Interpretable ECG-Based Classification of Accessory Pathway Laterality in Pediatric Wolff–Parkinson–White Using Cardiac Digital Twins

Sandra Perez-Herrero[1], Guadalupe García-Isla[1], Ferran Prats-Domenech[2], Ferran Roses-Noguer[2], Beatriz Trenor[1], Javier Saiz[1], Jorge Sánchez[1]

[1] Center for Research and Innovation in Bioengineering (Ci2B), Universitat Politècnica de València, Valencia, Spain

[2] Pediatric cardiology department, Hospital Universitario Vall d'Hebron, Barcelona, Spain

Abstract

Accessory pathway (AP) localization from the 12-lead ECG is clinically relevant in pediatric Wolff-Parkinson-White (WPW) syndrome, but remains challenging due to subjective interpretation and limited labeled datasets. We propose an interpretable, simulation-driven framework based on decision trees trained on automatically extracted ECG features from 696 simulated electrocardiogram signals generated across 174 AP locations under four pre-excitation timings. The model classified left-versus-right ventricular activation with high performance in simulated data (balanced accuracy 0.968, macro-F1 0.956), evaluated using cross-validation. On an independent clinical dataset (n=10, ages 5–14), the model achieved a balanced accuracy of 0.708 and macro-F1 of 0.697. These results indicate that ventricular laterality can be captured using a small set of quantitative ECG descriptors, supporting reproducible non-invasive pre-procedural assessment in pediatric WPW.

1. Introduction

Surface-ECG localization of accessory pathways (APs) in Wolff-Parkinson-White (WPW) syndrome remains clinically relevant, as it can guide pre-procedural planning, catheter access strategy, and identification of challenging septal locations associated with higher risk of conduction system injury [1]. These aspects are particularly important in pediatric populations, where minimizing procedure duration, fluoroscopy exposure, and mapping burden is critical. However, accurate AP localization from the standard 12-lead ECG remains challenging in children, as most algorithms were developed in adult or mixed populations and may not capture age-dependent differences in cardiac anatomy, ventricular size, conduction properties, and pre-excitation [2, 3].

ECG-based approaches are generally more reliable for coarse tasks such as laterality than for exact AP localization, with pediatric studies reporting modest accuracy for exact localization but more robust left-versus-right discrimination [3]. Although approaches such as EASY-WPW, SMART-WPW, and deep-learning models have improved performance in mixed populations [1, 4, 5], most rely on feature definitions derived from subjective clinical interpretation and are trained on limited datasets with uncertain ground-truth labels. Consequently, decision rules are often based on small cohorts, and true AP location is only confirmed invasively during ablation.

Computational simulations provide an alternative by enabling the generation of large, fully labeled datasets with known AP locations. Virtual-heart studies have shown that ECG morphology depends not only on AP location but also on anatomical and activation context, explaining the limitations of rule-based algorithms [6]. This framework enables systematic exploration of ECG-location relationships under controlled conditions and supports the development of data-driven and interpretable models.

In this study, we propose an interpretable, simulation-driven framework for pediatric AP localization based on decision trees trained on automatically extracted ECG features. Targets are defined using anatomical ground truth rather than clinical labels, aiming to reduce subjectivity, improve reproducibility, and provide clinically meaningful and interpretable decision rules.

2. Methods

2.1. Anatomical Model

A 3D anatomical model was reconstructed from computed tomography (CT) images of a healthy 15-year-old subject, including both ventricular geometry and torso. The ventricular anatomy was labeled using universal ventricular coordinates (UVC) [7], enabling a reproducible and anatomically consistent definition of spatial regions.

In addition, myocardial fiber orientation was assigned according to Streeter et al. [8], allowing the incorporation of anisotropic conduction properties into the model. Myocardial conduction velocities were set to 570 mm/s in the longitudinal direction, 239 mm/s in the transverse direction, and 2000 mm/s in the cardiac conduction system [6].

The cardiac conduction system was generated using a fully connected network obtained via the fractal growth algorithm proposed by Costabal et al. [9]. The network was divided into five functional regions to enable controlled physiological activation. One candidate His-bundle insertion point was defined within each region, and an optimization procedure was performed to identify the combination of insertion sites that minimized the error between simulated and clinical ECG signals, allowing the reproduction of physiological sinus rhythm.

2.2. Simulation framework and APs dataset generation

Electrical propagation was simulated using the Eikonal model combined with the cellular model proposed by Mitchell & Schaeffer [10] in openCARP [11].

After the sinus rhythm calibration, APs were introduced as ectopic stimuli applied at different locations along the atrioventricular ring (Fig. 1). A total of 174 distinct AP insertion sites were defined across both ventricles. For each location, four pre-excitation timings (50, 70, 90, and 110 ms) were simulated to account for inter-patient variability in the degree of ventricular pre-excitation. Twelve-lead ECG signals were subsequently computed using the lead-field method [12]. This resulted in a dataset of 696 simulated ECGs (174 locations $\times$ 4 pre-excitation timings, 120 left and 576 right), each associated with a known activation origin.

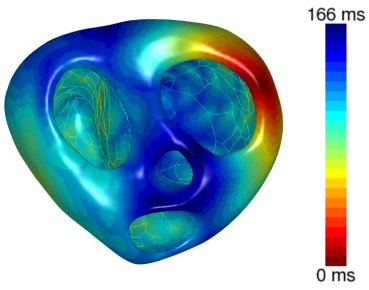

Figure 1. Local activation time map of ventricular activation during left lateral accessory pathway simulation.

Unlike clinical datasets, where AP location is inferred indirectly from signal interpretation, ground-truth labels were defined directly from the simulation. Ventricular chamber was determined from a UVC-derived coordinate encoding left-versus-right ventricular location, enabling objective classification and removing reliance on subjective clinical labeling.

2.3. ECG Signal Preprocessing and Feature Extraction

Simulated ECG signals underwent standardized preprocessing prior to feature extraction. The 12-lead QRS signals (350 ms) were resampled to 400 Hz and globally normalized to preserve relative amplitudes across leads. Baseline correction was then applied per lead using the median pre-QRS amplitude, followed by automatic QRS onset detection using a multi-lead energy-based criterion, as illustrated in Fig. 2.

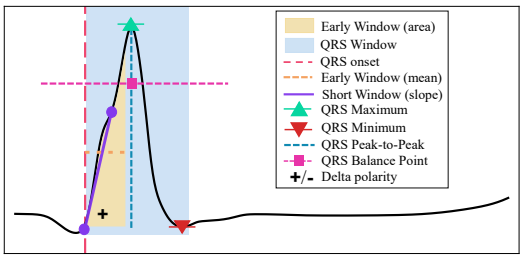

Figure 2. Illustrative ECG showing extracted features.

Features were extracted independently for each lead using windows defined as percentages of the post-onset signal. Early activation was characterized by delta-wave polarity and short-segment slope within the first 4% of the post-onset signal. Delta-wave polarity across all leads was also used to derive summary features for inferior leads (II, III, aVF), including both the number and specific identities of affected leads by negative delta polarity. Early QRS descriptors (first 8% of the post-onset signal) included integrated area and mean amplitude, while additional features over the first 18% of the post-onset signal captured maximum and minimum amplitudes, peak-to-peak amplitude, and a balance metric, computed as QRS maximum plus QRS minimum, to capture waveform asymmetry. The same pre-processing and feature extraction pipeline was applied to clinical ECGs following signal digitization.

2.4. Decision-tree Classification Models

The classification model was trained to identify the origin of AP activation. We focus on a model that discriminates between left and right ventricular activation (chamber classification). Training was performed using grouped 5-fold cross-validation, where all four pre-excitation timings from the same AP node were assigned to the same fold, ensuring no data leakage between training and validation sets with an 80/20 train-validation split in each fold. Decision trees were trained using the Gini impurity cri-

terion, with balanced class weighting to account for the skewed class distribution (120 left vs. 576 right), a maximum depth of 4, and a minimum of 5 samples per leaf to limit overfitting. Model performance was evaluated using accuracy, balanced accuracy, precision, recall, and macro F1-score, aggregated across all folds over the full dataset.

3. Results

3.1. Accessory Pathway Chamber Localization

The decision-tree model achieved high performance in discriminating left- and right-ventricular activation (accuracy 0.974, balanced accuracy 0.968, macro F1-score 0.956), with balanced accuracy confirming robustness to class imbalance (Table 1). Class-wise performance was slightly higher for right ventricular activation (precision 0.991, recall 0.977) than for left ventricular cases (precision 0.898, recall 0.958). Only 18 cases were misclassified overall (5 left-to-right, 13 right-to-left).

The resulting decision tree was compact and interpretable, relying on a reduced subset of features. The primary splitting variable was based on QRS balance in lead V1, followed by V2 short-window slope and lead III QRS balance, yielding simple and consistent decision rules (Fig. 3). Despite the high-dimensional feature space, classification was achieved using a small number of features and a tree of limited depth.

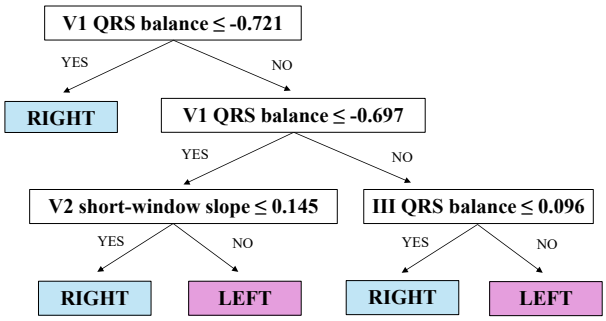

Figure 3. Decision tree for chamber classification.

3.2. Evaluation on Clinical Dataset

To assess real-world generalization, the model was evaluated on an independent clinical cohort of 10 pediatric WPW ECGs (ages 5–14 years) with ablation-confirmed accessory pathway location as the reference standard. On this external dataset, the model achieved an accuracy of 0.700, balanced accuracy of 0.708, and macro F1-score of 0.697, providing initial evidence that the framework can transfer from simulated data to clinically acquired ECGs. Performance varied across AP sides: left-sided pathways showed higher precision (0.800 vs. 0.600), whereas right-sided pathways showed higher recall (0.750 vs. 0.667).

Table 1. ECG chamber classifier performance.

Metric	ECG chamber classifier (simulated signals)	ECG chamber classifier (clinical signals)
Accuracy	0.974	0.700
Balanced acc.	0.968	0.708
Macro-F1	0.956	0.697
Precision (left)	0.898	0.800
Precision (right)	0.991	0.600
Recall (left)	0.958	0.667
Recall (right)	0.977	0.750

4. Discussion

The proposed decision-tree framework shows that left-versus-right ventricular activation can be accurately discriminated using a compact and interpretable set of ECG-derived features. In the simulated dataset, the model achieved strong performance (accuracy 0.974, balanced accuracy 0.968) despite its limited depth, indicating that ventricular laterality is captured by a small number of quantitative descriptors extracted from the ECG.

QRS balance, particularly in lead V1, emerged as the dominant discriminator, with additional contributions from V2 short-window slope and lead III QRS balance. This pattern is consistent with prior ECG-based AP localization studies highlighting lead V1, V2, and the inferior leads as informative for AP location and laterality [1, 4, 13]. In our model, V1 captured the strongest left-versus-right differences, V2 short-window slope reflected early QRS morphology, and lead III provided complementary inferior-axis information. Computational studies further support that AP location and pre-excitation timing shape ECG morphology across precordial and inferior leads [6]. Importantly, these relationships were learned directly from the simulated data, yielding simple and interpretable decision rules linked to the site of earliest activation.

Classical ECG-based algorithms rely on manual annotations of features such as delta-wave polarity and R/S ratios, introducing variability, especially in pediatric ECGs [1, 3, 4]. In contrast, the proposed approach replaces visual assessment with automatically extracted quantitative , enabling more reproducible and consistent characterization of ECG patterns.

The QRS balance can be interpreted as a continuous analog of the classical R/S ratio used in rule-based algorithms. As the R/S relationship reflects the global direction of ventricular depolarization and thus the activation

origin [13], the QRS balance generalizes this concept by integrating information over a broader portion of the QRS complex, providing a more robust descriptor.

When evaluated on clinical ECGs, performance decreased (accuracy 0.700, balanced accuracy 0.708). This reflects variability not captured in simulations, including noise, electrode placement, anatomical differences, and pre-excitation variability. Despite this, the model showed preliminary generalization on clinical ECGs, suggesting simulation-derived features remain relevant in clinical conditions, although validation on larger cohorts is required.

A major strength of this work is the use of simulations to generate a dataset with ground truth, enabling interpretable model development without uncertainty from clinically inferred labels. This combination of simulation-driven learning and interpretability bridges mechanistic understanding and data-driven approaches. Clinically, ventricular laterality is a first step toward AP localization, although finer discrimination, particularly of septal pathways, is required for procedural planning.

Limitations include the use of a single anatomical model, which may limit generalization, the absence of a formal analysis of feature importance and stability, and the small clinical dataset. Future work will address finer regional localization, improved feature extraction incorporating temporal and spatial information, and expanding simulated and clinical datasets to improve generalization and robustness.

5. Conclusions

This study presents an interpretable, simulation-driven framework for pediatric WPW laterality classification using automatically extracted ECG features. The model performed strongly on simulated data and showed moderate generalization to clinical ECGs, indicating that ventricular laterality is captured by a small set of physiologically meaningful descriptors. Overall, these results support simulation-based approaches as a reproducible and interpretable alternative to subjective clinical ECG assessment.

Acknowledgments

This work was partially supported by the Dirección General de Política Científica de la Generalitat Valenciana (CIPROM/2023/14) and CIAPOS/2024/033 co-funded by European Social Fund Plus (ESF+).

References

[1] Khalaph M, Trajkovska N, Didenko M, Braun M, Imnadze G, Akkaya E, et al. A novel ecg algorithm for accurate localization of manifest accessory pathways in both children and adults: Smart-wpw. Heart Rhythm 2025;.

[2] Wren C, Vogel M, Lord S, Abrams D, Bourke J, Rees P, et al. Accuracy of algorithms to predict accessory pathway location in children with wolff–parkinson–white syndrome. Heart 2012;98(3):202–206.

[3] Kurath-Koller S, Manninger M, Öffl N, Köstenberger M, Sallmon H, Will J, et al. Accuracy of algorithms predicting accessory pathway localization in pediatric patients with wolff-parkinson-white syndrome. Children 2022; 9(12):1962.

[4] El Hamriti M, Braun M, Molatta S, Imnadze G, Khalaph M, Lucas P, et al. Easy-wpw: a novel ecg-algorithm for easy and reliable localization of manifest accessory pathways in children and adults. Europace 2023;25(2):600–609.

[5] Hennecken J, Arends BK, Mast T, Dekker L, van Der Harst P, Blaauw Y, et al. Localization of accessory pathways in wolff-parkinson-white syndrome using ecg-based multitask deep learning. European Journal of Clinical Investigation 2025;55:e14385.

[6] Gillette K, Winkler B, Kurath-Koller S, Scherr D, Vigmond EJ, Bär M, et al. A computational study on the influence of antegrade accessory pathway location on the 12-lead electrocardiogram in wolff–parkinson–white syndrome. Europace 2025;27(2):euae223.

[7] Sánchez J, Llorente-Lipe I, Espinosa CB, Loewe A, Hernández-Romero I, Vicente-Puig J, et al. Enhancing premature ventricular contraction localization through electrocardiographic imaging and cardiac digital twins. Computers in Biology and Medicine 2025;190:109994.

[8] Streeter Jr DD, Spotnitz HM, Patel DP, Ross Jr J, Sonnenblick EH. Fiber orientation in the canine left ventricle during diastole and systole. Circulation research 1969; 24(3):339–347.

[9] Costabal FS, Hurtado DE, Kuhl E. Generating purkinje networks in the human heart. Journal of biomechanics 2016; 49(12):2455–2465.

[10] Mitchell CC, Schaeffer DG. A two-current model for the dynamics of cardiac membrane. Bulletin of mathematical biology 2003;65(5):767–793.

[11] Plank G, Loewe A, Neic A, Augustin C, Huang YL, Gsell MA, et al. The opencarp simulation environment for cardiac electrophysiology. Computer methods and Programs in Biomedicine 2021;208:106223.

[12] Potse M. Scalable and accurate ecg simulation for reaction-diffusion models of the human heart. Frontiers in physiology 2018;9:370.

[13] Taguchi N, Yoshida N, Inden Y, Yamamoto T, Miyata S, Fujita M, et al. A simple algorithm for localizing accessory pathways in patients with wolff-parkinson-white syndrome using only the r/s ratio. Journal of Arrhythmia 2014; 30(6):439–443.

Address for correspondence:

Sandra Perez-Herrero
Center for Research and Innovation in Bioengineering (Ci2B), Universitat Politècnica de València, Camino de Vera s/n, 46022 Valencia, Spain
sanpeher@ci2b.upv.es