Insights from Explainable Machine Learning on Biatrial Arrhythmia Vulnerability Assessment

Patricia Martínez Díaz[1], Pascal Maierhofer[1], Eric Invers-Rubio[2], Michael Beigl[3], Lluís Mont[2], Amir Jadidi[4], Olaf Dössel[1], Axel Loewe[1]

[1]Institute of Biomedical Engineering, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany
[2]Unitat de Fibrillació Auricular (UFA), Institut Clínic Cardiovascular (ICCV), Hospital Clínic, Universitat de Barcelona, Barcelona, Spain
[3]TECO, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany
[4]Department of Rhythmology-Electrophysiology, Lucerne Cantonal Hospital, Lucerne, Switzerland

Abstract

Despite recent advancements in comprehending atrial fibrillation pathophysiology, identifying individuals prone to developing arrhythmia remains difficult in clinical practice. Patient-specific computer models of the atria can aid in quantifying reentry inducibility in silico. We generated a cohort of 22 biatrial computer models and personalized patient-specific anatomy and substrate based on LGE-MRI data. We assessed arrhythmia vulnerability with a virtual S1S2 pacing protocol from stimulating points in both atria. Clinical features associated with arrhythmia propensity were extracted and used to train a random forest classifier to identify points inducing reentry. Features were assessed globally, and fibrotic features were further assessed locally. A total of 1438 reentries were induced. The random forest classifier achieved an area under the receiver operating characteristic curve of 0.74± 0.02. SHAP explainability was utilized to understand the interaction between model features and their effect on arrhythmia vulnerability. Fibrosis density in 10 mm proximity and conduction velocity were the features showing the highest impact on point inducibility prediction. The presented classifier may be a fast alternative for assessing arrhythmia vulnerability in silico without expensive computations of virtual pacing protocols, thus aiding the transition to clinical applications.

1. Introduction

Atrial fibrillation (AF) is a highly prevalent and complex arrhythmia resulting from the interplay of structural and electrophysiological characteristics, which vary across patients [1]. Generally, a trigger and a vulnerable substrate are the main determinants responsible for initiating and maintaining AF [2]. Ectopic trigger activity originating in the sleeves of the pulmonary veins (PVs) has long been recognized to initiate arrhythmia in the atria. Regarding vulnerable substrate, fibrosis, dilation, and remodeling have been associated with an increased AF propensity [3]. Due to the complex nature of AF, the specific arrhythmia vulnerability for individual patients remains unclear.

Personalized atrial computer models can serve as a valuable research tool for unraveling the intricate mechanisms of AF in an individualized manner. *In silico* assessment of arrhythmia vulnerability can assist in quantifying the ease of inducibility of an atrial model based on patient-specific characteristics. Multiple studies have assessed vulnerability in atrial patient-specific models [4–9]. However, evaluation of arrhythmia vulnerability is computationally expensive and restricts application to centers with access to high-performance computing. In this work, we used patient-specific biatrial modeling and simulation to train a machine learning (ML) model for predicting point-wise inducibility. SHAP explainability was utilized to understand the interaction between patient-specific model features and their effect on the prediction of arrhythmia vulnerability.

2. Methods

2.1. Patient-specific Modeling

We generated 22 patient-specific biatrial models from LGE-MRI data from Hospital Clinic (Barcelona, Spain), and University Heart Center (Bad Krozingen, Germany). Patients provided written informed consent and the ethical committee of both institutions approved the study. Biatrial anatomies were segmented from the MRI blood pool using ADAS 3D software v2.12.0 (Adas3D Medical SL, Barcelona, Spain). Substrate regions were identified using image intensity ratio (IIR>1.2) [10]. Biatrial bilayer models were generated using AugmentA [11].

2.2. Atrial Electrophysiology Modeling

Cellular electrophysiology of human atrial myocytes was modeled using the Courtemanche *et al.* mathematical model [12]. Ionic channel conductances were modified to represent persistent AF [13]. Electrical propagation heterogeneity in the atria was modeled based on anisotropy ratios defined previously [14]. To model substrate in regions of high LGE intensity (IIR>1.2), 70 % of the elements were set to have electrical remodeling, while the rest of the elements had electrical decoupling with $\sigma = 10^{-7}$ S/m [15]. Electrical propagation in the atria was modeled using the monodomain in openCARP [16]. We ran simulations under three CV (conduction velocity) scenarios. Monodomain tissue conductivities were tuned to achieve a CV in the longitudinal direction of 0.5, 0.7, and 1.0 m/s in the bulk atrial myocardium.

Arrhythmia vulnerability was evaluated with a virtual S1S2 inducing protocol from a set of stimulation points evenly distributed on the biatrial surface with a 2 cm interpoint distance, with 51±15 stimulation points per biatrial model. Vulnerability was defined as the ratio between the number of inducing points to the total number of stimulation points. Binary point inducibility was defined as reentrant activity maintained for at least 1 s [17].

2.3. Calculation of Features

We selected features associated with arrhythmia propensity: anatomical factors such as atrial volume, surface, sphericity, and point location; fibrosis characteristics including burden, density and entropy, and the electrophysiological factor CV. Features were quantified globally, and fibrotic features were further assessed locally. Global features are shown in Table 1. Sphericity ψ describes the similarity of an object to a sphere and is defined as the ratio of the nominal surface area of a sphere S_n, having the same volume V as the object, to the actual surface area of the object S:

$$\Psi = \frac{S_n}{S} = \frac{\sqrt[3]{36\pi V^2}}{S}\,. \tag{1}$$

Table 1. Clinical global features of patient cohort

Feature	LA	RA
Volume (cm^3)	108.2 ± 49.5	100.2 ± 39.7
Sphericity (%)	82.9 ± 5.1	79.4 ± 5.3
Surface (cm^2)	129.0 ± 34.0	128.7 ± 30.7
Fibrosis burden (%)	10.7 ± 8.0	8.4 ± 7.5
Fibrosis density (%)	6.7 ± 5.7	6.1 ± 6.0
Fibrosis entropy (%)	5.2 ± 3.8	3.3 ± 2.3

LA: left atrium, RA: right atrium. Values are shown as mean ± SD.

Fibrosis burden B was obtained from the normalized IIR derived from LGE-MRI. According to Zahid et al. [4], border zones with high fibrosis density D and entropy E promote reentrant activity. To calculate fibrosis features in each node, we identified subsets within a radius of 1 mm, i.e. maximum edge length. We quantified D as the ratio between the number of fibrotic nodes N_f, divided by the total number of nodes N in the subset, and assigned that value to the centre node. In addition, local E for each node i in the subset was calculated based on Shannon entropy:

$$E = \sum_{i=1}^{N} \frac{-p_i \cdot ln\,(p_i)}{N}\,, \tag{2}$$

where p_i is the fraction of nodes in the subset with a node type (fibrotic or non-fibrotic) different to the node in the centre. An example of the calculation of local fibrotic features is shown in Figure 1.

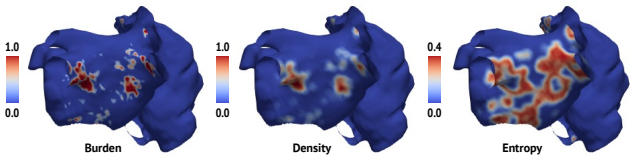

Figure 1. Biatrial fibrotic features for patient 1. Density and entropy were calculated for each node with a defined subset of radius equal to cell edge length (1 mm).

We established six reference regions around each stimulation point: three circles (conjunction) and three concentric rings (disjunction). The regions had radii of 5, 10, and 20 mm, representing near, mid, and far proximity. We calculated the mean E and B for each reference region. Finally, we included four positional features: the atrial chamber and three Laplace–Dirichlet-Rule-Based coordinates (ϕ_{ab}, ϕ_v, and ϕ_r) for each stimulation point [11].

2.4. Explainable Machine Learning

A random forest classifier (RF) was trained to identify stimulating points inducing reentry based on feature characteristics. The performance of different RF instances was evaluated using the area under the receiver operating characteristic curve (AUC ROC) metric and ten-fold cross-validation. The optimal hyperparameter setting was used to train the final RF classifier with an 80:20 training and test split. Data were split on a patient basis rather than individual points to avoid overfitting. To improve the performance of the RF algorithm, we evaluated sample normalization, modification of RF parameters, class weights and extraction of different feature subsets. For explainability, SHAP (SHapley Additive exPlanations) were selected to assess the impact of each feature on the prediction of inducibility as proposed before [8].

3. Results

A total of 1438 reentries were induced from 7131 stimulating points across the 22 biatrial models. The pacing protocol duration was 765±256.2 min per biatrial model, with each stimulation point lasting 15±2.4 min, utilizing 4 nodes×40 CPU cores (Intel Xeon Gold 6230 2.1GHz). The RF ML model was exposed to 5704 stimulating points with 27 features and required 0.6 s for training and 0.01 s for validation with 1427 stimulating points (Intel Core i5 3.1GHz). The RF classifier achieved an AUC of 0.74±0.02 (Figure 2). AUC with feature selection (top 5 features of SHAP from Figure 3) was 0.69±0.05, with balanced classes 0.72±0.05 and with sample normalization 0.66±0.05. SHAP identified the features showing the highest impact on the prediction of point inducibility (Figure 3). Among global features, CV had the greatest impact on inducibility prediction, whereas fibrosis density in the mid-proximity regions (10 mm) was the most influential local feature.

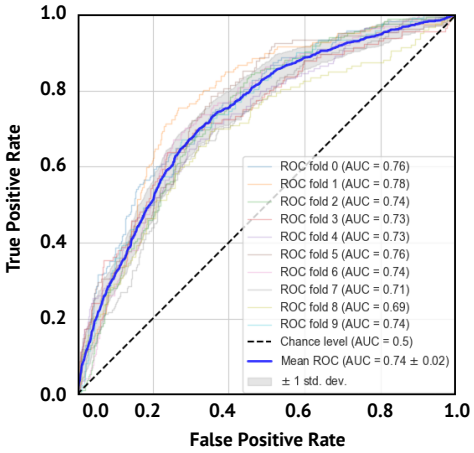

Figure 2. Receiver operating characteristic (ROC) of the random forest algorithm with ten-fold cross validation.

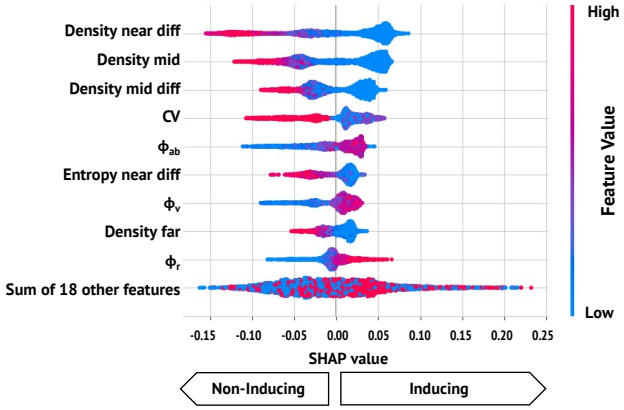

Figure 3. Beeswarm explainability plot of SHAP values showing the impact of each feature on point inducibility predictions. Dots represent features values for every stimulation point. Dots to the right of the midline indicate a predictive impact towards induction, while those to the left indicate a predictive impact towards non-induction. *CV: conduction velocity, conj: conjunction, disj: disjunction near, mid and far: 5, 10, and 20 mm.*

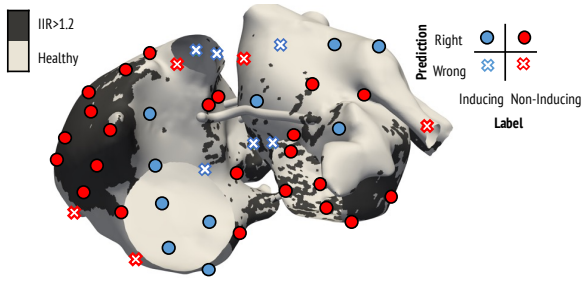

Figure 4. Analysis of model prediction, point inducibility, and fibrosis distribution. Color represents inducibility, with circles indicating correct predictions and crosses incorrect predictions. Areas of high intensity are shown in gray. Ground-truth inducibility from electrophysiological model.

4. Discussion

This study assessed arrhythmia vulnerability in 22 biatrial models with personalized anatomy and substrate informed by LGE-MRI. We trained and evaluated an ML RF classifier to predict point-wise inducibility with both global and local features. Modeling and simulation combined with ML algorithms can enhance mechanistic understanding of AF and provide digital solutions compatible with clinical timeframes. Previous studies demonstrated this potential by exploring the role of LA native fibrosis and post-ablation lesions on arrhythmia inducibility using explainable ML [8]. Our work is one of the first studies to assess feature impact on inducibility prediction in a biatrial cohort.

Fibrosis density at near and mid proximities had the highest impact on model inducibility prediction. Specifically, low fibrosis density in the surrounding 10 mm was associated with an increased inducibility. The study by Zahid *et al.* proposed that regions with higher fibrosis density aid reentrant activity [4]. Our results indicate that for inducibility, the areas surrounding the stimulation point should have a low fibrosis burden to facilitate impulse propagation. Among global features, CV had the highest impact on inducibility prediction. Previous studies have shown the impact of global CV on model vulnerability, however CV was not included as feature in their ML model [8]. A decrease in global CV increases vulnerability of biatrial models [9]. High CV was associated with re-

duced inducibility. In terms of positional features, ϕ_{ab} had the greatest influence on the model prediction. Stimulation points located closer to the interatrial septum (low ϕ_{ab}) were associated with increased inducibility, while those closer to the appendage and lateral walls were associated with non-inducibility. Interatrial connections may facilitate the formation of reentrant circuits between the LA and RA. Sphericity, LA volume, and RA volume showed minimal impact on the prediction. Our results demonstrate moderate prediction accuracy, as evidenced by AUC ROC. The ML classifier had a good performance in detecting true non-inducible points, as depicted in Figure 4. Our results showcase the complexity of point inducibility, suggesting that induction depends on the interplay of global and local factors, some of which might extend beyond the 20 mm proximity. We did not include other features associated with AF propensity, such as wall thickness, fat infiltration, and endo-epi dissociation. We did not incorporate personalized electrophysiology information. We utilized LGE-MRI data to get information on the substrate; however the discrepancy in the spatial distribution between low voltage areas and LGE-MRI is likely to influence vulnerability [1, 7]. To increase accuracy, expanding the sample size an incorporating further features could help to develop a computationally efficient and sufficiently accurate model. We showed that ML models offer a promising tool for predicting inducibility *in silico*, potentially reducing the need for expensive computations in virtual pacing protocols and bridging the gap for clinical application of patient-specific models with compatible clinical timings.

Acknowledgments

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 860974. This work was supported by the Leibniz Science-Campus "Digital Transformation of Research" with funds from the programme "Strategic Networking in the Leibniz Association". The authors acknowledge support by the state of Baden-Württemberg through bwHPC.

References

[1] Nairn D, Eichenlaub M, Müller-Edenborn B, et al. Differences in atrial substrate localization using late gadolinium enhancement-magnetic resonance imaging, electrogram voltage, and conduction velocity. Europace 2023; 25(9).

[2] Wijesurendra RS, Casadei B. Mechanisms of atrial fibrillation. Heart 2019;105(24):1860–1867.

[3] Goette A, Kalman JM, Aguinaga L, et al. EHRA/HRS/APHRS/SOLAECE expert consensus on atrial cardiomyopathies: definition, characterization, and clinical implication. Europace 10 2016;18(10):1455–1490.

[4] Zahid S, Cochet H, Boyle PM, et al. Patient-derived models link re-entrant driver localization in atrial fibrillation to fibrosis spatial pattern. Cardiovasc Res 2016;110:443–454.

[5] Boyle PM, Zghaib T, Zahid S, et al. Computationally guided personalized targeted ablation of persistent atrial fibrillation. Nat Biomed Eng 1 2019;3(11):870–879.

[6] Shade JK, Ali RL, Basile D, et al. Pre-procedure application of machine learning and mechanistic simulations predicts likelihood of paroxysmal atrial fibrillation recurrence following pulmonary vein isolation. Circ Arrhythm Electrophysiol 1 2020;.

[7] Azzolin L, Eichenlaub M, Nagel C, et al. Personalized ablation vs. conventional ablation strategies to terminate atrial fibrillation and prevent recurrence. Europace 2 2023; 25(1):211–222.

[8] Bifulco SF, Macheret F, Scott GD, et al. Explainable machine learning to predict anchored reentry substrate created by persistent atrial fibrillation ablation in computational models. J Am Heart Assoc 8 2023;12(16).

[9] Martínez Díaz P, Sánchez J, Fitzen N, et al. The right atrium affects in silico arrhythmia vulnerability in both atria. Heart Rhythm 1 2024;.

[10] Benito EM, Carlosena-Remirez A, Guasch E, et al. Left atrial fibrosis quantification by late gadolinium-enhanced magnetic resonance: a new method to standardize the thresholds for reproducibility. Europace 8 2017; 19(8):1272–1279.

[11] Azzolin L, Eichenlaub M, Nagel C, et al. Augmenta: Patient-specific augmented atrial model generation tool. Comput Med Imaging Graph 9 2023;108:102265–102265.

[12] Courtemanche M, Ramirez RJ, Nattel S. Ionic mechanisms underlying human atrial action potential properties: insights from a mathematical model. Am J Physiol Heart Circ Physiol 1998;275(1):H301–H321.

[13] Loewe A, Krueger MW, Platonov PG, et al. Left and right atrial contribution to the p-wave in realistic computational models. In Lecture Notes in Computer Science, volume 9126 of FIMH. 1 2015; 439–447.

[14] Krueger MW, Seemann G, Rhode K, et al. Personalization of atrial anatomy and electrophysiology as a basis for clinical modeling of radio-frequency ablation of atrial fibrillation. IEEE Trans Med Imaging 1 2013;32(1):73–84.

[15] Roney CH, Bayer JD, Zahid S, et al. Modelling methodology of atrial fibrosis affects rotor dynamics and electrograms. Europace 12 2016;18(suppl 4):iv146–iv155.

[16] Plank G, Loewe A, Neic A, et al. The openCARP simulation environment for cardiac electrophysiology. Comput Meth Progr Biomed 6 2021;208:106223.

[17] Azzolin L, Schuler S, Dössel O, et al. A reproducible protocol to assess arrhythmia vulnerability. Front Physiol 2021; 12:656411.

Address for correspondence:

Patricia Martínez Díaz, Institute of Biomedical Engineering
Karlsruhe Institute of Technology (KIT)
publications@ibt.kit.edu