Ensemble Learning on Body Surface Potential Mapping P-Waves: A Novel Approach to Predict Atrial Fibrillation Recurrence Post-Direct-Current Cardioversion

Fan Feng[1], Ekenedirichukwu N Obianom[2], Noor Qaqos[1], Ibrahim Antoun [1,4], Shamsu Idris Abdullahi[1], Abdulhamed M Jasim[1], G André Ng[2,3,4], Xin Li[1,3]

[1]University of Leicester, School of Engineering, Leicester, UK
[2]University of Leicester, Department of Cardiovascular Sciences, Leicester, UK
[3]National Institute for Health Research Leicester, UK
[4]Department of Cardiology, University Hospitals of Leicester NHS trust, Leicester, UK

Abstract

Background and Objective: *Atrial fibrillation (AF) frequently recurs after direct-current cardioversion (DCCV), and 12 lead ECG lacks spatial resolution. This study explored whether body surface potential mapping (BSPM) with ensemble learning improves recurrence prediction.*

Methods: *Fifty-two DCCV patients (29 with recurrence) were followed for 12 months. From 131-lead BSPM, 50 P-Wave features per subject were extracted across time, frequency, and wavelet domains. Seven classifiers, including tree-based models and a neural network, were evaluated. Ensemble strategies (voting, stacking) used fivefold cross-validation.*

Results: *Stacking achieved the best performance (AUC 0.795, F1 0.764), outperforming all base models. Key predictors included spectral and wavelet features like power_beta and wavelet entropy.*

Conclusion: *BSPM derived features combined with stacking improve AF recurrence prediction, supporting its use for personalized risk assessment after DCCV.*

1. Introduction

Atrial fibrillation (AF) is the most common sustained arrhythmia worldwide, affecting over 46 million people, and its prevalence increases with age [1].

Direct current cardioversion (DCCV) is widely used to restore sinus rhythm in persistent AF, but recurrence remains common, approaching 50% within a year and even higher in specific subgroups [2,3]. Recurrence is associated with increased risk of stroke, hospitalizations, and healthcare burden [4].

The 12-lead electrocardiogram (ECG) is used to extract P-Wave features such as duration, dispersion, and amplitude to predict recurrence. For example, prolonged filtered P-Wave duration (>137 ms) and reduced RMS voltage in the final 20 ms are reported as independent predictors [5].

However, the limited spatial resolution of 12-lead ECG impairs its ability to detect subtle atrial conduction abnormalities beyond its lead field [1].

Imaging modalities like echocardiography and cardiac MRI provide structural markers, e.g., left atrial volume index (LAVI), sphericity, and diastolic dysfunction [6,7], but are constrained by cost, equipment, and operator dependence, limiting routine or repeated use.

Body surface potential mapping (BSPM), using over 100 thoracic leads, enables high resolution detection of heterogeneous atrial depolarization not visible on standard ECG [8]. It allows torso wide P-Wave analysis, revealing spatial conduction dispersion linked to atrial remodeling.

Machine learning (ML) offers promise in ECG based risk prediction. While some ML models using single- or 12-lead ECG have been applied to recurrence prediction, they often suffer from limited features and generalizability [9].

Ensemble methods like Random Forests, XGBoost, and stacking provide better robustness and performance, particularly for high dimensional BSPM data [1,10].

This study integrates BSPM derived features with ensemble ML to develop a non-invasive, scalable model for AF recurrence prediction after DCCV, aiming to enable early and individualized rhythm risk stratification.

2. Materials and Method

A step-by-step overview of the data pipeline, including BSPM acquisition, P-Wave extraction, feature engineering, model training, and evaluation.

2.1. Dataset

This study prospectively enrolled 52 patients undergoing direct-current cardioversion (DCCV) for atrial fibrillation (AF) at a single tertiary center. Inclusion required successful cardioversion, complete BSPM recordings, and documented recurrence status over 12 months. 29 patients experienced AF recurrence; 23 maintained sinus rhythm.

Patients with structural heart disease, significant valvular abnormalities, poor ECG quality, or incomplete lead placement were excluded.



Figure 1. Workflow of the proposed prediction

Electrophysiological data were collected using a 131-lead BSPM system spanning the anterior and posterior thorax. Signals were recorded within 10 seconds post DCCV, a period capturing atrial stabilization, sampled at 1000 Hz with 16-bit resolution and bandpass filtered (0.05–150 Hz). Compared to 12-lead ECG, this setup offers superior spatial resolution.

Basic demographics and clinical metadata were also collected. Analysis focused on multi-lead P-Wave features linked to recurrence.

2.2. P-Wave Detection and Segmentation

After preprocessing, the lead with clearest atrial activity was manually selected per patient to serve as reference for P-Wave onset. An automatic onset detection algorithm based on slope and energy identified candidate points, which were manually refined. Final onset markers were propagated to all 131 leads for temporal alignment.
Fixed length windows (~120–140 ms before and after onset) were extracted around each onset to ensure consistent multi-lead P-Wave segmentation across beats and leads.

2.3. Feature Extraction

P-Wave features were extracted from each lead across three domains:

Time domain: Metrics included peak to peak amplitude, signal energy, inter-lead variability, duration, and skewness, reflecting conduction dispersion.

Frequency domain: Fast Fourier Transform (FFT) yielded power spectral density and low/high frequency energy ratios, indicating conduction heterogeneity.

Time frequency domain: Discrete Wavelet Transform (DWT) with Daubechies wavelets was used. Energy, entropy, and kurtosis of wavelet coefficients described transient signal irregularities.

Over 500 features were computed per patient. After standardization, mutual information ranking was applied. The top 50 features primarily from amplitude, spectral, and wavelet domains were used for modeling.

2.4. Machine Learning Models

Several classifiers were implemented to evaluate predictive performance:

Tree based ensembles: Random Forest, Gradient Boosting, HistGradientBoosting, XGBoost, LightGBM, and CatBoost.

These models capture non linear patterns and offer robustness to overfitting.

Variants differ in efficiency, regularization, and handling of categorical features.

Neural Network: A shallow feedforward network was used as a non-tree baseline.

Voting Ensemble: Combines predictions via majority rule.

Stacked Ensemble: A logistic regression meta-learner was trained on base model outputs to leverage their complementarity.

All models were trained on the same 50 features. Hyperparameter tuning used nested cross-validation. Class imbalance was addressed using weighted loss functions where applicable.

3. Results

3.1. Overall Performance of Classifiers

Figure 2 compares model performance across AUC, F1 score, sensitivity, and specificity. The stacking ensemble achieved the highest AUC (0.841), followed by Random Forest (0.826) and CatBoost (0.814). LightGBM and XGBoost reached 0.802 and 0.797, respectively.

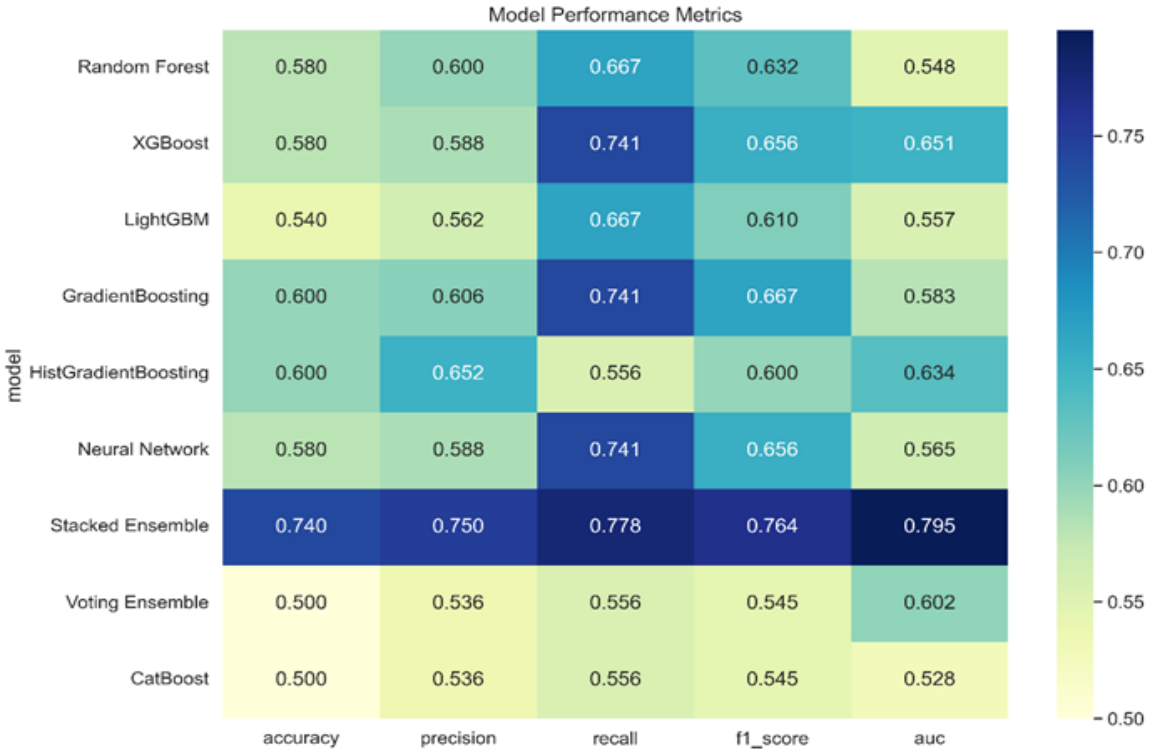


Figure 2. Performance comparison of classifiers across multiple metrics

Stacking also yielded the highest F1 score (0.779), outperforming all base classifiers. Neuralnetworks performed worst, with AUC 0.721 and F1 score 0.663. Sensitivity ranged from 0.688 (NN) to 0.789 (CatBoost), while stacking achieved the highest specificity (0.842).

3.2. Consensus Feature Occurrence

A total of 36 features appeared in at least one model's top 15 list. Among these, 12 features occurred in three or more classifiers. Figure 3 shows that the most frequent feature appeared in all five models, while several others were common to four models. These high frequency features are candidates for consensus-based biomarkers.

4. Discussion and Conclusion

This study demonstrated that body surface potential mapping (BSPM), when combined with machine learning, can effectively predict atrial fibrillation (AF) recurrence after direct current cardioversion (DCCV). Among the evaluated models, the stacking ensemble achieved the best performance across AUC, F1 score, sensitivity, and specificity, likely due to its ability to integrate complementary patterns from diverse base classifiers.

Feature analysis revealed consistent top predictors across models, particularly spectral and nonlinear features such as *power_beta*, *wavelet entropy*, and *ratio_alpha_theta*. These features appeared in four or five classifiers' top lists (Figure 3A), suggesting they capture robust and physiologically meaningful aspects of atrial conduction. This supports prior evidence that atrial remodeling is reflected in altered spectral and entropy characteristics rather than simple morphology.

Although some features overlapped, individual models showed distinct preferences: Random Forest emphasized amplitude-based metrics, while boosting models prioritized frequency and complexity features. This variability highlights the benefit of ensemble methods in leveraging diverse interpretations of the same signal set.

In conclusion, BSPM derived P-Wave features combined with ensemble learning, especially stacking, enable non-invasive, individualized prediction of AF recurrence. These results underscore the value of advanced signal analytics in arrhythmia risk stratification. Further validation in larger cohorts is needed to assess generalizability and support future clinical integration.

Acknowledgments

GAN and XL is supported by British Heart Foundation Research Excellence Award (RE/24/130031), Medical Research Council Biomedical Catalyst Developmental Pathway Funding Scheme (MR/S037306/1) and NIHR i4i grant

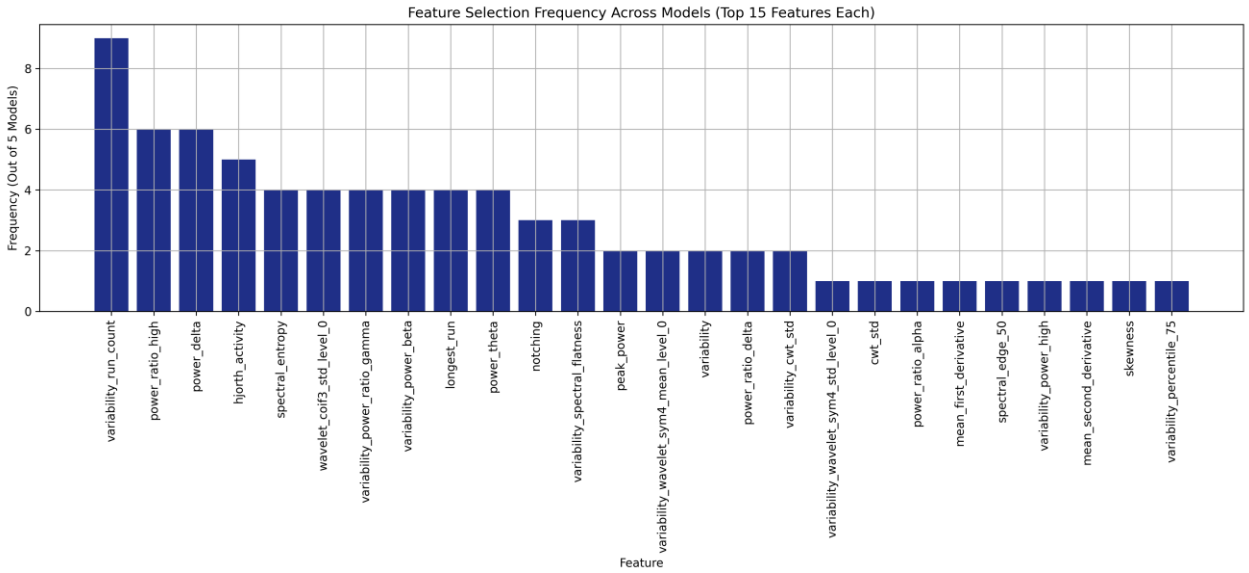


Figure 3. Frequency of top features across different classifiers

(NIHR204553). GAN is supported also supported by British Heart Foundation Programme Grant (RG/17/3/32774). This work was supported by a tuition fee waiver from the University of Leicester.

References

[1] Antoun M, Layton B. "Predicting the Outcomes of External Direct-Current Cardioversion for Atrial Fibrillation: A Narrative Review of Current Evidence." *Journal of Cardiovascular Development and Disease*, vol. 12, no. 5, 2025.

[2] Gurevitz O, Varadachari CJ, Ammash N, Malouf J, Rosales A, Herges R, et al. "The Effect of Patient Sex on Recurrence of Atrial Fibrillation Following Successful Direct-Current Cardioversion." *American Heart Journal*, vol. 152, no. 1, pp. 155.e9–13, 2006.

[3] Ulus T, Aliyev I, Arslan S, Çilingir O, Çolak E. "Genetic Pol ymorphism on Chromosome 4q25 (rs17570669) May Predict Recurrence After Successful Electrical Cardioversion in Patients with Persistent Atrial Fibrillation." *Archives of the Turkish Society of Cardiology*, vol. 51, no. 8, pp. 550–6, 2023.

[4] Melduni R, Cullen MW. "Role of Left Ventricular Diastolic Dysfunction in Predicting Atrial Fibrillation Recurrence After Successful Electrical Cardioversion." *Journal of Atrial Fibrillation*, vol. 5, no. 4, pp. 87–94, 2012.

[5] Opolski G, Scisło P, Stanisławska J, Górecki A, Steckiewicz R, Torbicki A. "Detection of Patients at Risk for Recurrence of Atrial Fibrillation After Successful Electrical Cardioversion by Signal-Averaged P-Wave ECG." *International Journal of Cardiology*, vol. 60, no. 2, pp. 181–5, 1997.

[6] Raniga D, Goda M, Hattingh L, Thorning S, Rowe M, Howes L. "Left Atrial Volume Index: A Predictor of Atrial Fibrillation Recurrence Following Direct-Current Cardioversion – A Systematic Review and Meta-Analysis." *International Journal of Cardiology: Heart & Vasculature*, vol. 51, pp. 101364, 2024.

[7] Osmanagic A, Möller S, Osmanagic A, Sheta H, Vinther KH, Egstrup K. "Left Atrial Sphericity Index Predicts Early Recurrence of Atrial Fibrillation After Direct-Current Cardioversion: An Echocardiographic Study." *Clinical Cardiology*, vol. 39, 2016.

[8] Sanna T, Sonaglioni A, Pieroni M, Russo A, Pelargonio G, Casella M, et al. "Baseline NT-Pro-BNP Levels and Arrhythmia Recurrence in Outpatients Undergoing Elective Cardioversion of Persistent Atrial Fibrillation: A Survival Analysis." *Indian Pacing and Electrophysiology Journal*, vol. 9, pp. 15–24, 2009.

[9] Chong JJH, Kumar S, Thomas L, Thomas SP. "Supraventric ular Ectopy and Recurrence of Atrial Fibrillation After Electrical Cardioversion." *Europace*, vol. 8, no. 5, pp. 341–4, 2006.

[10] Gurciková M, Kalužay J, Remišová S, Jurkovicova O, Ponťuch P. "Late Recurrences of Atrial Fibrillation in Patients After Direct-Current Cardioversion." *Vnitřní Lékařství*, vol. 54, no. 6, pp. 604–8, 2008.

Address for correspondence:

Fan Feng, Xin Li
University of Leicester, University Road, Leicester, United Kingdom. LE1 7RH.
ff98@leicester.ac.uk
xin.li@leicester.ac.uk