Machine Learning Models Enhance Prediction of Arrhythmogenic Right Ventricular Cardiomyopathy

Kwaku K Quansah[1–6]*, Sean A Murphy[1–6]*, Esther Kwon[1–6], Emma Anderson[1–6], Richard T Carrick[1,6], Cynthia A James[1,6], Hugh Calkins[1,6], Chulan Kwon[1–6]☒

1. Department of Cardiology, Johns Hopkins School of Medicine Baltimore, MD 21202
2. Department of Biomedical Engineering, Johns Hopkins University Baltimore, MD 21202
3. Cellular and Molecular Medicine Program, Johns Hopkins University Baltimore, MD 21202
4. Department of Cell Biology, Johns Hopkins University Baltimore, MD 21202
5. Institute for Cell Engineering, Johns Hopkins University Baltimore, MD 21202
6. Heart and Vascular Institute, Johns Hopkins Hospital, Baltimore, MD 21202

*equal contributions

☒corresponding author

Keywords

'Arrhythmogenic Right Ventricular Cardiomyopathy', 'Machine Learning', 'predictive modeling', 'Gradient Boosted Trees', 'matrix clinical data', 'clinical decision-making', 'diagnostic support'

Abstract

Arrhythmogenic Right Ventricular Cardiomyopathy (ARVC) is a leading contributor to sudden cardiac death worldwide, yet its diagnosis remains complex, expensive and time-consuming. Machine-learning (ML) classifiers offer a practical solution by delivering rapid, scalable predictions that can lessen dependence on expert interpretation and speed clinical decision-making. Here, we benchmarked eight ML algorithms for ARVC detection using area-under-the-curve (AUC) and accuracy as primary metrics. Gradient Boosted Trees outperformed all other models, achieving an accuracy of 94.34% after rigorous cross-validation. These results underscore the promise of Gradient Boosted Trees classifier as an effective decision-support tool within the ARVC diagnostic workflow, with potential to streamline evaluation and improve patient outcomes.

Introduction

Arrhythmogenic Right Ventricular Cardiomyopathy (ARVC) is a hereditary structural heart disease characterized by the progressive replacement of right ventricular myocardium with fibro-fatty tissue.[1,2] This pathological remodeling results in regional wall motion abnormalities and facilitates occurrence of ventricular arrhythmias, thereby elevating the risk of sudden cardiac death (SCD).[3] Early identification of ARVC is critical to prevent life-threatening ventricular arrhythmias and SCD, which often occur in young individuals and athletes.[4,5] The arrhythmogenic substrate in ARVC can be present well before overt structural changes appear on imaging, and exertion can precipitate malignant arrhythmias in susceptible individuals.[6–9] As a result, timely diagnosis allows for the implementation of life-saving interventions such as activity restriction, implantable cardioverter-defibrillator (ICD) placement, and family screening.[10–12] This emphasis on early recognition is reflected in the revised 2010 Task Force Criteria, which sought to standardize diagnostic thresholds by incorporating electrocardiographic, imaging, histological, and genetic parameters.[13,14] Furthermore, the position statement by the Heart Rhythm Society underscores

the importance of early detection and risk stratification to guide preventive therapy.[15] Longitudinal data from ARVC registries consistently show that delayed diagnosis is associated with increased arrhythmic burden and worse prognosis, highlighting the need for improved screening tools and early intervention pathways.[16–18]

Despite advances in diagnostic criteria, accurately diagnosing ARVC remains a significant clinical challenge.[19,20] Borderline and gene-elusive presentations often exhibit subtle, variable, or inconclusive findings, leading to high rates of referral for expert re-evaluation.[21–23] Specialized ARVC centers such as the Johns Hopkins ARVC center in the United States, serve as a referral hub, yet many referred patients ultimately do not meet diagnostic criteria.[19,24] This places a substantial burden on the limited number of expert centers and also exposes patients to unnecessary interventions, lifestyle restrictions, and psychological distress resulting from misdiagnosis.[25] The high volume of false-positive referrals and the potential for clinical mismanagement highlight the need for more efficient and accurate triage strategies.

Machine learning (ML) and deep learning (DL) approaches offer powerful tools to enhance diagnostic precision by identifying complex, nonlinear patterns across multimodal clinical data that may be missed by traditional methods.[26–28] Significant effort has been devoted to applying DL models to raw electrocardiogram (ECG) and imaging data to improve diagnostic accuracy.[29] For example, Carrick et al. recently developed an ECG-based deep learning (ECG-DL) model that interprets raw ECG waveforms to aid in ARVC diagnosis.[30] While this model achieved a modest but clinically meaningful improvement over expert ECG interpretation (c-statistic 0.87 vs. 0.85), diagnostic performance rose substantially (c-statistic 0.94) when ECG-DL outputs were combined with non-ECG clinical features from the 2010 TFC.[30] These findings emphasize the added value of multimodal data integration and suggest that future gains in diagnostic accuracy may come more from strategic model design and data fusion than from increasing algorithmic complexity alone.

Therefore, the aim of this study was to develop, validate, and compare a range of predictive models for ARVC using machine learning algorithms that span from traditional, interpretable approaches to modern, high-complexity techniques. Using a multimodal dataset derived from the Johns Hopkins ARVC Registry, we further compared test characteristics of the top-performing models on hold-out test cohort.

Methods

Study Population

For this project, a retrospective cohort was assembled from the Johns Hopkins ARVC Registry to serve as the basis for algorithm training and internal validation. The cohort comprised 688 patients who underwent evaluation for suspected ARVC at the Johns Hopkins ARVC Center between January 2011 and September 2019. Among these, 358 patients (52%) met diagnostic criteria for ARVC based on the 2010 Revised Task Force Criteria (TFC), as adjudicated by expert clinicians at the Johns Hopkins ARVC Center. The remaining 330 individuals received alternative

diagnoses, including other forms of cardiomyopathy.[19] For each patient, a comprehensive dataset comprising 32 features was constructed by integrating clinical, demographic, and genetic variables relevant to the diagnosis of ARVC.

Data Preprocessing

The initial dataset comprised 688 patient records, each containing 32 clinical, demographic, electrocardiographic, imaging, and genetic features extracted from the Johns Hopkins ARVC Registry. These features were derived from standardized diagnostic evaluations, including genetic testing, electrocardiograms (ECG), Holter monitoring, and cardiac magnetic resonance imaging (MRI). Categorical variables such as 'Ethnicity' were transformed into numerical input by one-hot encoding.

Approximately 34% of patients had missing values in at least one feature. To ensure robustness in downstream machine learning analyses, patients with incomplete data in critical variables were excluded. For features with non-critical missing values, imputation was performed using a k-nearest neighbors (KNN) algorithm with k = 1, enabling preservation of local data structure while maintaining dataset integrity. The final analytic dataset comprised a matrix of 688 patients with 32 complete or imputed feature sets.

To assess linear correlation among features, we conducted a Pearson correlation analysis across variables. Most features exhibited low-to-moderate correlation, with the highest observed correlation being 0.78 between certain ECG lead variables, (Table1). These findings supported the inclusion of all features in the modeling pipeline without redundancy reduction.

Outcomes

The primary outcome for all analyses was ARVC diagnosed or not, as adjudicated by ARVC experts at Johns Hopkins ARVC center.

Machine learning models

Eight distinct machine learning algorithms were employed, representing a range of model categories. These included ensemble methods—Gradient Boosted Trees, Random Forest, and Decision Trees—which are effective in capturing nonlinear relationships and complex feature interactions. Linear classifiers such as Logistic Regression, Lasso Regression, and Ordinary Least Squares Regression were used for their interpretability and efficiency. A probabilistic classifier, Gaussian Naive Bayes, was included for its capacity to handle high-dimensional, sparse data and class imbalance. Additionally, a deep learning–based model, TabNet, was implemented to explore attention-driven neural network performance on structured, tabular data.

All model development and analysis were conducted in Python version 3.10, using a suite of open-source libraries tailored for machine learning applications. We used scikit-learn (v1.2.2) for classical machine learning algorithms and model evaluation utilities, and PyTorch (v1.13.1) for neural network modeling. The TabNet model was implemented using the pytorch-tabnet library (v3.1.1). Additional libraries used for data processing and visualization included NumPy (v1.24.2)

and pandas (v1.5.3). To optimize model performance, hyperparameters for all algorithms except Naive Bayes, Logistic Regression and Ordinary Least Squares, were systematically tuned using GridSearchCV, a grid-based search algorithm from scikit-learn.

Model testing and analysis

Model evaluation was performed using an 80/20 train-test split, with 80% of the data allocated for training and cross-validation, and the remaining 20% reserved for independent testing or internal validation. Univariate analyses were performed to assess group differences across variables. Chi-square tests were used to compare categorical variables, while independent samples *t*-tests and analysis of variance (ANOVA) were employed to evaluate differences in continuous variables, as appropriate. Model discrimination was assessed by calculating the area under the receiver operating characteristic curve (AUROC) using the roc_auc_score function from the scikit-learn library. The classification task was framed as a binary outcome, where the model output represented the predicted probability of a patient meeting diagnostic criteria for ARVC. Feature importance for Gradient Boosting classifier was quantified using the Gini importance (also known as mean decrease in impurity), as implemented in the scikit-learn framework, to identify variables contributing most significantly to model predictions.

Confusion matrices were constructed for each model. From these, sensitivity, specificity, and overall accuracy were calculated along with their corresponding 95% confidence intervals (CIs). Sensitivity was defined as the proportion of true positive results among all individuals with confirmed ARVC, while specificity was defined as the proportion of true negative results among individuals without the disease. Diagnostic accuracy was calculated as the proportion of all cases correctly classified, regardless of disease status. Exact binomial confidence intervals were used to estimate sensitivity and specificity.

Results

Cohort and Feature Characteristics

For this study, a retrospective cohort was derived from the Johns Hopkins ARVC Registry to support algorithm development and internal validation. The cohort included 688 individuals suspected of having ARVC who were referred for second evaluation at the Johns Hopkins ARVC Center between January 2011 and September 2019. Of these, 358 patients (52%) were diagnosed with ARVC according to the 2010 Revised Task Force Criteria (TFC), as determined by expert clinical adjudication. The remaining 330 individuals were assigned alternative diagnoses, including other cardiomyopathy subtypes. A composite dataset encompassing all 688 cases was constructed, integrating 32 variables spanning clinical, demographic, and genetic features relevant to ARVC diagnosis **(Table 1).** An overview of the data processing and analysis workflow is presented in **Fig. 1**.

To guide feature selection, Pearson correlation analysis was conducted to assess linear relationships among all variables. Most features exhibited low to moderate correlations, with the highest observed value being 0.78 between specific ECG lead variables. This level of correlation is expected given the physiological overlap between adjacent ECG leads and reflects clinically meaningful signals rather than redundancy **(Fig.2).** Accordingly, all features were retained in the modeling pipeline, as no evidence of excessive collinearity or feature redundancy was identified.

Performance of Predictive Models and Feature Importance

A broad range of ML models were selected including generalized linear models, DL, and tree models. All eight predictive models demonstrated strong discriminatory capacity, with mean AUC values exceeding 0.90 **(Supplementary Fig.1).** Gradient Boosted Trees achieved the highest AUC at 0.943, closely followed by Random Forest (0.938) and Lasso Regression (0.935), indicating that ensemble-based methods and regularized linear models are better at capturing feature patterns associated with ARVC diagnosis **(Fig.3).**

Interestingly, both Logistic Regression and Ordinary Least Squares (OLS) models also achieved high AUCs of 0.935, suggesting that linear approaches can perform comparably to more complex models. Gaussian Naïve Bayes, a probabilistic model, achieved a slightly lower AUC of 0.929 but still demonstrated strong classification capability. These findings suggest that simpler, interpretable models may retain considerable diagnostic value while offering the additional benefit of transparency in clinical settings.

In contrast, Decision Trees and TabNet models showed relatively lower performance, with AUCs of 0.910 and 0.900, respectively. While still above the threshold typically considered acceptable for medical classification tasks, their performance lagged behind ensemble methods and linear models. Overall, these results underscore the importance of model selection in diagnostic tool development for ARVC, with ensemble and regularized linear models emerging as the most effective approaches in this study.

In terms of classification metrics, the Lasso Regressor yielded the highest sensitivity at 0.879 (95% CI: 0.779–0.937), indicating a strong ability to identify true positive cases. Random Forest and Gradient Boosted Trees also performed well, with sensitivities of 0.864 and 0.848, respectively. Notably, Gaussian Naïve Bayes exhibited the highest specificity at 0.901 (95% CI: 0.810–0.951), suggesting superior performance in correctly identifying true negatives, though this came at the cost of lower sensitivity (0.758). Logistic Regression and OLS Regression displayed balanced but slightly lower metrics across the board, with both models achieving AUC values of 0.935 and sensitivity/specificity near 0.83–0.85.

Overall accuracy across models was consistently high, ranging from 0.825 to 0.869. The Lasso Regressor attained the highest overall accuracy (0.869; 95% CI: 0.812–0.925), while Gradient Boosted Trees and Random Forest followed closely at 0.847 and 0.861, respectively. All models achieved statistically significant results with p-values < 0.001, reinforcing the robustness of their predictive performance **(Table 2).**

Feature Importance by Best Performing Model

The best performing model, Gradient Boosted Trees, identified T wave inversion in lead V3, PVC count from Holter monitoring, and T wave inversion in lead V2 as the most influential features for classifying ARVC, highlighting the strong predictive value of ECG-derived abnormalities **(Fig. 4).** Structural imaging markers such as right ventricular ejection fraction (RVEF) and late gadolinium enhancement (LGE) also contributed meaningfully, though to a lesser extent. Presence of genetic variants, including mutations in PKP2 and DSP, along with clinical variables like PR interval and age at first visit, showed comparatively lower importance. Overall, the model emphasizes that T wave inversions and arrhythmic burden carry the highest discriminatory power, consistent with their known early presence in ARVC pathogenesis.

Discussion

In this study, we tested eight ML models with a multimodal dataset derived from a large referral cohort and developed a tool that was the most effective in the classification of Arrhythmogenic Right Ventricular Cardiomyopathy (ARVC). This tool, a Gradient Boosted Trees classifier, achieved the highest overall performance, with an AUC of 0.943, outperforming both traditional linear models and more complex deep learning approaches. These results highlight the potential of ML-based decision support tools to improve diagnostic precision for ARVC, a disease where early identification is critical yet often elusive in standard clinical practice.

A key insight from our feature importance analysis was the prominent role of electrocardiographic markers, specifically, T wave inversion in precordial leads V2 and V3, and ventricular ectopy burden as measured by Holter monitoring. These findings are consistent with known disease mechanisms in ARVC and suggest that surface ECG features, which are both low-cost and readily available, may contain disproportionately high predictive value.[31] Structural imaging parameters, such as right ventricular ejection fraction and late gadolinium enhancement, contributed to a lesser extent, suggesting that electrical disturbances precede overt structural remodeling in many cases. Interestingly, genetic variants in *PKP2* and *DSP* displayed lower relative importance, which may reflect the variable penetrance of these mutations and the presence of gene-elusive forms of ARVC in the dataset.

Notably, several linear models, including Lasso and Logistic Regression, performed comparably to ensemble methods, achieving AUCs exceeding 0.93 while offering greater interpretability. This suggests that, in well-structured datasets, even relatively simple models can achieve clinically actionable performance. In contrast, the TabNet deep learning model underperformed despite its capacity to model attention across structured data, possibly due to limited sample size and the absence of temporally rich or high-dimensional input modalities. These results support the notion that model architecture should be carefully matched to the structure of available data and that increased complexity does not always yield superior performance. Overall, our findings support the use of ML models as adjunctive tools for streamlining ARVC diagnosis, particularly in settings where expert interpretation is limited or delayed.

This study has several limitations that warrant consideration. First, the data were sourced from a single high-volume referral center, which may limit external validity. The cohort likely reflects a more diagnostically challenging subset of patients, and findings may not generalize to broader, community-based populations. Second, while internal validation was robustly performed using cross-validation and a held-out test set, the absence of external validation across multiple institutions limits assessment of model transportability. Third, the dataset was composed entirely of structured, manually curated features. Unstructured data such as raw ECG signals or imaging sequences were not incorporated, which may have constrained the performance ceiling for models designed to learn from high-dimensional inputs. Furthermore, diagnoses were adjudicated using the 2010 Revised Task Force Criteria, which may not fully capture evolving clinical phenotypes, particularly in early-stage or left-dominant presentations. Finally, while feature importance metrics offer interpretability, they do not establish causality and may be affected by feature collinearity or model-specific biases.

Future work should prioritize external validation of these models in multi-institutional datasets, particularly those with greater ethnic, geographic, and clinical diversity. In addition, integrating unstructured data, including raw ECG waveforms, cardiac imaging, and free-text clinical notes, may enable development of hybrid models capable of capturing temporal dynamics and subtle phenotypic nuances. Prospective evaluation in real-world clinical workflows is also critical to assess whether ML-based tools improve diagnostic efficiency, reduce unnecessary referrals, and optimize patient outcomes. Furthermore, longitudinal models that incorporate time-series data may enhance risk stratification and predict disease progression in individuals with indeterminate or borderline features. Finally, as molecular and genomic testing become more integrated into routine cardiology care, future models should explore the utility of polygenic risk scores and transcriptomic signatures in improving classification performance, particularly in genetically elusive ARVC subtypes. Collectively, these efforts will support the safe and effective translation of ML approaches into clinical practice for heritable arrhythmia syndromes.

Data availability

Patient information was not able to be sufficiently deidentified and is protected. Therefore, is not available.

Code availability

Code is available on Github at https://github.com/Kwaku-quansah/ARVC_ML.git

Funding

This work was supported by grants from NIH/NHLBI (R01HL164936, R01HL171205) and AHA (23IPA1051956).

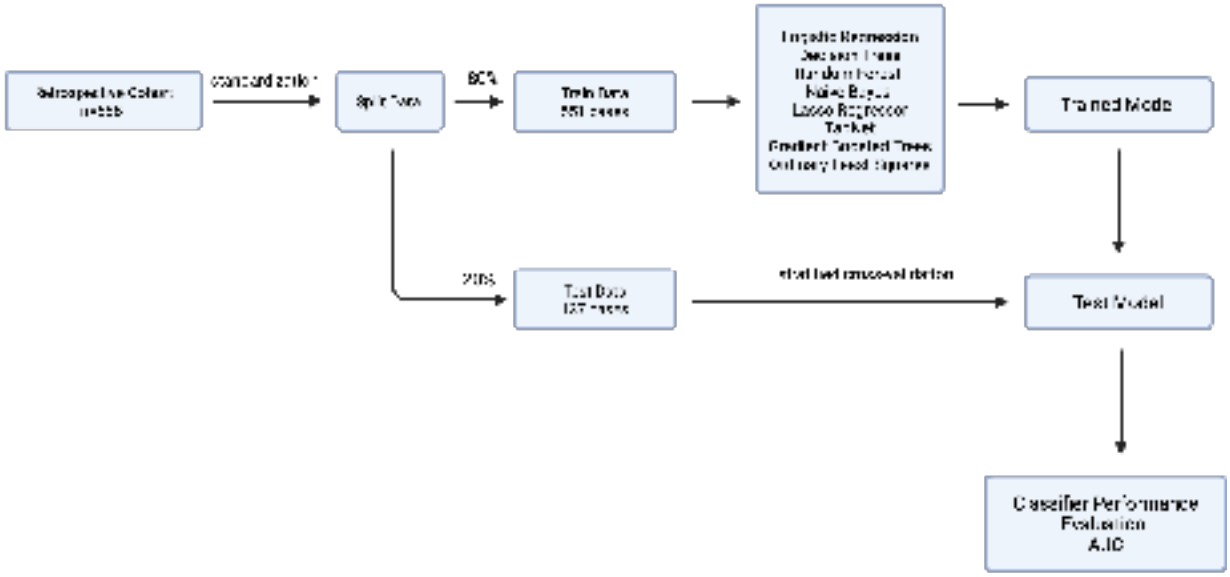


Figure 1. Diagram showing the data split for model training and creation of 8 machine learning models.

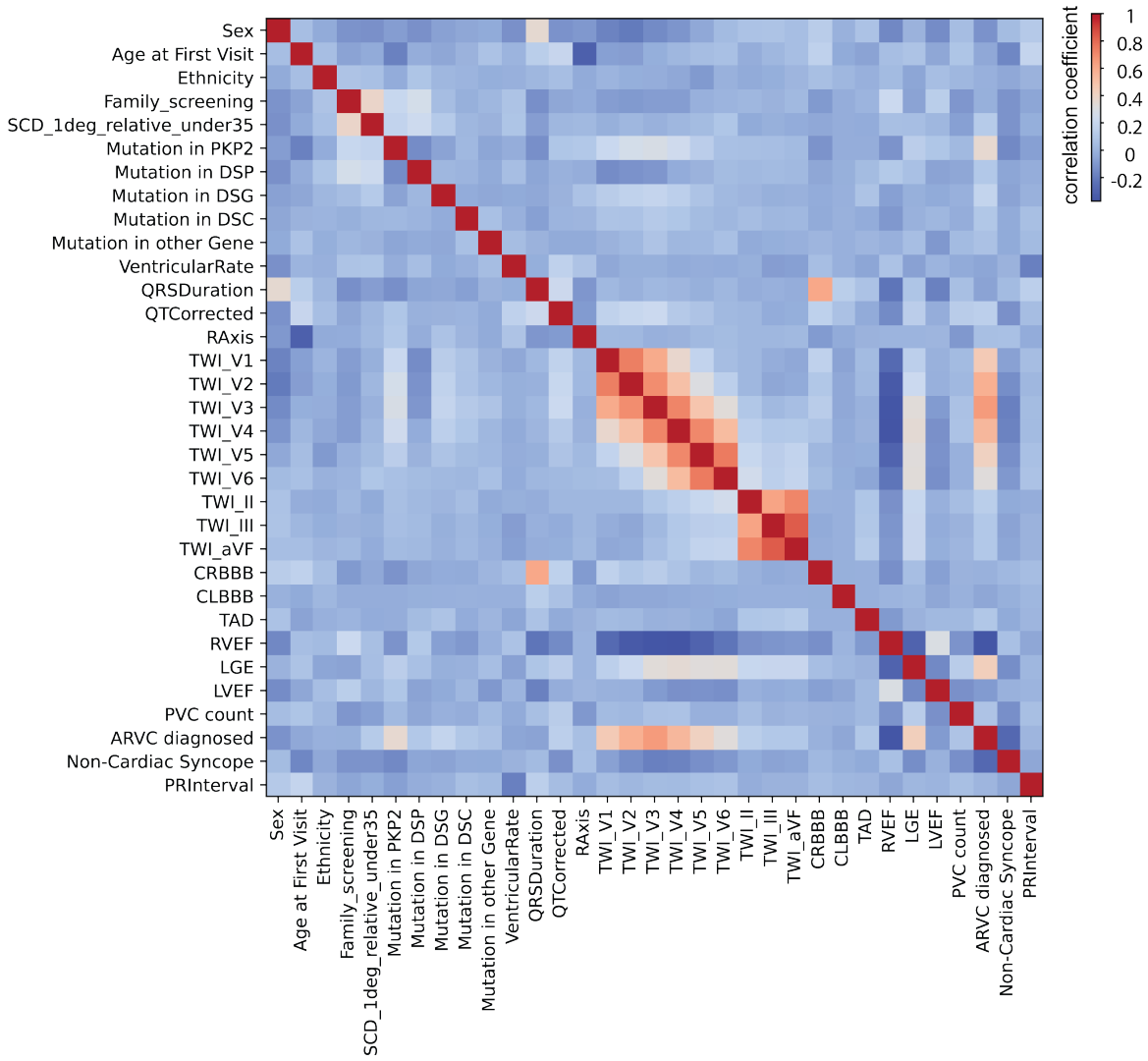


Figure 2. Pearson correlation heatmapwith scale bar ranges from red negative correlation to blue positive correlation. Abbreviations are available in the appendix.

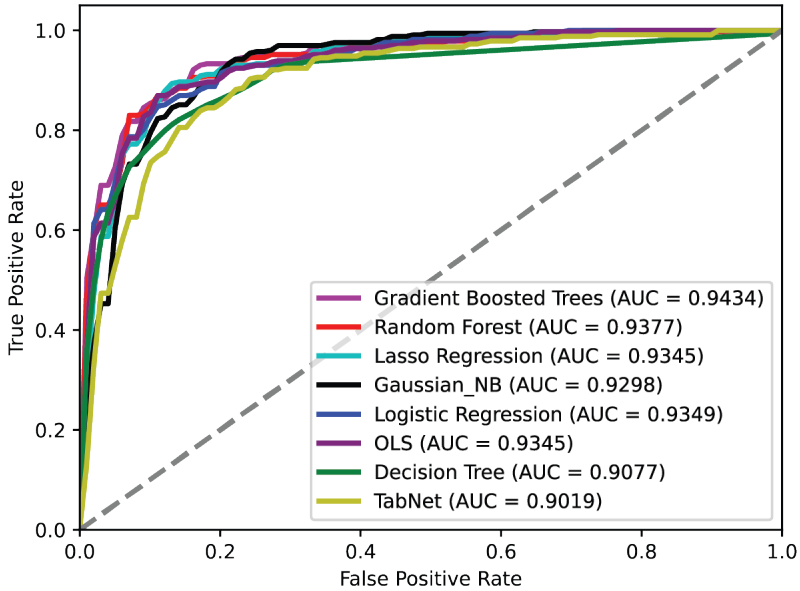


Figure 3. Comparison of ROC of all 8 machine learning models

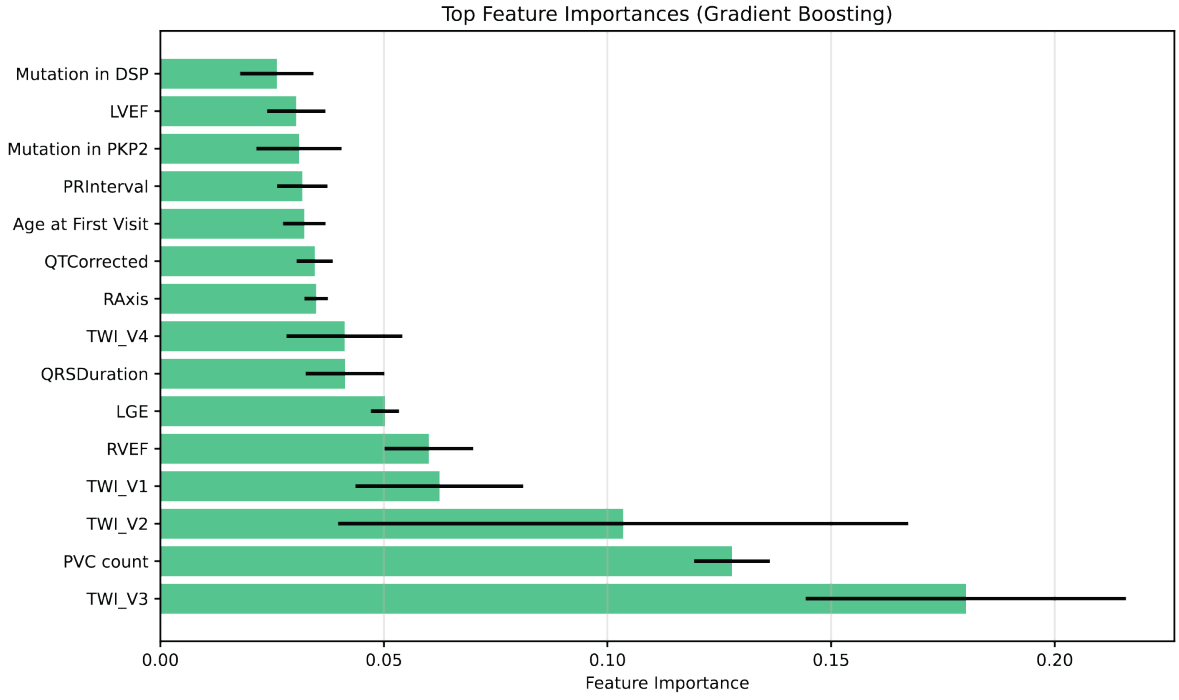
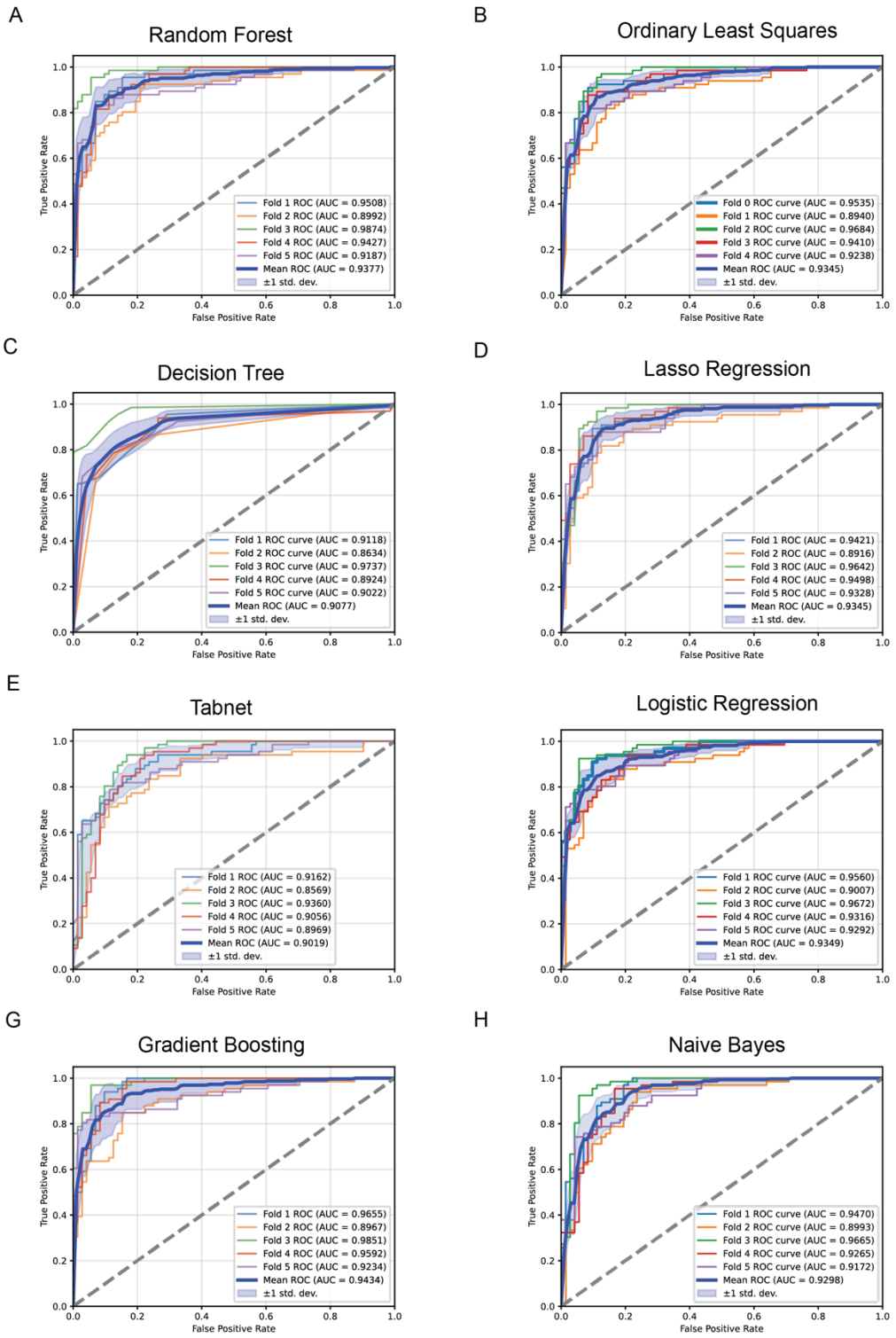


Figure 4. Feature Importance by Gradient Boosted Trees. Abbreviations are available in the appendix.



Supp Figure 1 ROC of each trained model A. random forest B. Ordinary Least Squares C. Decision Tree D. Lasso Regression E. Tabnet F. Logistic Regression G Gradient Boosting H. Naïve Bayes

Characteristic	All Cases	% Missing	ARVC cases	Non-ARVC cases	P-value
Sample size	688		329(47.8%)	359(52.2%)	chi-squared test for categorical data
Age	40.2±15.1	0	39.41 ± 14.69	40.84 ± 15.48	9.50E-04
Sex		0			0.0013
Male	394(57.3%)		167(50.76%)	227(63.23%)	
Female	294(42.7%)		162(49.24%)	132(36.77%)	
Ethnicity		0			0.8456
White	576(83.7%)		274(83.28%)	302(84.12%)	
Non-white	112(16.3%)		55(16.72%)	57(15.88%)	
SCD_1deg_relative_under35	118(17.15%)	0.15	73(21.6%)	45(12.4%)	0.0012
Mutation in PKP2	152(28.35%)	0	124(37.7%)	28(7.79%)	0
Mutation in DSP	49(7.12%)	0	33(10%)	16(4.4%)	0.0071
Mutation in DSG	31(4.5%)	0	28(8.5%)	3(0.8%)	0
Mutation in DSC	13(1.8%)	0	11(3.3%)	2(0.5%)	0.0163
Mutation in other Gene	35(5%)	0	22(6.7%)	13(3.62%)	0.0981
TWI_V1	391(56.91%)	0.15	265(80.55%)	126(35.20%)	0
TWI_V2	295(42.94%)	0.15	238(72.64%)	56(15.64%)	0
TWI_V3	288(41.92%)	0.15	248(75.38%)	40(11.17%)	0
TWI_V4	205(29.84%)	0.15	183(55.62%)	22(6.15%)	0
TWI_V5	126(18.34%)	0.15	115(34.95%)	11(3.07%)	0
TWI_V6	77(11.21%)	0.15	72(21.88%)	5(1.40%)	0
TWI_II	16(2.33%)	0.15	15(4.56%)	1(0.28%)	0.0005
TWI_III	35(5.09%)	0.15	25(7.60%)	10(2.79%)	0.0072
TWI_aVF	27(3.94%)	0.29	20(6.10%)	7(1.96%)	0.0096
CRBBB	49(7.13%)	0.15	23(6.99%)	26(7.26%)	1
CLBBB	2(0.29%)	0.29	0(0.00%)	2(0.56%)	0.5151
LGE	170(34.00%)	27.33	133(56.12%)	37(14.07%)	0
TAD	34(4.95%)	0.15	24(7.29%)	10(2.79%)	0.011
Ventricular_Rate	65.15 ± 14.36	0	64.29 ± 15.70	65.95 ± 12.99	1.33E-01
PR_Interval	158.93 ± 39.13	0	159.42 ± 43.92	158.48 ± 34.22	7.54E-01
QRS_Duration		0			9.63E-02
QT_Corrected	426.04 ± 33.92	0	431.39 ± 34.96	421.13 ± 32.22	7.38E-05
R_Axis	38.53 ± 53.34	0	41.80 ± 61.11	35.54 ± 44.94	1.29E-01
LVEF	56.66 ± 9.36	12.79	55.77 ± 10.14	57.48 ± 8.53	1.15E-01
PVC	3739.73 ± 7033.90	15.55	4386.20 ± 6567.72	3147.29 ± 7395.10	2.38E-02
RVEF	45.28 ± 10.86	31.4	41.12 ± 10.74	49.10 ± 9.49	2.26E-16

Table 1. Characteristics of the patient cohort

Model	AUC	TN	FP	FN	TP	Sensitivity(95% CI)	Specificity (95% CI)	Accuracy(95% CI)	P-Value
Gradient Boosted Trees	0.943	60	11	10	56	0.848(0.743-0.916)	0.845(0.743-0.911)	0.847(0.786-0.907)	<0.00
Random Forest	0.938	61	10	9	57	0.864(0.761-0.927)	0.859(0.76-0.922)	0.861(0.803-0.919)	<0.00
Lasso Regressor	0.935	61	10	8	58	0.879(0.779-0.937)	0.859(0.76-0.922)	0.869(0.812-0.925)	<0.00
Gaussian Naïve Bayes	0.929	64	7	16	50	0.758(0.642-0.845)	0.901(0.810-0.951)	0.832(0.77-0.895)	<0.00
Logistic Regression	0.935	57	14	10	56	0.848(0.743-0.916)	0.803(0.696-0.879)	0.825(0.761-0.888)	<0.00
OLS	0.935	60	11	11	55	0.833(0.726-0.904)	0.845(0.743-0.911)	0.839(0.778-0.901)	<0.00

Table 2. Test characteristics of top ARVC prediction models on hold-out test cohort

References

1. Sattar, Y., Abdullah, H. M., Neisani Samani, E., Myla, M. & Ullah, W. Arrhythmogenic Right Ventricular Cardiomyopathy/Dysplasia: An Updated Review of Diagnosis and Management. *Cureus* 11, e5381.
2. Corrado, D., Basso, C. & Thiene, G. Arrhythmogenic right ventricular cardiomyopathy: diagnosis, prognosis, and treatment. *Heart* 83, 588–595 (2000).
3. Olivetti, N. *et al.* Enhancing Arrhythmogenic Right Ventricular Cardiomyopathy Detection and Risk Stratification: Insights from Advanced Echocardiographic Techniques. *Diagnostics* 14, 150 (2024).
4. Moisa, S. M. *et al.* Arrhythmogenic Right Ventricular Cardiomyopathy in Children: A Systematic Review. *Diagnostics (Basel)* 14, 175 (2024).
5. Mariani, M. V. *et al.* Inherited Arrhythmias in the Pediatric Population: An Updated Overview. *Medicina (Kaunas)* 60, 94 (2024).
6. Chelko, S. P. *et al.* Therapeutic Modulation of the Immune Response in Arrhythmogenic Cardiomyopathy. *Circulation* 140, 1491–1505 (2019).
7. Lin, C.-Y. *et al.* Clinical significance of structural remodeling concerning substrate characteristics and outcomes in arrhythmogenic right ventricular cardiomyopathy. *Heart Rhythm O2* 3, 422–429 (2022).
8. Cruz, F. M. *et al.* Exercise Triggers ARVC Phenotype in Mice Expressing a Disease-Causing Mutated Version of Human Plakophilin-2. *JACC* 65, 1438–1450 (2015).
9. Chelko, S. P. *et al. Exercise Triggers CAPN1-Mediated AIF Truncation, Inducing Myocyte Cell Death in Arrhythmogenic Cardiomyopathy*. *Sci. Transl. Med* vol. 13 891 https://www.science.org (2021).
10. Corrado, D. *et al.* Treatment of arrhythmogenic right ventricular cardiomyopathy/dysplasia: an international task force consensus statement. *European Heart Journal* 36, 3227–3237 (2015).

11. Gasperetti, A. *et al.* Differences in underlying cardiac substrate among S-ICD recipients and its impact on long-term device-related outcomes: Real-world insights from the iSUSI registry. *Heart Rhythm* (2024) doi:10.1016/j.hrthm.2023.12.007.
12. Cadrin-Tourigny, J. *et al.* A new prediction model for ventricular arrhythmias in arrhythmogenic right ventricular cardiomyopathy. *Eur Heart J* 43, e1–e9 (2022).
13. Bosman, L. P. *et al.* Diagnosing arrhythmogenic right ventricular cardiomyopathy by 2010 Task Force Criteria: clinical performance and simplified practical implementation. *EP Europace* 22, 787–796 (2020).
14. Marcus, F. I. *et al.* Diagnosis of arrhythmogenic right ventricular cardiomyopathy/Dysplasia: Proposed modification of the task force criteria. *Circulation* 121, 1533–1541 (2010).
15. Towbin, J. A. *et al.* 2019 HRS expert consensus statement on evaluation, risk stratification, and management of arrhythmogenic cardiomyopathy. *Heart Rhythm* 16, e301–e372 (2019).
16. Groeneweg, J. A. *et al.* Clinical Presentation, Long-Term Follow-Up, and Outcomes of 1001 Arrhythmogenic Right Ventricular Dysplasia/Cardiomyopathy Patients and Family Members. *Circulation: Cardiovascular Genetics* 8, 437–446 (2015).
17. Rashid, S. *et al.* Validation of arrhythmogenic right ventricular cardiomyopathy risk calculator for sudden cardiac death: a systematic review. *Int J Arrhythm* 24, 25 (2023).
18. Muller, S. A. *et al.* Individualized Family Screening for Arrhythmogenic Right Ventricular Cardiomyopathy. *J Am Coll Cardiol* 82, 214–225 (2023).
19. Sampognaro, J. R. *et al.* Diagnostic pitfalls in patients referred for arrhythmogenic right ventricular cardiomyopathy. *Heart Rhythm* 20, 1720–1726 (2023).
20. Pinamonti, B., Brun, F., Mestroni, L. & Sinagra, G. Arrhythmogenic right ventricular cardiomyopathy: From genetics to diagnostic and therapeutic challenges. *World J Cardiol* 6, 1234–1244 (2014).

21. James, C. A. *et al.* International Evidence Based Reappraisal of Genes Associated With Arrhythmogenic Right Ventricular Cardiomyopathy Using the Clinical Genome Resource Framework. *Circulation: Genomic and Precision Medicine* 14, e003273 (2021).
22. Bosman, L. P. *et al.* Predicting arrhythmic risk in arrhythmogenic right ventricular cardiomyopathy: A systematic review and meta-analysis. *Heart Rhythm* 15, 1097–1107 (2018).
23. Oomen, A. W. G. J., Semsarian, C., Puranik, R. & Sy, R. W. Diagnosis of Arrhythmogenic Right Ventricular Cardiomyopathy: Progress and Pitfalls. *Heart, Lung and Circulation* 27, 1310–1317 (2018).
24. Vohra, J. Pitfalls in the diagnosis of arrhythmogenic right ventricular cardiomyopathy. *Heart Rhythm* 20, 1727–1728 (2023).
25. Sharma, A. *et al.* Misdiagnosis of ARVC leading to inappropriate ICD implant and subsequent ICD removal – lessons learned. *Journal of Cardiovascular Electrophysiology* 30, 2020–2026 (2019).
26. Safdar, S., Zafar, S., Zafar, N. & Khan, N. F. Machine learning based decision support systems (DSS) for heart disease diagnosis: a review. *Artif Intell Rev* 50, 597–623 (2018).
27. Abubaker, M. B. & Babayiğit, B. Detection of Cardiovascular Diseases in ECG Images Using Machine Learning and Deep Learning Methods. *IEEE Transactions on Artificial Intelligence* 4, 373–382 (2023).
28. Shameer, K., Johnson, K. W., Glicksberg, B. S., Dudley, J. T. & Sengupta, P. P. Machine learning in cardiovascular medicine: are we there yet? (2018) doi:10.1136/heartjnl-2017-311198.
29. Avula, V., Wu, K. C. & Carrick, R. T. Clinical Applications, Methodology, and Scientific Reporting of Electrocardiogram Deep-Learning Models. *JACC: Advances* 2, 100686 (2023).
30. Carrick, R. T. *et al.* Improved diagnosis of arrhythmogenic right ventricular cardiomyopathy using electrocardiographic deep learning. *Heart Rhythm* 22, 1080–1088 (2025).

31. Zhang, L., Liu, L., Kowey, P. R. & Fontaine, G. H. The Electrocardiographic Manifestations of Arrhythmogenic Right Ventricular Dysplasia. *Curr Cardiol Rev* 10, 237–245 (2014).

.