An Interpretable Machine Learning Classifier for Brugada Syndrome Detection on 12-Lead Electrocardiograms: A Baseline Study on the Brugada-HUCA Dataset

Sanjeet[1]

July 2026

[1] *Independent Researcher. Correspondence: sanjeet.jaysee@gmail.com*

Abstract

Background. Brugada syndrome is an inherited cardiac channelopathy associated with sudden cardiac death in structurally normal hearts. Its diagnostic electrocardiographic signature, ST segment elevation in the right precordial leads, is frequently subtle, intermittent, or absent on a resting electrocardiogram, and reliable unmasking currently requires a pharmacological provocation test that carries a small but recognized proarrhythmic risk. **Methods.** We developed an interpretable machine learning classifier using the Brugada-HUCA dataset (PhysioNet version 1.0.0, February 2026; 76 Brugada syndrome patients, 287 controls). Nine clinically grounded features (J point amplitude, ST segment amplitude at two offsets, ST slope, R and S amplitudes, QRS duration, T wave amplitude, and J to R ratio) were extracted from robust per lead median beats across all twelve leads, together with three global rhythm features, for a total of 111 features. An XGBoost classifier and a random forest comparator were trained and validated using patient level repeated stratified five fold cross validation with ten repeats. **Results.** The primary XGBoost model achieved an area under the receiver operating characteristic curve (AUROC) of 0.90 (95% confidence interval 0.80 to 0.97) and a sensitivity of 76% at a fixed specificity of 90%. SHAP (SHapley Additive exPlanations) attribution confirmed that the three most influential features were ST amplitude in lead V2, ST amplitude in lead V1, and J point amplitude in lead V2, consistent with the established right precordial signature of the disease. Error analysis demonstrated that sensitivity was 62% among patients whose baseline electrocardiogram was non pathological, compared with 83% among patients with an overtly abnormal baseline tracing, and that missed cases exhibited an intermediate degree of ST elevation between controls and correctly identified cases. **Conclusions.** An interpretable, feature based classifier achieves performance comparable to more complex deep learning approaches reported in the literature while remaining fully auditable. The model's errors concentrate on the concealed phenotype, the same population that motivates provocation testing in clinical practice,

indicating that its failure mode is physiologically coherent rather than arbitrary. This work is presented as a reproducible single center baseline and is not intended for clinical deployment.

Keywords: Brugada syndrome; electrocardiography; machine learning; XGBoost; interpretability; sudden cardiac death; SHAP

---

1 Introduction

1.1 Clinical Background

Brugada syndrome is an inherited cardiac channelopathy with an estimated prevalence of approximately 1 in 2,000, occurring more frequently in young to middle aged men, and it predisposes affected individuals to ventricular fibrillation and sudden cardiac death in the absence of structural heart disease [1]. Approximately one in ten cases of sudden cardiac death is attributable to an inherited arrhythmic condition such as Brugada syndrome [6]. Roughly 20% of clinically diagnosed probands carry an identifiable mutation in *SCN5A*, the gene encoding the principal cardiac sodium channel, although the majority of cases have no single identified causative variant, underscoring that Brugada syndrome remains fundamentally an electrocardiographic diagnosis [1,2].

The diagnostic hallmark of the disease is ST segment elevation in the right precordial leads (V1 and V2). A type 1, or coved, pattern with at least 2 mm of J point elevation and a downsloping ST segment is diagnostic in isolation, whereas a type 2, or saddleback, pattern is suggestive and requires further evaluation [1]. This electrocardiographic pattern is dynamic in nature. It may be present on one tracing and absent on the next, and it is known to be modulated by fever, autonomic tone, certain medications, and normal physiological variation [1,5]. Consequently, a single normal resting electrocardiogram does not exclude the diagnosis.

1.2 The Diagnostic Challenge

When a resting electrocardiogram is inconclusive, the established next step is pharmacological provocation testing, typically with ajmaline or flecainide, which unmasks the type 1 pattern in patients whose baseline tracing is otherwise unremarkable. In a large cohort of families evaluated following sudden unexplained death, ajmaline testing was required to establish the diagnosis in 97% of individuals ultimately found to have Brugada syndrome [3]. In the same cohort, 17% of patients with an initially concealed pattern subsequently developed a spontaneous type 1 pattern or a clinically significant arrhythmic event on follow up, indicating that concealed Brugada syndrome represents a genuine diagnostic difficulty rather than a benign variant [3]. Provocation testing, while generally safe when performed by experienced operators, is an invasive pharmacological stress test that carries a recognized, if small, risk of inducing the ventricular arrhythmias it is designed to unmask [4,5]. Consequently, it is reserved for patients in whom clinical suspicion already justifies this risk, and no established, non-intrusive method exists to identify which resting electrocardiograms warrant closer scrutiny in the first instance.

1.3 Prior Work in Artificial Intelligence Based Brugada Syndrome Detection

Several groups have applied machine learning methods to this problem. Melo et al. trained a deep neural network on raw electrocardiogram waveforms without provocation testing, reporting 88.4% accuracy and an AUROC of 0.934 in validation on an Italian cohort [6]. Zanchi et al. restricted their analysis to P wave morphology, reflecting an atrial phenotype associated with Brugada syndrome distinct from the classical

ventricular ST/J point signature, and achieved 81% accuracy (86% sensitivity, 73% specificity) using an AdaBoost classifier on 123 patients [7]. Ronan et al. applied a self supervised pretraining strategy (VICReg) to address the label scarcity characteristic of rare disease electrocardiogram classification, achieving an AUROC of 0.88 and an area under the precision recall curve (AUPRC) of 0.82, and further used the resulting model to identify previously undiagnosed cases within an institutional cohort [8]. A recent review summarizes this body of work and notes that artificial intelligence models have, in several settings, matched or exceeded the performance of trained cardiologists in detecting subtle Brugada patterns [9].

1.4 Study Objectives

None of the studies cited above used the Brugada-HUCA dataset, a twelve lead electrocardiogram dataset from Hospital Universitario Central de Asturias released on PhysioNet in February 2026 [10,11]. At the time of writing, no published classifier built on this dataset could be identified. The objective of this study is threefold: first, to establish an independent, fully reproducible classifier on Brugada-HUCA using interpretable, clinically grounded features under leakage free, patient level cross validation; second, to verify through SHAP attribution that the classifier relies on the established right precordial ST/J point signature rather than an incidental artifact of the dataset; and third, to characterize the classifier's error patterns in relation to the concealed phenotype described above.

2 Materials and Methods

2.1 Dataset

Table 1. Brugada-HUCA dataset summary.

Property	Value
Source	Brugada-HUCA, PhysioNet v1.0.0 (February 2026) [11]
Subjects	363 (76 Brugada syndrome, 287 healthy controls)
Brugada subtypes	69 type 1, 7 type 2
Prevalence	20.9%
Leads	12 standard (I, II, III, aVR, aVL, aVF, V1 to V6)
Sampling rate	100 Hz
Duration	12 s (1,200 samples per lead)
Records per patient	1 (confirmed; no risk of train/test leakage)
Metadata fields	patient identifier, basal pattern, sudden death flag, Brugada label

The dataset contains a single, consistent lead ordering with no missing samples and exactly one electrocardiogram per patient, so that any patient level cross validation split is inherently leakage free. The Brugada label takes three values (0, control; 1, type 1; 2, type 2); the two Brugada subtypes were collapsed into a single binary positive class for the primary classification task. The basal pattern flag, which indicates whether the baseline resting electrocardiogram was already read as pathological, was withheld from the model and used exclusively in the post hoc error analysis described in Section 3.4.

Figure 1. Representative right precordial waveforms: Brugada type 1 case (patient 188981) versus control (patient 251972).

2.2 Signal Preprocessing

Each lead was band pass filtered (0.5 to 40 Hz, fourth order Butterworth, zero phase forward backward filtering) to remove baseline wander and high frequency noise. R peaks were detected on lead II. Quality control indicated no flat leads, no R peak detection failures, and physiologically plausible heart rates across the entire dataset (39 to 182 beats per minute, median 69 beats per minute). No records were excluded. Figure 2 illustrates this pipeline for the two representative patients introduced in Figure 1.

Figure 2. Preprocessing quality control for the representative Brugada and control cases.

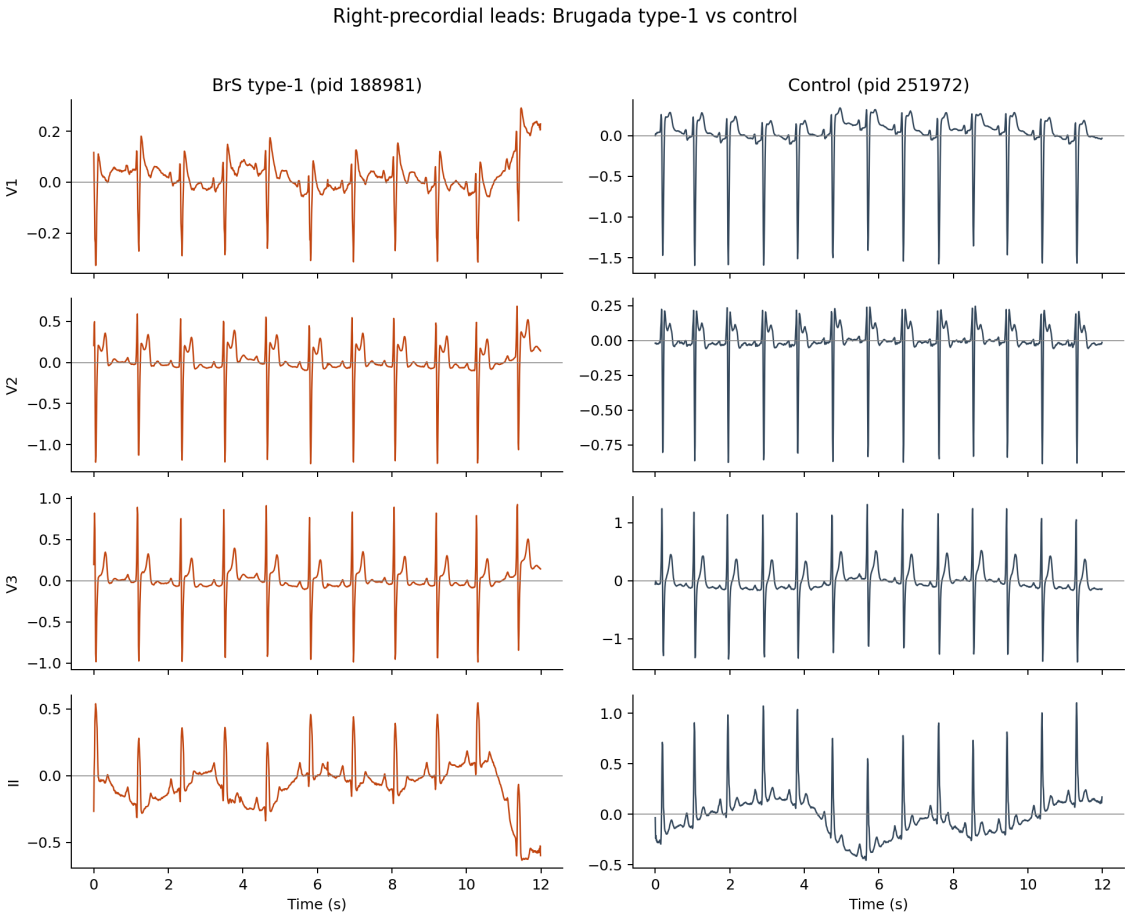

Figure 1: Right precordial leads for a confirmed Brugada type 1 case compared with a control. The Brugada case shows the coved, elevated ST/J point morphology in V1 and V2 characteristic of the type 1 pattern, with a downsloping return to baseline. The control shows a sharp, unremarkable return to baseline in the same leads.

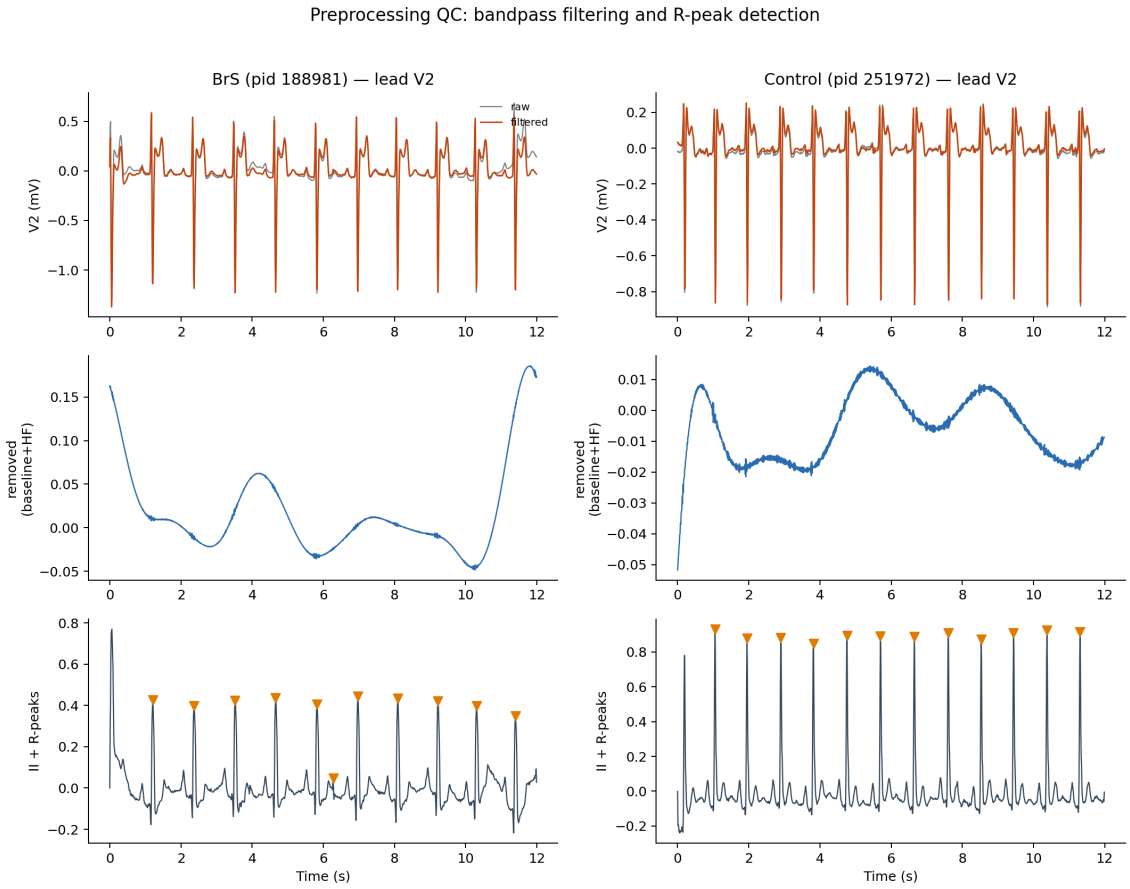

Figure 2: Preprocessing quality control. Top row: raw versus band pass filtered lead V2. Middle row: the component removed by filtering (baseline wander and high frequency noise). Bottom row: detected R peaks (markers) overlaid on filtered lead II. Left column: Brugada case. Right column: control.

2.3 Feature Extraction

For each subject and lead, a robust median beat was constructed by aligning all detected beats on the R peak and computing the sample wise median across beats, a procedure that suppresses single beat noise and motion artifact without requiring manual selection of a representative beat. From each median beat, nine interpretable features were extracted: J point amplitude, ST amplitude at J plus 40 ms and J plus 80 ms, ST slope, R and S amplitudes, QRS duration, signed T wave amplitude, and the J to R amplitude ratio. Across twelve leads this yields 108 features, to which three global rhythm features were appended (median heart rate, median RR interval, and RR interval variability), for 111 features in total with no missing values.

2.4 Statistical Feature Screening

Prior to model development, univariate discrimination between Brugada syndrome and control groups was assessed for every feature using Cohen's d and the Mann-Whitney U test with Bonferroni correction. This step serves as an independent, model free confirmation of which features should drive a downstream classifier if that classifier is learning genuine physiology rather than noise. Forty three of 111 features survived Bonferroni correction. The strongest discriminators, summarized in Table 2 and Figure 3, correspond directly to the leads and intervals implicated in the established Brugada syndrome signature.

Table 2. Top discriminating features by Cohen's d (Bonferroni significant).

Feature	Cohen's d	Interpretation
V2 ST amplitude (J+40 ms)	+1.08	Right precordial ST elevation
V1 ST amplitude (J+40 ms)	+1.00	Right precordial ST elevation
V4 QRS duration	+0.90	QRS prolongation
V3 QRS duration	+0.90	QRS prolongation
V2 QRS duration	+0.87	QRS prolongation
V2 T wave amplitude	-0.85	Attenuated positive T wave
V2 J point amplitude	+0.84	J point elevation

The V2 T wave effect reflects an attenuation of a normally positive T wave in Brugada syndrome (median +0.18 mV in cases versus +0.31 mV in controls) rather than a group level inversion to a negative T wave; only approximately 28% of Brugada syndrome records in this dataset exhibit a truly negative V2 T amplitude.

Figure 3. Distributions of the six most discriminating features between Brugada syndrome and control groups.

2.5 Model Development

An XGBoost classifier was trained as the primary model (300 trees, maximum depth 3, learning rate 0.05, with the scale_pos_weight parameter set to correct for class imbalance), together with a random forest comparator using balanced class weighting. Class imbalance (20.9% prevalence) was addressed through

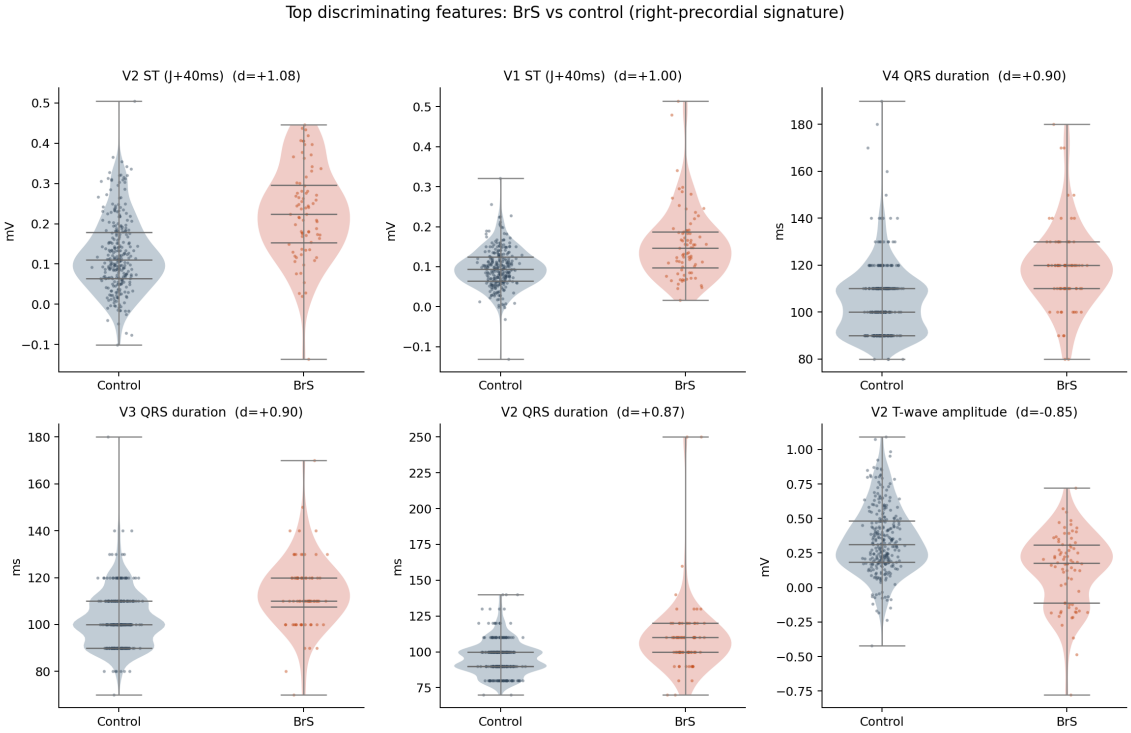

Figure 3: Distributions of the six strongest discriminating features. Each panel compares control and Brugada syndrome groups, with individual patients jittered over a violin density estimate; horizontal lines mark the 25th, 50th, and 75th percentiles.

class weighting rather than synthetic oversampling, since naive application of oversampling methods such as SMOTE across cross validation folds risks leaking synthesized samples between training and test partitions.

2.6 Cross Validation Strategy

Both models were validated using repeated stratified five fold cross validation with ten repeats, for 50 folds in total. Because Brugada-HUCA contains exactly one electrocardiogram per patient, every resulting split is inherently patient level and leakage free.

2.7 Performance Metrics

Reported metrics include AUROC, AUPRC, sensitivity, specificity, and sensitivity at a fixed specificity of 90%. Accuracy is reported for completeness but is de-emphasized as a primary outcome, since a classifier that predicts the majority class for every subject already achieves 79% accuracy on this dataset without providing any clinical utility.

3 Results

3.1 Discriminating Features

Univariate screening (Section 2.4) identified right precordial ST and J point elevation, together with QRS prolongation in leads V2 through V4, as the strongest discriminators between groups, consistent with the established electrocardiographic definition of Brugada syndrome.

3.2 Classification Performance

Table 3. Cross validated performance, mean over 50 folds with 95% interval in brackets.

Metric	XGBoost (primary)	Random forest
AUROC	0.900 [0.80, 0.97]	0.910 [0.84, 0.97]
AUPRC	0.777 [0.57, 0.92]	0.789 [0.64, 0.91]
Sensitivity (threshold 0.5)	0.693 [0.42, 0.87]	0.725 [0.47, 0.88]
Specificity (threshold 0.5)	0.944 [0.88, 1.00]	0.923 [0.88, 0.98]
Sensitivity at 90% specificity	0.763 [0.53, 0.94]	0.751 [0.53, 0.93]
Accuracy	0.891	0.881

The two models did not differ meaningfully in AUROC. The XGBoost model was selected as primary on the basis of superior probability calibration, shown in Figure 4. As an independent check, the full cross validation procedure was re-executed from the raw data checkpoints during manuscript preparation, reproducing an AUROC of 0.900 [0.80, 0.97] and an AUPRC of 0.775 [0.59, 0.92], closely matching the original result and indicating that the finding is not dependent on a single random seed.

At the default classification threshold of 0.5, the model achieves high specificity (94.4%) but misses approximately 30% of true Brugada syndrome cases. Adjusting the threshold to maintain 90% specificity recovers sensitivity to 76.3%. The choice between these operating points depends on the relative clinical cost of a missed case versus a false alarm, a decision that lies outside the scope of this study.

Figure 4. Receiver operating characteristic, precision recall, and calibration curves for the primary XGBoost model.

3.3 Model Interpretability

SHAP attribution was computed on the fitted XGBoost model to determine whether its predictions are driven by clinically meaningful features rather than incidental correlations specific to this dataset. The three features with the highest mean absolute SHAP value were ST amplitude in lead V2, ST amplitude in lead V1, and J point amplitude in lead V2, matching the univariate screen in Table 2 feature for feature. Five of the ten most important features by SHAP value originate in leads V1 through V3. This agreement between an independent

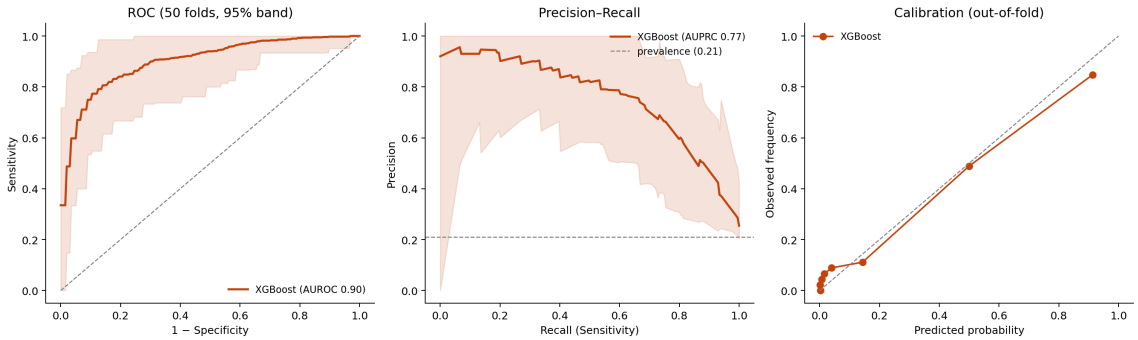

Figure 4: Left: receiver operating characteristic curve with a 95% band across the 50 cross validation folds. Center: precision recall curve relative to the 21% prevalence baseline. Right: out of fold calibration, showing predicted probabilities tracking the diagonal.

statistical screen and the trained model's attribution indicates that the classifier has learned the established right precordial Brugada signature rather than an artifact specific to this hospital's recording equipment or patient population.

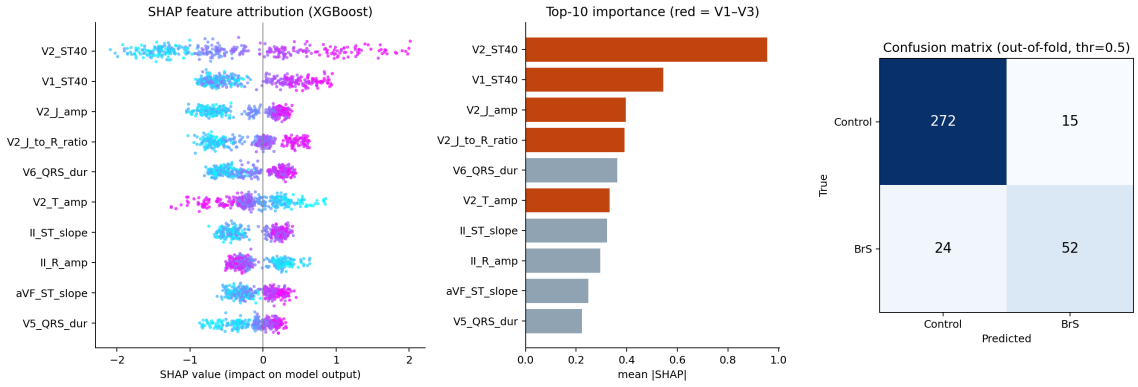

Figure 5: Left: SHAP beeswarm plot, in which each point represents one patient, color encodes that patient's feature value, and horizontal position indicates the magnitude and direction of that feature's contribution to the prediction. Center: the same ten features ranked by mean absolute SHAP value, with right precordial leads (V1 to V3) shown in a distinct color. Right: out of fold confusion matrix at the default threshold.

Figure 5. SHAP feature attribution and out of fold confusion matrix (272 true negatives, 52 true positives, 24 false negatives, 15 false positives).

3.4 Error Analysis

The 24 false negative cases were not distributed randomly across the Brugada syndrome group; they concentrated in the concealed phenotype. Sensitivity was 62% (33 of 53) among patients whose baseline electrocardiogram was non pathological, compared with 83% (19 of 23) among patients with an overtly abnormal baseline tracing. Sensitivity by subtype was 67% (46 of 69) for type 1 and 6 of 7 for type 2, the latter estimate being too small a sample to support any independent conclusion and reported here for completeness only.

Missed cases exhibited an intermediate degree of right precordial ST elevation: a mean V2 ST amplitude

of 0.174 mV among false negatives, compared with 0.119 mV among controls and 0.252 mV among true positives. The classifier's errors therefore concentrate on patients whose resting electrocardiogram does not fully express the Brugada signature, the same population that routine clinical interpretation struggles with and the same population that sodium channel blocker provocation testing is designed to identify.

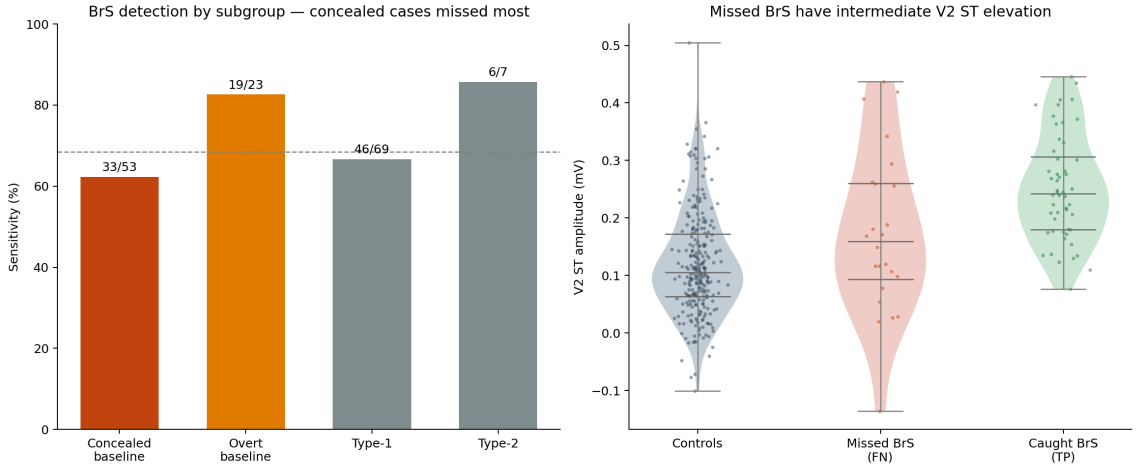

Figure 6: Left: sensitivity by clinical subgroup. Concealed baseline cases are detected substantially less reliably than overt baseline cases; the dashed line marks overall sensitivity. Right: V2 ST amplitude by outcome group. Missed Brugada syndrome cases (false negatives) fall between controls and correctly identified cases (true positives).

Figure 6. Subgroup sensitivity and V2 ST amplitude distribution by classification outcome.

4 Discussion

4.1 Comparison with Prior Work

Table 4 situates the present results alongside the studies discussed in Section 1.3. This comparison is presented as context rather than as a controlled benchmark. The underlying datasets, patient populations, provocation status, input signal type, and evaluation metrics differ substantially across studies, and no row to row comparison should be interpreted as establishing superiority in either direction.

Table 4. Machine learning approaches to Brugada syndrome detection in context.

Study	Data and signal	Model	Headline result
Melo et al., 2023 [6]	Raw waveform, no provocation	Deep neural network	88.4% accuracy, AUC 0.934
Zanchi et al., 2023 [7]	P wave morphology, n=123	AdaBoost	81% accuracy, 86% sensitivity, 73% specificity
Ronan et al., 2025 [8]	Raw waveform, self supervised pretraining	VICReg plus classifier	AUROC 0.88, AUPRC 0.82
This study	Interpretable per lead features, n=363	XGBoost	AUROC 0.90, 76% sensitivity at 90% specificity

An interpretable, feature based model applied to a newly released dataset achieves performance in the same range as more complex deep learning and self supervised approaches developed on other cohorts. This finding should be read as evidence that interpretable methods remain competitive for this task, not as a claim of superiority over any individual prior study.

4.2 Clinical Implications of the Error Analysis

The concentration of classification errors within the concealed phenotype is arguably the most clinically informative finding of this study. A classifier whose errors are distributed randomly would offer little guidance regarding where additional clinical scrutiny is warranted. A classifier whose errors concentrate specifically on patients with intermediate ST elevation, falling between the control and clearly positive distributions, mirrors the diagnostic uncertainty that motivates provocation testing in clinical practice. This represents a physiologically coherent failure mode rather than an arbitrary one, and it suggests that model confidence scores in this intermediate range may themselves carry clinical meaning worth further investigation.

4.3 Limitations

This study has several limitations that constrain the interpretation and generalizability of its findings.

First, all data originate from a single center using a single equipment configuration, and no external validation cohort was available. The reported performance metrics likely overstate accuracy under real world domain shift arising from differences in recording equipment, patient population, or lead placement convention. External validation against an independent Brugada syndrome cohort represents the most important direction for future work.

Second, the sample size is small (76 Brugada syndrome patients, of whom only 7 are type 2), which limits statistical power and widens the confidence intervals reported throughout this study. The type 2 sensitivity estimate (6 of 7) should be regarded as descriptive rather than as a reliable estimate of population performance.

Third, the 100 Hz sampling rate is relatively coarse for fine grained analysis of QRS and J point morphology. Several ST slope and notch related descriptors are likely quantized more severely than they would be at a higher clinical grade sampling rate.

Fourth, the basal pattern flag indicates that a subset of confirmed Brugada syndrome patients had a non-diagnostic resting electrocardiogram at the time of recording. This establishes an inherent ceiling on the sensitivity achievable by any classifier restricted to resting electrocardiogram data, independent of model quality.

Fifth, and most importantly, this classifier is a research baseline. It has not been externally validated, has not been calibrated for clinical deployment, and has not undergone regulatory review. It is not intended to inform clinical decision making in its current form.

5 Conclusion

This study presents an interpretable machine learning classifier for Brugada syndrome detection on the Brugada-HUCA dataset, achieving an AUROC of 0.90 and a sensitivity of 76% at 90% specificity. SHAP attribution confirms that the classifier relies on the clinically established right precordial ST and J point signature, and error analysis demonstrates that classification failures concentrate on the concealed phenotype that also challenges human electrocardiogram interpretation and motivates pharmacological provocation testing. Priority directions for future work include external validation on an independent Brugada syndrome cohort, comparison against raw waveform deep learning approaches on matched cross validation splits, focused investigation of the concealed phenotype subgroup incorporating outcome data such as sudden death history, and threshold selection informed by the clinical cost structure of missed diagnoses relative to false alarms.

Data and Code Availability

The underlying data are available from the Brugada-HUCA dataset, PhysioNet version 1.0.0 [11]. Analysis was performed in Python using wfdb, neurokit2, scikit-learn, XGBoost, and SHAP. Feature tables, cross validation metrics, effect size statistics, and per patient error analysis are provided as supplementary files alongside this manuscript, and the complete analysis codebase is available in an accompanying repository.

References

1. Krahn AD, Behr ER, Hamilton R, Probst V, Laksman Z, Han HC. Brugada Syndrome. *JACC Clin Electrophysiol.* 2022;8(3):386 to 405. https://doi.org/10.1016/j.jacep.2021.12.001

2. Kapplinger JD, Tester DJ, Alders M, et al. An international compendium of mutations in the SCN5A encoded cardiac sodium channel in patients referred for Brugada syndrome genetic testing. *Heart Rhythm.* 2010;7(1):33 to 46. https://doi.org/10.1016/j.hrthm.2009.09.069

3. Papadakis M, Papatheodorou E, Mellor G, et al. The Diagnostic Yield of Brugada Syndrome After Sudden Death With Normal Autopsy. *J Am Coll Cardiol.* 2018;71(11):1204 to 1214. https://doi.org/10.1016/j.jacc.2018.01.031

4. Amir M, Kabo P, Rahma I. Provocative testing using low dose oral flecainide for diagnosis of Brugada syndrome: a report of two cases. *Eur Heart J Case Rep.* 2022;6(12):ytac460. https://doi.org/10.1093/ehjcr/ytac460

5. Tisdale JE, Chung MK, Campbell KB, et al. Drug Induced Arrhythmias: A Scientific Statement From the American Heart Association. *Circulation.* 2020;142(15):e214 to e233. https://doi.org/10.1161/CIR.000000000000

6. Melo L, Ciconte G, Christy A, Vicedomini G, Anastasia L, Pappone C, Grant E. Deep learning unmasks the ECG signature of Brugada syndrome. *PNAS Nexus.* 2023;2(11):pgad327. https://doi.org/10.1093/pnasnexus/pgad327

7. Zanchi B, Faraci FD, Gharaviri A, Bergonti M, Monga T, Auricchio A, Conte G. Identification of Brugada syndrome based on P wave features: an artificial intelligence based approach. *Europace.* 2023;25(11):euad334. https://doi.org/10.1093/europace/euad334

8. Ronan R, Tarabanis C, Chinitz L, Jankelson L. Self supervised VICReg pre-training for Brugada ECG detection. *Sci Rep.* 2025;15(1):9396. https://doi.org/10.1038/s41598-025-94130-x

9. Sahebnasagh M, Farjoo MH. Artificial intelligence for Brugada syndrome diagnosis and gene variants interpretation. *Am J Cardiovasc Dis.* 2025;15(1):1 to 12. https://doi.org/10.62347/YQHQ1079

10. Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PC, Mark RG, Mietus JE, Moody GB, Peng CK, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation.* 2000;101(23):e215 to e220. https://doi.org/10.1161/01.cir.101.23.e215

11. Brugada-HUCA: 12-lead ECG dataset for Brugada syndrome detection. PhysioNet v1.0.0 (February 2026). Hospital Universitario Central de Asturias.