for contingency data for variables with more than two categories. The relationship between electrocardiographic and clinical parameters with outcomes was assessed using univariate Cox proportional-hazards model. Variables with $p < 0.05$ were incorporated into a multivariate model, as well as a scoring system. Briefly pertaining to the scoring system, a point assigned to a variable was equivalent to the halved value of the hazard ratio, rounded up to the nearest integer. Statistical analysis was performed using Stata (Version 13.0, StataCorp, College Station, Texas, United States). A two-sided p-value < 0.05 was considered statistically significant.

2.4 Development of a Machine Learning Survival Learning Model

A non-parametric machine learning survival analysis model was developed to predict incident VT/VF, new-onset HFrEF and all-cause mortality in ARVC/D patients. The underlying motivation for the implementation of machine learning survival analysis models stemmed from the ability of these algorithms to better capture nonlinear and interactive patterns within survival data compared to traditionally used Cox regression models, which assume the existence of a hazard function between survival data and censored outcomes. A major problem pertaining to the use of Cox regression models is the assumption of a linear relationship between covariates and the time of event occurrence. Many modifications have been proposed aiming to circumvent this limitation, namely by generalizing the Cox regression model to take into account the corresponding nonlinear and interactive relations between covariates and the time of event. Survival trees [14] and random survival forest (RSF) [15] models were developed on the premonition that tree-based models, after being combined with baseline models (e.g., decision trees), can generate the best survival predictions. Recently, a weighted random survival forest (wRSF) [16] model was proposed as an efficient modification of RSF models by replacing the standard procedure of averaging used for the estimation of RSF hazard function with a weighted averaging strategy, wherein the weights are assigned to every tree and can be viewed as training parameters computed by maximizing Harrell's concordance index (C-index).

The present study introduced a wRSF model for the prediction of incident VT/VF, HFrEF and all-cause mortality after first presentation of ARVC/D. The most important

variables for outcome prediction were derived with a variable importance ranking approach of the wRSF model. The ranked results were subsequently used to construct a machine learning-based electronic frailty index with prognostic value in assessing the incidence of the three aforementioned outcomes. The survival prediction performance of wRSF, RSF and multivariate Cox models in discriminating incident VT/VF, HFrEF and all-cause mortality were compared using several evaluation measures, including precision, recall, area under the receiver operating characteristics curve (AUC), and Harrell's C-index. The comparative experiments were conducted based on the input of significant univariate predictors identified by the initial univariate Cox proportional-hazards model. R packages, including *randomForestSRC* (Version 2.1.5, https://cran.r-project.org/web/packages/randomForestSRC/index.html), *randomForestSRC* (Version 2.9.3, https://cran.r-project.org/web/packages/randomForestSRC/index.html), *survival* (Version 2.42-3, https://cran.r-project.org/web/packages/survival/index.html) and *ggplot2* (Version 3.3.2, https://cran.r-project.org/web/packages/ggplot2/index.html), were used to generate the survival prediction results.

3. Results

In this ARVC/D cohort (n = 109), the median age was 61 [46–71] years and 63 (58%) were male. The baseline characteristics are presented in Table 1, with patients stratified according to the development of incident VT/VF. The median ages of the VT/VF (n = 49) and non-VT/VF (n = 60) groups were 65 [45–71] years and 59 [46.5–71.5] years, respectively with similar ages at diagnosis of ARVC/D. Patients who developed incident VT/VF tended to present more often with right ventricular dilatation and systolic dysfunction. This group also demonstrated a significantly longer QRS duration, which took a median value of 108.0 [96.0–129.0] ms, as well as a significantly greater proportion of subjects developing epsilon waves (n = 13; 27%).

Anti-arrhythmic therapy was prescribed in the form of amiodarone (n = 48), sotalol (n = 34) and mexiletine (n = 4), with majority of patients taking these medications after the first episode of VT/VF (amiodarone: n = 28; sotalol: n = 21; mexiletine: n = 4). Some subjects who had received anti-arrhythmic therapy experienced subsequent VT/VF episodes (amiodarone: n = 40; sotalol: n = 27; mexiletine: n = 3), HFrEF (amiodarone: n = 17; sotalol: n = 10; mexiletine: n = 1) or cardiovascular-related mortality (amiodarone: n = 7; sotalol: n = 5; mexiletine: n = 0) post-therapy. A total of 58 patients received implantable cardioverter-defibrillator (ICD) placement either for VT/VF prophylaxis following ARVC/D diagnosis (n = 32) or to prevent VT/VF recurrence after the first episode (n = 26). After ICD implantation, a total of 30 patients experienced at least one episode of VT/VF and six suffered cardiovascular-related mortality.

3.1 Predictors of Adverse Outcomes

In the following ARVC/D cohort, 49 patients and 24 patients developed incident VT/VF and new-onset HFrEF, respectively. A total of 5 patients underwent cardiac transplantation, and 14 patients passed away during follow-up, 10 of which suffered from ARVC/D-related complications, namely SCD and HFrEF, whereas the remaining 4 suffered non-cardiovascular-related deaths. The results of univariate Cox proportional-hazards regression analysis for predicting incident VT/VF, new-onset HFrEF and all-cause mortality are reported in **Supplementary Tables 1–3**, respectively, with the corresponding Kaplan-Meier survival curves shown in Fig. 1.

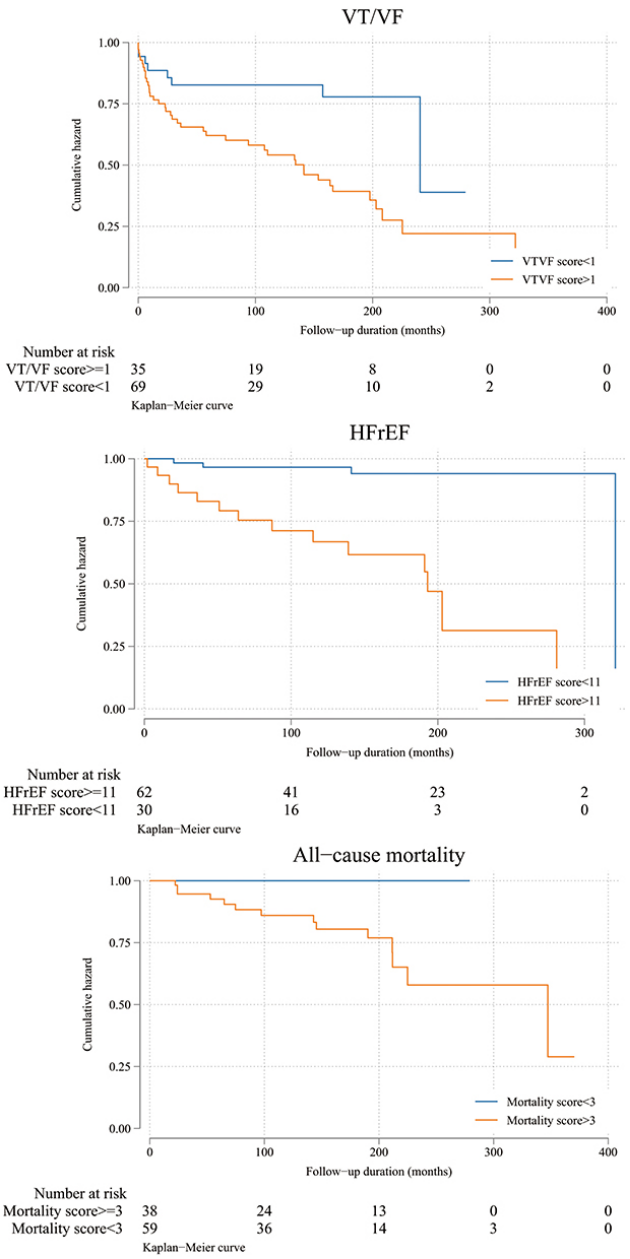

Fig. 1. Kaplan-Meier survival curves for incident VT/VF, new-onset HFrEF and all-cause mortality.

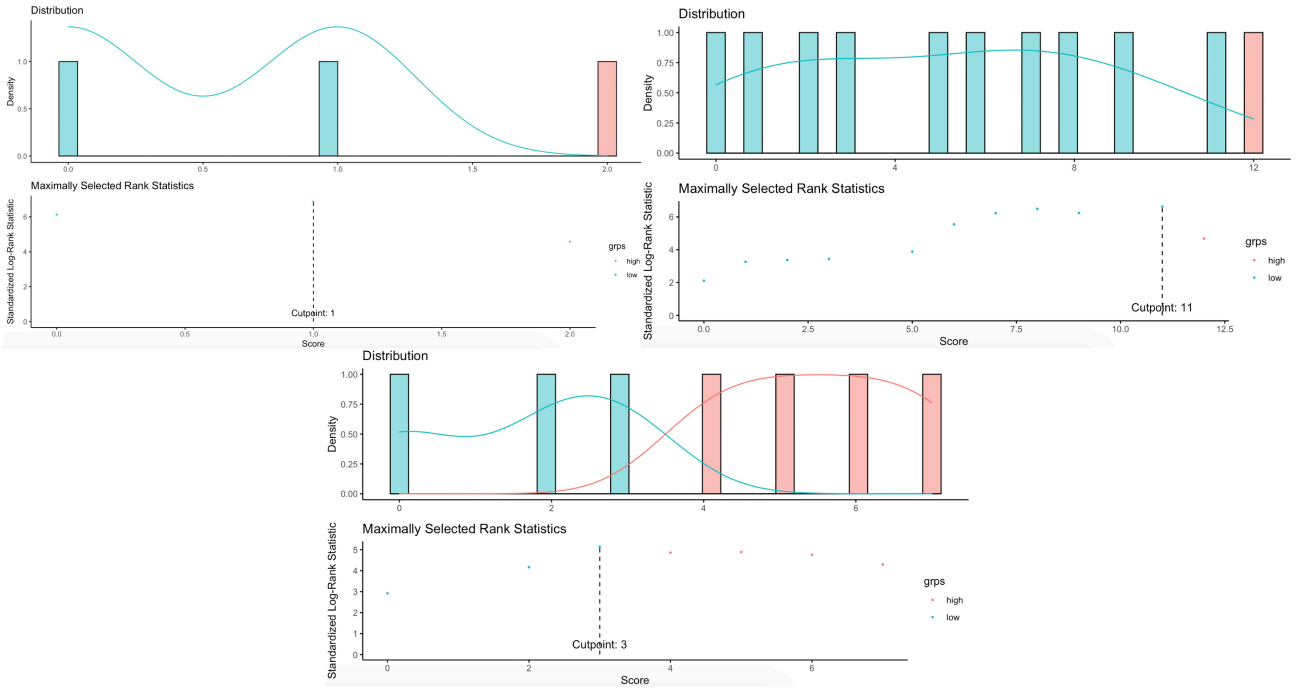

Fig. 2. Optimal cut-off for the scoring system for incident VT/VF, new-onset HFrEF and all-cause mortality with the maximal rank statistics approach.

Incident VT/VF (**Supplementary Table 1**) was associated with prolonged QRS duration, presence of epsilon waves, and syncope, with only the foremost retaining significance after multivariate adjustment ($p < 0.05$). Regarding secondary outcomes, univariate predictors of new-onset HFrEF (**Supplementary Table 2**) included female gender, prolonged QTc duration, presence of epsilon waves and TWI in any lead except aVR/V1, all of which retained significance in the multivariate model ($p < 0.05$). Likewise, all-cause mortality (**Supplementary Table 3**) was similarly associated with all univariate predictors of new-onset HFrEF, in addition to prolonged QRS duration ($p < 0.05$). Resultant significant parameters in multivariate Cox regression included female gender, prolonged QRS duration, prolonged QTc duration, and presence of epsilon waves ($p < 0.05$).

3.2 Scoring System for New-Onset VT/VF in ARVC/D

Significant clinical and electrocardiographic parameters in univariate Cox regression ($p < 0.05$) were used to design a scoring system to predict new-onset VT/VF in ARVC/D. Receiver operator characteristics (ROC) curves were used to determine optimal cut-offs for significant continuous variables. The optimal cutoff value for QRS duration was 98.5 ms (AUC: 0.69; sensitivity = 72%; specificity = 67%). After categorization, this parameter retained significance in univariate prediction of incident VT/VF in ARVC/D, and was therefore eligible for inclusion. QRS duration $>$ 98.5 ms, along with presence of syncope and epsilon waves were subsequently used to form the final scoring system (**Supplementary Table 4a**). Subjects who developed VT/VF presented with a median score that was 0.5 points higher than those who remained free of VT/VF. Cox proportional-hazards analysis revealed that patients with a per unit increase in the score had a 74% higher risk of incident VT/VF (HR: 1.74; 95% CI: 1.30–2.33; $p < 0.001$) (**Supplementary Table 4b**). Categorization of the VT/VF score using the maximal rank statistics approach (Fig. 2) revealed an optimal cut-off of 1 point. Subsequent Cox proportional-hazards analysis demonstrated that patients with a score $\geq$1 point had an approximate 2-fold increase in risk of new-onset VT/VF (**Supplementary Table 4c**).

3.3 Scoring System for New-Onset HFrEF in ARVC/D

A similar process was conducted for the creation of a score for HFrEF. The optimal cutoff value for QTc duration was 437.5 ms (AUC: 0.76; sensitivity = 88%; specificity = 65%), which remained significant in univariate prediction of VT/VF in ARVC/D. Overall, four binary parameters were ultimately included in the final scoring system, including female gender, presence of epsilon waves, presence of TWI in any lead except aVR/V1, and QTc $>$437.5 ms (**Supplementary Table 5a**). Patients with HFrEF presented with a median score that was 6 points higher compared to those without HFrEF (**Supplementary Table 5b**). Cox proportional-hazards analysis revealed that patients with a higher score, when used as a continuous variable, had a 48% higher risk of new-onset HFrEF (HR: 1.48; 95% CI: 1.27–1.74; $p < 0.001$) (**Supplementary Table 5c**). Like-

wise, categorization of the HFrEF score using the maximal rank statistics approach (Fig. 2) revealed an optimal cut-off of 11 points, using which it was shown that those with a score $\geq$11 points had a more than 15-fold increase in risk of new-onset HFrEF (**Supplementary Table 5c**).

3.4 Scoring System for All-Cause Mortality in ARVC/D

As it pertains to all-cause mortality, the optimal cut-off values for (i) QRS duration was 122.5 ms (AUC = 0.71; sensitivity = 57%; specificity = 85%) and (ii) QTc duration was 448.5 ms (AUC = 0.73; sensitivity = 79%; specificity = 63%), both of which retained significance in univariate prediction of all-cause mortality in ARVC/D. Per unit decrease in LVEF was also a predictor of all-cause mortality, but failed to retain significance following categorization and was therefore not included in the scoring system. All in all, a total of four binary parameters were subsequently included in the final scoring system, including gender, presence of epsilon waves, QRS $>$122.5 ms, and QTc $>$448.5 ms (**Supplementary Table 6a**). Patients who suffered death presented with a median score that was 4.5 points higher compared to those who survived (**Supplementary Table 6b**). Cox proportional-hazards analysis revealed that patients with a per unit increase in the score had a 65% higher risk of all-cause mortality (HR: 1.65; 95% CI: 1.33–2.06; $p < 0.001$) (**Supplementary Table 6c**). With an optimal cut-off of 3 points determined by the maximal rank statistics approach (Fig. 2), those with a score $\geq$3 points had an almost 23-fold increase in risk of mortality (**Supplementary Table 6c**).

3.5 Machine Learning to Predict Incident VT/VF, New-Onset HFrEF and All-Cause Mortality

wRSF models were used to predict primary and secondary outcomes with the identified significant parameters in univariate analysis as input variables (p value < 0.05). The optimal tree number used to build each wRSF model was selected by the error rate minimization with iteration approach. The tree number selected to predict VT/VF, new-onset HFrEF and all-cause mortality was 250, 200, and 400 respectively (Fig. 3). The derived importance value ranking of the variables is shown in Table 2. QRS duration $>$98.5 ms was the most predictive variable for incident VT/VF, followed by presence of syncope and presence of epsilon waves. In contrast, prolonged QTc duration demonstrated the strongest predictivity for HFrEF, followed by TWI in any lead except aVR/V1, presence of epsilon waves, and female gender. As it pertains to all-cause mortality, QRS $>$122.5 ms was the most important predictor, followed by presence of epsilon waves, QTc $>$448.5 ms, and female gender. As shown in Table 3, the ability of the wRSF model to predict the primary outcomes was compared with an RSF model and multivariate Cox model, based on evaluation metrics of precision, recall, AUC and C-index. Findings indicate that the wRSF models performed best in the prediction of all three outcomes based on the significant univariate predictors.

Table 2. Variable importance ranking generated by wRSF models for primary and secondary outcomes.

Incident VT/VF	
Variable	Importance Value
QRS $>$98.5	0.0778
Syncope	0.0054
Epsilon wave	0.0006
New-onset HFrEF	
Variable	Importance Value
QTc $>$437.5 ms	0.1180
TWI in any lead except aVR/V1	0.0347
Epsilon wave	0.0147
Female Gender	0.0066
All-cause mortality	
Variable	Importance Value
QRS $>$122.5	0.0700
Epsilon wave	0.0377
QTc $>$448.5	0.0291
Female Gender	0.0251

Table 3. Performance comparison between wRSF, RSF, and multivariate Cox model. Italicized values indicate the best model for that particular performance metric.

Incident VT/VF				
Model	Precision	Recall	AUC	C-index
wRSF model	*0.8352*	*0.8478*	*0.8341*	*0.8202*
RSF model	0.8283	0.8372	0.8161	0.8116
Multivariate Cox model	0.7493	0.7793	0.7524	0.7835
New-onset HFrEF				
Model	Precision	Recall	AUC	C-index
wRSF model	*0.8290*	*0.8305*	*0.8363*	*0.8182*
RSF model	0.8053	0.8172	0.8184	0.8021
Multivariate Cox model	0.7493	0.7641	0.7620	0.7770
All-cause mortality				
Model	Precision	Recall	AUC	C-index
wRSF model	*0.7322*	*0.7395*	*0.7549*	*0.7481*
RSF model	0.7071	0.6922	0.6984	0.7012
Multivariate Cox model	0.6734	0.6808	0.6853	0.6744

4. Discussion

The present study is novel as it is among the first territory-wide investigations of ARVC/D patients in Hong Kong, allowing for the development of the first risk scores that include both electrocardiographic and clinical parameters for predicting incident VT/VF, new-onset heart failure as well as mortality in ARVC/D. In addition, to our knowledge, this is also the first investigation to apply machine learning algorithms to assess ARVC/D prognosis, in turn

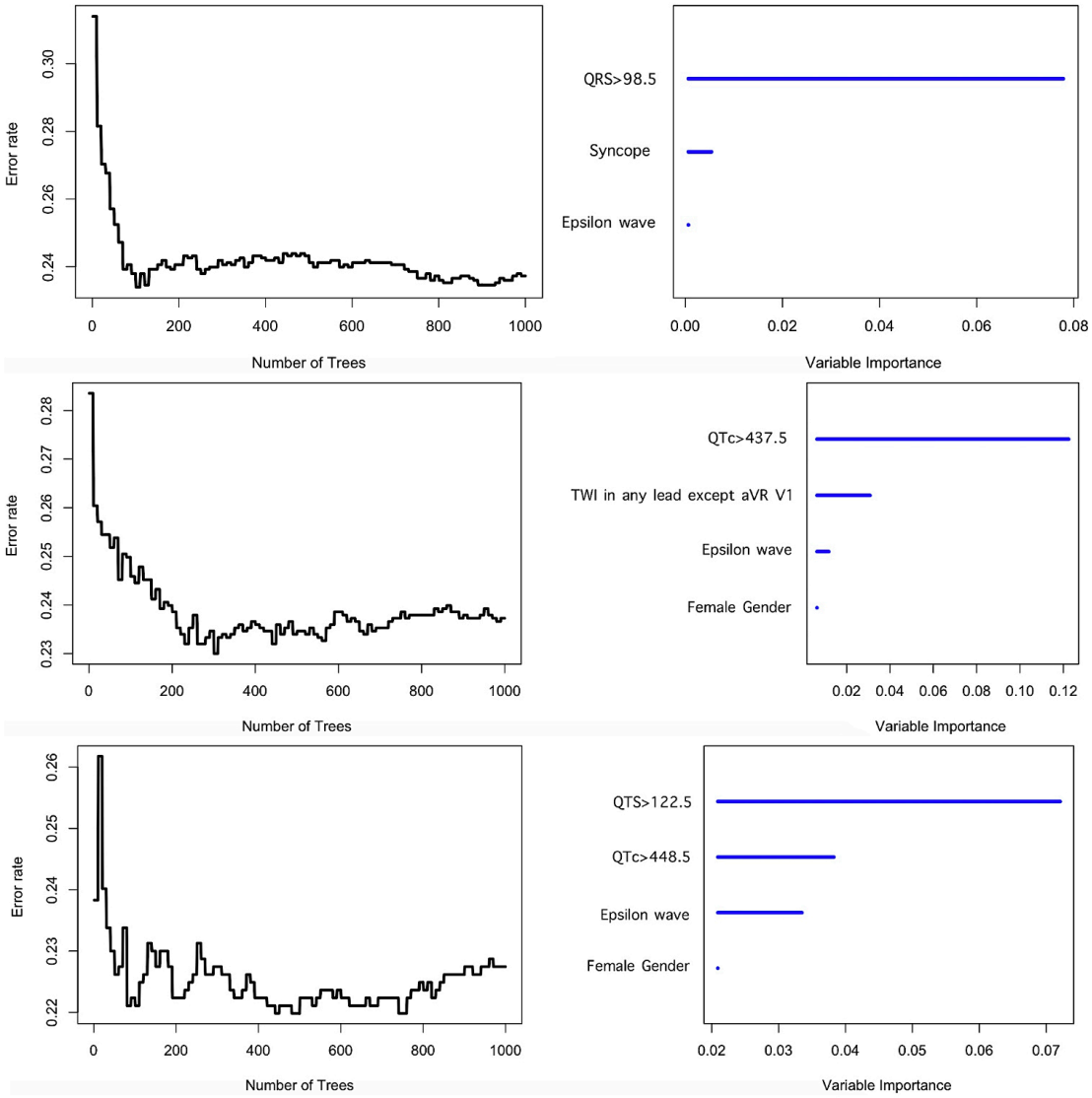

Fig. 3. Optimal tree number of wRSF model and variable importance ranking to predict incident VT/VF, new-onset HFrEF and all-cause mortality.

demonstrating enhanced risk prediction for outcomes with such algorithms when compared to other analytical models.

The substrate for arrhythmogenesis in ARVC/D is a combination of conduction and repolarization abnormalities associated with structural alterations in the right ventricle [17]. However, recent work has demonstrated that non-classical forms of arrhythmogenic cardiomyopathy, namely left dominant or biventricular forms, are more prone to ventricular arrhythmias than classical ARVC/D [18]. Moreover, the atria, in addition to the ventricles, are also abnormal in ARVC/D [19] with complications such as sinoatrial arrests and atrial fibrillation [20,21]. Different ECG indices have been identified as risk markers of ventricular arrhythmias [22], amongst which the epsilon wave is the classical pathognomonic feature of ARVC/D [23,24]. Repolarization criteria, including TWI in inferior leads, a precordial QRS amplitude ratio of $\leq$0.48, and QRS fragmentation also constitute valuable variables for predicting adverse outcomes in this disease [25]. Repolarization abnormalities, such as TWI, are important, and electroanatomic mapping areas have shown to be proportional to extent of TWI on 12-lead ECG [26,27]. Strain imaging by speckle-tracking echocardiography has been used to risk-stratify patients in heart failure [28] and recent work has shown that incorporation of mechanical dispersion can further improve risk prediction [29], as heart failure is typically under-recognized in ARVC/D [30].

Ventricular arrhythmias are a common occurrence amongst patients with so-called inherited cardiac arrhythmias, including ARVC/D, Brugada syndrome, long QT-syndrome and short QT-syndrome, amongst which ARVC/D cohorts have shown to present with the highest rates of frequent ventricular premature complexes, (non)sustained ventricular tachycardias, and malignant

ventricular tachyarrhythmias [31]. Several large-scale studies have reported on the clinical characteristics and predictors of ventricular arrhythmias in ARVC/D patients. In 131 definite ARVC/D patients, spontaneous sustained ventricular arrhythmias, cardiac syncope, male gender, proband, and inducibility in electrophysiology study were all significant predictors of incident sustained VT/VF and SCD [32]. The same group further studied the phenotype of ARVC/D patients with late presentation, demonstrating that this subpopulation does not confer a benign prognosis and has a high arrhythmic risk [33]. Another study of 135 patients identified prolonged QRS duration on signal-averaged ECG, non-sustained VT on 24 h-ECG, and the absence of negative T waves in lead aVR on a 12-lead surface ECG as significant predictors of recurrent sustained ventricular arrhythmias and hospitalization due to ventricular arrhythmias in ARVC/D [34]. Moreover, an investigation of 137 patients from France found that low LVEF, positive electrophysiological studies and physical activity >6 h/week were shown to be independently associated with the development of ventricular arrhythmias [35]. The findings of the aforementioned studies clearly demonstrate the high incidence of ventricular arrhythmias within ARVC/D cohorts, thereby necessitating the use of prophylactic antiarrhythmics to reduce their occurrence [36], along with radiofrequency catheter ablation for patients who end up developing these rhythm abnormalities [37].

Recently, a systematic review and meta-analysis summarized the current literature, identifying consistently predictive risk factors in patients with definite ARVC across different studies [38]. These were male sex, syncope, TWI in lead V3, right ventricular dysfunction, and prior (non)sustained VT/VF. Our present work extends these findings by demonstrating that longer QRS duration, the presence of epsilon waves and TWI in 2/3 inferior leads were significantly associated with incident VT/VF, albeit only QRS remained a significant predictor after multivariate adjustment. Several parameters retained significance in multivariate prediction of new-onset HFrEF, including longer QTc duration, presence of epsilon waves, TWI in any lead except aVR/V1 and female gender. Likewise, longer QRS duration, presence of epsilon waves, LVEF and age at diagnosis of ARVC/D were all significantly associated with all-cause mortality in multivariate analysis. It was then possible to further enhance risk prediction through the application of wRSF model analysis, which we have recently used to better risk prediction in acquired long QT syndrome [39] as well as Brugada syndrome [40]. The wRSF model was able to improve the risk stratification for incident VT/VF, new-onset HFrEF and all-cause mortality in this ARVC/D cohort.

Furthermore, the clinical heterogeneity typically observed amongst populations with ARVC/D makes the use of scoring algorithms a potentially useful method to amalgamate the different patient parameters for the purposes of risk stratification. Such an approach has been adopted previously in a large cohort of 528 ARVC/D patients to predict the long-term risk of ventricular arrhythmias. The model constructed, which included age, male gender, cardiac syncope in the prior 6 months, prior non-sustained VT, number of PVCs in 24 h, number of leads with TWI and right ventricular ejection fraction, demonstrated an improved ability to estimate risk of ventricular arrhythmias and guide decision-making in ICD implantation for such patients [41]. A meta-analysis identified the following 11 variables as the most important factor for predicting arrhythmic events: (1) male gender, (2) presyncope, (3) left ventricular dysfunction, (4) TWI in inferior leads, (5) proband status, (6) late potentials, (7) syncope, (8) inducibility at electrophysiological study, (9) right ventricular dysfunction, (10) epsilon waves, and (11) premature ventricular contractions greater than 1000/24 h [42]. To our knowledge, such scoring algorithms have not been used to investigate outcomes beyond VT/VF in ARVC/D cohorts. As such, the present study also developed two multi-parametric scores for predicting new-onset HFrEF and all-cause mortality, respectively, both of which demonstrated efficacy in assessing ARVC/D patient prognosis.

5. Limitations

This investigation has limitations that should be noted. Firstly, data is primarily based on patients of Chinese ethnicity and therefore lacks the subject variability needed for a comprehensive evaluation of ARVC/D, which itself presents with a heterogeneous phenotype. Secondly, several subjects were prescribed amiodarone and/or sotalol as treatment, both of which have been previously shown to influence certain ECG parameters, for instance by lengthening QTc interval. This could have potentially influenced the reported relationship between QTc duration and new-onset HFrEF as well as all-cause mortality. Finally, the adverse ECG findings, such as prolongation of QRS duration or QTc duration and the presence of epsilon waves, are likely linked to a greater underlying disease severity that in turn leads to malignant ventricular arrhythmias. As such, these parameters possibly only serve as markers, as opposed to outright predictors, of VT/VF, albeit further study is required to confirm this.

6. Conclusions

The phenotypic variability and adverse prognosis potentially associated with ARVC/D necessitates a multimodality approach for risk stratification that includes both clinical and electrocardiographic parameters. The findings of the current investigation are the first to be demonstrated in an Asian population, thereby extending the generalizability of pre-existing data to that of Asian cohorts. Moreover, the present study is among the first to not only demonstrate the use of scoring systems comprising of both electrocardiographic and clinical parameters in the assessment of

long-term outcomes in ARVC/D, but also to employ combinatorial methods involving machine learning algorithms to evaluate prognosis. Such algorithms are able to account for underlying inter-variable interactions, thereby improving overall event and survival prediction. As a result, with further study into their use, machine learning techniques could possibly provide an alternative, more effective avenue to assess patient prognosis in such heterogeneous disease conditions.

Author Contributions

Conceptualization—IL and GT; methodology—IL, JZ and QZ; software—IL, JZ and QZ; validation—IL and JZ; formal analysis—IL and JZ; investigation—IL, JZ, SL, KLi, KLe, JH, YL, GL, TL, WW, IW, NM, CM, QZ, GT; resources—GT, TL, QZ; data curation—IL and GT; writing—original draft preparation—IL, JZ, SL, KLi, KLe, JH, YL, GL, TL, WW, IW, NM, CM, QZ, GT; writing—review and editing—IL, JZ, SL, KLi, KLe, JH, YL, GL, TL, WW, IW, NM, CM, QZ, GT; visualization—IL and JZ; supervision—TL, QZ and GT; project administration—QZ and GT; funding acquisition—not applicable. All authors have read and agreed to the published version of the manuscript.

Ethics Approval and Consent to Participate

This study was approved by The Joint Chinese University of Hong Kong – New Territories East Cluster Clinical Research Ethics Committee (Approval number: 2019.422). The need for patient consent was waived by the Committee owing to the retrospective and observational nature of the study.

Acknowledgment

Not applicable.

Funding

This research received no external funding.

Conflict of Interest

The authors declare no conflict of interest. Gary Tse, Sharen Lee and Tong Liu are serving as one of the Guest Editors of this journal. We declare that Gary Tse, Sharen Lee and Tong Liu had no involvement in the peer review of this article and has no access to information regarding its peer review. Full responsibility for the editorial process for this article was delegated to Stefan Peters.

Supplementary Material

Supplementary material associated with this article can be found, in the online version, at https://doi.org/10.31083/j.rcm2307231.

References

[1] Idris A, Shah SR, Park K. Right ventricular dysplasia: management and treatment in light of current evidence. Journal of Community Hospital Internal Medicine Perspectives. 2018; 8: 101–106.

[2] Bazoukis G, Letsas KP, Xia Y, Tse G, Li KHC. A novel desmin mutation causing severe left ventricular arrhythmogenic cardiomyopathy/dysplasia. Journal of Thoracic Disease. 2018; 10: S3100–S3102.

[3] van Lint FHM, Murray B, Tichnell C, Zwart R, Amat N, Lekanne Deprez RH, *et al*. Arrhythmogenic Right Ventricular Cardiomyopathy-Associated Desmosomal Variants are Rarely De Novo. Circulation: Genomic and Precision Medicine. 2019; 12

[4] Marcus FI, McKenna WJ, Sherrill D, Basso C, Bauce B, Bluemke DA, *et al*. Diagnosis of arrhythmogenic right ventricular cardiomyopathy/dysplasia: proposed modification of the Task Force Criteria. European Heart Journal. 2010; 31: 806–814.

[5] Shaheen F, Iqbal K, Hafeez I, Choh NA, Tramboo NA, Lone A, *et al*. Clinico-radiological profile of arrhythmogenic right ventricular dysplasia at a tertiary care center: Two year experience. Journal of the Saudi Heart Association. 2013; 25: 79–84.

[6] Ackerman MJ, Priori SG, Willems S, Berul C, Brugada R, Calkins H, *et al*. HRS/EHRA Expert Consensus Statement on the State of Genetic Testing for the Channelopathies and Cardiomyopathies: this document was developed as a partnership between the Heart Rhythm Society (HRS) and the European Heart Rhythm Association (EHRA). Europace. 2011; 13: 1077–1109.

[7] Tabib A, Loire R, Chalabreysse L, Meyronnet D, Miras A, Malicier D, *et al*. Circumstances of Death and Gross and Microscopic Observations in a Series of 200 Cases of Sudden Death Associated with Arrhythmogenic Right Ventricular Cardiomyopathy and/or Dysplasia. Circulation. 2003; 108: 3000–3005.

[8] Sevinç Şengül F, Tunca Şahin G, Özgür S, Kafalı HC, Akıncı O, Güzeltaş A, *et al*. Clinical features and arrhythmic complications of patients with pediatric-onset arrhythmogenic right ventricular dysplasia. The Anatolian Journal of Cardiology. 2019; 22: 60–67.

[9] Corrado D, van Tintelen PJ, McKenna WJ, Hauer RNW, Anastastakis A, Asimaki A, *et al*. Arrhythmogenic right ventricular cardiomyopathy: evaluation of the current diagnostic criteria and differential diagnosis. European Heart Journal. 2020; 41: 1414-1429.

[10] Nasir K, Bomma C, Tandri H, Roguin A, Dalal D, Prakasa K, *et al*. Electrocardiographic Features of Arrhythmogenic Right Ventricular Dysplasia/Cardiomyopathy According to Disease Severity: a need to broaden diagnostic criteria. Circulation. 2004; 110: 1527–1534.

[11] Marcus FI, McKenna WJ, Sherrill D, Basso C, Bauce B, Bluemke DA, *et al*. Diagnosis of arrhythmogenic right ventricular cardiomyopathy/dysplasia: proposed modification of the task force criteria. Circulation. 2010; 121: 1533–1541.

[12] Tse G, Zhou J, Lee S, Liu T, Bazoukis G, Mililis P, *et al*. Incorporating Latent Variables Using Nonnegative Matrix Factorization Improves Risk Stratification in Brugada Syndrome. Journal of the American Heart Association. 2020; 9: e012714.

[13] Lee S, Zhou J, Liu T, Letsas KP, Hothi SS, Vassiliou VS, *et al*. Temporal Variability in Electrocardiographic Indices in Subjects With Brugada Patterns. Frontiers in Physiology. 2020; 11: 953.

[14] Gordon L, Olshen RA. Tree-structured survival analysis. Cancer Treatment Reports. 1985; 69: 1065–1069.

[15] Ishwaran H, Kogalur UB, Blackstone EH, Lauer MS. Random survival forests. The Annals of Applied Statistics. 2008; 2: 841–860.

[16] Utkin LV, Konstantinov AV, Chukanov VS, Kots MV,

Ryabinin MA, Meldo AA. A weighted random survival forest. Knowledge-Based Systems. 2019; 177: 136–144.

[17] Li KHC, Bazoukis G, Liu T, Li G, Wu WKK, Wong SH, *et al*. Arrhythmogenic right ventricular cardiomyopathy/dysplasia (ARVC/D) in clinical practice. Journal of Arrhythmia. 2018; 34: 11–22.

[18] Casella M, Gasperetti A, Gaetano F, Busana M, Sommariva E, Catto V, *et al*. Long-term follow-up analysis of a highly characterized arrhythmogenic cardiomyopathy cohort with classical and non-classical phenotypes–a real-world assessment of a novel prediction model: does the subtype really matter. EP Europace. 2020; 22: 797–805.

[19] Li G, Fontaine GH, Fan S, Yan Y, Bode PK, Duru F, *et al*. Right atrial pathology in arrhythmogenic right ventricular dysplasia. Cardiology Journal. 2019; 26: 736–743.

[20] Saguner AM, Ganahl S, Kraus A, Baldinger SH, Medeiros-Domingo A, Saguner AR, *et al*. Clinical Role of Atrial Arrhythmias in Patients with Arrhythmogenic Right Ventricular Dysplasia. Circulation Journal. 2014; 78: 2854–2861.

[21] Burghouwt DE, Kammeraad JA, Knops P, du Plessis FA, de Groot NM. Bradyarrhythmias: first Presentation of Arrhythmogenic Right Ventricular Cardiomyopathy? Journal of Clinical Medicine Research. 2015; 7: 278–281.

[22] Tse G, Yan BP. Traditional and novel electrocardiographic conduction and repolarization markers of sudden cardiac death. EP Europace. 2017; 19: 712–721.

[23] Li G, Saguner AM, Fontaine GH, Frank R. Epsilon waves: Milestones in the discovery and progress. Annals of Noninvasive Electrocardiology. 2018; 23: e12571.

[24] Li G, Saguner AM. Updated definition of epsilon waves. Pacing and Clinical Electrophysiology. 2018; 41: 1271–1272.

[25] Saguner AM, Ganahl S, Baldinger SH, Kraus A, Medeiros-Domingo A, Nordbeck S, *et al*. Usefulness of Electrocardiographic Parameters for Risk Prediction in Arrhythmogenic Right Ventricular Dysplasia. The American Journal of Cardiology. 2014; 113: 1728–1734.

[26] Ustunkaya T, Desjardins B, Wedan R, Chahal CAA, Zimmerman SL, Saju N, *et al*. Epicardial Conduction Speed, Electrogram Abnormality, and Computed Tomography Attenuation Associations in Arrhythmogenic Right Ventricular Cardiomyopathy. JACC: Clinical Electrophysiology. 2019; 5: 1158–1167.

[27] Kubala M, Pathak RK, Xie S, Casado Arroyo R, Tschabrunn CM, Hayashi T, *et al*. Electrocardiographic Repolarization Abnormalities and Electroanatomic Substrate in Arrhythmogenic Right Ventricular Cardiomyopathy. Circulation: Arrhythmia and Electrophysiology. 2018; 11: e005553.

[28] Tse G, Zhou J, Woo SWD, Ko CH, Lai RWC, Liu T, *et al*. Multi-modality machine learning approach for risk stratification in heart failure with left ventricular ejection fraction $\leq$45% ESC Heart Failure. 2020; 7: 3716–3725.

[29] Lie Øyvind H, Rootwelt-Norberg C, Dejgaard Lars A, Leren IS, Stokke MK, Edvardsen T, *et al*. Prediction of Life-Threatening Ventricular Arrhythmia in Patients with Arrhythmogenic Cardiomyopathy. JACC: Cardiovascular Imaging. 2018; 11: 1377–1386.

[30] Gilotra NA, Bhonsale A, James CA, te Riele ASJ, Murray B, Tichnell C, *et al*. Heart Failure is Common and under-Recognized in Patients with Arrhythmogenic Right Ventricular Cardiomyopathy/Dysplasia. Circulation: Heart Failure. 2017; 10: e003819.

[31] Sitorus GDS, Ragab AAY, Houck CA, Lanters EAH, Heida A, van Gastel VE, *et al*. Ventricular Dysrhythmias during Long-Term Follow-up in Patients with Inherited Cardiac Arrhythmia. The American Journal of Cardiology. 2019; 124: 1436–1441.

[32] Wang W, Cadrin-Tourigny J, Bhonsale A, Tichnell C, Murray B, Monfredi O, *et al*. Arrhythmic outcome of arrhythmogenic right ventricular cardiomyopathy patients without implantable defibrillators. Journal of Cardiovascular Electrophysiology. 2018; 29: 1396–1402.

[33] Bhonsale A, te Riele ASJM, Sawant AC, Groeneweg JA, James CA, Murray B, *et al*. Cardiac phenotype and long-term prognosis of arrhythmogenic right ventricular cardiomyopathy/dysplasia patients with late presentation. Heart Rhythm. 2017; 14: 883–891.

[34] Vischer AS, Castelletti S, Syrris P, Bastiaenen R, Miles C, Akdis D, *et al*. Risk score for the exclusion of arrhythmic events in arrhythmogenic right ventricular cardiomyopathy at first presentation. International Journal of Cardiology. 2019; 290: 100–105.

[35] Maupain C, Badenco N, Pousset F, Waintraub X, Duthoit G, Chastre T, *et al*. Risk Stratification in Arrhythmogenic Right Ventricular Cardiomyopathy/Dysplasia without an Implantable Cardioverter-Defibrillator. JACC: Clinical Electrophysiology. 2018; 4: 757–768.

[36] Corrado D, Wichter T, Link MS, Hauer RNW, Marchlinski FE, Anastasakis A, *et al*. Treatment of Arrhythmogenic Right Ventricular Cardiomyopathy/Dysplasia: An International Task Force Consensus Statement. Circulation. 2015; 132: 441–453.

[37] Burg AEB, Groot NMS, Erven L, Bootsma M, Wall EE, Schalij MJ. Long-Term Follow-up after Radiofrequency Catheter Ablation of Ventricular Tachycardia: a Successful Approach? Journal of Cardiovascular Electrophysiology. 2002; 13: 417–423.

[38] Bosman LP, Sammani A, James CA, Cadrin-Tourigny J, Calkins H, van Tintelen JP, *et al*. Predicting arrhythmic risk in arrhythmogenic right ventricular cardiomyopathy: a systematic review and meta-analysis. Heart Rhythm. 2018; 15: 1097–1107.

[39] Chen C, Zhou J, Yu H, Zhang Q, Gao L, Yin X, *et al*. Identification of important risk factors for all-cause mortality of acquired long QT syndrome patients using random survival forests and non-negative matrix factorization. Heart Rhythm. 2021; 18: 426–433.

[40] Lee S, Zhou J, Li KHC, Leung KSK, Lakhani I, Liu T, *et al*. Territory-wide cohort study of Brugada syndrome in Hong Kong: predictors of long-term outcomes using random survival forests and non-negative matrix factorisation. Open Heart. 2021; 8: e001505.

[41] Cadrin-Tourigny J, Bosman LP, Nozza A, Wang W, Tadros R, Bhonsale A, *et al*. A new prediction model for ventricular arrhythmias in arrhythmogenic right ventricular cardiomyopathy. European Heart Journal. 2019; 40: 1850–1858.

[42] Bazoukis G, Letsas KP, Thomopoulos C, Tse G, Korantzopoulos P, Naka KK, *et al*. Predictors of Adverse Outcomes in Patients with Arrhythmogenic Right Ventricular Cardiomyopathy: A Meta-Analysis of Observational Studies. Cardiology in Review. 2019; 27: 189–197.