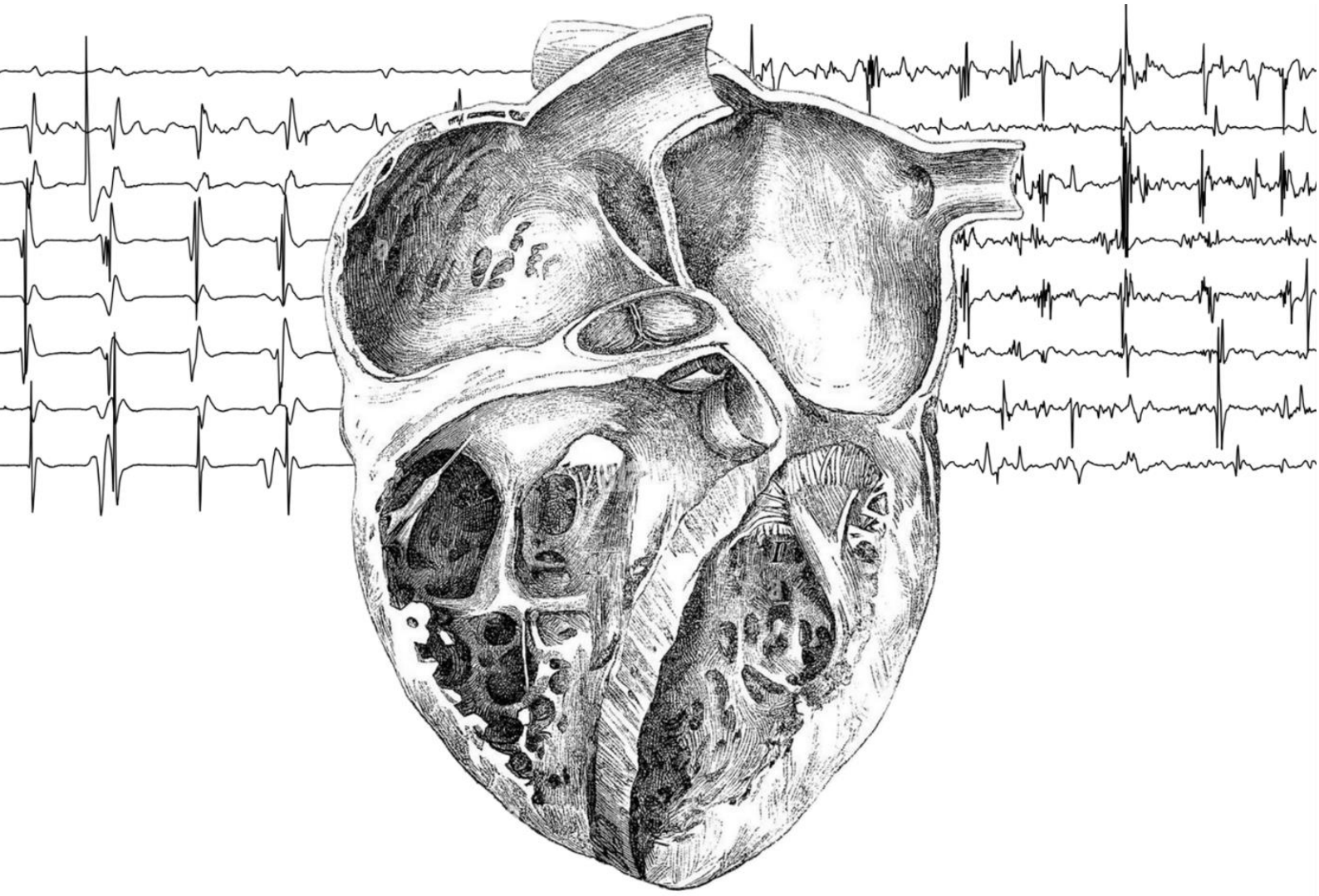
Optimalisation of Potential Spatio-temporal Dispersion-based Software for Identification of Drivers from Atrial Fibrillation

Stan Benjamins

Master Technical Medicine ~ Medical Sensing and Stimulation
Department of Cardiology, Isala Zwolle

March 13, 2025

UNIVERSITY OF TWENTE. | isala



ABSTRACT

Introduction: Pulmonary vein isolation (PVI) is the cornerstone intervention in atrial fibrillation (AF), particularly for patients with symptomatic, drug-refractory paroxysmal or (longstanding) persistent AF. Nevertheless, the success rate of PVI in patients with persistent AF ranges from 43% to 67%, resulting in many patients continuing to experience AF recurrences. Interestingly, 56% of patients undergoing redo-ablation do not show PV reconnection, indicating other non-PV contributing AF drivers within the atria. Recent studies have focused on targeting non-PV drivers, such as spatio-temporal dispersions (STDs), detected using intra-atrial electrograms (IAEs) during AF. VX1, an AI-based software, standardizes STD recognition to improve outcomes, though long-term effectiveness remains uncertain due to limited operator experience and short follow-ups.

Objective: Evaluate the potential of the novel VX1-software as a standardized approach for optimizing ablation strategies in patients with AF, compared to the patient-specific approach of the electrophysiologist.

Method: This thesis is based on the cohort of patients from a retrospective, single-center, observational registry at the Isala Hospital in Zwolle, the Netherlands, between May 2022 to December 2023. IAEs were labeled as STD-based or non-STD-based by both the electrophysiologist and the VX1-software. A total of 24 features were extracted to quantify IAE complexity in both temporal and spatial domains. Machine learning (ML) models, including principal component analysis (PCA) with K-Means clustering and extreme gradient boosting (XGB), were used to classify IAEs and evaluate their performance. In addition, patient-specific complexity scores were calculated, and discrepancies in intra-atrial areas identified by the VX1-software and electrophysiologist were assessed.

Results: A total of 52 patients and 110 left intra-atrial electroanatomic maps (EAMs) were analyzed, yielding 2792 STD-based IAEs by the VX1-software, 2516 by the electrophysiologist, and 11614 non-STD-based reference IAEs. The VX1-software and electrophysiologist showed strong agreement in differentiating STD-based IAEs, with similar temporal and spatial characteristics. However, the VX1-software placed more emphasis on fractionation, high-frequency activation, and voltage asymmetry, whereas the electrophysiologist focused on waveform repetitiveness, fractionation, and activation pattern consistency. ML models based on the STD-based IAEs from the VX1-software outperformed those based on electrophysiologist-identified IAEs, though the identification of IAEs proved complex. Additionally, patient-specific complexity scores correlated with AF severity and progression, with lower complexity scores linked to greater discrepancies in STD-based intra-atrial regions identified by both methods.

Conclusion: The VX1-software shows promise in standardizing STD-based IAE identification with strong ML performance and accurate detection of STD-based intra-atrial areas. However, it emphasizes fractionation, high-frequency activation, and voltage asymmetry, while electrophysiologists focus on waveform repetitiveness and activation pattern consistency. Additionally, the VX1-software lacks a patient-specific approach, operating within a generalized framework that does not fully address individual complexities. These findings underscore the need for a more tailored approach that combines the standardization of the VX1-software with the patient-specific insights of the electrophysiologist. Future research should focus on integrating this tailored approach into the VX1-software to improve its accuracy, adaptability, and ultimately enhance patient outcomes while reducing recurrence rates.

Keywords: Atrial fibrillation, pulmonary vein isolation, spatio-temporal dispersions, artificial intelligence software

ACKNOWLEDGEMENTS

This master's thesis, titled *"Optimization of Potential Spatiotemporal Dispersion-Based Software for Identification of Drivers in Atrial Fibrillation",* represents nearly a year of dedicated work. It has been written to fulfill the graduation requirements of the Master's program in Technical Medicine, within the Sensing and Stimulation track, at the University of Twente in Enschede. My graduation internship and research took place in the Cardiology Department at Isala Hospital in Zwolle, where I had the opportunity to expand my knowledge of both the clinical and technical aspects of this field. This journey would not have been possible without the guidance and support of my supervisors, to whom I am deeply grateful.

I would like to express my sincere gratitude to ***Jaap Jan Smit*** and ***Karlijn Aarnink*** for their insightful discussions and constructive feedback. Your continuous availability to answer my questions, assist with data collection, and support my clinical goals was invaluable. Additionally, both of you played a key role in interpreting the clinical implications of my results, significantly improving the quality of my work and simplifying the research process. Furthermore, I would like to thank ***Martine Lagerweij*** for her guidance with the practical aspects in the hospital and her invaluable support throughout the research process.

I would also like to thank *Arlene John* for her technical supervision, whose guidance was crucial in refining my approach to analyzing and presenting the results, helping me clarify my findings and strengthen the overall quality of my conclusions.

I extend my sincere appreciation to ***Marleen Groenier*** for her invaluable guidance and support in both my personal and professional development over the years.

I would also like to thank ***Jurgen Fütterer*** for fulfilling the role of chairman throughout my thesis process. Your presence and support were greatly appreciated. Additionally, I am grateful to ***José Manuel Rivera Arbelaez*** for serving as my external committee member and for evaluating my thesis.

Finally, I would like to express my heartfelt gratitude to my ***family*** for their unwavering support and encouragement throughout this journey. Your belief in me has been a constant source of motivation, and I am truly thankful for everything you have done to help me reach this milestone.

This thesis marks the end of my academic journey at the University of Twente, but I hope it also serves as the beginning of meaningful contributions to the field of cardiology.

GRADUATION COMMITTEE

Prof. Dr. J.J. Fütterer	Chairman	University of Twente
Dr. J.J.J. Smit	Medical supervisor	Isala Hospital Zwolle
Dr. A. John	Technical supervisor	University of Twente
MSc. K.M. Aarnink	Technical Medicine supervisor	Isala Hospital Zwolle
Dr. M. Groenier	Process supervisor	University of Twente
Dr. Ir. J.M. Rivera Arbelaez	External member	University of Twente

OPTIMALISATION OF POTENTIAL SPATIO-TEMPORAL DISPERSION-BASED SOFTWARE FOR IDENTIFICATION OF DRIVERS FROM ATRIAL FIBRILLATION

AUTOR: STAN BENJAMINS

UNIVERSITY OF TWENTE. | isala

Educational institution:
University of Twente, Master Technical Medicine - Medical Sensing & Stimulation

Clinical institution:
Isala Hospital Zwolle, Department of Cardiology

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AF	Atrial fibrillation
ARI	Adjusted rand index
AUC-PR	Area under the precision recall curve
AUC-ROC	Area under the receiver operating characteristic curve
DBI	Davies-Bouldin Index
EAM	Electroanatomic map
EP-STD	Spatio-temporal dispersion identified by the electrophysiologist
FIRM	Focal impulse and rotor modulation
IAE	Intra-atrial electrogram
LOO-CV	Leave-one-out cross-validation
LVA	Low-voltages areas
ML	Machine learning
PCA	Principal component analysis
PV	Pulmonary vein
PVI	Pulmonary vein isolation
PWI	Posterior wall isolation
REF	Non-STD-based reference intra-atrial electrogram
RQA	Recurrence quantification analysis
SHAP	Shapley additive explanations
S-Score	Silhouette score
STD	Spatio-temporal dispersion
VX-STD	Spatio-temporal dispersion identified by the VX1-software
XGB	Extreme gradient boosting

1. General Introduction

Pulmonary vein isolation (PVI) is the cornerstone intervention in atrial fibrillation (AF), particularly for patients with symptomatic, drug-refractory paroxysmal or (longstanding) persistent AF [1,2]. The pulmonary veins (PVs) are well-known as critical contributors and drivers of AF, primarily originating from the myocardial sleeves surrounding the PVs [3]. These myocardial sleeves contain cells with pacemaker-like properties, generating ectopic electrical activity that trigger AF [3,4]. Additionally, the disordered architecture and electrophysiological characteristics of the PVs create an arrhythmogenic substrate, aiding in both the initiation and perpetuation of AF [3]. Nevertheless, the success rate of PVI in patients with persistent AF ranges from 43% to 67%, resulting in many patients continuing to experience AF recurrences [5]. The recurrences might be caused by PV reconnection, which occurs when the PVs re-establish electrical connections with the atrial tissue due to atrial remodeling or inadequate ablation lesions, effectively negating the effects of the initial ablation procedure [6,7]. Continuous advancements in imaging and mapping technologies are enhancing the gap identification, resulting in more precise ablation and improved outcomes with reduced recurrence rates caused by PV reconnection [7,8]. However, 56% of patients undergoing redo-ablation do not exhibit PV reconnection, indicating other non-PV contributing AF drivers within the atria [8]. These additional drivers of arrhythmogenic activity also contribute to AF recurrence, emphasizing the complexity of developing an optimal, generalized ablation procedure strategy, and highlighting the necessity for novel individual approaches [9,10].

Recent studies have investigated several strategies beyond traditional PVI to address the limitations associated with AF recurrence, including the targeting of non-PV drivers implicated in the maintenance of AF [11]. The CAPLA trial compared PVI combined with posterior wall isolation (PWI) to PVI alone in patients with persistent AF, reporting a freedom from recurrent AF of 52.4% for the combined approach and 53.6% for PVI alone [12]. These findings suggest that both strategies offer similar effectiveness in preventing recurrent AF. The STAR-AF II trial compared PVI combined with ablation of complex fractionated atrial electrograms (CFAEs) or linear ablation to PVI alone. The results showed that 59% of patients in the PVI-only group remained free from recurrent AF, compared to 49% in the PVI plus CFAE-based ablation group and 46% in the PVI plus linear ablation group [13]. Nevertheless, subsequent studies yielded inconsistent results, most probable caused by inconsistencies in the definition of CFAEs used

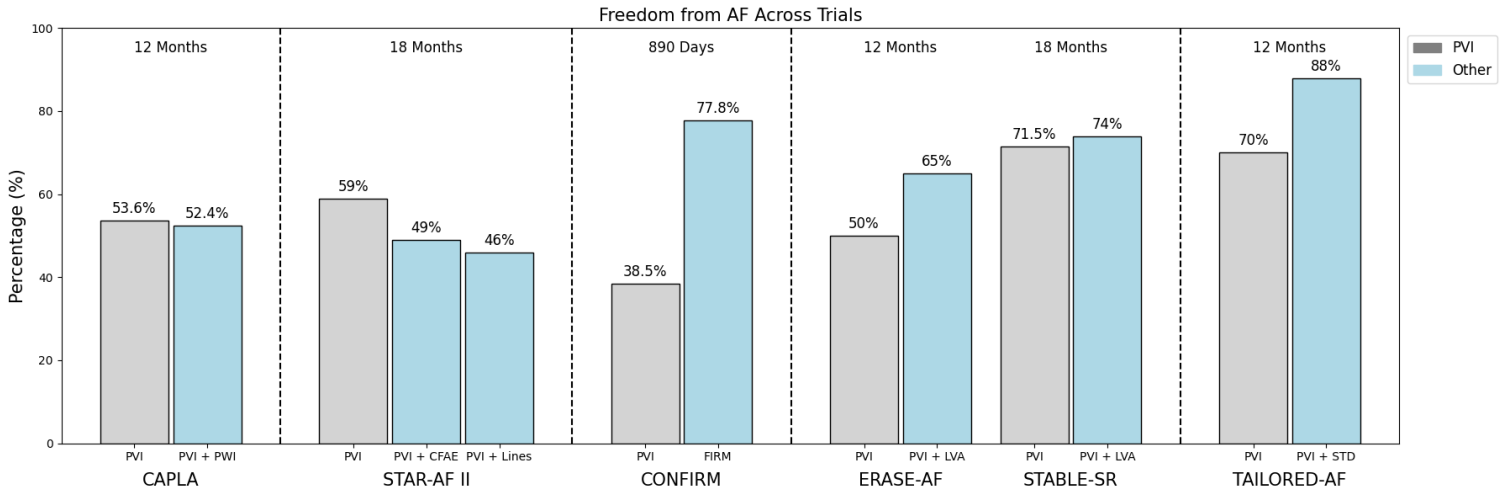


Figure 1: Overview of key outcomes from recent trials assessing various ablation strategies for reducing AF recurrences [12-14,17,19,24]. The figure compares the freedom from AF of different approaches, including PVI and PVI combined with adjunctive ablation techniques. However, it is important to note that inconsistent results were observed in subsequent studies of FIRM-guided ablation and CFAE-based strategies, as highlighted in the text.

throughout the studies [11,13]. The CONFIRM trial, a non-randomized study, assessed focal impulse and rotor modulation (FIRM)-guided ablation for AF, finding a long-term freedom from AF rate of 77.8% for FIRM-guided ablation versus 38.5% for conventional PVI [14]. However, subsequent studies reported poor long-term outcomes with FIRM-guided ablation, highlighting the need for randomized trials to better evaluate the efficacy of FIRM-guided ablation [11,15,16]. The ERASE-AF trial randomized ablation of low-voltage areas (LVAs) in combination with PVI to PVI alone, showing that AF recurrence occurred in 50% of patients who underwent PVI alone, compared to 35% in those who received LVA ablation in addition to PVI [17]. Meta-analysis further supported these findings, confirming that combining LVA ablation with PVI is more effective [18]. Conversely, the variability in patient populations, including both paroxysmal and (longstanding) persistent AF, may affect outcomes, emphasizing the need for a patient-specific approach that considers the presence of LVAs [18]. In contrast, the STABLE-SR trial, a randomized-controlled trial specifically focused on a non-paroxysmal patient population, also compared ablation of PVI with selective LVAs during sinus rhythm to PVI alone. However, this trial found no significant differences in freedom from atrial arrhythmias, with 74% in the PVI with selective LVA ablation group and 71.5% in the PVI alone group [19]. Although recent trials have explored various strategies to reduce AF recurrences, the differing results, as summarized in Figure 1, underscore the need for further research. This highlights the importance of durable PVI and refined techniques, as well as a more tailored ablation approach to effectively identify and target AF drivers [11].

A novel ablation approach to identify non-PV drivers of AF involves detection and targeting intra-atrial areas of spatio-temporal dispersions (STDs) using intra-atrial electrograms (IAEs) during AF [20]. A STD refers to both the spatial (across intra-atrial areas) and temporal (over the course of the cardiac cycle) variation of electrical activity, as shown in Figure 2, reflecting the heterogeneity of the atrial electrical conduction and activation patterns [20]. Integrating the CARTO™ 3 System *(Biosense Webster, Johnson and Johnson Medical, Irvine, California,*

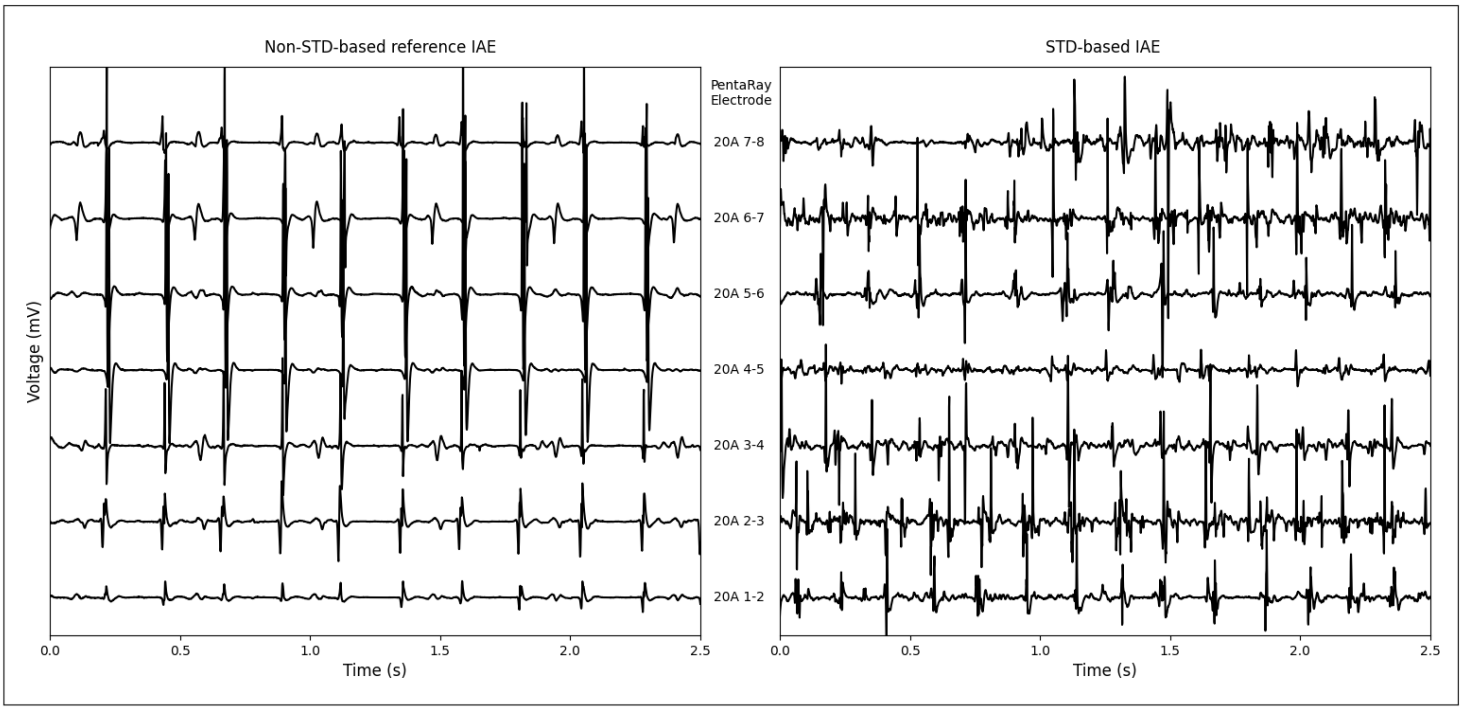


Figure 2: Comparison of multiple bipole electrodes for a non-STD-based reference IAE and an STD-based IAE, with the specific bipoles from the multielectrode PentaRay® catheter displayed in between, highlighting spatial and temporal variations.

United States), IAEs of interest corresponding to STDs can be directly mapped onto a 3D atrial model. The CARTO™ 3 system enables detailed visualization and real-time intra-atrial navigation, allowing precise localization and targeting of STD-related regions [21]. This approach overlays IAE data onto the 3D electroanatomic map (EAM), providing a comprehensive visualization of electrical activity patterns and their intra-atrial spatial distribution [21]. Nevertheless, the criteria for the selection of intra-atrial areas and the specific cut-off values for timing and morphology of the intracardiac signals for ablation remain unclarified [20,22].

The use of artificial intelligence to identify STD-based IAEs is one of the few approaches that personalize ablation strategies to target AF drivers and reduce recurrence rates [20]. Seitz et al [20]. introduced VX1 *(Volta Medical, Marseille, France*), an AI-based, expert-trained software, that standardizes the recognition of STD-based IAEs, ensuring consistent center-to-center ablation outcomes. Ablation of STD-based regions with the VX1-software resulted in an average freedom from documented AF of 86% after a single procedure and 89% after an average of 1.3 procedures per patient, with a follow-up duration of 13.5 ± 3.2 months [23]. Additionally, the TAILORED-AF trial demonstrated that using the VX1-software resulted in a freedom from documented AF of 88% after a single procedure, compared to 70% with PVI alone, as shown in Figure 1, with a 12-month follow-up [24]. However, despite the promising role of the VX1-software in optimizing ablation strategies for AF, the potential for standardization and consistency of long-term outcomes across complex patient populations remains uncertain [21,25,26]. The studies by Seitz et al [20]. and Deisenhofer et al [24]. showed higher success rates, possibly because only 26% and 18% of patients, respectively, had longstanding persistent AF, a condition that may require more complex and patient-specific approaches [25,27]. Additionally, while the VX1-software targets complex IAEs, the lack of specificity could lead to unnecessary ablation resulting in potentially pro-arrhythmic zones of electrical heterogeneity [28]. Although the short-term outcomes are promising, concerns about the long-term sustainability of these results persist, especially with the limited experience of operators and relatively short follow-up periods [25]. Therefore, further research is essential to evaluate the potential of the VX1-software in accurately identifying STD-based IAEs, refining ablation strategies, and addressing these challenges to ensure improved, consistent, and patient-specific outcomes in AF treatment.

1.1 OBJECTIVES

The general aim of this thesis is to evaluate the potential of the novel VX1-software as a standardized approach for optimizing ablation strategies in patients with AF, compared to the patient-specific approach of the electrophysiologist. The primary objectives are outlined as follows:

- *To determine the spatio-temporal characteristic differences between STD-based IAEs identified by the VX1-software, those identified by the electrophysiologist, and non-STD-based reference IAEs.*
- *To investigate the performances of existing machine learning models in differentiating between IAEs, based on the feature importances of both the VX1-software and the electrophysiologist.*
- *To assess the performance of the VX1-software, as standardized approach, in identifying STD-based IAEs across patients with varying IAE complexities.*

Consequently, the objectives of this thesis provide insights into whether the VX1-software can improve the identification of STD-based IAEs as primary drivers of AF, contributing to more consistent and patient-specific ablation approach with reducing recurrence rates.

1.2 STRUCTURE OF THE THESIS

This thesis is structured to systematically examine the objectives related to evaluating the potential of the VX1-software as a standardized approach for optimizing ablation strategy in AF. **Chapter 2** compares the spatial and temporal characteristics of IAEs across three categories: (1) STD-based IAEs identified by the VX1-software, (2) STD-based IAEs identified by the electrophysiologist, considered as gold standard, and (3) non-STD-based reference IAEs. **Chapter 3** investigates the feature importances of (1) STD-based IAEs identified by the VX1-software compared to non-STD-based reference IAEs, and (2) STD-based IAEs identified by the electrophysiologist compared to non-STD-based reference IAEs, to evaluate the performances of existing machine learning (ML) models to differentiate between STD-based and non-STD-based reference IAEs. **Chapter 4** assesses the performance of the VX1-software in identifying STD-based IAEs using a standardized approach, comparing its effectiveness to the patient-specific analysis of the electrophysiologist, examining discrepancies in STD-based areas, and evaluating its consistency and reproducibility across patients with different stadia and complexities of AF. **Chapter 5** synthesizes the overall findings as explored throughout the thesis, emphasizing the potential of the VX1-software in identifying STD-based IAEs and its role in optimizing ablation strategies for AF. It also discusses limitations and provides recommendations for future research on its potential as a standardized ablation approach. Finally, **Chapter 6** concludes the thesis by summarizing the key findings and reinforcing the potential of the VX1-software as a standardized approach in AF ablation.

2. Comparison of Spatio-temporal Characteristics of Intra-atrial Electrograms

2.1 Abstract

Introduction: The identification of STD-based IAEs is a promising advancement in optimizing AF ablation strategies, subsequent to PVI. The VX1-software offers a standardized method for identifying these STD-based IAEs, providing objective and reproducible ablation targets as compared to manual appraisal of the IAEs. However, the identification of STD-based IAEs is complex due to spatio-temporal variations in electrical activity, and the proprietary nature of the VX1-software raises concerns about its transparency and reliability.

Objective: This study compares the spatio-temporal characteristics of STD-based IAEs between the VX1-software and the electrophysiologist, regarded as the gold standard, with non-STD-based reference IAEs.

Method: This study is based on a cohort of patients enrolled in a retrospective, single-center observational registry at Isala Hospital in Zwolle, the Netherlands, from May 2022 to December 2023. Patients who underwent a redo-ablation procedure and had a left intra-atrial EAM generated using the CARTO™ 3 System were selected. Additionally, the IAEs were labeled as STD-based or non-STD-based by electrophysiologists and the VX1-software. The analysis of the IAEs involves extracting 24 features to quantify the complexity of the IAEs in both time and spatial domain. Statistical significance was assessed using non-parametric tests, including the Kruskal-Wallis and Dunn's tests, with a Bonferroni correction applied.

Results: In total, 52 patients including 110 left intra-atrial EAMs consisting of 2792 STD-based IAEs by the VX1-software, 2516 by the electrophysiologist and 11614 non-STD-based reference IAEs**.** The results reveal significant agreement and divergence in the spatio-temporal characteristics of STD-based IAEs by the VX1-software and electrophysiologist compared to non-STD-based reference IAEs. Both methods consistently distinguish STD-based IAEs, with strong agreement in key temporal and spatial features, supporting the reliability of the VX1-software. However, differences in specific spatio-temporal characteristics suggest that the VX1-software prioritizes localized voltage features, while electrophysiologists focus more on waveform repetitiveness and organization. Despite these discrepancies, overlapping interquartile ranges indicate a high degree of consistency between the two methods, reinforcing the potential of the VX1-software as a clinical tool.

Conclusion: The study underscore the potential of the VX1-software to offer a standardized approach to objective and consistent of IAE identification, focusing on localized voltage characteristics and fractionation. However, further refinement of its algorithm and signal processing is needed to fully integrate its strengths with those of the electrophysiologist, enhancing its clinical utility and accuracy.

2.2 INTRODUCTION

The identification of STD-based IAEs represent a promising contemporary advancement in optimizing AF ablation strategies, complementing the efficacy of PVI [20,24,29]. Nevertheless, accurate identification of these STD-based IAEs is essential for refining ablation strategies and improving patient outcomes. The VX1-software, developed by Seitz et al [20]., offers a proprietary and standardized approach to identify STD-based IAEs, distinguishing it from traditional methods that depend on the expertise and subjective interpretation of electrophysiologists. The introduction of this novel tool has the potential to standardize STD-based IAE identification, providing more objective, reproducible results and improving consistency in clinical practice [20,24].

However, the identification of STD-based IAEs is inherently complex, as it involves both spatial (across different intra-atrial regions) and temporal (throughout the cardiac cycle) variations in electrical activity [20]. Consequently, the full scope of STD-based IAE characteristics remains partially understood, as existing literature typically examine only a limited set of spatio-temporal features [22,30,31]. This complexity is further compounded by the proprietary methodology of the VX1-software to identify STD-based IAEs. The unique features and underlying processes of the VX1-software remain undisclosed by *Volta Medical*, raising concerns about its transparency [20]. This lack of clarity makes it challenging to evaluate the accuracy and reliability of the VX1-software, potentially resulting in a lack of specificity that could lead to unnecessary ablation and pro-arrhythmic zones of electrical heterogeneity [28]. Therefore, further investigation is necessary to explore the spatio-temporal characteristics of STD-based IAEs more comprehensively, to better understand their role in AF ablation and improve the precision and clinical applicability of the VX1-software.

The aim of this study is to analyze the spatial and temporal characteristics of STD-based IAEs identified by the VX1-software, and compare them with those identified through visual analysis by the electrophysiologist, as well as with non-STD-based reference IAEs. The electrophysiologist is considered the gold standard for identifying and analyzing IAEs, as the subsequent ablation procedure relies exclusively on the STD-based IAEs identified by the electrophysiologist. The objective of this study is:

- *To determine the spatio-temporal characteristic differences between STD-based IAEs identified by the VX1-software, those identified by the electrophysiologist*, *and non-STD-based reference IAEs.*

This objective evaluates the agreement between STD-based IAEs identified by the VX1-software and those identified through the visual analysis of the electrophysiologist, as well as their differences compared to non-STD-based reference IAEs. The similarities and differences in the IAE characteristics will provide insights into the accuracy, reliability and limitations of the VX1-software in identifying STD-based IAEs. Furthermore, this comparison will contribute to a deeper understanding of the consistency and potential clinical applications of the VX1-software as a novel AF ablation strategy.

The hypothesis is that the VX1-software identifies STD-based IAEs with spatial and temporal characteristics that are comparable to those identified by the electrophysiologist, considered as gold standard, suggesting a high level of agreement and supporting its reliability and potential clinical applicability in AF ablation strategies. However, significant differences in spatio-temporal characteristics are expected between STD-based and non-STD-based IAEs, highlighting the unique properties of STD-based IAEs. Additionally, due to the inherent variability of IAEs, no strict boundary values are expected for each spatio-temporal characteristic, underscoring the need for a robust and adaptable identification method.

2.3 METHOD

This study is based on the cohort of patients enrolled in a retrospective, single-center, observational registry at the Isala Hospital in Zwolle, the Netherlands, with an inclusion period from May 2022 to December 2023. The subsequent sections, visualized in Figure 3, provide a detailed description of the methodology, including the criteria for patient selection, procedures for data collection, and the steps involved in data preprocessing and processing.

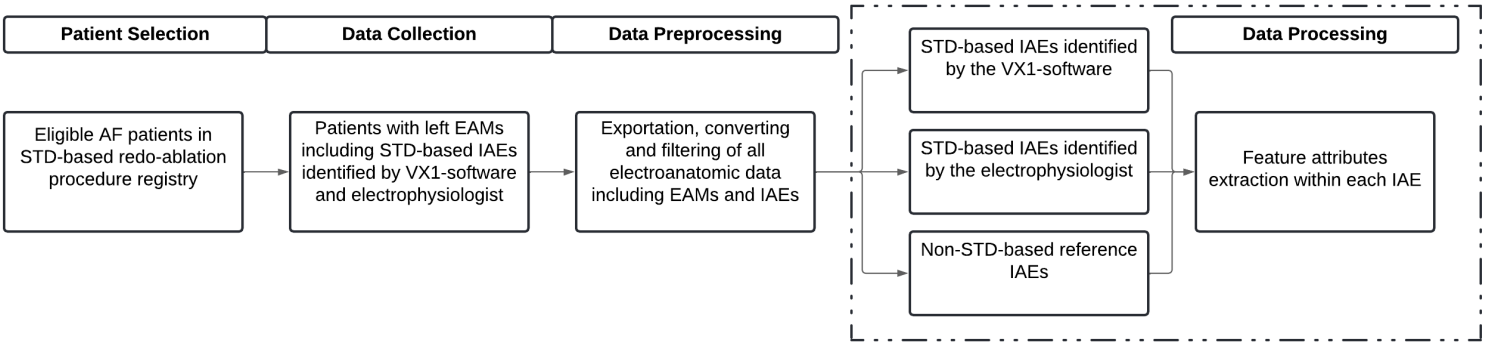


Figure 3: Schematic overview of the methodology of this study. This figure outlines the main steps of the research process, including patient selection criteria, data collection procedures, and the processes involved in data preprocessing and analysis.

2.3.1 PATIENT SELECTION

This study incorporates specific inclusion criteria that allow for the participation of patients with AF who underwent an ablation procedure. These criteria include: (1) undergoing a redo-ablation procedure conform EHRA guidelines [32], (2) presence of a left intra-atrial EAM, and (3) labeled STD-based IAEs.

The first criterion includes patients who previously underwent primary PVI using either conventional radiofrequency or cryoballoon ablation [32]. A redo-ablation procedure is indicated in cases of symptomatic recurrent AF following the initial unsuccessful ablation procedure, with the aim of reducing symptoms, recurrence, and the progression of AF. This is recommended in both the *2024 ESC Guideline* (Class IIA, Level B) and *2023 ACC/AHA/ACCP/HRS Guideline* (Class I, Level B-NR) [33,34]. The second and third criteria focus on utilizing advanced mapping and identification techniques to analyze specific intra-atrial areas. The second criterion involves mapping the left atrium using the CARTO™ 3 System, providing a detailed visualization, real-time navigation, and precise targeting of intra-atrial areas of interest [35]. The majority of AF drivers originates in the left atria, while the right atria is rarely a significant source of these drivers [36]. Therefore, this study includes only patients with a left intra-atrial EAM as most patients undergo ablation and mapping of the left atria rather than the right. The CARTO™ 3 System overlays IAEs onto the EAM, enabling a comprehensive visualization of intra-atrial electrical activity patterns [35]. The third criterion incorporates the labeling of STD-based IAEs by the VX1-software, and those identified by visual analysis of the electrophysiologist. Together, these criteria integrate the labeled STD-based IAEs onto the EAM, allowing for precise localization and targeting of STD-related intra-atrial areas. Baseline patient and redo-ablation procedure data was noted. Further details on these criteria are provided in the following subsection.

2.3.2 DATA COLLECTION

The collection of the left intra-atrial EAMs with corresponding labeled STD-based IAEs for each patient with a redo-ablation procedure, both of which are essential for further analysis, utilized different methods.

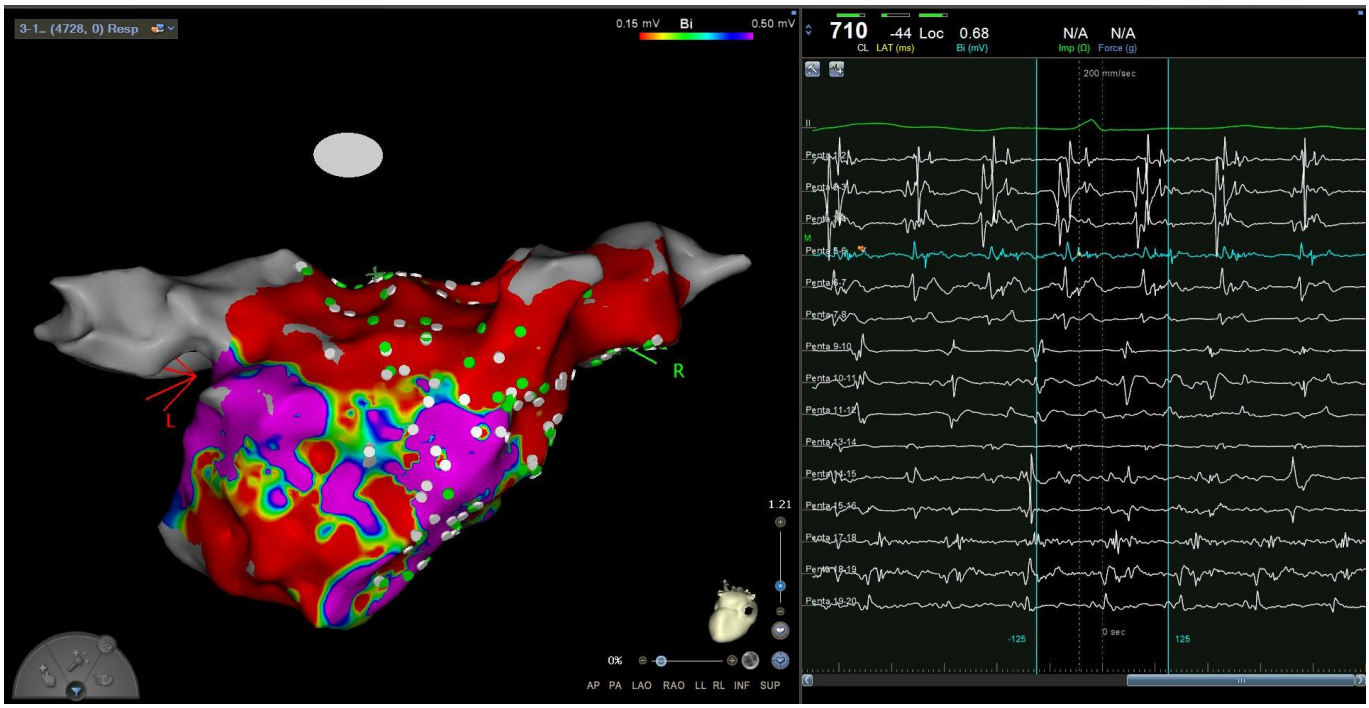


Figure 4: Real-time visualization of intra-atrial electroanatomic mapping with analyzing IAEs. On the EAM are STD-based IAEs shown by the VX1-software (grey) and the electrophysiologist (green) with the voltage characteristics of the left atria. On the right, the real-time IAEs of the multielectrode PentaRay® catheter are shown.

Intra-atrial Electroanatomic Mapping

The collection of the left intra-atrial EAMs, including all intracardiac electrophysiological data, was conducted using the CARTO™ 3 System [35]. This system is specifically designed for navigating through the heart, generating detailed three-dimensional maps, and accurately positioning catheters during diagnostic and therapeutic procedures for patients with cardiac arrhythmias [35,37]. Intra-atrial electroanatomic mapping was accomplished using the multielectrode PentaRay® catheter *(Biosense Webster, Johnson and Johnson Medical, Irvine, California, United States)* which is systematically placed in different atrial regions. The catheter provides three-dimensional electroanatomic mapping and captures sequential bipolar IAEs in a 2.5-second window [37]. During the intra-atrial electroanatomic mapping, IAEs were analyzed and labeled simultaneously as either STD-based IAE or non-STD-based reference IAE. This real-time process is visualized in Figure 4. After the redo-ablation procedure, all intracardiac electrophysiological data, including IAEs along with their corresponding labels, locations and voltages on the intra-atrial EAM, were exported for further analysis.

Patients may require multiple left EAMs during a redo-ablation procedure due to the need for detailed and repeated mapping of various intra-atrial areas. In some cases, the ablation may induce changes in electroanatomic characteristics of the atria, such as altered voltage and conduction patterns, that the original EAM does not fully represent, necessitating additional mapping for accurate visualization [38]. Furthermore, changes in arrhythmia mechanisms may require additional EAMs for thorough assessment and optimal targeting [39]. These considerations are particularly relevant in complex arrhythmias, where the changing characteristics demand multiple EAMs to refine the ablation procedure and enhance patient outcomes [39]. In this study, all unique IAEs from all EAMs are utilized, with the VX1-software labeling only the IAEs in the first EAM, while the electrophysiologist manually annotates the IAEs in all subsequent EAMs.

INTRA-ATRIAL ELECTROGRAM LABELING

The analysis of IAEs to classify them as an STD-based IAE was performed using two different methods: (1) the electrophysiologist and (2) the VX1-software. These methods are described below.

An experienced electrophysiologist identified and labeled STD-based IAEs through visual analysis of the intracardiac signals displayed in the BARD® EP Recording System *(Boston Scientific Corporation, Boston, Massachusetts, United States).* This advanced system is a software-driven tool designed for the acquisition, display and analysis of intracardiac electrophysiological data [40,41]. The intracardiac signals processed by the BARD® EP Recording System are filtered using a bandpass-filter (30–100 Hz) with a Notch-filter (50 Hz) to remove high frequency noise and power-line interference [40]. This ensures clean, interpretable IAEs, enabling accurate visual assessment. Afterward, the electrophysiologist relies on both the visual characteristics of STD-based IAEs and the personal clinical experience to label the IAEs accurately. The electrophysiologist spends an average of 3 seconds at a specific intra-atrial area to determine whether intracardiac signals with STDs are present and consistently rechecks whether a specific intra-atrial area continues to show STDs at a later time. All the IAEs were exclusively analyzed and labeled solely by an experienced electrophysiologist, ensuring accuracy, consistency and prevent mislabeling of the IAEs. Therefore, the electrophysiologist is considered the gold standard for identifying STD-based IAEs, as the redo-ablation procedure depends entirely on these STD-based IAEs. Subsequently, STD-based IAEs identified by the electrophysiologist using the BARD® EP Recording System were manually cross-labeled with corresponding IAEs within the CARTO™ 3 System to ensure alignment and consistency across both systems.

The VX1-software classifies and labels STD-based IAEs utilizing various machine and deep learning algorithms, each pre-trained offline on a dataset of annotated STD-based IAEs. This AI-based and expert-trained software is specifically designed for real-time adjudication of multipolar electrograms on STDs [20]. The VX1-software integrates also with the BARD® EP Recording System, and the same filtering process (bandpass filter of 30–100 Hz and Notch filter of 50 Hz) that eliminates high frequency noise and power-line interference. The analysis of the IAEs by the VX1-software is not bound to the predetermined importance of a single IAE parameter, instead considering multiple adjacent IAEs for a more comprehensive assessment [20,23]. The methodology to identify STD-based IAEs, including the specific spatio-temporal characteristics of the STD-based IAEs, remains exclusive and proprietary to *Volta Medical* [23]. Similar to the method of the electrophysiologist, STD-based IAEs identified by the VX1-software and recorded in the BARD® EP Recording System were manually cross-labeled with corresponding IAEs within the CARTO™ 3 System. This process enables precise alignment and validation of data between the two systems, enhancing the reliability and applicability of the analysis in practice.

2.3.3 DATA PREPROCESSING

The collected data, particularly the IAEs, are preprocessed, following the preprocessing steps to enable analysis in Python *(Python Software Foundation, Wilmington, Delaware, USA).*

EXPORTING ELECTROANATOMIC MAPS FROM THE CARTO™ 3 SYSTEM

The initial step in preprocessing each EAM from patients, including the IAEs and other intracardiac electrophysiological data, involves exporting the data per patient from the CARTO™ 3 System into the hard drive. This process is essential for subsequent analysis, enabling the data to be imported into Python for more detailed examination. However, this step was time-consuming due to the enormous amount of generated data, approximately 50.000

files with a total size of 25 GB, contained in each redo-ablation procedure in the CARTO™ 3 System. The exportation process typically took around 0.5 hours to complete for each patient.

CONVERTING ELECTROANATOMIC DATA TO MATLAB-FILE

The second step in preprocessing each EAM from patients involves converting the electroanatomic data into a MATLAB-file. This is accomplished using the OpenEP framework implemented in MATLAB *(MATLAB R2023b, The Mathworks, Inc., Natick, Massachusetts, United States).* The OpenEP framework is a comprehensive and freely available toolkit designed to support data parsing and analysis for cardiac electrophysiology research [42]. Besides, there are functions provided for analysis of chamber geometry, activation mapping, conduction velocity mapping, voltage mapping, ablation sites, and IAEs [42]. This process was also time-consuming due to the enormous amount of data contained in each EAM, typically taking around 5 hours to complete. Nevertheless, once the data is converted into the MATLAB-file, it can be processed further in Python for more advanced analysis.

FILTERING INTRA-ATRIAL ELECTROGRAMS

Filtering of the IAEs in the EAM, exported from the CARTO™ 3 System, is an essential preprocessing step in analyzing IAEs. The primary goal of filtering is to remove unwanted noise and artifacts while preserving the key electrophysiological features, particularly those related to the STDs. To achieve this, a bandpass filter with a frequency range of 30–250 Hz was applied to each IAE. This filter range was selected to preserve the high-frequency components that capture the rapid electrical dynamics of atrial activation, while removing low-frequency drift and high-frequency noise [30]. The preservation of these rapid dynamics is crucial for studying the STD of electrical activity, as it directly impacts the analysis of atrial conduction and activation patterns [30,31]. By retaining the 30–250 Hz frequency range, the filter ensures that important signal components reflecting the propagation of electrical waves across the atrium are maintained [31,43]. In addition to the bandpass filter, a notch filter was used to remove powerline interference, which typically occurs at 50 Hz. This external interference can distort the IAE signal and affect the accuracy of subsequent analyses. The combination of the bandpass (30–250 Hz) and notch filter (50 Hz) effectively cleans the IAEs, preserving the rapid spatio-temporal dynamics of atrial electrical signals while removing noise and interference [43]. This approach allows for accurate mapping of electrical wave propagation and the analysis of IAEs on STDs. The application of the bandpass- and notch filter in this study closely follows the filtering methods used in the CARTO™ 3 System, ensuring that the preprocessed data aligns with standard clinical practices and remains comparable to the clinical analyses conducted by the VX1-software and the electrophysiologist. The bandpass filter was adjusted from a range of 30-100 Hz to 30-250 Hz to better preserve rapid dynamics, as it directly influences the analysis of atrial conduction and activation patterns.

2.3.4 DATA PROCESSING

The preprocessed data utilized is processed to analyze the spatio-temporal characteristics of STD-based IAEs identified by the VX1-software, and compare them with those identified through visual analysis by the electrophysiologist, as well as with non-STD-based reference IAEs. The different processing steps are further explained in the following subsections.

INTRA-ATRIAL ELECTROGRAM FEATURE EXTRACTION

The analysis of the IAEs involves assessing various features, as documented in literature [31,44–47], to quantify the complexity of the IAEs in both time and spatial domain. The calculated features are divided into several categories, with an overview of the assessed 24 features shown in Table 1. The six attributes, commonly used in the literature, are selected to provide a comprehensive overview of the key characteristics of both STD-based and non-STD-based reference IAEs.

According to literature [23], the VX1-software does not rely on the predetermined importance of a single IAE. Instead, it considers multiple adjacent IAEs (n=3) for a more comprehensive analysis. Therefore, the features are calculated using the STD-based IAE along with the two IAEs measured from the adjacent electrodes, averaging the feature values across all three IAEs. However, the voltage attributes are calculated using all IAEs, simultaneously measured at each time point with the multielectrode PentaRay® catheter. This includes both the STD-based IAEs by the VX1-software and those by the electrophysiologist.

Table 1: Overview of all features with abbreviation calculated from the IAE, describing the spatial and temporal domain, in the different categories.

Categories	***Features (+ Abbreviation)***				***Domain***
Entropy Attributes	Sample Entropy	*SampEn*	Shannon Entropy	*ShEn*	Temporal
Frequency Attributes	Dominant Frequency	*DF*	Organization Index	*OI*	Temporal
Recurrence Attributes	Recurrence Rate	*RR*			Temporal (+ Spatial)
	Determinism	*DET*	Laminarity	*LAM*	
	Avg. Diagonal Line	D_{avg}	Trapping Time	*TT*	
	Max. Diagonal Line	D_{max}	Max. Vertical Line	V_{max}	
	Entropy Diagonal Line	D_{ent}	Entropy Vertical Line	V_{ent}	
Wave Similarity Attributes	Avg. Cycle Length	*CL*	Std. Cycle Length	CL_{std}	Temporal
	Avg. Peak-to-peak	*PP*	Std. Peak-to-peak	PP_{std}	
	Wave Similarity Index	*WSI*			
Interval Attributes	Int. Confidence Level	*ICL*	Shortest Complex Int.	*SCI*	Temporal
	Average Complex Int.	*ACL*			
Voltage Attributes	Voltage Absolute Values	*VAV*	Kurtosis	*KUR*	Temporal + Spatial
	Skewness	*SKE*			
Total Features:	24 Features				

Note: The Voltage Attributes are calculated analyzing all the measured IAEs simultaneously measured at each time point with the multielectrode PentaRay® catheter. Abbreviations: Max: maximal; Avg: average; Std: standard deviation; Int: interval;

In short, a comprehensive overview of all 24 calculated features, along with their formulas, is provided in Appendix 8.1: *"Literature Review of Intra-atrial Electrogram Features".* The following provides a brief description of each category and the corresponding features, as visualized in Figure 5, where the calculated features are shown for both a single STD-based IAE and a non-STD-based reference IAE.

- *Entropy attributes: sample entropy (SampEn) and Shannon entropy (ShEn).*

The SampEn quantifies the unpredictability by measuring the complexity and irregularity of an IAE, assessing the repetitiveness of patterns in the IAE to characterize fractionation [44,48]. Moreover, the ShEn quantifies the irregularity of the IAE by measuring the uncertainty or randomness in the probability distribution of the IAE, reflecting inherent unpredictability and indicating intra-atrial rotational activation patterns, such as reentry circuits [45,49,50]. Together, these metrics provide insights into cardiac dynamics, particularly in arrhythmias like AF, capturing the complexity and variability of intra-atrial electrical activity. The combination enhances the ability to quantify arrhythmia severity and localize reentry circuits, or focal drivers, contributing to persistent arrhythmias [22,30,50].

- *Frequency attributes: dominant frequency (DF) and organization index (OI).*

The DF quantifies the local activation rate by identifying the most prominent frequency within the IAE frequency spectrum, representing regions with rapid electrical activity often linked to remodeled atrial substrate and primary drivers of AF persistence [22,51]. Furthermore, the OI

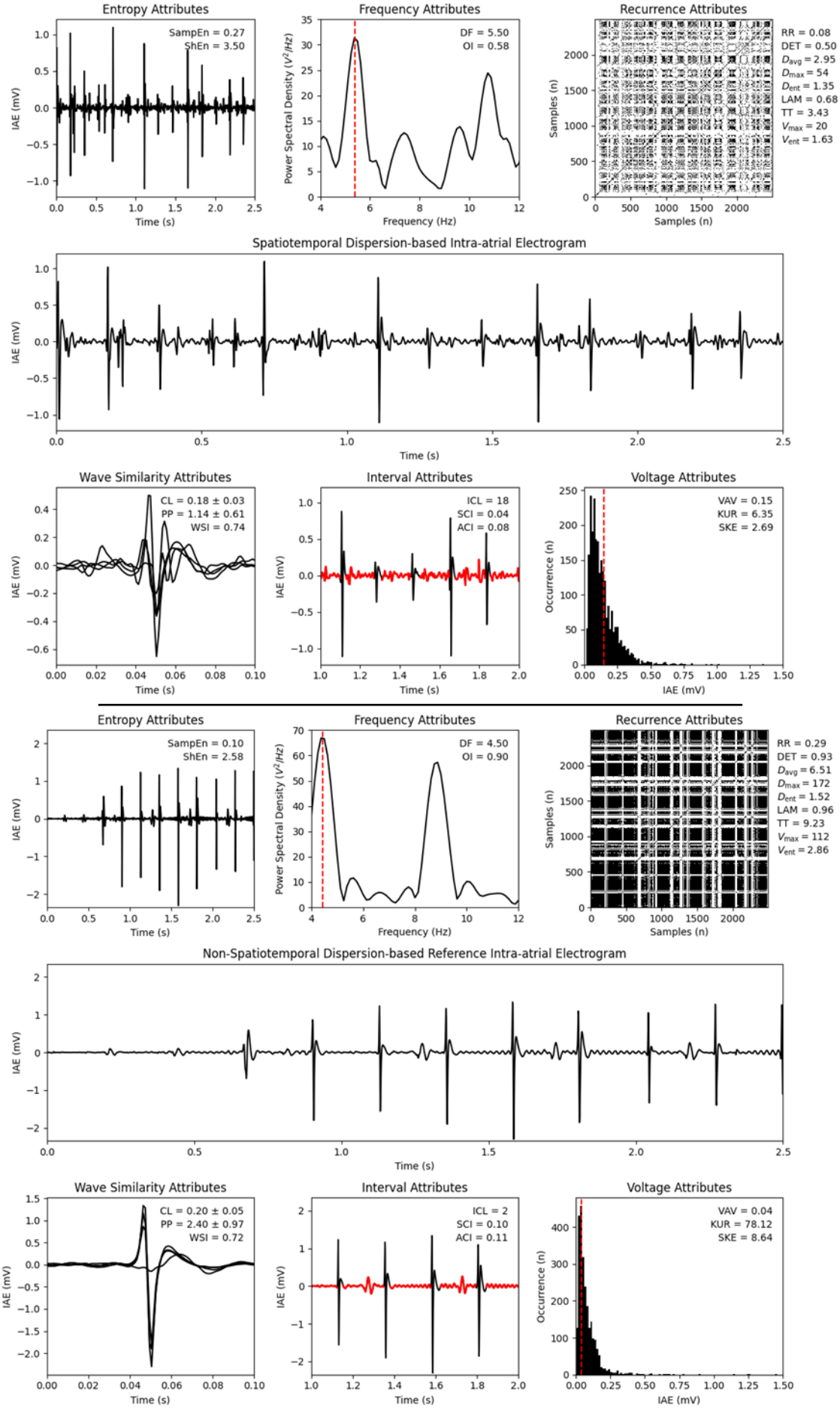


Figure 5: Overview of feature extraction for a single STD-based IAE and a single non-STD-based reference IAE, encompassing six different attributes and a total of 24 calculated features to provide insights into the spatio-temporal domain complexities of the IAEs.

quantifies the organization of the IAE relative to the DF by evaluating the spectral dispersion, assessing the periodicity and variability of frequency components [51]. More periodic and organized atrial activity is potentially more effective in driving the overall rhythm and may display a greater tendency to exhibit recurring spatial patterns across atrial regions. Together, these metrics provide valuable information on AF conduction complexity, spatial heterogeneity, and the degree of atrial electrical remodeling. By combining DF and OI, critical high-frequency, well-organized regions have been identified as key to maintaining AF, making them potential ablation targets to enhance success rates of AF termination [22].

- *Recurrence attributes: recurrence rate (RR); determinism (DET); average, maximal and entropy of diagonal line (D_{avg}, D_{max}, D_{ent}); laminarity (LAM); trapping time (TT); maximal and entropy of vertical line (V_{max}, V_{ent}).*

The recurrence attributes, derived through recurrence quantification analysis (RQA), are metrics for identifying recurring patterns that reveal the non-linear and potentially unstable periodic dynamics of AF [52]. RQA evaluates the relationship between the atrial rate and spectral center frequency, enabling the classification of fractionated or temporally unstable IAEs. It is particularly effective at identifying local intra-atrial dynamics, such as regions of stable or repetitive driver activity with high recurrence patterns, distinguishing these critical atrial substrates from passive regions exhibiting chaotic or erratic signals [46,52]. These metrics provide valuable insights into the spatio-temporal structure, predictability, and complexity of IAEs, highlighting regions of stability, irregularity, and transitions within the phase-space trajectory. By quantifying state recurrence and persistence, RQA metrics provide insights in the stability and complexity of the rhythm, offering a deeper understanding of the dynamic behaviors and mechanisms underlying AF [46,52]. A detailed overview of the derived RQA metrics are described in Appendix 8.1: *"Literature Review of Intra-atrial Electrogram Features"*..

- *Wave similarity attributes: intra-atrial cycle length (CL) with standard deviation (CL_{std}); peak-to-peak amplitude (PP) with standard deviation (PP_{std}); and wave similarity index (WSI).*

The CL measures the average interval between successive activations, while CL_{std} captures the variability in these intervals, reflecting the temporal organization and rhythm stability of atrial activation [53,54]. Similarly, PP quantifies the average amplitude of activation waves, representing the strength of atrial depolarization, with PP_{std} assessing the variability in these amplitudes to assess the consistency of conduction. The WSI complements these by evaluating waveform repetitiveness and organization, identifying regions with regular, consistent activation, such as rotors or focal drivers, which are critical for sustaining AF [55]. By combining these metrics, it becomes possible to distinguish critical sites of high-frequency, organized activity from passive regions with disorganized signals. This comprehensive analysis supports the spatio-temporal characterization of IAEs, enabling precise localization of arrhythmogenic drivers and providing valuable guidance for targeted ablation and improved treatment outcomes [53–55].

- *Interval attributes: interval confidence level (ICL); shortest complex skewness (SCI); and average duration of complex intervals (ACI).*

The ICL quantifies the number of complex intervals within the IAE, reflecting the degree of fractionation present. The SCI identifies the fastest atrial conduction interval, serving as a marker for rapid atrial activation and high-frequency activity, while the ACI measures the average duration of complex intervals, indicating the overall consistency of activation across the atrial myocardium [47,56]. Together, these interval attributes, provide valuable insights into

the temporal and electrical conduction properties of IAEs. It assesses fractionation, activation rapidity, and conduction regularity, offering critical guidance for understanding arrhythmic substrates and designing targeted ablation strategies to improve outcomes [47,56].

- *Voltage attributes: voltage absolute value (VAV); kurtosis (KUR); and skewness (SKE).*

The VAV represents the average peak voltage over time, reflecting the strength of electrical activity at a specific site. The KUR quantifies the tailedness of the voltage distribution, highlighting the presence of extreme values or outliers that may indicate irregular or highly fragmented activation patterns. The SKE measures the asymmetry of the voltage distribution, revealing whether the voltage time series skews toward higher or lower values, which can point to uneven activation or abnormal electrical conduction [31]. Together, these attributes provide valuable insights into the electrophysiological behavior of atrial tissue, identifying regions with varying activity, abnormal conduction, and potential arrhythmogenic drivers [31,57].

Overall, the 24 features provide a comprehensive understanding of the spatio-temporal characteristics of IAEs. These features are designed to capture various aspects of the electrical activity, allowing for a detailed comparison between STD-based IAEs and non-STD-based reference IAEs.

EVALUATION METRICS

The distribution of the spatio-temporal characteristics of the IAEs is visualized through boxplots, which provide a clear overview of the variability, central tendency, and spread of these features. Each box represents the interquartile range, with the median value marked by a line inside the box, while the whiskers indicate the range of the data, excluding outliers. Outliers are shown as individual points outside the whiskers. This visualization facilitates comparison across IAEs, revealing potential patterns and inconsistencies in IAE identification.

To assess significant differences between the IAEs (STD-based IAE by the VX1-software, STD-based IAE by the electrophysiologist, and the non-STD-based reference IAE), the normality of each feature was first determined. Since none of the features followed a normal distribution, non-parametric tests were directly applied, which are appropriate for analyzing non-normal data distributions.

The Kruskal-Wallis test was used to determine significant differences among the IAEs for each feature. This non-parametric alternative to the one-way ANOVA compares the ranks of the data rather than the raw values. It tests the null hypothesis that the distributions of the different IAEs are identical [58]. A significant result from the Kruskal-Wallis test indicates that at least one IAE differs from the other IAEs. To identify the specific pairs of IAEs that were significantly different, Dunn's test was performed as a post hoc analysis [59].

The Dunn’s test is designed for pairwise comparisons following a significant Kruskal-Wallis test result. It compares the ranks of the data between each pair of IAE to determine which pairs differ significantly [59]. It is suitable for non-normal data and incorporates a p-value adjustment, in this case the Bonferroni correction, to control the risk of Type I errors (false positives). The Bonferroni correction ensures a minimized likelihood of false conclusions when multiple comparisons are made. Significant pairwise differences were identified based on a p-value threshold of 0.05, ensuring a more reliable understanding of how the IAEs differed with respect to each feature, while minimizing the risk of Type I errors across multiple comparisons [59].

In conclusion, this approach combines the Kruskal-Wallis test, Dunn's test, and Bonferroni correction to ensure statistical analysis by minimizing the risk of false conclusions. Applying non-parametric tests and making appropriate adjustments for multiple comparisons ensures reliable conclusions regarding the significant differences between the IAEs for each features.

2.4 RESULTS

This study incorporates specific inclusion criteria that allow for the participation of 108 patients who underwent an ablation procedure. Exclusion of 56 patients ensures available left intra-atrial EAMs with labeled STD-based IAEs for analysis. The reasons for exclusion were unavailability of labeling for STD-based IAEs by the VX1-software (n=46), the presence of only a right intra-atrial EAM with labeled STD-based IAEs (n=5), as well as unsuccessful exportation from the CARTO™ 3 System (n=5). As a result, the left intra-atrial EAMs of 52 patients were utilized, which includes STD-based IAEs identified by the VX1-software (VX-STD) and those identified by the electrophysiologist (EP-STD), and non-STD-based reference IAEs (REF).

Consequently, the exported left intra-atrial EAMs of 52 patients were used for further analysis, comprising all intracardiac electrophysiological data. These 52 patients includes 110 left intra-atrial EAMs consisting a total of 2792 VX-STD, 2516 EP-STD and 11614 REF IAEs. The REF IAEs are defined as an IAE, within one centimeter of the surrounding intra-atrial area without VX-STD nor EP-STD IAE. IAEs are 2.5 seconds in duration, with a sampling frequency of 1000 Hz. The patient baseline and redo-ablation procedure characteristics are shown in Table 2.

Table 2: Patient baseline characteristics, including redo-ablation procedure characteristics.

Patients baseline characteristics		**All patients (n=52)**	
Demographics			
Male sex, no. (%)		37 (71,15%)	
Age, median [IQR] (year)		65 [58,25 – 71,25]	
CHA_2DS_2-VASC score, median [IQR]		2 [1 – 3]	
Comorbidities			
Hypertension, no. (%)		30 (57.69%)	
Obesity, no. (%)		11 (21.15%)	
Sleep apnea, no. (%)		16 (30.77%)	
Diabetes mellitus, no. (%)		4 (7.69%)	
Stroke/TIA, no. (%)		9 (17.31%)	
PM/ICD, no. (%)		2 (3.85%)	
Endocardiographic			
LVEF, no. (%)	Normal Mild Moderate Severe	36 (69.23%) 8 (15.39%) 6 (11.54%) 2 (3,84%)	
LAVI, median [IQR] (ml/m^2)		38 [32 – 48]	
AF type			
Paroxysmal, no. (%)		23 (44.23%)	
Persistent, no. (%)		14 (26.92%)	
Longstanding persistent, no. (%)		15 (28.85%)	
AF (maximal) duration			
AF history, year [IQR]		6 [3 – 11]	
All, months [IQR]		5 [0 – 12]	
Longstanding persistent, months [IQR]		12 [10 – 25]	
Redo-ablation procedure characteristics			
EAMs, no.		110	
VX-STD IAEs, no. (% of IAEs)		2792 (16.50%)	Total: 16922
EP-STD IAEs, no. (% of IAEs)		2516 (14.87%)	
REF IAEs, no. (% of IAEs)		11614 (68.63%)	

Note: Values are mean ± standard deviation, median [interquartile range], or no. (%). Abbreviations: EP, electrophysiologist; ICD, implantable cardioverter defibrillators; LA, left atrium; left atrium volume index; LVEF, left ventricular ejection fraction; PM, pacemakers; REF, reference; TIA, transient ischaemic attack; VX, VX1-software.

2.4.1 TEMPORAL CHARACTERISTICS

Figure 6 shows most of the temporal characteristics of: (1) VX-STD IAEs, (2) EP-STD IAEs, and (3) REF IAEs. The comparison between VX-STD and REF IAEs reveals significant differences in nearly all temporal characteristics, except for OI and ACI (*p>0.05*). Similarly, comparing the EP-STD and REF IAEs, significant differences are shown for almost all temporal characteristics, except the PP, PP_{std} and ACI (*p>0.05*). Moreover, the comparison between VX-STD and EP-STD IAEs showed significant differences in temporal characteristics, except the SampEn, ShEn, CL, CL_{std} and ACI (*p>0.05*).

Additionally, there is agreement between the VX1-software and the electrophysiologist, with both showing significant differences compared to REF IAEs for the temporal characteristics SampEn, ShEn, CL and CL_{std} (*p<0.05*). However, discrepancies are noted for the DF, WSI, ICL and SCI (*p<0.05*), where the two methods do not align in identifying STD-based IAEs, although both still showing significant differences when compared to REF IAEs.

Furthermore, the distributions of temporal characteristics between STD-based and REF IAEs shows overlap, with the exceptions of the WSI and ICL, where the interquartile ranges do not overlap. In contrast, the distributions of all temporal characteristics between the VX1-software and the electrophysiologist exhibit overlapping interquartile ranges. Moreover, the medians for the temporal characteristics WSI and SCI are more deviant for the electrophysiologist than for the VX1-software, when compared to the REF IAEs. Conversely, for the temporal characteristics DF and ICL, the opposite is true. Besides, the median values of the SampEn, ShEn, DF, CL_{std}, and ICL are significantly lower for REF IAEs compared to EP-STD IAEs, the gold standard. In contrast, the median values of the OI, CL, WSI, and SCI are significantly higher for REF IAEs than for EP-STD IAEs.

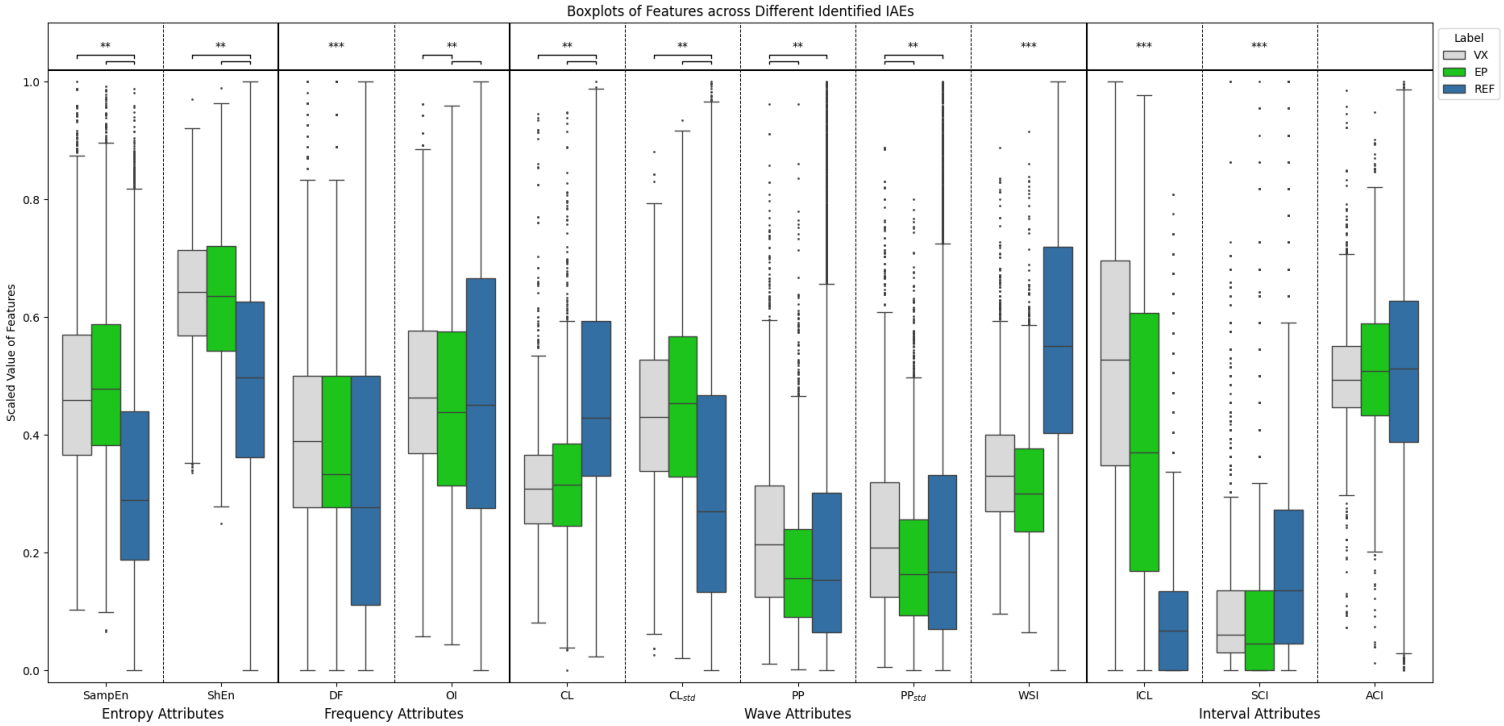


Figure 6: Visualization of the temporal characteristics of VX-STD IAEs (grey), EP-STD IAEs (green), considered as gold standard, and REF IAEs (blue). Note: Significant differences between the groups are indicated at the top: (***) denotes significant differences across all IAEs, while (**) highlights significant differences between the IAEs linked by brackets. The values of the features are scaled using a min-max normalization to enhance their clarity and comparison in the figure. Abbreviations: EP, EP-STD IAEs; REF, REF IAEs; VX, VX-STD IAEs.

2.4.2 SPATIAL CHARACTERISTICS

Figure 7 shows the spatial characteristics of: (1) VX-STD IAEs, (2) EP-STD IAEs, and (3) REF IAEs. Comparing both the VX-STD and EP-STD IAEs, with the REF IAEs, significant differences are shown for all spatial characteristics. Furthermore, comparing the VX-STD IAEs with the EP-STD IAEs, significant differences are determined for the spatial characteristics, except the RR and LAM ($p>0.05$).

Moreover, there is agreement between the VX1-software and the electrophysiologist, with both showing significant differences compared to REF IAEs for the spatial characteristics RR and LAM ($p<0.05$). However, discrepancies are noted for the DET, D_{avg}, D_{max}, D_{ent}, TT, V_{max}, V_{ent}, VAV, KUR, and SKE, where the two methods diverge in identifying STD-based IAEs, though both still demonstrate significant differences compared to REF IAEs.

Regarding the distributions, all spatial characteristics between STD-based and non-STD-based reference IAEs show overlap in interquartile ranges. Similarly, the distributions of all spatial characteristics between the VX1-software and the electrophysiologist exhibit overlapping interquartile ranges. Additionally, the medians for all the spatial characteristics are more deviant for the electrophysiologist than for the VX1-software, except for the VAV, KUR, and SKE, when compared to the non-STD-based reference IAEs. Besides, the median value of only the VAV is significantly lower in REF IAEs compared to the gold-standard EP-STD IAEs. In contrast, the median values of almost all spatial features (RR, DET, D_{avg}, D_{max}, D_{ent}, LAM, TT, V_{max}, V_{ent}, KUR, and SKE) are significantly higher in REF IAEs than in EP-STD IAEs.

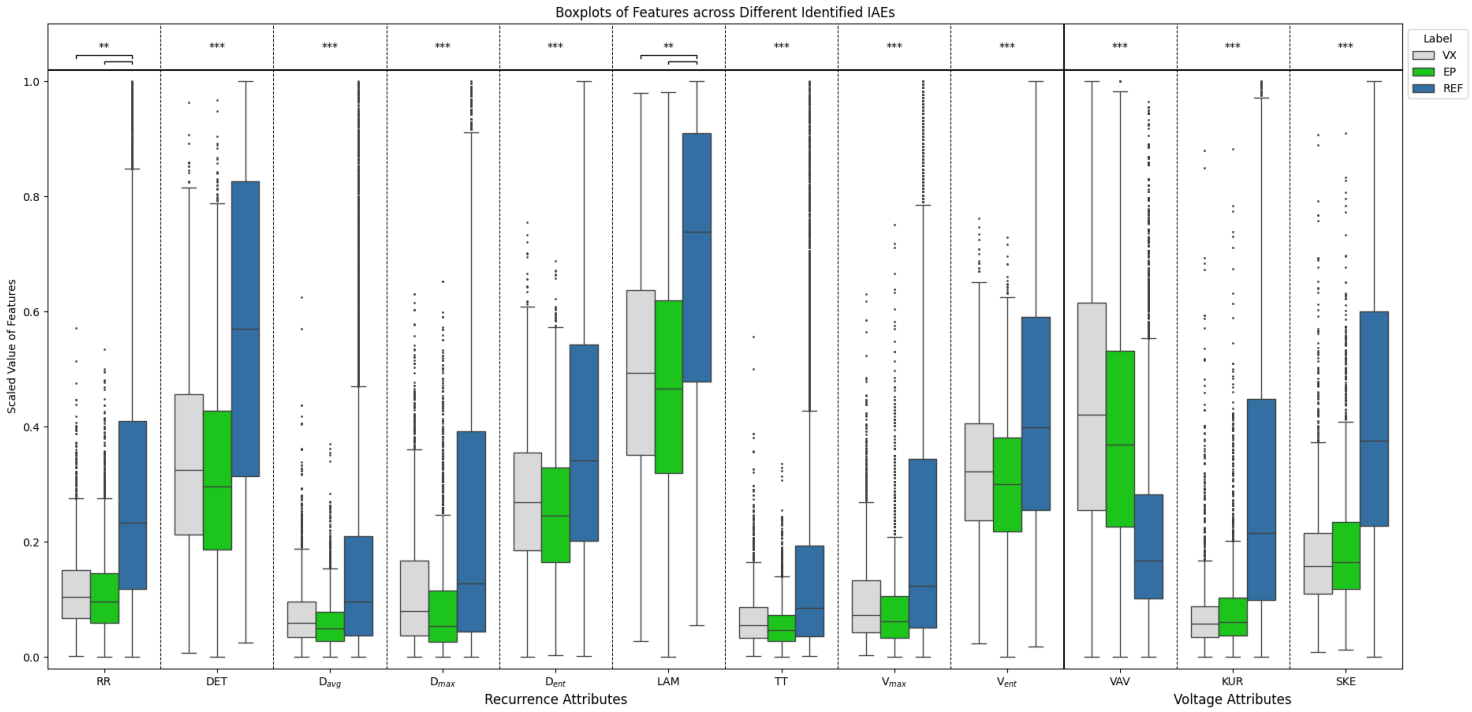


Figure 7: Visualization of the spatial characteristics of VX-STD IAEs (grey), EP-STD IAEs (green), considered as gold standard, and REF IAEs (blue). Note: Significant differences between the groups are indicated at the top: (***) denotes significant differences across all IAEs, while (**) highlights significant differences between the IAEs linked by brackets. The values of the features are scaled using a min-max normalization to enhance their clarity and comparison in the figure. Abbreviations: EP, EP-STD IAEs; REF, REF IAEs; VX, VX-STD IAEs.

2.5 DISCUSSION

In the discussion, the findings are presented alongside the limitations of the study, followed by an exploration of the clinical relevance of the results and concluding with recommendations for future research.

2.5.1 FINDINGS

The findings, shown in Figures 6 and 7, reveal several significant patterns of agreement and divergence in both temporal and spatial characteristics when comparing VX-STD and EP-STD IAEs, to REF IAEs. The REF IAEs differ significantly from STD-based IAEs, including both VX-STD and EP-STD, across nearly all spatio-temporal characteristics, highlighting the unique characteristics of STD-based IAEs. These findings emphasize the ability of both methods to consistently identify the unique characteristics of STD-based IAEs, distinguishing them from non-STD-based reference IAEs.

Furthermore, the VX1-software and the electrophysiologist demonstrated significant agreement in STD-based IAEs, particularly for key temporal and spatial characteristics such as SampEn, ShEn, CL, CL_{std}, RR and LAM. This consistent alignment supports the reliability of the VX1-software in capturing core features of STD-based IAEs, validating its use as a tool for identifying these IAEs with accuracy comparable to the electrophysiologist, the gold standard. This suggests that the VX1-software can effectively complement or replace manual identification, particularly for calculated temporal and spatial features. However, the spatio-temporal characteristics that showed significant differences between the VX1-software and the electrophysiologist, namely the WSI, SCI, DET, D_{max}, D_{ent}, TT, V_{max} and V_{ent}, are more prominent for the electrophysiologist. In contrast, the spatio-temporal characteristics DF, ICL, VAV, KUR, and SKE are more prominent for the VX1-software. These discrepancies suggest that the electrophysiologist and the VX1-software are also analyzing different aspects of the IAEs, possibly due to differences in their method of analyzing IAEs. The areas for potential improvement highlight the need for combining the different aspects of both methods of the VX1-software and the electrophysiologist, which could enhance the performance of the VX1-software, while also providing the electrophysiologist with insights to improve their analysis and understanding of STD-based IAEs.

Additionally, in terms of distributions, the VX1-software and the electrophysiologist often demonstrated overlapping interquartile ranges for most spatio-temporal characteristics, indicating a high degree of consistency in their identification of STD-based IAEs. This overlap further suggests that the VX1-software is a reliable tool for recognizing the overall characteristics of STD-based IAEs, making it a promising tool in clinical practice, unless the significant differences with the electrophysiologist in certain spatio-temporal characteristics. Nevertheless, most spatio-temporal characteristics of STD-based (VX-STD and EP-STD) and REF IAEs also exhibit overlapping interquartile ranges, indicating that the differences in IAE characteristics are subtle. However, the WSI and ICL do not show overlap in interquartile ranges between the electrophysiologist with the REF IAEs, suggesting that the electrophysiologist places more emphasis on waveform repetitiveness and organization, as well as the degree of fractionation. In contrast, the VX1-software does not show overlap in interquartile ranges for the ICL, KUR, and SKE, indicating that it prioritizes the voltage characteristics in the area around a specific point in IAE analysis, as well as the degree of fractionation in the IAE.

These findings highlight the nuanced differences between the two methods, with each excelling in certain aspects and presenting challenges in others, while the similarities suggest opportunities for integration to improve overall analysis.

2.5.2 LIMITATIONS

Despite the findings of this study highlighting the potential of the VX1-software in identifying STD-based IAEs, several limitations should be acknowledged. One limitation is that the analysis of the IAEs by the VX1-software is not bound to the predetermined importance of a single IAE, instead considering multiple IAEs from adjacent electrodes, as stated in literature [20]. In contrast, the electrophysiologist analyzes the IAE using the data from a single electrode. This discrepancy could lead to inconsistencies in IAE identification between the VX1-software and the electrophysiologist, especially in areas without signals on adjacent bipoles of the PentaRay® catheter. Specifically, some STD-based IAEs may not be identified by the VX1-software if adjacent IAEs do not exhibit the necessary spatio-temporal characteristics. Conversely, the VX1-software might identify more STD-based IAEs if adjacent IAEs show STDs that are not present in the single electrode analyzed by the electrophysiologist. This might explain why there are more VX-STD IAEs compared to EP-STD IAEs (2792 vs. 2516). As a result, this difference in analysis introduces variability that may impact the accuracy and consistency of comparisons between the VX1-software and the electrophysiologist.

Additionally, in this study, each feature was calculated using a standardized method, which may not fully capture the variability inherent in different types of IAEs. For instance, low-voltage IAEs with less distinct peaks complicates accurate calculation of features, especially when the IAE lacks well-defined peaks. Spatio-temporal features such as DF, OI, PP, WSI, ICL, SCI, and ACI are especially affected, as their calculation relies on the identification of these peaks within the IAE. This is in contrast to high-voltage IAEs, where the peaks are more prominent and therefore easier to detect. The challenges presented by low-voltage IAEs may lead to discrepancies in feature calculation, as the standardized method applied in this study may not be optimized for such signals. Previous studies have identified areas of low-voltage IAEs as the primary location for STDs [20,24,60]. This limitation highlights the need for further refinement in the feature calculation process, particularly for low-voltage IAEs, to enhance accuracy and ensure more reliable feature values across a wider range of IAE signal characteristics.

Moreover, another limitation is the lack of transparency regarding the specific features utilized by the VX1-software for IAE identification. Since the features employed by the VX1-software are not fully disclosed, some of the observed significant differences between the VX1-software and electrophysiologist may arise from the reliance of the software on features that are excluded from the manual identification process. Although these significant differences may appear to indicate incorrect identifications by the VX1-software compared to the electrophysiologist, it may instead be a result of the design, where certain features are intentionally not included for IAE analysis. Consequently, without clarity on the features used, it becomes challenging to determine whether these differences arise from the feature selection of the VX1-software or represent fundamental discrepancies in IAE identification. This limitation underscores the importance of transparency in the algorithms used by the software and may warrant further investigation into the specific features employed by the VX1-software to fully assess its performance.

2.5.3 CLINICAL RELEVANCE

Accurate identification and characterization of STD-based IAEs are critical for the management of AF, as they directly impact treatment decisions and patient outcomes. Understanding the spatio-temporal characteristics of IAEs is essential for evaluating AF substrates, as they can provide insights into the mechanisms driving the arrhythmia and help guide treatment strategies. The findings from this study highlight the potential of the VX1-software as a valuable tool for identifying key spatio-temporal features of STD-based IAEs, offering a promising

alternative or complement to manual identification by electrophysiologists. However, the electrophysiologist has a greater focus on waveform repetitiveness and fractionation, while the VX1-software prioritizes voltage characteristics and localized fractionation in IAE analysis. Combining both methods of the VX1-software and the electrophysiologist could enhance the performance of the VX1-software, while also providing the electrophysiologist with insights to improve their analysis and understanding of STD-based IAEs. Consequently, by providing objective, automated measurements, the VX1-software could improve diagnostic accuracy, reduce clinician workload, and streamline clinical workflows, ultimately leading to more efficient and precise arrhythmia mapping and better patient outcomes.

Furthermore, the observed overlap in the feature distributions between STD-based and non-STD-based IAEs adds an additional layer of complexity to the identification of IAEs. This overlap suggests the presence of considerable variability within the IAEs themselves, which can make it more challenging to consistently distinguish between different types of IAEs. This finding aligns with existing literature [30,31,46,61], where a high degree of variability and overlap between IAE patterns has been noted, particularly in relation to complex fractionated AF substrates. Although this variability complicates the identification process, it also offers valuable insights into the heterogeneous nature of AF and the various arrhythmic substrates that sustain its persistence.

Additionally, the values of the spatio-temporal features observed in this study are consistent with those reported in literature [30,31,61]. Non-STD-based reference IAEs are characterized by low SampEn and ICL, and high WSI and DET, while the opposite is true for STD-based IAEs. These findings align with the study of Nadim et al [30]., further validating the use of these feature values in distinguishing between the two types of IAEs. This alignment with existing literature underscores the reliability of the VX1-software in identifying clinically relevant patterns and highlights its potential to support consistent, evidence-based decision-making.

2.5.4 FUTURE PERSPECTIVES

Although current methodologies offer valuable insights, there are several challenges that need to be addressed in order to enhance the accuracy and clinical applicability of STD-based IAE analysis. Future studies could focus on investigating the patterns and interactions between spatio-temporal features, as this could enhance the accuracy and robustness of the VX1-software. Additionally, exploring the heterogeneity of AF substrates, as reflected by the overlap in IAE characteristics, is crucial for understanding the diverse mechanisms driving AF persistence and enabling more tailored approaches in personalized treatment. This deeper understanding could refine the VX1-software and help identify more precise treatment strategies, ultimately improving patient outcomes.

Furthermore, enhancing the transparency of the features used by the VX1-software for IAE analysis is another area of focus. Future research should examine how feature selection impacts IAE identification, particularly in addressing whether discrepancies between the VX1-software and electrophysiologist arise from differences in feature sets or challenges with complex IAEs.

Finally, the challenges of detecting low-voltage IAEs, which have less distinct peaks, must be addressed. Refining algorithms to better capture low-voltage features, possibly through advanced filtering or signal enhancement, will improve peak detection and feature accuracy. In summary, addressing these challenges will enhance the accuracy, clinical applicability, and reliability in IAE analysis of the VX1-software, further establishing its role in clinical practice.

2.6 CONCLUSION

The study demonstrates the strengths of both the VX1-software and the electrophysiologist in consistently identifying STD-based IAEs, with the VX1-software emerging as a valuable tool that could standardize IAE identification, offer a promising alternative or complement to the electrophysiologist, and provide more objective, reproducible results to improve consistency in clinical practice. Conversely, the differences in feature emphasis indicate that electrophysiologists prioritize waveform repetitiveness and organization, whereas the VX1-software focuses on localized voltage characteristics and fractionation. Integrating the unique strengths of both methods could enhance the performance of the VX1-software while providing the electrophysiologist with additional insights, ultimately improving the accuracy and consistency of STD-based IAE identification. Future research should focus on advancing algorithm development and signal processing to fully realize the clinical potential of the VX1-software as an optimized ablation strategy for AF, following PVI.

3. Comparative Evaluation of Existing Machine Learning Models for Intra-atrial Electrogram Differentiation

3.1 Abstract

Introduction: Accurately identifying STD-based IAEs is challenging due to their complex and variable spatio-temporal characteristics, making differentiation from non-STD-based reference IAEs difficult. Additionally, discrepancies between the VX1-software and electrophysiologists arise from differences in how each identifies STD-based IAEs, further complicating the process. Machine learning (ML) models offer a promising solution by analyzing high-dimensional IAEs to uncover patterns that enhance classification and highlight areas for refining the VX1-software, ultimately improving IAE detection and ablation accuracy.

Objective: This study compares the performances of existing ML models in IAE differentiation based on the feature importances of both the VX1-software and the electrophysiologist.

Method: This study applies existing ML models to differentiate IAEs using a dataset with 24 features per IAE. Feature importances are assessed using an extreme gradient boosting (XGB) model, with Shapley additive explanations (SHAP) enhancing interpretability by quantifying individual feature impact. Unsupervised ML, utilizing principal component analysis (PCA) and K-Means clustering, along with the supervised XGB model, is used to classify IAEs and identify underlying patterns. Model performance evaluation employs multiple metrics, including the silhouette score (S-Score), Davies-Bouldin index (DBI), adjusted rand index (ARI), F1-score, area under the receiver operating curve (AUC-ROC), and area under the precision-recall curve (AUC-PR).

Results: Feature importance analysis reveals that while the VX1-software relies on the degree of fractionation, high-frequency activation, and asymmetry of the local voltage distribution, electrophysiologists focus on waveform repetitiveness, degree of fractionation, and consistency of activation patterns. The superior clustering and classification performance of the K-Means and XGB models, driven by the feature importance of the VX1-software, highlights the complexity of IAE identification by the electrophysiologist. This underscores the need for the VX1-software to adopt a more refined, multi-faceted approach to effectively capture intricate STD-based IAE patterns.

Conclusion: The study underscore the strengths of existing ML models in identifying patterns and classifying IAEs, as well as the generalization capabilities of the VX1-software in IAE identification. Nevertheless, further refinement is needed to integrate patient-specific patterns through a multi-faceted approach, enhancing the accuracy and reliability of the VX1-software as a comprehensive IAE identification tool.

3.2 INTRODUCTION

Accurately identifying STD-based IAEs remains a significant challenge in AF ablation procedures [60,62]. This challenge arises from an incomplete understanding of their spatio-temporal characteristics, complicating differentiation from non-STD-based reference IAEs [22,30,31]. Although notable differences exist, several factors such as patient-specific anatomy, arrhythmia dynamics, and electrode positioning further influence the variability in feature distribution, contributing to the complexity of the spatio-temporal characteristics [20,30,61]. Additionally, the nature of IAEs can change throughout the ablation procedure, adding another layer of complexity to their identification [20,63]. Consequently, the identification of STD-based IAEs requires a more comprehensive approach.

The emergence of various patterns of STD-based IAEs further necessitates the use of multiple features to reliable differentiate them from non-STD-based reference IAEs, emphasizing the importance of a comprehensive, multifaceted methodology to improve classification accuracy and clinical outcomes [30,31,61]. Furthermore, discrepancies between the VX1-software and the electrophysiologist arise from variations in how each method identifies STD-based IAEs. This results in suboptimal targeting of STD-based AF drivers, causing unnecessary ablation, potentially pro-arrhythmic zones, an increased likelihood of complications, and ultimately reduced treatment efficacy [28]. These challenges underscore the need for accurate identification of IAEs through improved alignment between the VX1-software and the electrophysiologist to optimize treatment outcomes, ultimately enhancing AF ablation success.

Machine learning (ML) models provide a powerful solution to these challenges by analyzing complex, high-dimensional data and uncovering patterns that traditional methods cannot detect [60,64]. The use of unsupervised ML models focus on independently identifying IAEs, providing insights into their patterns. In contrast, supervised ML models directly compare the effectiveness of IAE classification between the VX1-software and the electrophysiologist [64]. This approach not only evaluates IAE classification effectiveness but also reveals underlying IAE patterns. These capabilities position ML models as an ideal tool to improve IAE classification and give insights in IAE complexity [23,60,65].

The aim of this study is to analyze the performance differences of existing ML models in IAE identification, focusing on the feature importances from both the VX1-software and the electrophysiologist. The electrophysiologist is regarded as the gold standard for identifying IAEs, as the ablation procedure depends solely on these IAEs. The objective of this study is:

- *To investigate the performances of existing machine learning models in differentiating between IAEs, based on the feature importances of both the VX1-software and the electrophysiologist.*

This objective aims to evaluate how existing ML models can enhance the development of more reliable and effective methods for identifying STD-based IAEs, ultimately improving the VX1-software. Additionally, it seeks to provide insights into distinct IAE patterns, further advancing the understanding and clinical management of AF.

The hypothesis is that the existing ML models will demonstrate superior performance in differentiating between IAEs when using the features derived from the VX1-software compared to the electrophysiologist. This is due to the standardized approach of the VX1-software, in contrast to the patient-specific analysis performed by the electrophysiologist, which may introduce greater feature variability. Additionally, the ML models are expected to uncover IAE patterns that will enhance the classification of STD-based IAEs, ultimately improving IAE identification for both the VX1-software and the electrophysiologist in AF ablation procedures.

3.3 METHOD

This study utilizes the same cohort of patients as referenced in the previous Chapter 2: "*Comparison of Spatio-temporal Characteristics of Intra-atrial Electrograms*". The dataset, comprising 24 features corresponding to each IAE, is reused here for further analysis. The subsequent sections, and Figure 8, provide a comprehensive overview of the methodology used in this study to enable the advanced processing and analysis of the data.

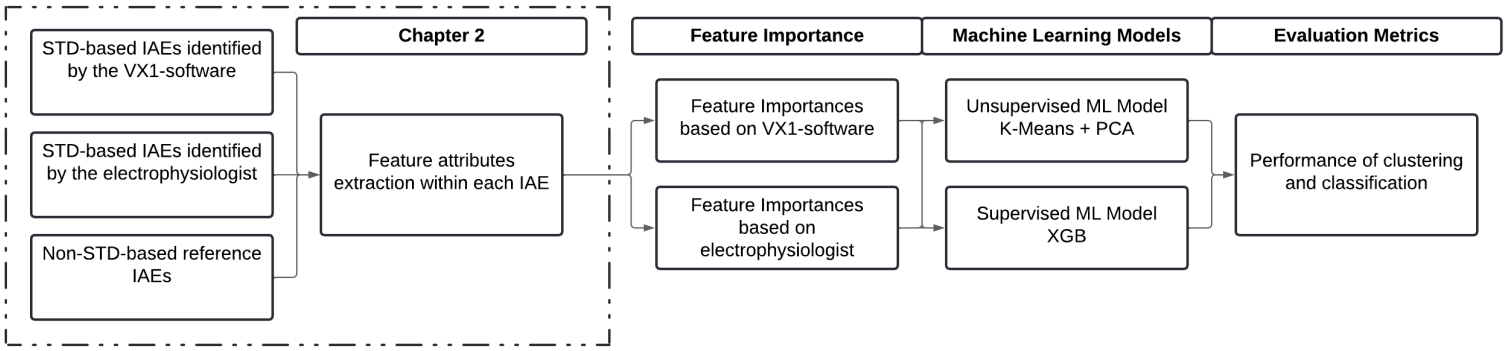


Figure 8: Overview of the methodology of this study. Abbreviations: XGB, extreme gradient boosting; PCA, principal component analysis.

3.3.1 FEATURE IMPORTANCE ANALYSIS

Feature importance analysis is essential for identifying the most relevant spatio-temporal features for IAE identification, enhancing ML model efficiency by reducing dimensionality while preserving key information [66,67]. Given that each IAE is represented in 24 features, this analysis examines the contribution of each feature to the predictive accuracy of the ML model. The goal is to provide insight in the most significant features for differentiating between STD-based and non-STD-based reference IAEs, thereby laying the foundation for the further development of both supervised and unsupervised ML models. This methodology is crucial for optimizing model performance and ensuring that the selected features are both highly relevant and impactful in improving predictive accuracy.

Feature importances scores are computed using an extreme gradient boosting (XGB) model, known for its effectiveness in complex, high-dimensional datasets [68]. The XGB model determines feature importance through an assessment in how much each feature contributes to reducing the loss, or error, across all iterations of the model. Features are ranked according to their contribution to predictive accuracy, utilizing metrics such as gain, coverage, and the frequency of splits in decision trees [68,69]. To further enhance model interpretability, Shapley additive explanations (SHAP) values are calculated to quantify the contribution of each feature to the output of the model for each individual prediction [67,69]. This offers a granular understanding of how specific features influence the outcome of the model, providing detailed insights into the decision-making process for each prediction, and complementing the global feature importance derived from the XGB model [69]. Features that significantly influence the output of the model are given higher importance scores, indicating their greater impact on predictive performance [68,69].

Additionally, this process incorporates cross-validation with 250 folds to validate the robustness and reliability of feature selection [67,68]. Cross-validation minimizes the risk of overfitting, enhancing the generalizability of the selected features to new, unseen data. This ranking of features helps pinpoint the most relevant variables for further analysis, guiding the identification of a smaller, more efficient subset of features for subsequent modeling. The

subset of selected features not only improve the interpretability of the model, but also optimize computational efficiency, resulting in faster training and classification of IAEs [68,69].

Thus, the process of selecting the most important features based on this analysis helps to eliminate redundant or less informative features, thereby reducing the risk of overfitting and improving overall model performance [68,69]. Features are ranked separately in two ways: (1) according to the VX-STD and REF IAEs, and (2) based on the EP-STD and REF IAEs. This results in two distinct feature importance rankings, which form the basis for both subsequent supervised and unsupervised ML models designed to differentiate between STD-based and non-STD-based IAEs.

3.3.2 UNSUPERVISED MACHINE LEARNING MODEL

The unsupervised ML approach in this study combines principal component analysis (PCA) followed by K-Means clustering to differentiate between STD-based and non-STD-based IAEs while also identifying distinct patterns within the IAEs. PCA is employed as a dimensionality reduction technique to transform high-dimensional data into lower-dimensional representation while retaining the most significant variance [30,60,66]. The transformation of the data onto a new set of orthogonal principal components reduces feature redundancy, enhances computational efficiency, and improves interpretability. This transformation is particularly useful for visualizing complex datasets, mitigating noise, and improving clustering performance [60,66].

Following PCA, the K-Means clustering algorithm is applied to group IAEs based on their shared spatio-temporal characteristics [30,64,66]. K-Means iteratively assigns data points to the nearest cluster center, known as centroids, and updates the centroids based on the mean of assigned points. This process continues until convergence, where the cluster assignments stabilize [30,64]. While the primary objective is to distinguish between STD-based and non-STD-based reference IAEs, multiple clusters are also used to capture patterns within all the IAEs. This allows for the identification of potential subtypes or variations in IAE characteristics, providing a more comprehensive understanding of IAE morphology. The selection of the optimal number of clusters *(k)* is guided by evaluation metrics such as the silhouette analysis, as explained in subsequent section, to ensure meaningful and interpretable groupings [30,49].

The combination of PCA and K-Means effectively reduces dimensionality while preserving essential information, enabling the identification of both broad IAE classes (STD-based vs. non-STD-based reference IAEs) and underlying patterns within the IAEs [30,60,64]. This enhances the ability of the model to distinguish between IAE classes, improving both the performance and interpretability of clustering.

3.3.3 SUPERVISED MACHINE LEARNING MODEL

The supervised ML model employed in this study is XGB, a state-of-the-art gradient boosting algorithm known for its high efficiency, scalability, and predictive accuracy [70,71]. XGB is an ensemble learning approach, where multiple weak learners, typically decision trees, are sequentially combined to reduce prediction errors. Each successive tree is trained to correct the errors made by the previous tree, improving the accuracy of the model by minimizing a predefined loss function [71–73].

A major strength of XGB model lies in its built-in regularization techniques, which prevent overfitting by penalizing excessively complex models. This regularization contributes to the robustness of the model and ability to generalize well to unseen data. Moreover, XGB enhances computational efficiency by supporting parallelization during both the tree construction and evaluation processes, making it particularly effective for handling large datasets. Additionally, XGB employs a sophisticated tree pruning mechanism, which removes

unnecessary branches from the decision trees to reduce overfitting, ultimately improving the generalization of the model [71]. The core process of the XGB model involves adding decision trees sequentially, where each tree is weighted by a learning rate. This learning rate controls the contribution of each tree, allowing the model to gradually adjust predictions and avoid overfitting by refining them over time. During the training process, the model minimizes a loss function while simultaneously applying regularization to manage model complexity [70,72].

To ensure optimal performance, key hyperparameters of the XGB model, such as learning rate, maximum tree depth, number of trees, and regularization parameters, are carefully tuned. Hyperparameter optimization is essential for achieving the best performance, as these parameters directly influence the accuracy and efficiency of the model [67,71]. Once tuned, this model is used to differentiate between VX-STD or EP-STD IAEs, and REF IAEs.

3.3.4 Evaluation Metrics

In this study, both unsupervised and supervised ML models use different evaluation metrics to assess their performance. The following sections provide a description of these evaluation metrics.

Unsupervised Machine Learning Model

In evaluating the performance of the K-Means model, the understanding of how well the model groups similar data points, or IAEs, is a critical step [74,75]. Several metrics have been developed to quantify the quality of clustering without relying on labeled data. The used evaluation metrics are the Silhouette Score (S-Score), Davies-Bouldin Index (DBI), and Adjusted Rand Index (ARI), each of which offers different insights into clustering quality [75].

The S-Score is a widely used metric that measures both the cluster cohesion, indicating how closely data points within a cluster are grouped, and cluster separation, reflecting how well one cluster is differentiated from another [75,76]. The S-Score ranges from -1 to 1, where a score close to 1 signifies well-clustered points that are near their own centroid and far from the decision boundary, while a score near -1 indicate poor clustering, with points assigned to a wrong cluster [74,75]. Besides, the S-Score is valuable for selecting the optimal number of clusters, offering a clear and interpretable evaluation of the clustering configuration. [30,74].

Furthermore, the DBI is another metric that measures the cohesion and separation of clusters. However, unlike the S-Score, it quantifies the average similarity between each cluster and its most similar counterpart [74,75]. This similarity is computed as the sum of within-cluster scatter, which is the average distance between data points and their cluster centroid, to the distance between cluster centroids [75]. In this context of distinguishing between (non-)STD-based IAEs, the DBI provides insight into how well these two types of IAEs are grouped. A higher DBI suggests poor clustering, indicating that the IAEs are not well distinguished, while a lower DBI signifies better clustering, where IAEs are compact within their clusters and well-separated from one another [75,76].

Moreover, the ARI measures the similarity between two clustering results while adjusting for random chance [74]. In this study, the ARI is used to evaluate the effectiveness of K-Means in grouping IAEs into STD-based or non-STD-based reference IAE clusters. The ARI corrects for the fact that random clustering can result in high similarity scores despite a lack of meaningful structures [74]. The ARI ranges from -1 to 1, where a value of 1 indicates perfect agreement between the two groupings, 0 suggests no more agreement than random chance, and -1 implies that the groupings disagree more than would be expected by chance [74,75].

Together, these metrics provide a comprehensive toolkit for assessing the clustering performance of K-Means, as unsupervised ML model. This multi-metric approach ensures a

robust evaluation of clustering performance and provides valuable insights into the differences between the VX1-software and the electrophysiologist.

SUPERVISED MACHINE LEARNING MODEL

Evaluating the performance of the XGB model is essential for understanding how effectively it distinguishes between different labels, specifically STD-based and non-STD-based reference IAEs. Several key metrics are used to evaluate classification performance, including the F1-score, area under the receiver operating characteristic (AUC-ROC) curve, and area under the precision-recall (AUC-PR) curve. These metrics are especially important for imbalanced datasets, such as in our study, which includes less STD-based IAEs compared to non-STD-based reference IAEs [77–79].

The F1-score is a metric that measures the harmonic mean of precision and recall, providing a single value that balances both [77]. Precision evaluates how many of the predicted positive instances are actually true positives, while recall focuses on how many of the actual positive instances are correctly identified by the model [77,78]. The F1-score ensures that a model maintains a balance between these two metrics, avoiding situations where one is optimized at the expense of the other. A higher F1-score indicates better overall performance, with a score of 1 representing perfect accuracy and 0 indicating the worst possible result [77].

Additionally, the AUC-ROC measures the ability of the model to differentiate between positive and negative classes, like STD-based and non-STD-based reference IAEs, independent of the classification threshold. The ROC-curve plots the true positive rate (sensitivity) against the false positive rate (1-specificity) from where the AUC-ROC is calculated [77,78]. In cases of the imbalanced dataset, the AUC-ROC offers a useful overall indication of model performance. An AUC-ROC of 1 indicate excellent discrimination between the two classes of IAEs, while an AUC-ROC of 0.5 suggest that the model is not performing better than random guessing [77,78].

Besides, the AUC-PR measures the ability of the model to correctly classify instances of the minority class, like STD-based IAEs, making it particularly valuable the imbalanced dataset. The PR-curve plots the precision against the recall, from where the AUC-PR is calculated [79]. The AUC-PR often provides more useful information than the AUC-ROC, as it focuses on the effectiveness of the model in identifying the minority class. A higher AUC-ROC closer to 1 indicates better performance in identifying the positive minority class, while a score of 0.5 suggests random guessing [79].

Overall, these metrics provide a comprehensive evaluation of the performance of the XGB model, particularly in the presence of class imbalance, ensuring an accurate assessment of its ability to classify both the majority (non-STD-based IAE) and minority (STD-based IAE) class.

SPATIO-TEMPORAL CHARACTERISTIC METRICS

As described in Chapter 2: "*Comparison of Spatio-temporal Characteristics of Intra-atrial Electrograms*", the distribution of the spatio-temporal characteristics of the IAEs are visualized using boxplots, and the statistical methods, including the Kruskal-Wallis test, Dunn's test, and Bonferroni correction, were applied to assess significant differences between the IAEs. These same methods are used in this study to ensure reliable conclusions regarding the differences between the IAE clusters, taking into account the non-normal distribution of the data.

3.4 RESULTS

The following subsections evaluate the feature importances of both VX1-software and electrophysiologist, the patterns of IAEs, and ML model performances, with separate analyses for the VX1-software and the electrophysiologist. The used dataset, comprising 24 features corresponding to each IAE, consist of 2792 VX-STD IAEs, 2516 EP-STD IAEs, and 11614 REF IAEs.

3.4.1 FEATURE IMPORTANCES

In Figure 9, the feature importances, based on SHAP values, are visualized for all spatio-temporal characteristics derived from both VX-STD and EP-STD IAEs. The feature importances based on the VX1-software indicate that the most important features are the ICL, SCI and SKE, with corresponding SHAP values of 0.203, 0.095 and 0.072, respectively. On the contrary, the feature importances based on the electrophysiologist highlight the WSI, ICL,

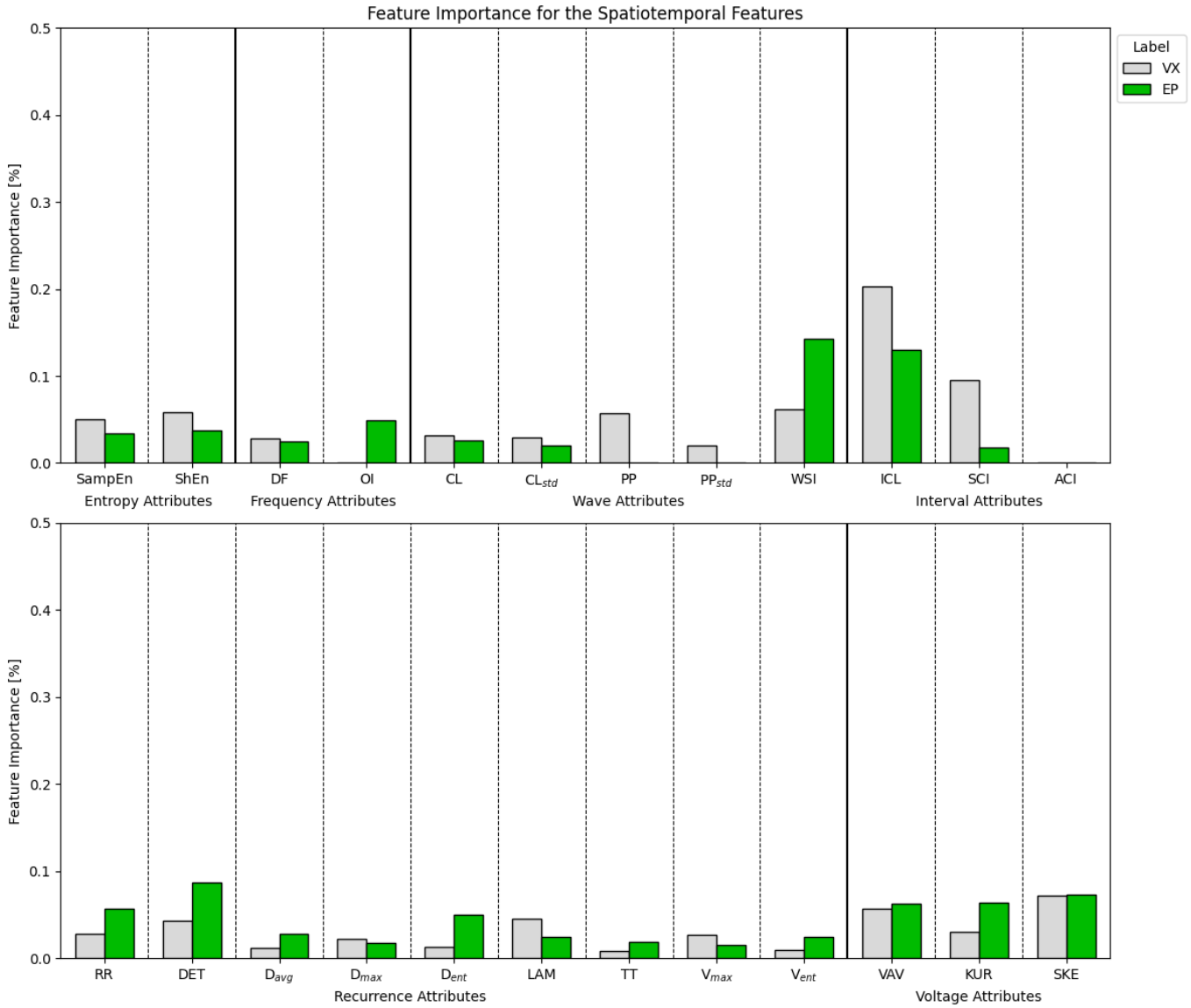


Figure 9: Feature Importances based on SHAP values for each spatio-temporal characteristic of the IAE. The label 'VX' means that the feature importance is based on the STD-based IAEs identified by the VX1-software, and the non-STD-based reference IAEs. The label 'EP' means that the feature importance is based on the STD-based IAEs identified by visual analysis of the electrophysiologist, and the non-STD-based reference IAEs. Note: Some characteristics are not used in the feature importances because there was no significant difference between the STD-based IAEs and non-STD-based reference IAEs. Abbreviations: EP, electrophysiologist; VX, VX1-software.

and DET as the most important features, with corresponding SHAP values of 0.143, 0.130, and 0.087, respectively. Overall, the feature importances are predominantly represented by the interval (29.9%), recurrence (20.5%), and wave (20.1%) attributes for the VX1-software, while for the electrophysiologist, the most influential attributes are recurrence (31.9%), voltage (19.8%), and wave (18.9%). A notable difference in feature importance between the VX1-software and the electrophysiologist is observed with the WSI, which holds more than twice the importance for the electrophysiologist compared to the VX1-software, and the SCI, which is more important for the VX1-software. Since no significant differences were observed for multiple features (OI, PP, PP_{std} and ACI) between the STD-based and non-STD-based reference IAEs in Chapter 2.4: *"Results"*, these features were excluded from the feature importances.

In Figure 10, the cumulative importance, based on SHAP values, of the including spatio-temporal features are visualized, highlighting the relative contribution of each feature to the overall performance of the model. The curve shows that the majority of the importance is concentrated in the top-ranked features, with a steeper initial rise in cumulative importance for both VX1-software and electrophysiologist. After these top features, the curve starts to plateau. Interestingly, the cumulative importance rises slightly more steeply for the feature importance of the electrophysiologist. Nonetheless, both approaches exhibit a similar pattern, crossing the cumulative importance thresholds of 90% and 95% after the same number of features (16 and 18 features), respectively. The cumulative importance threshold of 100% is reached after 21 features for the VX1-software and after 22 features for the electrophysiologist.

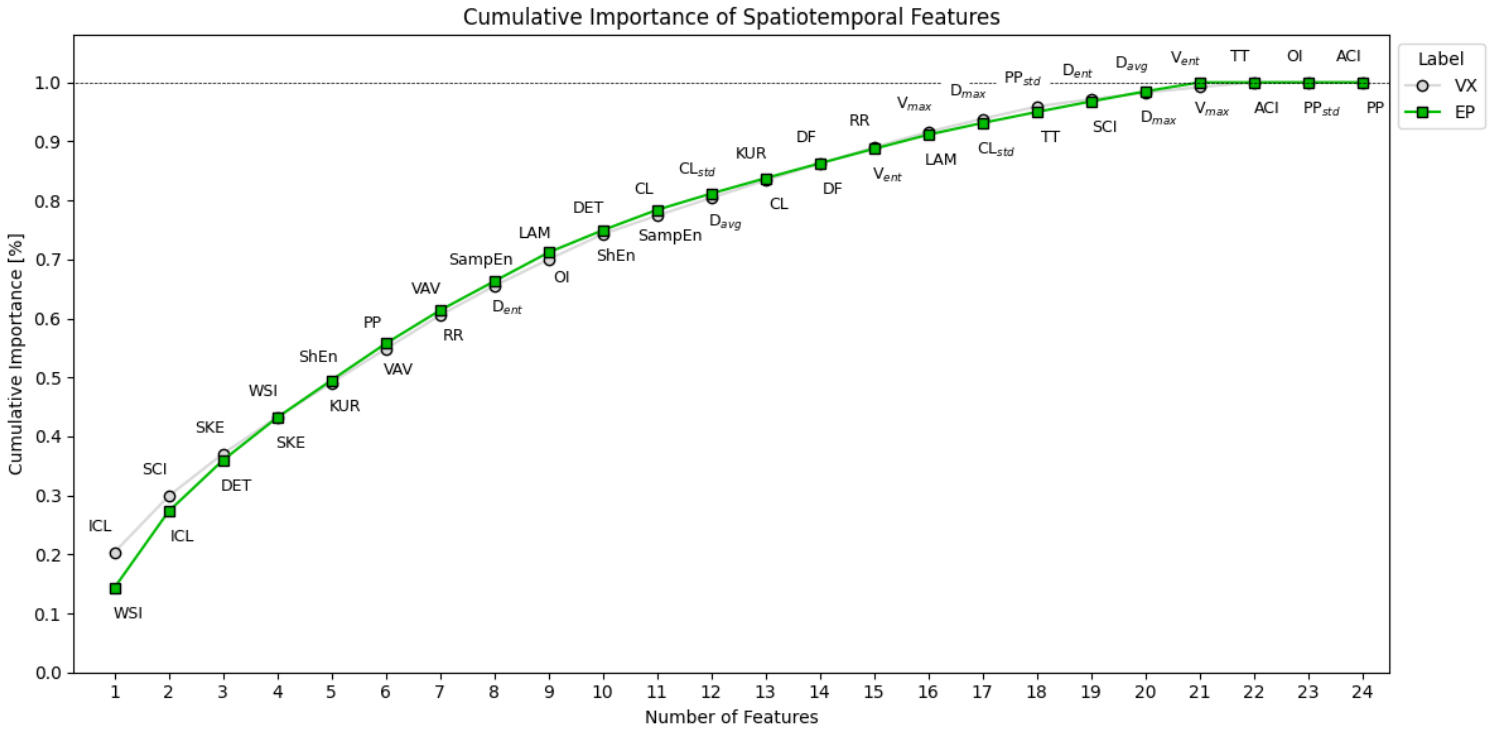


Figure 10: Cumulative importance of SHAP values of the included spatio-temporal features of the IAEs. The label 'VX' means that the feature importance is based on the VX-STD IAEs and the REF IAEs. The label 'EP' means that the feature importance is based on the EP-STD IAEs and the REF IAEs. Note: The feature names are displayed below the line for the electrophysiologist and above the line for the VX1-software, because of the overlap. Abbreviations: EP, electrophysiologist; VX, VX1-software.

3.4.2 PATTERNS OF INTRA-ATRIAL ELECTROGRAMS

In Figure 11, the optimal S-Score, DBI, and ARI values for different cluster configurations using K-Means are presented. These values are derived from the feature importance determined by the electrophysiologist and are solely clustered based on the EP-STD IAEs and the REF IAEs. The ARI reaches its maximum value of 0.301 at an optimal configuration of 6 clusters and 2 included features. However, the corresponding S-Score and DBI are 0.414 and 0.788, respectively, which are not at their optimal values. A local maximum of the ARI (0.288) is observed at 6 clusters and 5 features, with the DBI reaching its optimal minimum of 0.970 at corresponding number of features. Nevertheless, the S-Score at this configuration is 0.328, slightly lower than the maximum at 5 clusters (0.348). After including 5 features, the clustering performance reaches the highest ARI, with 6 clusters emerging as the most optimal cluster configuration based on both ARI and DBI.

In Figure 12, the distributions, consistency, and characteristics of IAEs across six clusters based on five selected features are shown. The STD-based IAEs are primarily distributed in clusters 1 to 3, accounting for 59.46%, 28.06%, and 9.38%, respectively, while non-STD-based reference IAEs contribute 3.21%, 28.00%, and 22.70% in these clusters. However, the distribution across clusters 4 to 6 shows that the STD-based IAEs account for 2.19%, 0.64%, and 0.28%, while the non-STD-based reference IAEs are more prominent, comprising 13.37%, 16.42%, and 16.30%, respectively. The cluster consistency is demonstrated by the distribution of STD-based and non-STD-based reference IAEs across all the clusters. Cluster 1 predominantly contains STD-based IAEs (80.04%), while the remaining clusters show lower proportions: 17.84%, 8.22%, 3.42%, 0.83%, and 0.37%, respectively. In contrast, clusters 5 and 6 are almost entirely composed of non-STD-based reference IAEs, with 99.17% and 99.63%, respectively. Cluster sizes vary, with clusters 2 and 3 containing the highest number of IAEs (3958 and 2872), while the remaining clusters each include approximately 1900 IAEs (1869, 1608, 1923, and 1900). The characteristics of the IAEs across the clusters show significant differences between the cluster characteristics for almost all clusters. Cluster 1 is characterized by a low WSI, high ICL, low DET, low SKE, and low KUR. Conversely, cluster 6 exhibits a high WSI, low ICL, high DET, high SKE, and high KUR. Clusters 2 to 5 display intermediate feature values, each following specific patterns. There is no significant difference in the ICL between clusters 2 and 5, 2 and 6, or 5 and 6. Similarly, no significant difference in the DET is found between clusters 5 and 6.

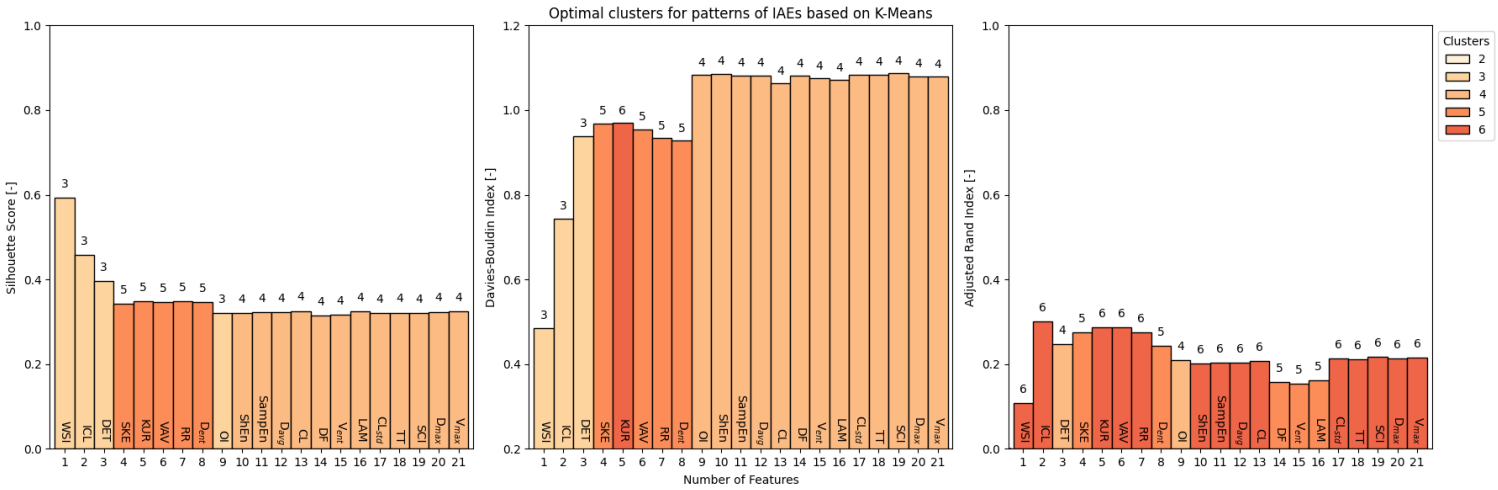


Figure 11: Clustering metrics (S-Score, DBI, and ARI) for different K-Means cluster configurations to identify patterns in IAEs. Note: For each feature, the maximum (S-Score and ARI) and minimum (DBI) values are indicated with the corresponding cluster configuration.

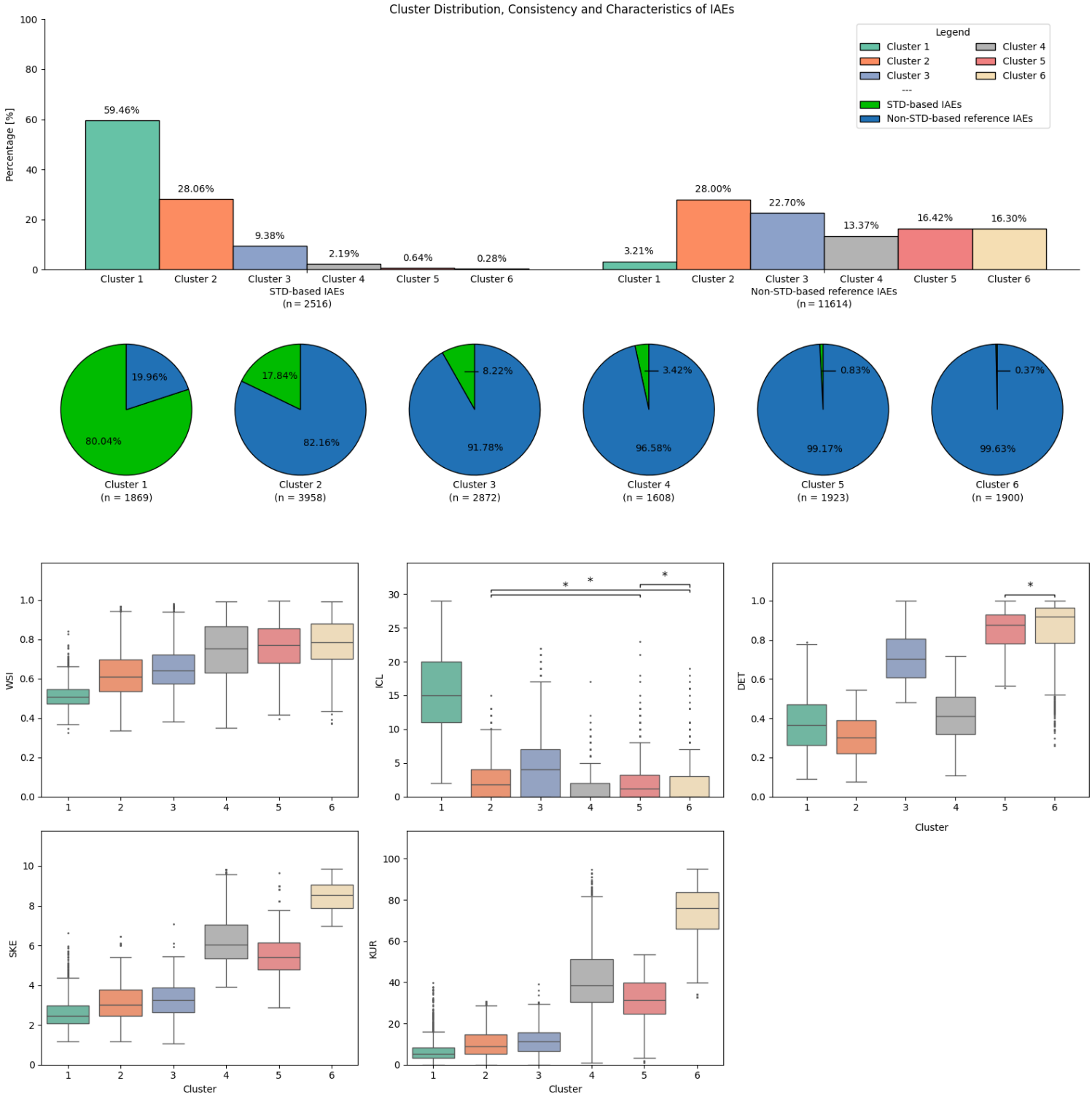


Figure 12: Distributions, consistency, and characteristics of the IAEs across the 6 clusters based on 5 included features. The 5 selected features correspond to the top-ranked features based on feature importance, which is calculated using SHAP values derived from STD-based IAEs by the electrophysiologist and non-STD-based reference IAEs. Note: No significant differences between the clusters are indicated at the top with (*) linked by brackets.

3.4.3 PERFORMANCE OF UNSUPERVISED K-MEANS MODEL

The performance of the unsupervised K-Means model is evaluated separately using the feature importance from both the VX1-software and the electrophysiologist. The clusters of the K-Means model are first determined using the VX-STD and REF IAEs, with EP-STD IAEs validated within these clusters. This process is then repeated using EP-STD and REF IAEs, validating the VX-STD IAEs in the resulting clusters.

PERFORMANCE OF K-MEANS BASED ON FEATURE IMPORTANCE OF THE VX1-SOFTWARE

In Figure 13, the optimal S-Score, DBI, and ARI values are presented for the specific cluster configuration *(k = 2)*. The ARI reaches its maximum value of 0.491 when only one feature (ICL) is included. The corresponding S-Score and DBI are 0.704 and 0.480, respectively, both representing their optimal values. Additionally, a local maximum in the ARI of 0.193 is observed at 7 included features, with a corresponding global minimum of 0.352 for the S-Score and a global maximum of 1.080 for the DBI. Beyond this point, including more features leads to an increase in the S-Score while the DBI and ARI continue to decrease. Consequently, with the two-cluster configuration, clustering performance reaches its optimal state, achieving the highest S-Score, lowest DBI, and highest ARI when only the ICL feature is included.

In Figure 14, the distribution and consistency of the clusters *(k = 2)* based on only the ICL feature are visualized. The VX-STD IAEs are predominantly distributed in cluster 2, representing 77.33%, while the remaining 22.67% are assigned to cluster 1. Similarly, the EP-STD IAEs are also mainly distributed in cluster 2 at 54.37%, while 45.63% are found in cluster 1. Conversely, the REF IAEs are predominantly distributed in cluster 1, accounting for the majority of instances. Specifically, 94.18% of the REF IAEs are assigned to cluster 1, while the remaining 5.82% are found in cluster 2. The cluster consistency indicates that cluster 1 consists of 4.98% VX-STD IAES, 9.03% EP-STD IAEs, and 86.00% REF IAEs. In contrast, cluster 2 is made up of 51.37% VX-STD IAEs, 32.55% EP-STD IAEs, and 16.08% REF IAEs. Additionally, cluster 1 contains 12,719 IAEs (75.16%), while cluster 2 contains 4,203 IAEs (24.84%).

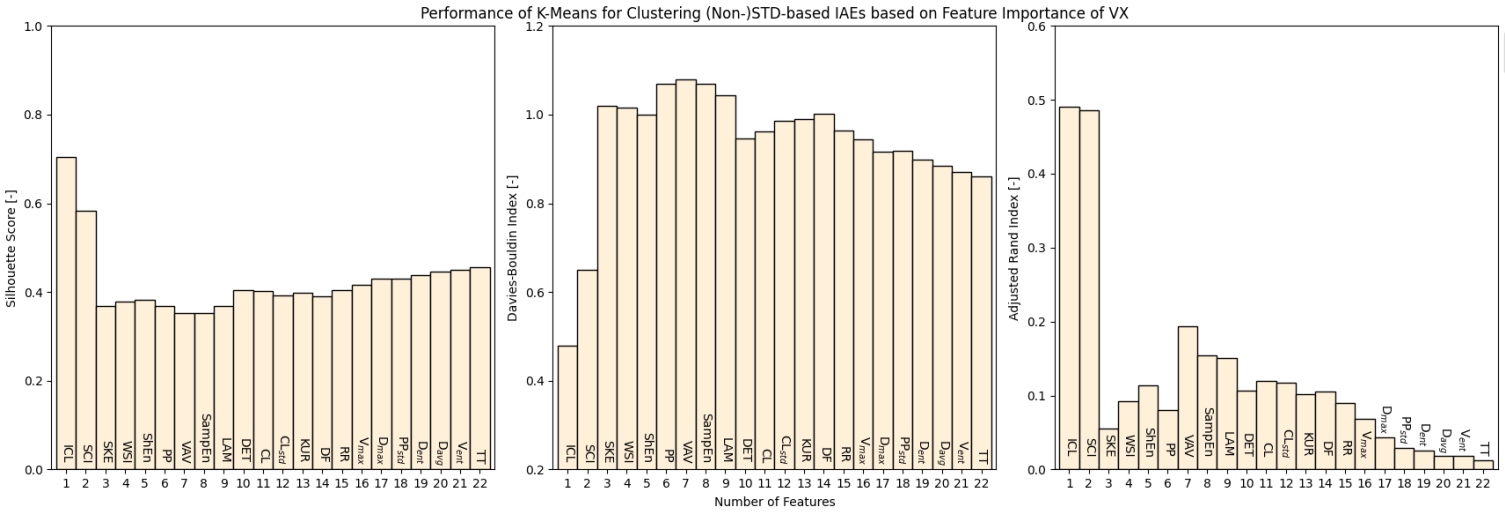


Figure 13: Clustering metrics (S-Score, DBI, and ARI) for the performance of K-Means with 2 clusters to distinguish between (non-)STD-based (reference) IAEs, based on the feature importance of the VX1-software. Abbreviations: VX, VX1-software.

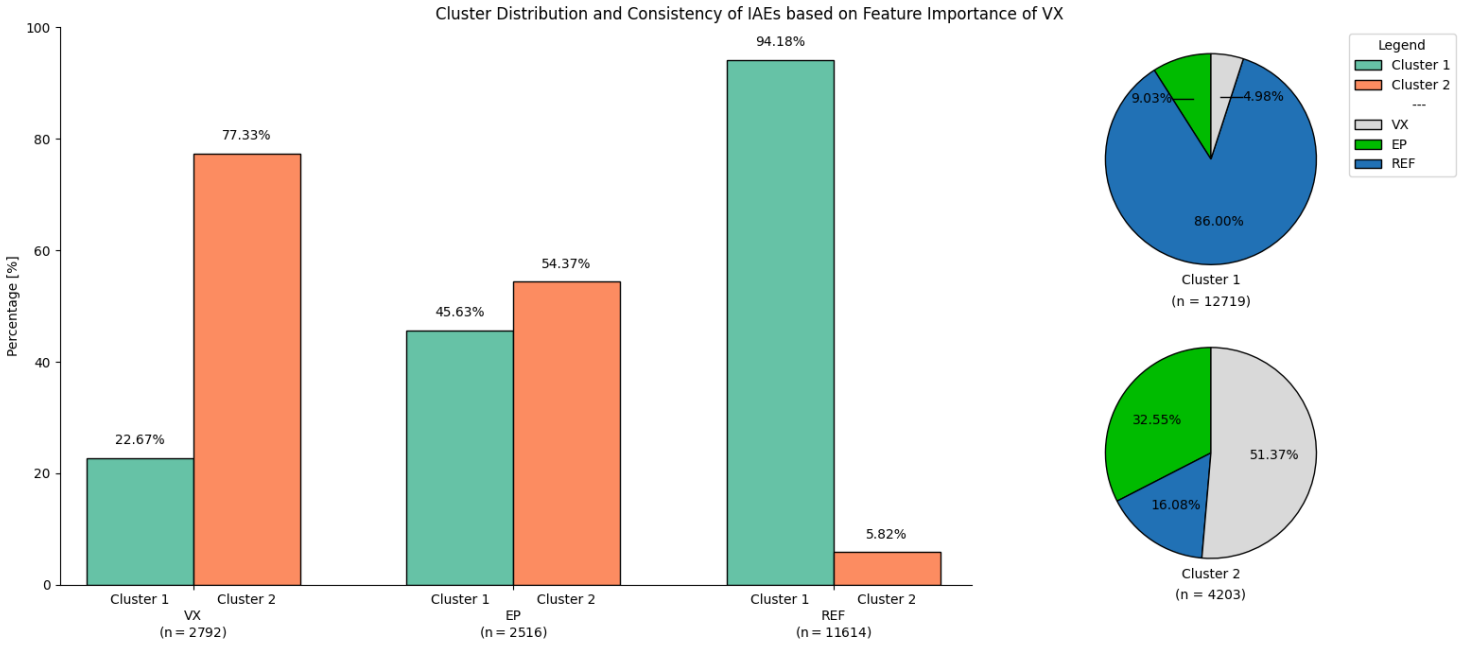


Figure 14: Distribution and consistency across the two clusters based on including only one feature (ICL). Abbreviations: EP, STD-based IAEs by the electrophysiologist (EP-STD IAEs); REF, non-STD-based reference IAEs (REF IAEs); VX, STD-based IAEs by the VX1-software (VX-STD IAEs).

PERFORMANCE OF K-MEANS BASED ON FEATURE IMPORTANCE OF THE ELECTROPHYSIOLOGIST

In Figure 15, the optimal performance metrics, including S-Score, DBI, and ARI, are shown for a specific cluster configuration (k = 2). The S-Score remains relatively stable around 0.420, while both the DBI and ARI follow a similar pattern. The ARI starts at 0.027 when only the WSI feature is included, coinciding with the global maximum of the S-Score (0.613) and the global minimum of the DBI (0.532). The ARI reaches its highest value of 0.121 when two features are included, with a corresponding S-Score of 0.387 and DBI of 0.903. The clustering performance achieves the optimal state with the two-cluster configuration, yielding the highest ARI and relatively a low DBI when both the WSI and ICL features are included.

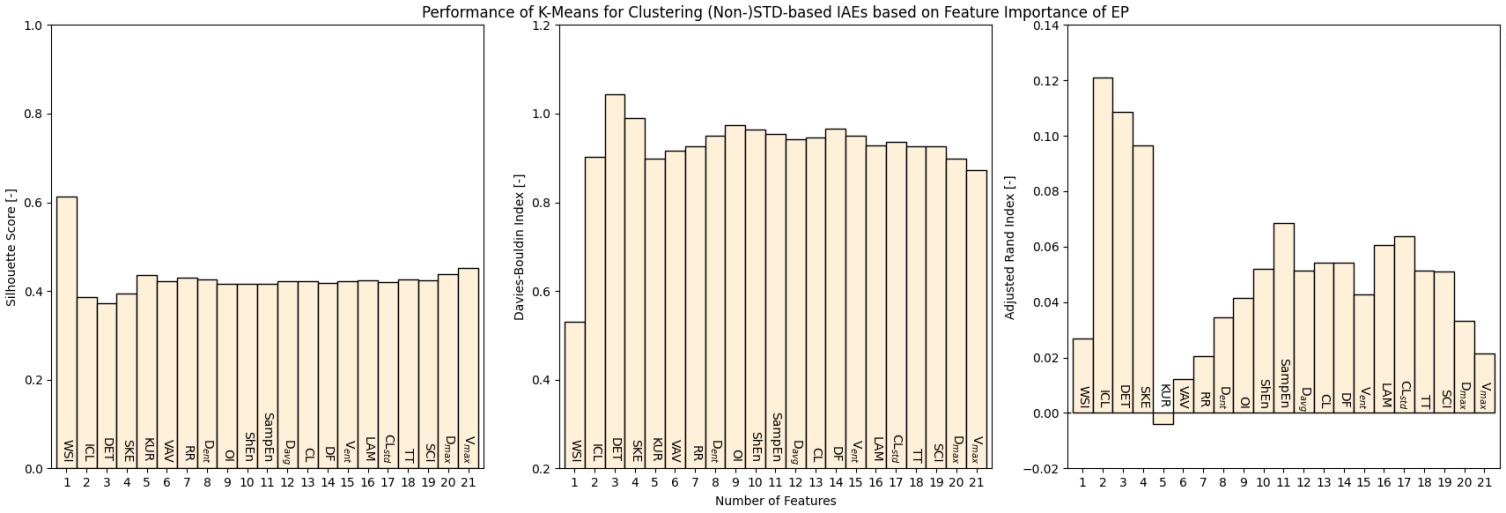


Figure 15: Clustering metrics (S-Score, DBI, and ARI) for the performance of K-Means with 2 clusters to distinguish between (non-)STD-based (reference) IAEs, based on the feature importance of the electrophysiologist. Abbreviations: EP, electrophysiologist

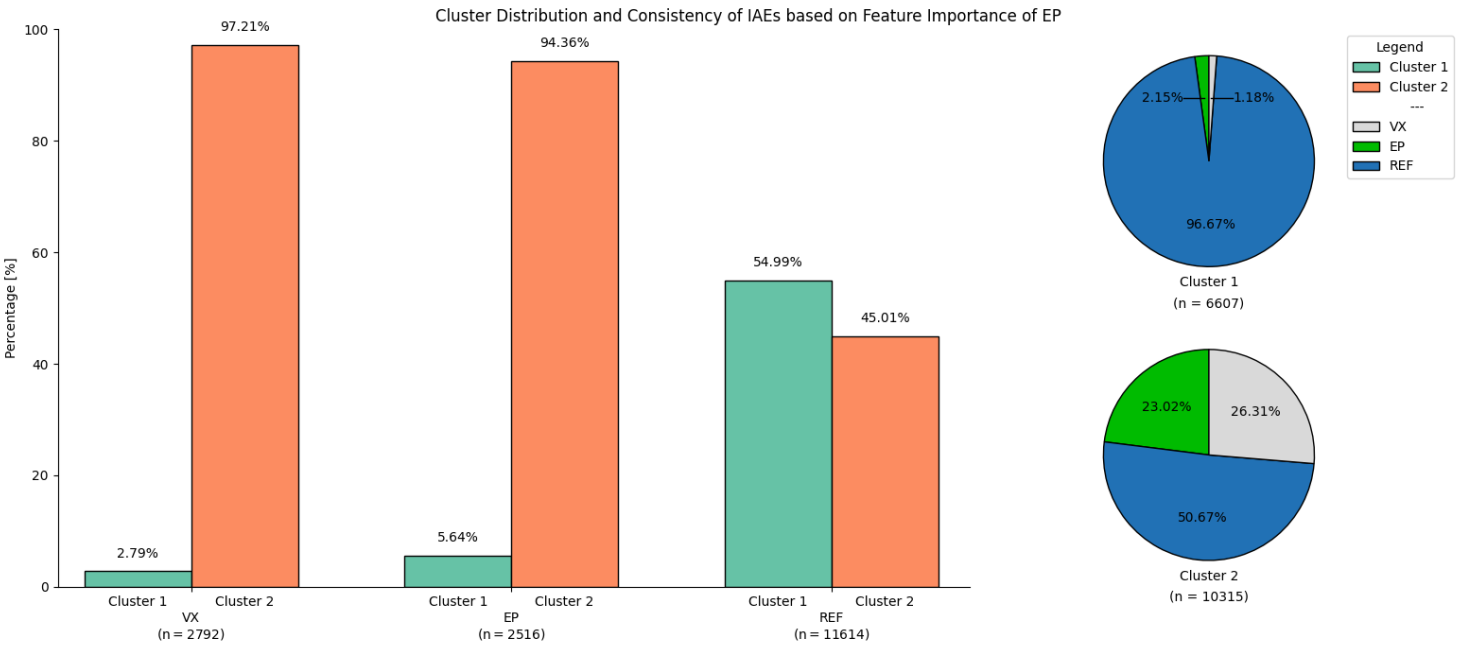


Figure 16: Distribution and consistency across the two clusters based on including two features (WSI and ICL)). Abbreviations: EP, STD-based IAEs by the electrophysiologist (EP-STD IAEs); REF, non-STD-based reference IAEs (REF IAEs); VX, STD-based IAEs by the VX1-software (VX-STD IAEs).

Figure 16, the distribution and consistency of the clusters (k = 2) based on the WSI and ICL features are visualized. The VX-STD and EP-STD IAEs are mostly assigned to cluster 2, comprising 97.21% and 94.36%, respectively, with the remaining 2.79% and 5.64% allocated to cluster 1. In contrast, the REF IAEs are predominantly placed in cluster 1, accounting for 54.99%, while 45.01% are distributed in cluster 2. The cluster consistency reveals that cluster 1 contains 1.18% of VX-STD IAEs, 2.15% of EP-STD IAEs, and 96.67% of REF IAEs. Conversely, cluster 2 consists of 26.31% of VX-STD IAEs, 23.02% EP-STD IAEs, and 50.67% of REF IAEs.

3.4.4 PERFORMANCE OF SUPERVISED EXTREME GRADIENT BOOSTING MODEL

The performance of the supervised XGB model is also evaluated separately using the feature importance from both the VX1-software and the electrophysiologist. The XGB model is initially trained and validated on VX-STD and REF IAEs, then evaluated using EP-STD IAEs. The process is then repeated with training on EP-STD and REF IAEs, followed by evaluation using VX-STD IAEs.

PERFORMANCE OF XGB BASED ON FEATURE IMPORTANCE OF VX1-SOFTWARE

In Figure 17, the F1-score, AUC-ROC, and AUC-PR values are shown for the performance of the XGB model, based on the feature importances from the VX1-software. A similar trend is seen across the F1-score, AUC-ROC, and AUC-PR values as more features are incorporated. The AUC-PR starts at 0.927 with the inclusion of one feature, ICL, and increases as additional features are added, reaching a local maximum of 0.930 when two features are used. This local maximum is also observed in the F1-score (0.852) and AUC-ROC (0.883). The optimal performance of the XGB model occurs with an F1-score of 0.850, AUC-ROC of 0.883, and AUC-PR of 0.944 when 17 features are included.

In Figure 18, the classification distribution and consistency for the IAEs are presented based on the feature importance from the VX1-software, reflecting the optimal classification results of the XGB model. The VX-STD IAEs are classified as STD-based IAEs in 78.08%, compared to 46.46% of the EP-STD IAEs and 0.79% of the REF IAEs. On the other hand, 21.92% of the VX-STD IAEs are classified as non-STD-based reference IAEs, compared to 53.54% of the EP-STD IAEs and 99.21% of the REF IAEs. The IAEs labeled as non-STD-based reference IAEs by the XGB model (*n=4381*) consist of 3.49% VX-STD IAEs, 30.75% EP-STD IAEs, and 65.76% REF IAEs. In contrast, the IAEs labeled by the XGB model as STD-based IAEs (*n=1737*) include 31.38% VX-STD IAEs, 67.30% EP-STD IAEs and 1.32% non-REF IAEs.

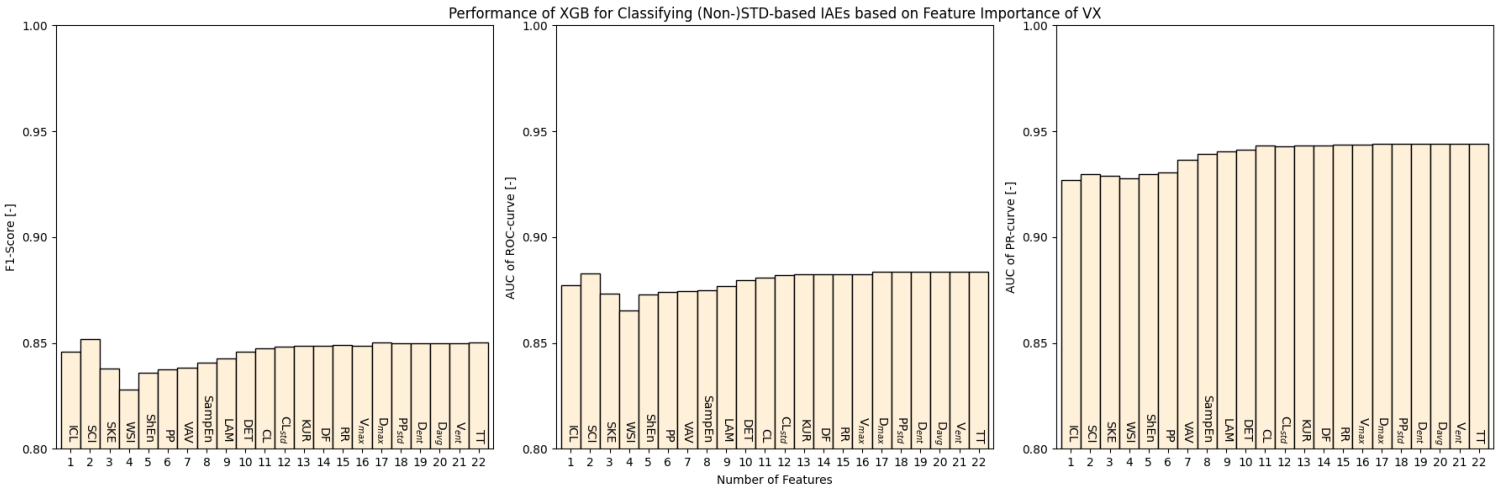


Figure 17: Classification metrics (F1-score, AUC-ROC and AUC-PR), for the performance of the XGB model in classification of the (non-)STD-based (reference) IAEs, based on the feature importance of the VX1-software. Abbreviations: VX, VX1-software.

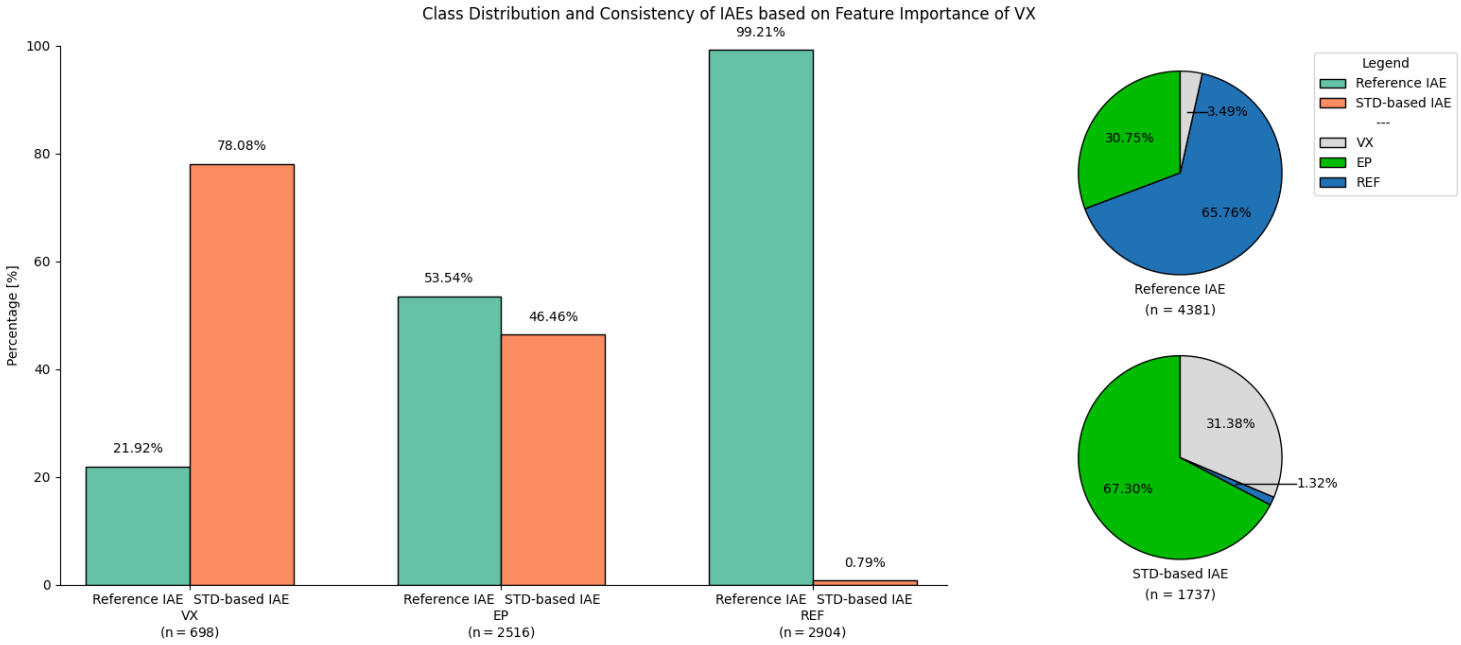


Figure 18: Distribution and consistency of the two classes (STD-based and non-STD-based reference IAE) classified by the XGB model, based on the feature importance of the VX1-software for 17 included features. Abbreviations: EP, STD-based IAEs by the electrophysiologist (EP-STD IAEs); REF, non-STD-based reference IAEs (REF IAEs); VX, STD-based IAEs by the VX1-software (VX-STD IAEs).

PERFORMANCE OF XGB BASED ON FEATURE IMPORTANCE OF ELECTROPHYSIOLOGIST

In Figure 19, the F1-score, AUC-ROC, and AUC-PR values for the performance of the XGB model are visualized, based on the feature importances from the electrophysiologist. The metrics improve as more features are added, with the most notable increase occurring when the ICL feature is included in addition to the WSI feature. The WSI, which has the highest feature importance based on SHAP values, results in an XGB model performance with an F1-score of 0.008, AUC-ROC of 0.502, and AUC-PR of 0.500. The optimal performance of the XGB model, with global maxima, occurs when 14 features are included, yielding an F1-score of 0.676, AUC-ROC of 0.765, and AUC-PR of 0.843.

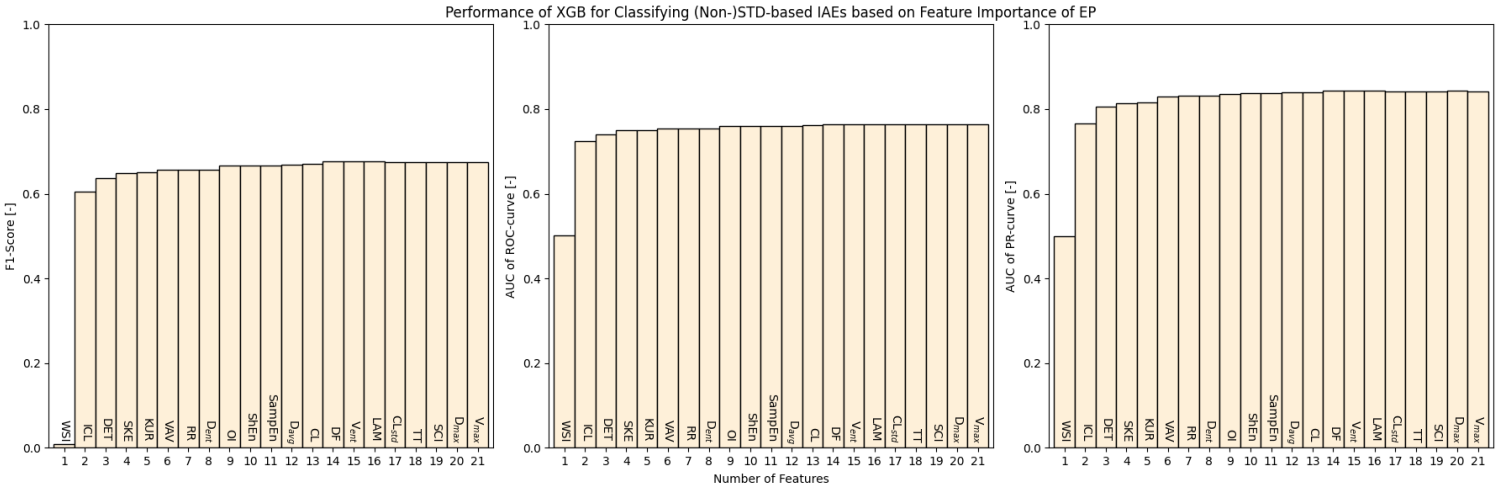


Figure 19: Classification metrics (F1-score, AUC-ROC and AUC-PR), for the performance of the XGB model in classification of the (non-)STD-based (reference) IAEs, based on the feature importance of the electrophysiologist. Abbreviations: EP, electrophysiologist.

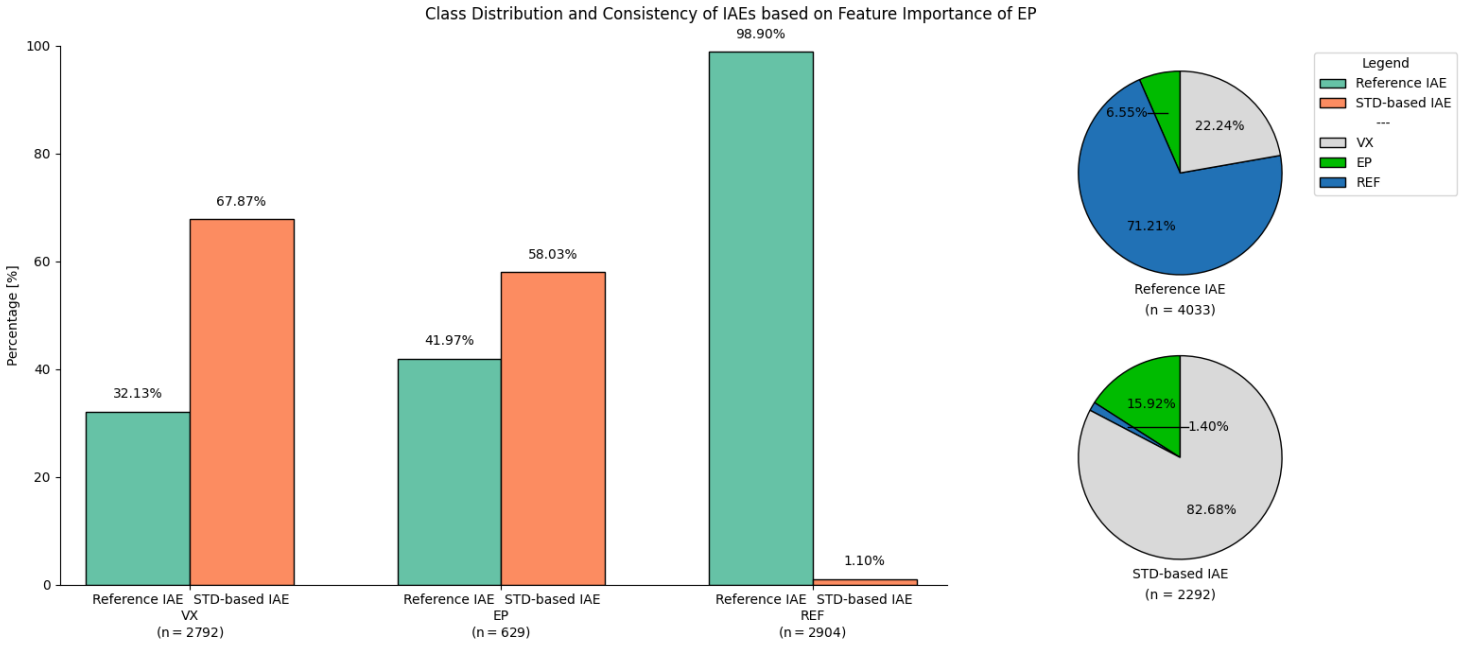


Figure 20: Distribution and consistency of the two classes (STD-based and non-STD-based reference IAE) classified by the XGB model, based on the feature importance of the electrophysiologist for 14 included features. Abbreviations: EP, STD-based IAEs by the electrophysiologist (EP-STD IAEs); REF, non-STD-based reference IAEs (REF IAEs); VX, STD-based IAEs by the VX1-software (VX-STD IAEs).

In Figure 20, the classification distribution and consistency for the IAEs are presented based on the feature importance from the electrophysiologist, reflecting the optimal classification results of the XGB model. The VX-STD IAEs are classified as STD-based IAEs in 67.87%, while 58.03% of the EP-STD IAEs are categorized as STD-based IAEs, and 1.10% of the REF IAEs fall into this category. In contrast, 32.13% of the VX-STD IAEs classified as non-STD-based reference IAEs, compared to 41.97% of the EP-STD IAEs and 98.90% of the REF IAEs. The IAEs labeled as non-STD-based reference IAEs by the XGB model (n = 4033) consist of 22.24% VX-STD IAEs, 6.55% EP-STD IAEs, and 71.21% REF IAEs. Conversely, the IAEs labeled as STD-based IAEs by the XGB model (n = 2292) are composed of 82.68% VX-STD IAEs, 15.92% EP-STD IAEs, and 1.40% REF IAEs.

3.5 DISCUSSION

The discussion presents the findings along with study limitations, explores their clinical relevance, and concludes with recommendations for future research.

3.5.1 FINDINGS

The findings of the feature importances based on the SHAP values, shown in Figures 9 and 10, reveal notable differences in feature prioritization between the VX1-software and the electrophysiologist. The VX1-software prioritizes the ICL, followed by the SCI and SKE, whereas the electrophysiologist emphasizes the WSI, ICL and DET. Specifically, the VX1-software analyzes IAEs based on the degree of fractionation, high-frequency activation, and asymmetry of the local voltage distribution, while the electrophysiologist focuses on waveform repetitiveness, degree of fractionation, and the consistency of activation patterns. Nevertheless, there are also similarities, which show a comparable pattern in the cumulative importance of the features. These findings highlight the distinct methodologies of each approach in IAE classification, impacting feature interpretation and decision-making, while offering valuable insights into their strengths and potential improvements for clinical practice.

Additionally, the optimal clustering performance, as shown in Figures 11 and 12, suggests that six patterns of IAEs based on five features, named WSI, ICL, DET, SKE, and KUR, provide the most stable results, effectively balancing the ARI, DBI and S-Score. The results underscore the challenge of IAE identification, as both STD-based and non-STD-based reference IAEs appear across all patterns. However, the pattern with the highest concentration of STD-based IAEs is characterized by a low WSI, DET, SKE, and KUR, alongside a high ICL, indicating highly variable, disorganized conduction with fragmented and irregular activation patterns. The opposite pattern, predominantly composed of non-STD-based reference IAEs, is characterized by a high WSI, DET, SKE, and KUR, alongside a low ICL, indicating more organized but intermittently unstable conduction, with periodic high-amplitude bursts. These overlapping characteristics across patterns highlight the continuum of conduction properties rather than a strict dichotomy, reinforcing the complexities of differentiating IAEs. This underscores the need for a multi-faceted approach to IAE classification that incorporates various techniques and perspectives to better capture the nuances of IAE patterns.

Furthermore, the performance of the K-Means model, as shown in Figures 13-16, demonstrates that clustering based on the VX1-software results in higher ARI values, with S-Score and DBI values being comparable to the electrophysiologist. This suggests that the VX1-software ensures more consistent identification of STD-based IAEs, whereas the electrophysiologist introduces greater variability, making it more difficult for K-Means to distinctly separate clusters. Conversely, the ICL strengthens clustering accuracy, enhancing the ability of the K-Means model to separate STD-based and non-STD-based IAEs, resulting in higher ARI values for both methods and highlighting its importance in IAE identification. Moreover, the cluster distribution reveal that the K-Means model based on the VX1-software prioritizes specificity, clustering non-STD-based IAEs more distinctly, while the model based on the electrophysiologist prioritizes sensitivity, achieving better separation of STD-based IAEs. However, these findings suggest that K-Means clustering does not fully capture the complexity of the data, as more nuanced patterns exist within the IAEs, indicating that clustering may not be the optimal approach for IAE identification.

Besides, the performance of the XGB model, as shown in Figures 17-20, reveals that the classification results are better when based on the VX1-software compared to the electrophysiologist. The XGB model, based on the VX1-software, effectively identifies and classifies VX-STD IAEs as STD-based IAEs with minimal misclassification, demonstrating strong performance in IAE identification. Nevertheless, the majority of the EP-STD IAEs are

misclassified, indicating poor IAE identification by the electrophysiologist. Conversely, despite that the performance of the XGB model is lower based on the electrophysiologist, both VX-STD and EP-STD IAEs are mostly classified as STD-based IAEs, with better results for VX-STD IAEs. These findings suggest that the XGB model based on the VX1-software does not fully capture the complexity of EP-STD IAEs, pointing to more intricate patterns of STD-based IAEs. This further suggests that STD-based IAE identification may be patient-specific and challenging to generalize, as the VX1-software is designed to attempt.

3.5.2 LIMITATIONS

Despite the findings of this study highlighting the performance of existing ML models to differentiate IAEs, several limitations should be acknowledged. One limitation of this study is that the imbalance between STD-based and non-STD-based reference IAEs affects ML model performance in IAE identification but also makes it more challenging to identify distinct IAE patterns. The limited number of STD-based IAEs in the dataset restricts both K-Means and XGB model from effectively learning and generalizing meaningful IAE patterns due to reduced exposure to their characteristics. In the K-Means model, this imbalance particularly impacts clustering, as the algorithm tends to favor the majority class, which can result in poor separation between the two groups [80,81]. As a result, the model may struggle to form well-defined clusters that effectively distinguish STD-based from non-STD-based IAEs. Similarly, while the XGB model shows stronger classification capabilities, the imbalance may still cause misclassification of STD-based IAEs, affecting the overall effectiveness of IAE differentiation. The use of F1-score, AUC-ROC, and AUC-PR helps mitigate this issue by providing a balanced evaluation. However, additional techniques such as data resampling, weighting strategies, synthetic data generation, and incorporating more data from a wider range of patients could further enhance model performance and improve the reliability of IAE classification [81].

Moreover, another limitation of this study is that the calculation of feature importance using SHAP values involves averaging over multiple folds, which does not explicitly capture variability across folds. This approach provides a robust estimate of feature importance and helps reduce bias from class imbalances in individual folds, but it may also obscure inconsistencies in how certain features contribute to the performance of the model in specific data splits. Although WSI is considered as the most important feature for the electrophysiologist, the performance of both the K-Means and XGB models is limited when WSI is used alone. This suggests that the averaging process could obscure these inconsistencies, resulting in a less detailed understanding of how features perform across different subsets of data. Therefore, considering individual fold results or using techniques like feature selection or ensemble methods could better capture feature variability and provide a more nuanced understanding of their impact.

Additionally, the combination of PCA and K-Means clustering may result in the unintended removal of important features critical for IAE identification by PCA, particularly when complex, non-linear interactions exist. In addition to this, the K-Means model assumes spherical clusters of approximately equal size, which may not align with the inherent structure of the data, potentially leading to suboptimal clustering results for IAE identification. Together, these limitations may reduce the effectiveness of the overall model in accurately classifying IAEs. This limitation highlights the need for alternative dimensionality reduction or clustering techniques that can better capture complex feature relationships and more accurately reflect the true structure of the data.

3.5.3 CLINICAL RELEVANCE

Superior performance of the VX1-software in the identification of STD-based IAEs is important for providing a generalized method that complements the expertise of electrophysiologists in

identifying and characterizing IAEs. The findings from this study highlight the different methodologies between the VX1-software and the electrophysiologist in prioritization of features, offering valuable insights into clinical decision-making. However, the consistency of the important features, such as WSI, ICL, and SKE, reinforces the reliability of spatio-temporal markers critical for AF ablation. This consistency in feature prioritization and classification could help reduce variability in clinical diagnoses, leading to more standardized and reproducible treatment strategies for patients with arrhythmias.

Furthermore, the identified patterns of IAEs characterize STD-based IAEs as highly variable, with disorganized conduction and fragmented, irregular activation patterns, as also noted in literature [30,31,60,62]. Despite, there is a lot of overlap in characteristics between the patterns of IAEs, which underline the complexity of IAE identification. Additionally, the VX1-software emphasizes the degree of fractionation, high-frequency activation, and asymmetry of local voltage distribution, which aids in identifying specific patterns of STD-based IAEs. However, this focus may also lead to the exclusion of other potential STD-based IAEs that do not display these particular characteristics. This insight underscores the necessity of using a multi-faceted approach to IAE classification, combining both the VX1-software and electrophysiologist to account for the full spectrum of spatio-temporal characteristics.

Additionally, the better performance of the VX1-software in clustering and classification highlights its potential to refine IAE identification and assist in real-time arrhythmia classification, offering electrophysiologists a powerful decision support tool that reduces cognitive load and decision-making time. However, the performances also underscores the complexity of IAE identification by the electrophysiologist, pointing to more intricate patterns of STD-based IAEs. Therefore, it is important to acknowledge the challenges in fully capturing the nuances of patient-specific arrhythmic patterns, underscoring the need for ongoing refinement of the VX1-software to better accommodate variability across individual cases.

3.5.4 FUTURE PERSPECTIVES

Although current methodologies offer valuable insights, there are several challenges that need to be addressed in order to enhance the accuracy and clinical applicability of the VX1-software. Future research should focus on addressing the key challenge of dataset imbalance identified in this study, particularly the disproportionate representation of non-STD-based reference IAEs compared to STD-based IAEs. This imbalance may hinder the ability to accurately distinguish between the two types of IAEs, as models trained on imbalanced data are more likely to favor the majority class. Therefore, increasing the number of STD-based IAEs, either through the collection of additional clinical data or by employing techniques such as oversampling or synthetic data generation, will allow for a more accurate representation of their characteristics, thereby improving the ability to generalize and enhance classification accuracy.

Besides, a critical area for future improvement is the refinement of the VX1-software. The current version tends to classify STD-based IAEs in a relatively uniform way, which may miss the diversity of IAE patterns across different patients. Future research should investigate advanced techniques, such as combining machine and deep learning models, capable of identifying complex, non-linear feature interactions. These advancements would enable the VX1-software to more accurately capture the diverse spatio-temporal characteristics of IAEs, providing electrophysiologists with more precise and detailed information for guiding ablation procedures. Refining the VX1-software to detect a wider variety of patient-specific IAE patterns will enable clinicians to make more informed decisions, ultimately enhancing patient outcomes during AF ablation.

3.6 CONCLUSION

The study demonstrates the strengths of existing ML models in identifying patterns and classifying IAEs, as well as the generalization capabilities of the VX1-software in IAE identification. Conversely, the methodological differences in IAE identification suggest that while the VX1-software relies on the degree of fractionation, high-frequency activation, and asymmetry of the local voltage distribution, electrophysiologists focus on waveform repetitiveness, degree of fractionation, and consistency of activation patterns. Consequently, the VX1-software excludes other STD-based IAEs that do not exhibit these specific characteristics. Furthermore, the performance of the VX1-software underscore the complexity of IAE identification by the electrophysiologist, emphasizing the need for it to address more intricate STD-based IAE patterns through a refined, multi-faceted approach. Future research should focus on investigating patient-specific IAE patterns to improve the ability of the VX1-software to identify and differentiate IAE patterns, ultimately providing a more comprehensive and reliable tool for IAE identification in AF ablation procedures.

4. Assessing the Performance of the VX1-software: Standardized vs. Patient-specific Approach

4.1 Abstract

Introduction: A standardized approach to IAE differentiation is essential for effective generalized ablation strategies, and while the VX1-software enhances classification accuracy, its analysis with existing ML models reveals both strengths and limitations in capturing the complexities of STD-based IAEs. Optimizing outcomes across AF progression stages and IAE complexities requires balancing standardized algorithms with personalized insights, emphasizing the need for tailored approaches to ensure accuracy and clinical relevance.

Objective: This study compares the performance of the VX1-software, as standardized approach, in identifying STD-based IAEs across patients with varying IAE complexities.

Method: This study analyzes patient-specific complexities in IAE identification using a dataset of 52 patients, with 24 features for each IAE. Complexity scores are computed based on the mean and median absolute differences between STD-based and non-STD-based reference IAEs, incorporating feature importance weights. Patients are clustered using K-Means to evaluate the performance of the VX1-software across varying IAE complexities. The XGB model is applied for IAE classification, utilizing a leave-one-out cross-validation (LOO-CV) approach to account for cluster variability. Furthermore, discrepancies in intra-atrial areas are assessed by categorizing IAEs based on proximity and analyzing their impact on XGB model performance.

Results: Patient-specific complexity scores reveal the connection with the severity and progression of AF, providing valuable insights for tailoring treatment strategies. Higher complexity scores, indicating lower complexity, suggest that more standardized approaches may be effective, while lower scores, reflecting higher complexity, highlight the need for more advanced strategies. These scores also uncover key differences between STD-based and non-STD-based IAEs, enhancing IAE identification and informing ablation decisions. Additionally, discrepancies in intra-atrial areas between the VX1-software and the electrophysiologist point to limitations in identifying STD-based regions, underscoring the need for personalized approaches to optimize treatment and reduce AF recurrence.

Conclusion: The study emphasizes the need for personalized AF management, providing valuable insights into the severity and progression of AF. The VX1-software shows strong potential in IAE identification with minimal discrepancies in STD-based intra-atrial areas, but the correlation between lower complexity and higher discrepancies highlights the necessity for patient-specific approaches in IAE classification. Future refinement of the VX1-software, integrating advanced methods to improve the identification of complex IAE patterns, is crucial for optimizing ablation strategies and enhancing patient outcomes.

4.2 INTRODUCTION

A standardized approach to differentiate IAEs is essential for the development of effective, generalized ablation strategies targeting non-PV AF drivers. As the understanding of AF complexity continues to evolve, the ability to accurately classify IAEs becomes increasingly important for optimizing treatment outcomes. A consistent, standardized ablation method ensures reliable and reproducible results across diverse patient populations. The VX1-software offers a significant advantage in this area, enhancing IAE classification accuracy and consistency, which leads to higher success rates [20,24]. Nevertheless, analyzing the VX1-software with existing ML models demonstrated superior performance compared to the patient-specific approach of the electrophysiologist, highlighting the advantages of VX1-software while also emphasizing its limitations in fully capturing the complexities of STD-based IAEs.

A key finding of the TAILORED-AF study was the higher clinical success rate observed in a patient cohort where only 18% had longstanding persistent AF, a subgroup known for its more complex treatment needs [20,24,25,27]. Longstanding persistent AF, an advanced stage characterized by sustained irregular atrial activity lasting more than 12 months, induces significant structural and electrical remodeling of atrial tissue, including fibrosis, atrial enlargement and disrupted ion channel function [11,82]. These pathological changes create a highly challenging substrate for ablation, as the extensive fibrosis and STD of electrical signals interfere with standard ablation approaches [11]. However, this raises the possibility that IAE mimicking STDs could stem from different underlying pathologies, complicating the reliance on a generalized approach. As such, patients with longstanding persistent AF often require advanced, personalized strategies to target STD-based IAEs [11,25]. Therefore, further research is necessary to assess how effectively the VX1-software can identify STD-based IAEs in more complex patient populations. Achieving optimal outcomes across different stages of AF progression requires balancing standardized algorithms with personalized insights. This variability underscores the need for tailored approaches to ensure both accuracy and clinical relevance for each unique condition.

This study aims to evaluate the performance of the VX1-software in identifying STD-based IAEs in varying patient complexities. Specifically, the objective of this study is:

- *To assess the performance of the VX1-software, as standardized approach, in identifying STD-based IAEs across patients with varying IAE complexities.*

This objective assesses the effectiveness of the VX1-software compared to the patient-specific analysis of the electrophysiologist, focusing on discrepancies in STD-based areas and the performance of the VX1-software across patients with varying complexities of AF. It will also provide insights into how consistently and reproducibly the VX1-software identifies IAEs, highlighting the potential advantages of a standardized approach in contrast to the patient-specific analysis of the electrophysiologist.

The hypothesis is that, while the patient-specific expertise of the electrophysiologist offers superior accuracy in identifying STD-based IAEs, especially in patients with complex arrhythmic signals, the VX1-software provides a consistent and objective method that ensures reproducibility across diverse patient populations. By integrating the strengths of the standardized framework of the VX1-software with the nuanced, patient-specific approach of electrophysiologist, diagnostic precision can be significantly enhanced, resulting in more effective and reliable ablation strategies for AF patients.

4.3 METHOD

This study builds upon the cohort of patients described in Chapter 2: "*Comparison of Spatio-temporal Characteristics of Intra-atrial Electrograms*". The dataset, containing 24 features for each IAE, is reused for further analysis. Additionally, the supervised XGB model from Chapter 3: "*Comparative Evaluation of Existing Machine Learning Models for Intra-atrial Electrogram Differentiation*", is applied once again for this analysis. The following sections, along with Figure 21, offer a detailed explanation of the methodology used to facilitate advanced processing and analysis of the data.

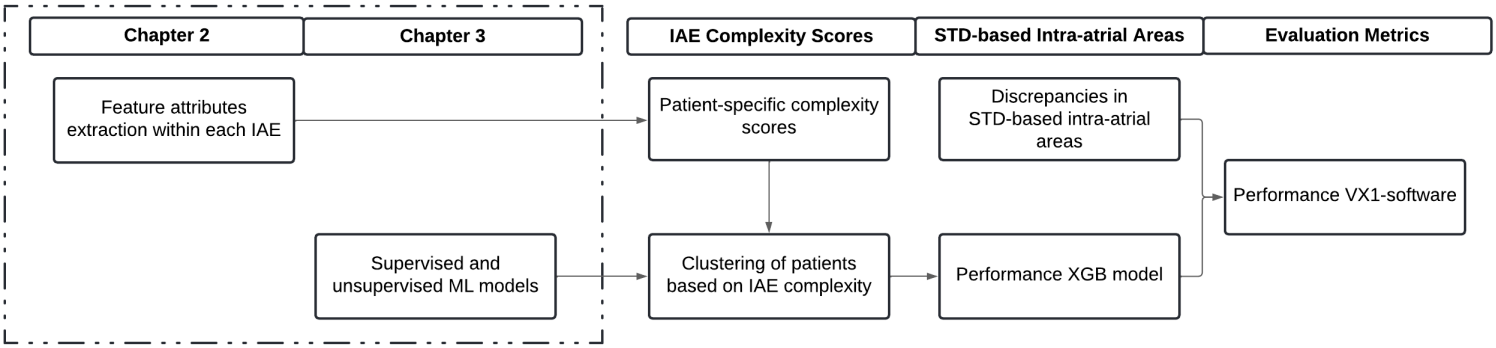


Figure 21: Overview of the methodology of this study. Abbreviations: IAE, intra-atrial electrogram; ML, machine learning; STD, spatio-temporal dispersion; XGB, extreme gradient boosting.

4.3.1 CLUSTERING OF PATIENTS ON PATIENT-SPECIFIC COMPLEXITY SCORES

Clustering patients based on IAE complexity levels enables a more detailed assessment of the performance of the VX1-software across a spectrum of complexities. The first step in this process involves computing complexity scores for each patient using the 24 spatio-temporal IAE characteristics.

To ensure comparability across patients, each extracted IAE characteristic was standardized. This was achieved using min-max normalization, which scaled each feature to a range between 0 and 1. The normalization process followed the formula:

$$Feature_{norm} = \frac{Feature_{value} - Feature_{min}}{Feature_{max} - Feature_{min}} \quad (1)$$

, where $Feature_{max}$ and $Feature_{min}$represent the maximum and minimum values of the feature across all patients, respectively, and $Feature_{norm}$ is the normalized value. This approach ensured uniformity in feature scaling, preventing potential biases caused by varying measurement units and allowing each spatio-temporal characteristic to contribute proportionately to the final complexity score.

After standardizing the IAE characteristics, the mean and median of each feature for the labeled STD-based and non-STD-based reference IAEs were calculated for each patient, along with the absolute difference in mean and median between the two groups. This allowed for a comparison between the two types of IAEs, providing insights into their distribution and central tendency. Additionally, the complexity score based on the mean and median values is calculated using the weights in feature importances, as outlined in Chapter 3: "*Comparative Evaluation of Existing Machine Learning Models for Intra-atrial Electrogram Differentiation*". The formulas for the weighted scores are as follows:

$$S_{mean,i} = w_1 * |(EP_{mean} - REF_{mean})|_1 + \cdots + w_n * |(EP_{mean} - REF_{mean})|_n \quad (2)$$

$$S_{median,i} = w_1 * |(EP_{median} - REF_{median})|_1 + \cdots + w_n * |(EP_{median} - REF_{median})|_n \quad (3)$$

, where $|(EP_{mean,median} - REF_{mean,median})|_n$ is the absolute difference in mean and median between the STD-based (EP-STD) and non-STD-based reference (REF) IAEs, w_n are the weights assigned to each spatio-temporal characteristic n, and $S_{mean,median,i}$ is the complexity score based on the mean or median for each patient i. Higher values indicate a larger discrepancy between the STD-based and non-STD-based IAEs, suggesting greater variation in patient-specific IAE characteristics and, therefore, less complex IAE identification. Conversely, lower values suggest a closer alignment between the two IAEs, reflecting a more uniform or similar IAE profile for the patient, and thus more complex IAE identification.

As a result, for each patient, two complexity scores were calculated, one based on the mean and the other on the median. These scores were then subjected to K-Means clustering, a technique used to group patients based on their similarities in IAE complexity. K-Means enabled the identification of distinct patient clusters, revealing patterns and relationships within the data and offering insights into different IAE profiles across the patient population. The evaluation of the clustering was based on the S-Score and the DBI for determining the optimal number of clusters (*k = 3*). For more detailed information, the methodology of the K-Means clustering is outlined in Chapter 3: "*Comparative Evaluation of Existing Machine Learning Models for Intra-atrial Electrogram Differentiation*".

4.3.2 EXTREME GRADIENT BOOSTING MODEL ANALYSIS

The evaluation of the VX1-software for STD-based IAE identification was conducted using the XGB model. This model was selected for its superior performance and capability to handle complex, high-dimensional data, as outlined in Chapter 3: "*Comparative Evaluation of Existing Machine Learning Models for Intra-atrial Electrogram Differentiation*". The XGB model is used to assess how the model would classify the VX-STD IAEs, which were identified by the VX1-software, by comparing the classifications of the model with the labeled VX-STD IAEs.

In each cluster, a leave-one-out cross-validation (LOO-CV) approach was applied. This indicates that the XGB model is trained and validated with cross-validation based on the IAEs of all other clusters, along with the non-STD-based reference IAEs from the excluded cluster. The model is then tested on the VX-STD IAEs of the excluded cluster. This hybrid approach ensures that the training data reflects the variability across the different patient clusters, preventing the model from overfitting to the specific characteristics of any single cluster. Incorporation of the non-STD-based reference IAEs from the excluded cluster ensures for better generalization of the model across diverse IAE types. Additionally, this methodology allows for a separate, direct comparison of the assessment performance between the VX1-software and the XGB model, providing insights into the relative accuracy and effectiveness of both approaches in identifying STD-based IAEs.

To assess the performance of the XGB model, evaluation metrics such as the F1-score, ROC-AUC, and PR-AUC were used, given the imbalanced nature of the dataset. The F1-score provides a balance between precision and recall, while ROC-AUC and PR-AUC evaluate the ability of the model to distinguish between the IAE classes. These metrics offer a comprehensive view of the performance, highlighting their effectiveness in handling the challenges of imbalanced data, with a more detailed explanation provided in the previous chapter.

4.3.3 SPATIO-TEMPORAL DISPERSION-BASED INTRA-ATRIAL AREAS

To examine the discrepancies in STD-based intra-atrial areas between the VX1-software and the electrophysiologist, and to assess how these discrepancies affect the performance of the

XGB model, the IAEs are classified into distinct groups based on specific criteria. The VX-STD IAEs are categorized with the following criteria:

- *Criteria I:* Presence of an VX-STD IAE within one centimeter of the surrounding intra-atrial area without EP-STD IAE.
- *Criteria II:* Presence of an VX-STD IAE within one centimeter of the surrounding intra-atrial area with more than two EP-STD IAEs.
- *Criteria III:* Presence of VX-STD IAE within one centimeter of the surrounding intra-atrial area with multiple VX-STD and EP-STD IAEs.

The EP-STD IAEs are categorized with the same criteria:

- *Criteria I:* Presence of an EP-STD IAE within one centimeter of the surrounding intra-atrial area without VX-STD IAE.
- *Criteria II:* Presence of an EP-STD IAE within one centimeter of the surrounding intra-atrial area with more than two VX-STD IAEs.
- *Criteria III:* Presence of EP-STD IAE within one centimeter of the surrounding intra-atrial area with multiple VX-STD and EP-STD IAEs.

The non-STD-based reference IAEs, known as REF IAEs, are categorized based on the following criterion:

- Criteria *I:* Presence of an REF IAE, within one centimeter of the surrounding intra-atrial area with no VX-STD or EP-STD IAEs present.

The one-centimeter threshold was selected based on clinical expertise, suggesting that this spatial proximity adequately captures the influence of neighboring intra-atrial areas without introducing excessive overlap. This ensures more accurate classification of IAEs while minimizing the effect of minor spatial variations that may not be clinically relevant. To measure the distance between IAEs, the geodesic distance is employed using the EAM provided by the CARTO™ 3 System. Unlike Euclidean distance, which measures straight-line distance, geodesic distance calculates the shortest path between two points on a curved surface, specifically the complex, irregular surface of the atrium. This approach accounts for the curvature of the atria, providing a more accurate representation of spatial relationships. The geodesic distance is computed using methods such as the fast marching method, which follows the intrinsic geometry of the atrial surface. This method is essential for accurately modeling the electrophysiological properties of the atrium, reflecting true anatomical distances and improving our understanding of spatial interactions between regions of electrical activity.

Consequently, it is crucial to examine how specific IAEs, under each criterion, influence the performance of the XGB model. This evaluation is carried out by analyzing the precision of the XGB model in classifying VX-STD IAEs as STD-based IAEs for each cluster. By providing insight into the ability of the XGB model to correctly identify these IAEs, it can be assessed how the discrepancies in the intra-atrial areas impact the accuracy in classification of the XGB model, offering a deeper understanding of its performance across different patient clusters. This approach helps highlight the strengths and weaknesses in distinguishing STD-based IAEs with non-STD-based reference IAEs, further refining the clinical applicability of the VX1-software.

4.4 RESULTS

The following subsections evaluate the patient-specific complexity scores, the performance of the XGB model on these patient-specific complexities, and the importance of the discrepancies in STD-based intra-atrial areas. The used dataset, comprising 24 features corresponding to each IAE, consist of 2792 VX-STD IAEs, 2516 EP-STD IAEs, and 11614 REF IAEs.

4.4.1 PATIENT-SPECIFIC COMPLEXITY SCORES

In Figure 22, the complexity scores based on both the mean and median values are visualized for each patient, alongside the specific clusters and the distribution of patients by type of AF. The optimal clustering result from the K-Means analysis is achieved with three clusters. Cluster 1 shows S_{mean} = [0,14] and S_{median} = [0,15], Cluster 2 has S_{mean} = [14,25] and S_{median} = [15,25], and Cluster 3 is characterized by S_{mean} = [25,40] and S_{median} = [25,40]. The distribution of patients across the three clusters varies by AF type, with 66.7% (n=10) of patients with longstanding persistent AF in Cluster 1, 13.6% (n=3) in Cluster 2, and 13.3% (n=2) in Cluster 3. For persistent AF, 20.0% (n=3) of patients are in Cluster 1, 40.9% (n=9) in Cluster 2, and 13.3% (n=2) in Cluster 3. Similarly, for paroxysmal AF, 13.3% (n=2) of patients are in Cluster 1, 45.5% (n=10) in Cluster 2, and 73.3% (n=11) in Cluster 3. The total number of patients in each cluster is as follows: Cluster 1 contains 15 patients, Cluster 2 contains 22 patients, and Cluster 3 contains 15 patients.

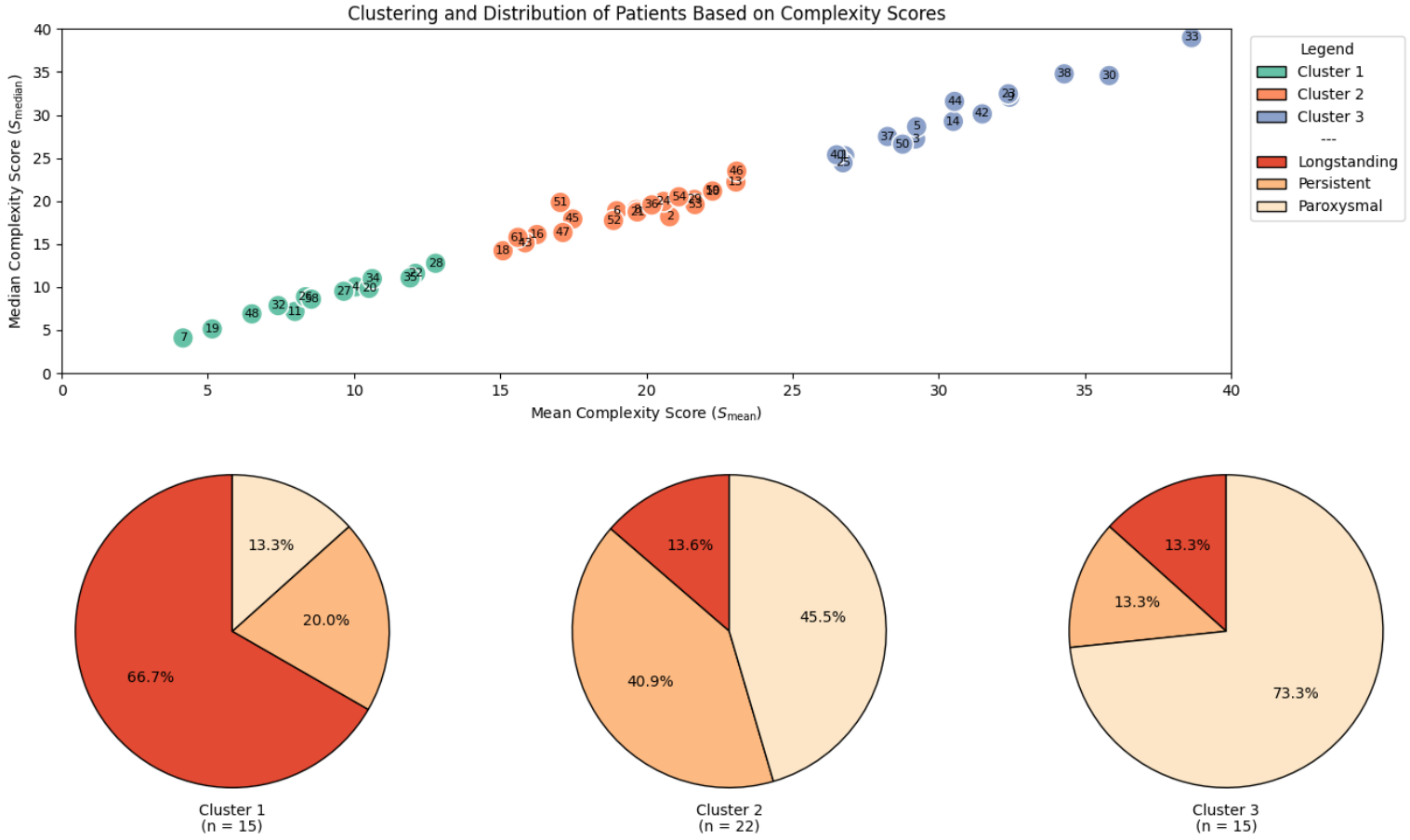


Figure 22: Complexity scores (S_{mean} and S_{median}) for each patient with clustering results and distribution on AF type: longstanding persistent, persistent, or paroxysmal. Note: In the top plot, the specific patient IDs are plotted at each point in the scatter plot.

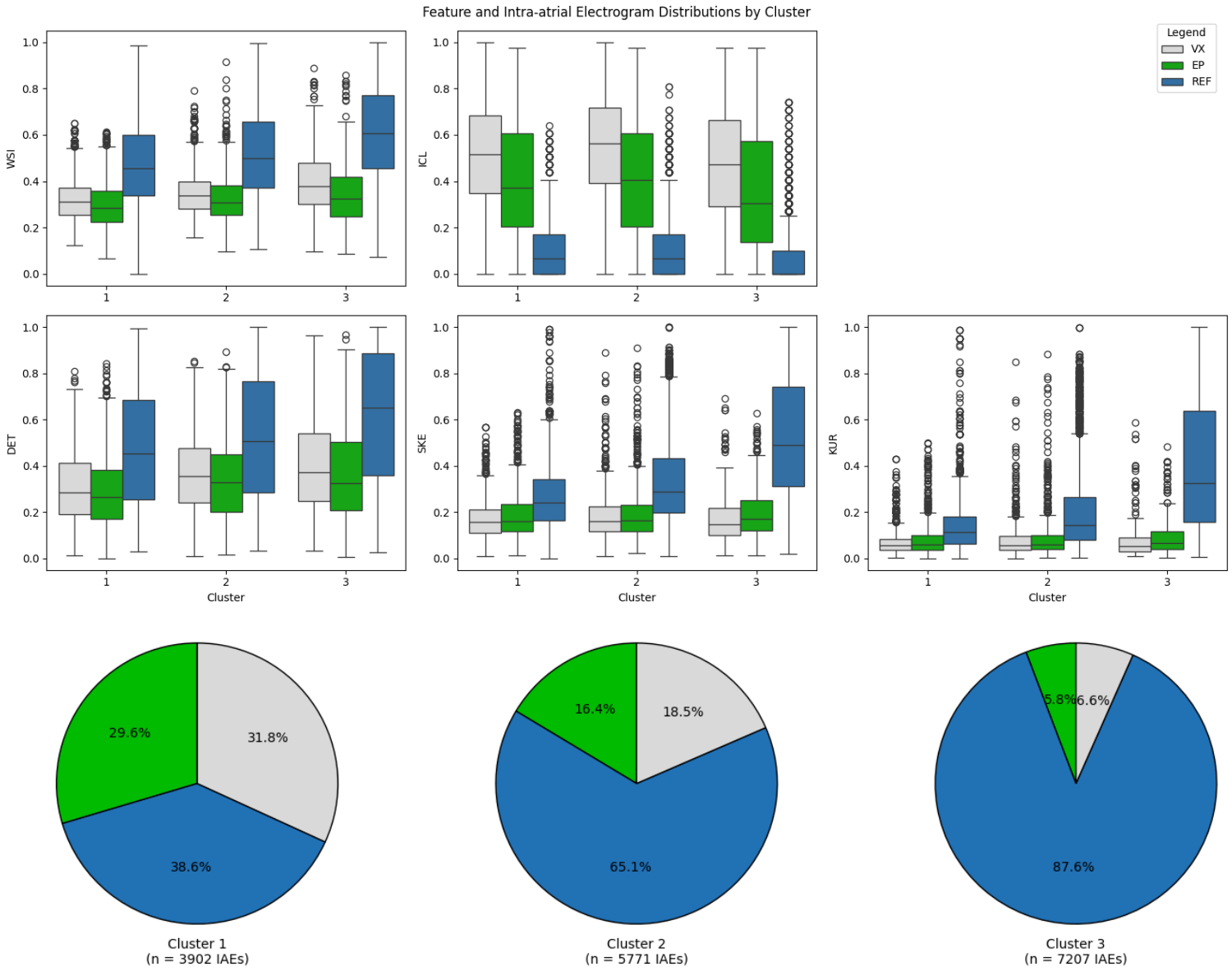


Figure 23: The top five features, based on the feature importance determined by the electrophysiologist, are plotted alongside the corresponding IAE distribution in each cluster. Abbreviations: EP, STD-based IAEs by the electrophysiologist (EP-STD IAEs); REF, non-STD-based reference IAEs (REF IAEs); VX, STD-based IAEs by the VX1-software (VX-STD IAEs).

In Figure 23, the top five features, selected based on the feature importances of the electrophysiologist, are shown with the corresponding IAE distribution for each cluster. The characteristics of the IAEs within the clusters reveal significant differences across all the various IAEs in each cluster. The VX-STD and EP-STD IAEs are characterized by lower WSI, DET, SKE, and KUR, and higher ICL, compared to the REF IAEs in each cluster. Additionally, the difference in interquartile ranges between REF IAEs and VX-STD or EP-STD IAEs is greater in Cluster 3 compared to Cluster 1. The interquartile ranges of the VX-STD IAEs show greater deviation for ICL, SKE, and KUR than the EP-STD IAEs, relative to the REF IAEs. Furthermore, the percentage of different IAEs in each cluster varies, with 29.6% of IAEs in Cluster 1 being EP-STD IAE, 38.6% being REF IAE, and 31.8% being VX-STD. In Cluster 2, 16.4% are EP-STD IAE, 65.1% are REF IAE, and 18.5% are VX-STD IAE, while in Cluster 3, 5.8% are EP-STD IAE, 87.6% are REF, and 6.6% are VX-STD IAE. The total number of IAEs in each cluster is 3902, 5771 and 7202, respectively.

4.4.2 SPATIO-TEMPORAL DISPERSION-BASED INTRA-ATRIAL AREAS

In Figure 24, the distribution of various criteria for VX-STD and EP-STD IAEs is presented, highlighting discrepancies and similarities in intra-atrial areas for each cluster. In all three clusters, most VX-STD and EP-STD IAEs are classified under criteria III, occurring in similar intra-atrial areas, with 92.43% in Cluster 1, 91.00% in Cluster 2, and 77.41% in Cluster 3 for VX-STD IAEs, and 92.47%, 88.47%, and 83.41% for EP-STD IAEs, respectively. In contrast, criteria I, occurring in different intra-atrial areas, is more prevalent in Cluster 3, representing 22.18% of VX-STD and 15.38% of EP-STD IAEs. It accounts for 7.00% of VX-STD and 6.75% of EP-STD IAEs in Cluster 1, and 8.43% of VX-STD and 10.48% of EP-STD IAEs in Cluster 2. Nevertheless, criteria II is minimally represented across all clusters, with less than 1.25% for both VX-STD and EP-STD IAEs.

Regarding the overall distribution, Cluster 1 contains the highest number of IAEs, with a total of 1242 VX-STD and 1155 EP-STD IAEs. Cluster 2 includes a total of 1067 VX-STD and 945 EP-STD IAEs, while Cluster 3 has the lowest representation, with a total of 478 VX-STD and 416 EP-STD IAEs.

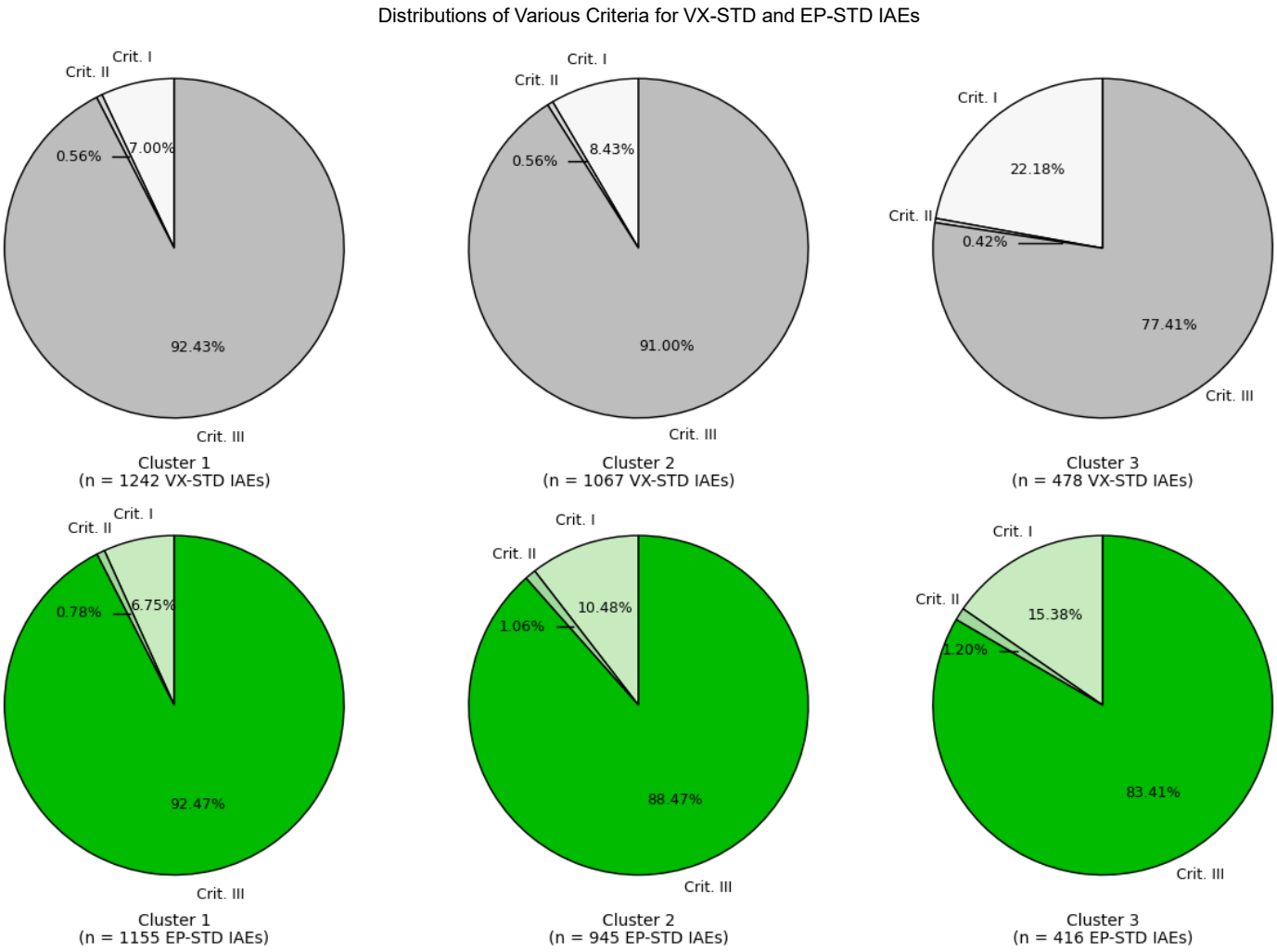


Figure 24: Distribution of VX-STD and EP-STD IAEs based on categorization criteria across the clusters. Abbreviations: Crit., criteria.

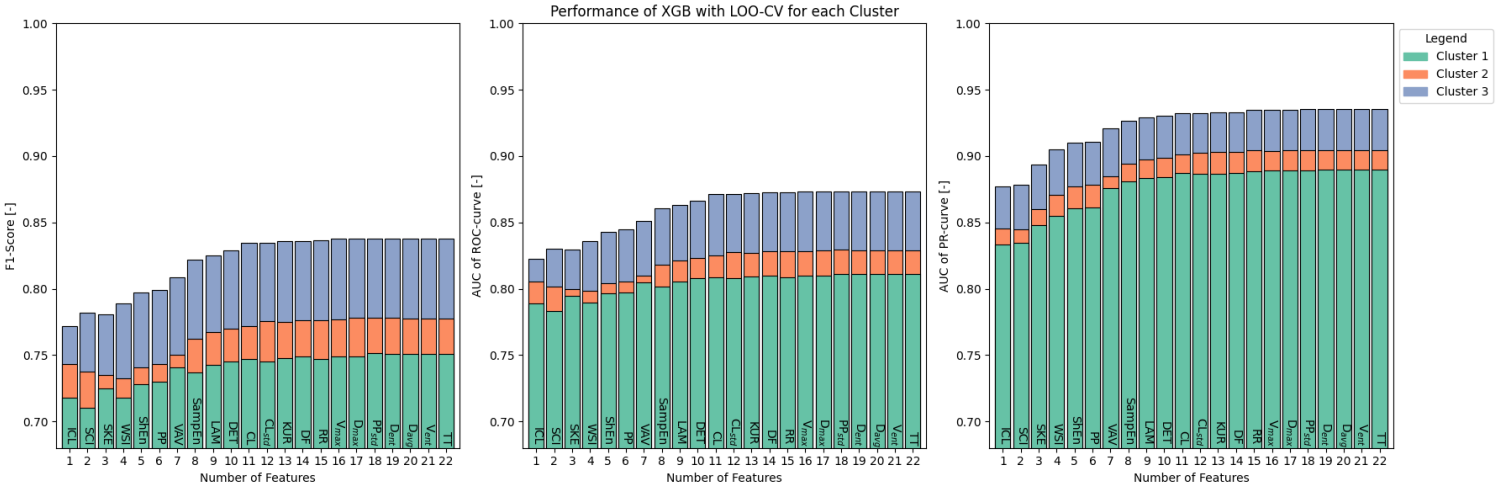


Figure 25: Classification metrics (F1-score, AUC-ROC, and AUC-PR) for the performance of LOO-CV across clusters in the XGB model, based on the feature importance of VX1-software. Note: The performance of each specific cluster indicates that the cluster is not used for training and validation of the XGB model. Abbreviations: LOO-CV, leave-one-out cross-validation.

4.4.3 PERFORMANCE OF EXTREME GRADIENT BOOSTING MODEL

In Figure 25, the F1-score, AUC-ROC, and AUC-PR for the performance of the XGB model with LOO-CV across clusters is shown. The results reveal how the XGB model performed when that specific cluster was not used for training and validation. A similar trend is observed across the F1-score, AUC-ROC, and AUC-PR values for all clusters as more features are incorporated. The highest performance of the XGB model, based on the F1-score, AUC-ROC, and AUC-PR, is achieved when it is trained and validated on Clusters 1 and 2, with a maximum F1-score of 0.838 at 17 features, AUC-ROC of 0.873 at 17 features, and AUC-PR of 0.935 at 20 features. However, after 11 features, the increase in these metrics becomes marginal. The lowest performance occurs when the model is trained and validated on Clusters 2 and 3, with a maximum F1-score of 0.751 at 18 features, AUC-ROC of 0.811 at 18 features, and AUC-PR of 0.890 at 20 features.

In Figure 26, the accuracy of the XGB model across different clusters is shown, evaluated on the EP-STD IAEs with various categorization criteria. The performance of the XGB model is assessed using EP-STD IAEs classified under Criteria I (EP_I), Criteria II (EP_{II}), and Criteria III (EP_{III}). In Cluster 1, the EP_{II} achieved the highest accuracy across all features, outperforming

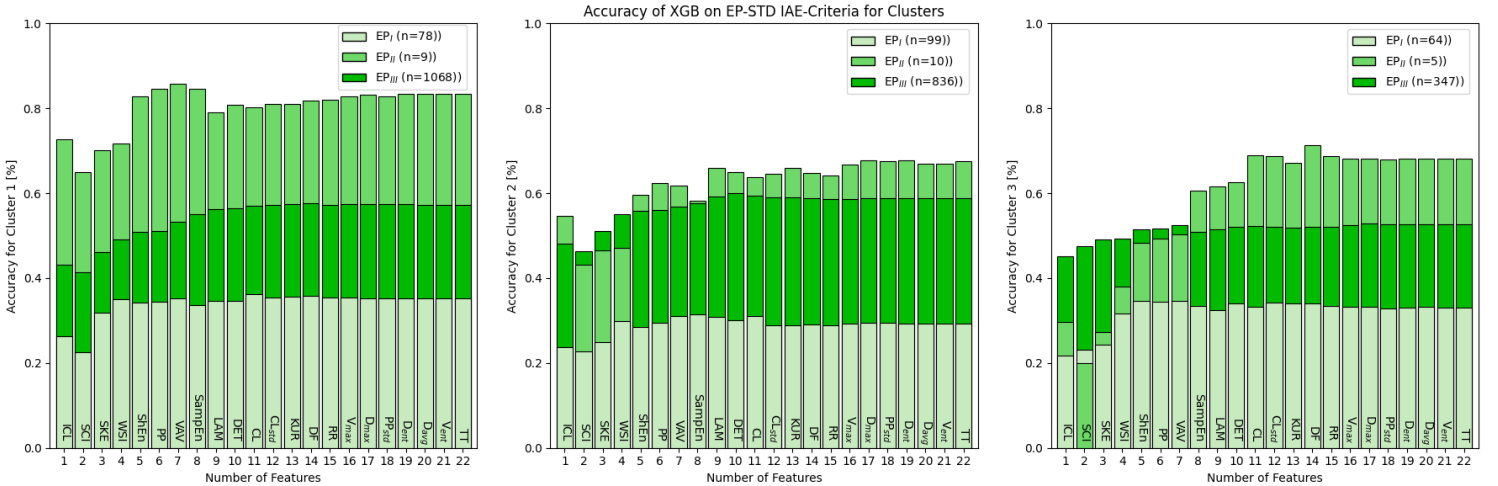


Figure 26: Accuracy of the XGB model for the different clusters, tested on the EP-STD IAEs with various categorization criteria. Abbreviations: EP_I, EP-STD IAEs with criteria I; EP_II, EP-STD IAEs with criteria II; EP_III, EP-STD IAEs with criteria III.

both to EP_I and EP_{III}. The peak accuracy for EP_{II} is 0.859, achieved with 7 included features. In comparison, the highest accuracy for EP_I and EP_{III} is 0.363 at 11 features and 0.576 at 14 features, respectively. In Cluster 2, the trends for all the three categories are similar. The accuracy starts at a minimum with 2 included features, then increases gradually to reach a plateau. The maximal accuracies for EP_I, EP_{II} and EP_{III} are 0.314, 0.677 and 0.599, with 8, 17, and 10 included features, respectively. In Cluster 3, EP_{III} achieves the highest accuracy up to 7 included features. After that, EP_{II} surpasses EP_{III} in accuracy. The EP_{II} shows significant increase until 10 included features, whereas the EP_I and EP_{III} show minimal improvement. The maximal accuracies for EP_I, EP_{II} and EP_{III} are 0.343, 0.714 and 0.528, at 12, 14, and 17 included features respectively. Overall, the accuracy in all clusters is consistently higher for the EP_{III} compared to EP_I.

In Figure 27, the accuracy of the XGB model across different clusters is presented, but this time evaluated on the VX-STD IAEs with various categorization criteria. The performance of the XGB model is assessed using VX-STD IAEs classified under Criteria I (VX_I), Criteria II (VX_{II}), and Criteria III (VX_{III}). In Cluster 1, the VX_{II} achieves the highest accuracy across nearly all features, with performance similar to the VX_{III}, but with lower accuracy for VX_I. The peak accuracy for VX_{II} is 0.996, attained with 1 included feature. In comparison, the highest accuracies for VX_I and VX_{III} are 0.724 at 1 feature and 0.790 at 18 features, respectively. In Cluster 2, the VX_{III} achieved the highest accuracy across all features, while the VX_I shows lower accuracy and VX_{II} has the lowest accuracy. All categories initially exhibit the highest accuracy with a decrease followed by a plateau. The maximal accuracies for VX_I, VX_{II} and VX_{III} are 0.815, 0.673 and 0.829, at 2, 1, and 2 included features, respectively. In Cluster 3, the same pattern as in Cluster 2 is shown for the three categories. However, in this cluster, the maximum accuracies are all achieved with 2 included features, with accuracies of 0.716 for VX_I, 0.835 for VX_{II}, and 0.784 for VX_{III}. Once again, the accuracy across all clusters is consistently higher for the VX_{III} compared to VX_I.

The comparison between the VX-STD IAEs and EP-STD IAEs reveals several key differences. The EP-STD IAEs, generally show lower accuracy compared to the VX-STD IAEs. However, in both VX-STD and EP-STD IAEs, the group with Criteria I (VX_I and EP_I) yield lower accuracies than the group with Criteria III (VX_{III} and EP_{III}). Additionally, the VX-STD IAEs often achieve similar or higher accuracies with fewer features (1 or 2) compared to the EP-STD IAEs, where higher accuracies require more features (11–17).

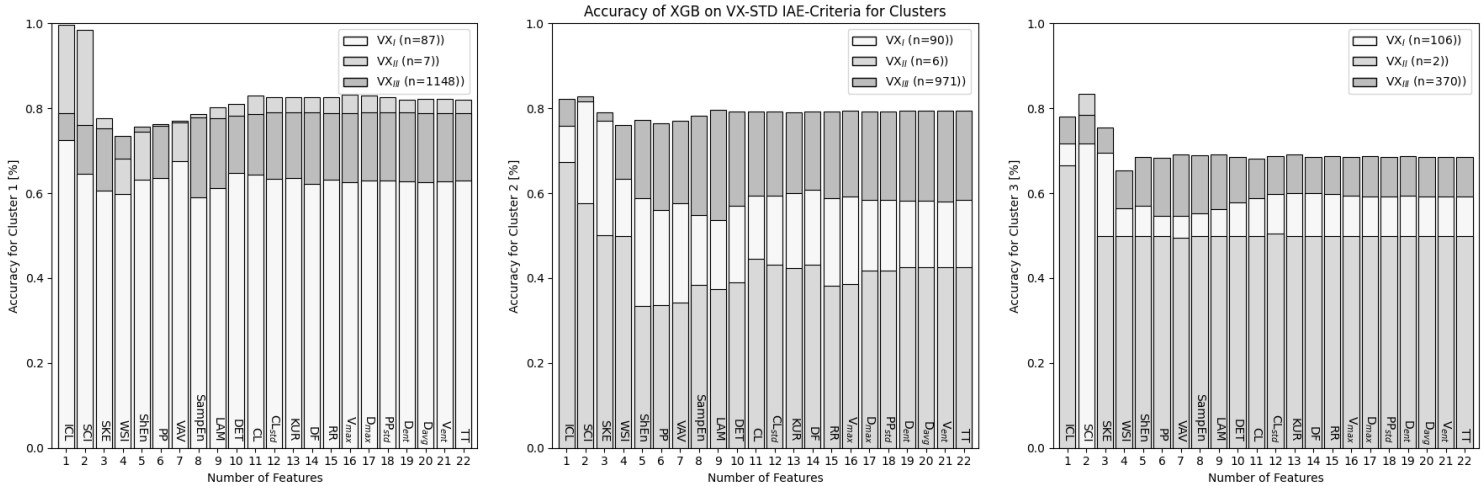


Figure 27: Accuracy of the XGB model for the different clusters, tested on the VX-STD IAEs with various categorization criteria. Abbreviations: VX_I, VX-STD IAEs with criteria I; VX_II, VX-STD IAEs with criteria II; VX_III, VX-STD IAEs with criteria III.

4.5 DISCUSSION

The discussion highlights the findings, acknowledges study limitations, explores clinical relevance, and offers recommendations for future research.

4.5.1 FINDINGS

The findings of the patient-specific complexity scores, shown in Figures 22 and 23, reveal a clear stratification of patients based on their AF type and overall complexity. The K-Means clustering identified three distinct complexity clusters. Cluster 1 had lower scores, suggesting a closer alignment between STD-based and non-STD-based IAEs, while Cluster 3 had higher scores, indicating greater discrepancies between IAEs. The distribution of patients in the three clusters suggests that the complexity scores offer valuable insights into the severity and progression of AF. Specifically, patients with longstanding persistent AF tend to fall into the lower-score category, while those with paroxysmal AF are more likely to be found in the higher-score category. Furthermore, Cluster 1 showed a higher proportion of STD-based IAEs, whereas Cluster 3 had fewer, highlighting the connection between AF progression and the need for advanced ablation strategies. These findings underscore the importance of using patient-specific complexity scores to inform a personalized approach to AF management, ultimately improving treatment outcomes and guiding tailored interventions.

Moreover, the similarities and discrepancies in STD-based intra-atrial areas, as shown in Figure 24, demonstrates that most VX-STD and EP-STD IAEs are in similar intra-atrial regions, indicating alignment between the VX1-software and the electrophysiologist. However, the proportion of IAEs classified under Criteria I, indicating discrepancies in STD-based intra-atrial areas, increased with each cluster. This suggests that greater discrepancies between IAEs, particularly in Custer 3, are associated with greater spatial heterogeneity between the VX1-software and the electrophysiologist, which may lead to suboptimal ablation outcomes. Nevertheless, these results align with expectations, as less severe AF in Cluster 1 is characterized by fewer STD-based intra-atrial areas, leading to more potential discrepancies between the VX1-software and the electrophysiologist. Besides, the minimal representation of VX-STD and EP-STD IAEs under Criteria II across all clusters highlights that the two approaches show minimal discrepancies regarding the significance of STD-based intra-atrial areas. These findings suggest a strong alignment between the VX1-software and the electrophysiologist in identifying STD-based intra-atrial areas, highlighting their reliability in pinpointing key regions for ablation. However, it also underscore the need for precise ablation strategies, especially in cases with less severe or progressed AF, where discrepancies could result in suboptimal treatment outcomes.

Additionally, the performance of the XGB model with LOO-CV, as shown in Figure 25, reveals similar trends across the clusters, suggesting that excluding a complexity cluster from the training and validation process has minimal impact on the overall performance. This could indicate that AF-related complexity is more predictable in patients with lower complexity scores, where consistent patterns lead to better model performance, whereas the increased heterogeneity in IAEs within Cluster 3 introduces challenges for accurate classification. Alternatively, it is possible that the inclusion of more STD-based IAEs enhances performance, as evidenced by the model performing worst when Cluster 1, which has the most STD-based IAEs, is excluded, and best when Cluster 3, with the fewest STD-based IAEs, is excluded. Notably, after incorporating around 11 features, the performance of the XGB model showed only marginal improvements, suggesting that a smaller, more focused set of key features is sufficient to achieve robust predictions.

Furthermore, the accuracy of the XGB model between the various categorization criteria of VX-STD and EP-STD IAEs, as shown in Figures 26 and 27, consistently reveals higher

accuracy for VX-STD IAEs compared to EP-STD IAEs. This suggests that VX-STD IAEs contain more distinguishable patterns, which enhance classification performance, aligning with previous findings. Conversely, both VX-STD and EP-STD IAEs show lower accuracy for Criteria I (VX_I and EP_I) compared to Criteria III (VX_{III} and EP_{III}), indicating that discrepancies between the VX1-software and the electrophysiologist in IAE identification lead to higher misclassifications of STD-based IAEs. This suggests that areas with discrepancies are likely non-STD-based reference IAEs rather than correctly labeled STD-based IAEs. Additionally, the accuracy of Criteria II varies across the clusters, possibly due to the small number of these IAEs. Furthermore, the VX-STD IAEs often achieved similar or higher accuracies with fewer features compared to the EP-STD IAEs, which required more features to reach higher accuracies. These findings suggest that VX-STD IAEs tend to be more efficient identified and in achieving higher accuracy with fewer features, while EP-STD IAEs may require a larger number of features to achieve similar or better performance. Overall, these findings highlight the need for patient-specific approaches in IAE classification, emphasizing the importance of selecting the right categorization criteria and feature sets, and underscoring the need for tailored diagnostic models to address individual patient complexities for optimal outcomes.

4.5.2 LIMITATIONS

Despite the findings of this study highlighting the performance of the VX1-software across patients with varying complexities of AF, several limitations should be acknowledged. A limitation of this study is the limited distribution of STD-based and non-STD-based reference IAEs across the different clusters, which could influence the performance of the XGB model. Specifically, Cluster 1, primarily composed of longstanding persistent AF patients, has a limited number of non-STD-based reference IAEs. This is because the atria of these patients are predominantly characterized by multiple STD-based areas, resulting in fewer non-STD-based IAEs being identified and underrepresented in this cluster. Conversely, Cluster 3, consisting mainly of paroxysmal AF patients, has a limited number of STD-based IAEs, as their atria are characterized by fewer STD-based regions. The scarcity of both STD-based and non-STD-based reference IAEs in these clusters can lead to issues in the ability of the XGB model to learn from a diverse and representative set of IAEs. With fewer IAEs in these clusters, the XGB model has less information to accurately classify STD-based IAEs. This limitation can cause the model to overfit to the available data, resulting in inaccurate predictions when applied to new, unseen IAEs. Moreover, the lack of diversity in the IAEs may lead to biased learning, where the model fails to generalize well across the full spectrum of AF complexities. This can ultimately skew the classification, as the model might not fully capture the variability in intra-atrial regions between different AF types. Consequently, this underrepresentation of IAEs may reduce the accuracy of the XGB model, especially in IAE identification for patients with AF subtypes that are less represented in the clusters.

Furthermore, another limitation of this study is the one-centimeter threshold, which is highly dependent on the quality of the EAM of the atria. While the threshold was chosen based on clinical expertise to capture relevant spatial proximity, inaccuracies in the EAM, such as poor spatial resolution or incomplete mapping, could affect the precision of geodesic distance calculations. The resolution and accuracy of the EAM are influenced by factors such as the number of points collected, the resolution of the mapping system, and the catheter used, which can lead to variations in atrial geometry and impact the classification of intra-atrial areas. If the EAM is imprecise, the one-centimeter threshold may overestimate or underestimate the proximity between STD-based intra-atrial areas, causing errors in categorization and reducing the reliability of the findings. Therefore, the quality of the EAM is crucial for accurate spatial measurements, and any deficiencies in the mapping process could limit the overall performance of the XGB model.

4.5.3 CLINICAL RELEVANCE

Patient-specific complexity scores play a key role in personalizing AF management by stratifying patients according to their overall IAE complexity. The findings of this study emphasize the connection between these scores and the severity and progression of AF, providing valuable insights for a tailored approach that can improve treatment outcomes and guide tailored interventions, as supported by existing literature [83,84]. Specifically, higher scores, indicating lower complexity, may benefit from more standardized approaches, while lower scores, reflecting higher complexity, may require more nuanced or advanced approaches. Additionally, the scores reveal differences between STD-based and non-STD-based reference IAEs in terms of their spatio-temporal characteristics, further advancing the understanding and identification of IAEs. These scores assist in guiding decisions on ablation strategies, enhancing procedural outcomes, and ultimately optimizing long-term AF management.

Furthermore, the accuracy of IAE identification for both VX-STD and EP-STD IAEs, with discrepancies in intra-atrial areas, highlights the limitations of both the VX1-software and electrophysiologists in identifying STD-based intra-atrial regions. Specifically, lower complexity correlates with higher discrepancies in STD-based intra-atrial regions for both the VX1-software and electrophysiologists. This is clinically significant because misclassifying non-STD-based IAEs as STD-based could lead to suboptimal ablation targeting, potentially overlooking critical atrial regions. Consequently, patients may experience less effective treatment, with an increased risk of AF recurrence or incomplete ablation. These findings underscore the need for patient-specific approaches in IAE classification, emphasize the importance of selecting appropriate categorization criteria and feature sets, and highlighting the necessity of tailored diagnostic models to address individual patient complexities for optimal outcomes.

4.5.4 FUTURE PERSPECTIVES

Although current methodologies offer valuable insights, there are several challenges that need to be addressed in order to enhance the accuracy and clinical applicability of this study. Future research should examine the differences in IAE complexity using larger datasets that include more non-STD-based IAEs in patients with longstanding persistent AF and more STD-based IAEs in those with paroxysmal AF. Including a broader range of patients, particularly from these two groups, would enhance the understanding of how AF complexity influences IAE patterns. This could offer insights into how different AF types contribute to the presence of non-STD-based versus STD-based IAEs and how these variations impact ablation outcomes. A more comprehensive dataset would also aid in refining ablation strategies tailored to the unique characteristics of each AF subtype, resulting in more personalized and effective treatments.

Additionally, future research should provide insight in the discrepancies between the VX1-software with the electrophysiologists in STD-based intra-atrial areas. Understanding the spatio-temporal characteristics of these discrepancies and how this relate to overlapping STD-based regions are crucial for refining ablation strategies. It is essential to explore whether these discrepancies represent true anatomical or electrogram differences, or if they arise from limitations in the ability of the VX1-software to distinguish certain regions. This research could clarify the underlying causes of misclassifications and lead to improvements in the accuracy of the VX1-software, ensuring that it aligns more closely with electrophysiologist assessments. Addressing these discrepancies may enhance the role of the VX1-software in guiding more precise and effective ablation procedures, ultimately improving patient outcomes.

4.6 CONCLUSION

The study highlights the strengths of patient-specific IAE complexity scores in facilitating patient stratification, underscoring the need for personalized AF management while providing insights into the severity and progression of AF. In contrast, the VX1-software demonstrates strong potential in IAE identification with minimal discrepancies in STD-based intra-atrial areas compared to the electrophysiologist, indicating that the VX1-software offers a more generalized approach, while the approach of the electrophysiologist is more intricate. However, the correlation between lower complexity and higher discrepancies emphasizes the need for patient-specific approaches in IAE classification. This highlights the importance of selecting appropriate categorization criteria and feature sets and underscores the necessity of tailored diagnostic models to address individual patient complexities for optimal outcomes. Future research should focus on addressing these discrepancies, incorporating advanced methods to enhance the ability of the VX1-software to identify complex IAE patterns and optimize ablation strategies, ultimately improving patient outcomes.

5. GENERAL DISCUSSION

The general discussion provides an overview of the key findings from this thesis, discusses the limitations, and offers future perspectives for enhancing STD-based IAE identification as a standardized approach for optimizing ablation strategies in patients with AF.

5.1 CLINICAL RELEVANCE

The clinical relevance of this thesis highlights the evolving landscape of ablation procedure strategies in patients with AF, with a focus on the strengths and potential of the VX1-software in identification of STD-based IAEs as a standardized approach, compared to the patient-specific approach of the electrophysiologist. The VX1-software demonstrates significant overlap with the electrophysiologist in the key spatio-temporal characteristics of STD-based IAEs, underscoring the emerging value of the VX1-software as a tool that could standardize IAE identification. Conversely, the differences in spatio-temporal characteristics indicate that electrophysiologist prioritize waveform repetitiveness and organization, whereas the VX1-software focuses on localized voltage characteristics and fractionation.

Furthermore, the generalization capabilities of the VX1-software are supported by the superior performance of existing ML models based on the STD-based IAEs of the VX1-software. These capabilities provide valuable insights into the methodological differences in IAE identification, highlighting the importance in focus of the VX1-software on fractionation, high-frequency activation, and voltage asymmetry, while the electrophysiologist emphasize waveform repetitiveness, fractionation, and activation pattern consistency. Consequently, the performance of the VX1-software underscore the complexity of IAE identification by the electrophysiologist, emphasizing the need for it to address more intricate STD-based IAE patterns through a refined, multi-faceted approach.

Additionally, the patient-specific nature of IAE identification adds another layer of complexity. Although the VX1-software, as a generalized approach, demonstrates strong potential with minimal discrepancies in identifying STD-based intra-atrial areas compared to the electrophysiologist, limitations remain. The approach of the electrophysiologist is inherently patient-specific, considering not only general patterns but also the individual spatio-temporal characteristics of each patient. In contrast, the VX1-software, while capable of providing reliable and reproducible results, operates within a generalized framework that does not yet fully address the personalized nature of IAE complexity. Therefore, a more tailored approach to STD-based IAE identification is needed, one that considers the unique patterns of IAEs found in different patients to better capture the complexity of IAE identification.

Ultimately, combining the strengths of both the VX1-software and the electrophysiologist could provide a more comprehensive solution, bridging the gap between a generalized and patient-specific approach in STD-based IAE identification. The VX1-software could be enhanced by a more nuanced, multifaceted approach that leverages the expertise of the electrophysiologist in interpreting complex STD-based IAE patterns and variability. This integration has the potential to improve the accuracy of IAE identification, refine treatment strategies, and ultimately lead to better patient outcomes with the VX1-software, offering a standardized approach for optimizing ablation strategies in patients with AF.

5.2 LIMITATIONS

Despite the valuable contributions of this thesis, certain limitations must be acknowledged. In this thesis, the manual identification of STD-based IAEs was conducted by a single electrophysiologist, introducing subjectivity and variability in IAE interpretation. The experience

and expertise of the electrophysiologist could have influenced the classification of IAEs, potentially leading to differences in feature recognition and discrepancies between the VX1-software and the electrophysiologist. Additionally, there is an inherent learning curve associated with identifying STD-based IAEs, as this field was not yet well-established in clinical practice at the time the patients were included. The decision to involve only one electrophysiologist was primarily due to the limited availability of specialists with extensive experience in STD-based IAE identification, which constrained the scope of this thesis.

Furthermore, a limitation of this thesis is the limited number of available STD-based and non-STD-based reference IAEs, which may impact the robustness of the findings and their generalizability to a wider patient population. Additionally, the relatively small proportion of patients across different types of AF restricts the variability in IAE characteristics within these AF subtypes. This limitation also affects the evaluation of the performance of the VX1-software and the accuracy of the electrophysiologist in identifying IAEs. These constraints were primarily due to the availability of eligible patients within the study period, which resulted in a restricted dataset with limited diversity in terms of both patients and IAEs.

5.3 FUTURE PERSPECTIVES

To address the limitations, future studies should include multiple electrophysiologists from different centra to reduce the impact of individual variability and improve the consistency of IAE identification. A broader range of experts would allow for a more comprehensive evaluation of the VX1-software, and contribute to understanding whether discrepancies arise from the VX1-software itself or from the inherent subjectivity of IAE identification. In addition, involving multiple electrophysiologists would improve the generalizability of the findings and strengthen the clinical relevance of the VX1-software.

Additionally, future research should aim to include a larger and more diverse patient cohort, with a more balanced representation of different AF types. This would provide valuable insights into the performance of the VX1-software across various patient groups, as well as how spatio-temporal characteristics of IAEs evolve with the progression of AF. Such an approach would deepen the understanding of disease progression and offer greater clarity on the potential role of IAEs in long-term AF management, ultimately guiding both individualized and generalized treatment strategies.

Furthermore, the incorporation of the spatial involvement of IAEs in the analysis must be analyzed. In this thesis, most spatio-temporal characteristics were analyzed using a single IAE, which provided valuable insights into the temporal dynamics of AF. However, spatial heterogeneity within the atria should also be considered to fully understand the role of STD-based IAEs. A more detailed exploration of how STD-based IAEs are distributed across atrial regions could reveal specific patterns contributing to AF initiation and progression. Therefore, future studies should leverage multiple ML and deep learning models to capture the complex relationships between spatio-temporal characteristics of IAEs. Combining various models would not only improve the accuracy of IAE detection but also provide deeper insights into AF mechanisms, potentially guiding more targeted and personalized ablation strategies.

Moreover, investigating the impact of ablating STD-based regions on the recurrence of AF, particularly with longer follow-up, could reveal which specific patterns of STD-based IAEs are most linked to AF recurrence. Identifying these key patterns would help determine which IAEs contribute most to the persistence or recurrence of AF. This deeper understanding could guide more targeted ablation strategies, focusing on how the spatio-temporal dynamics of these IAEs interact in AF progression, ultimately leading to more personalized and precise treatment strategies.

6. GENERAL CONCLUSION

The VX1-software shows significant potential in standardizing STD-based IAE identification by aligning with the electrophysiologist in key spatio-temporal characteristics, demonstrating strong generalization capabilities through superior ML model performance, and accurately identifying STD-based intra-atrial areas. Nevertheless, the VX1-software differs in its emphasis on fractionation, high-frequency activation, and voltage asymmetry, whereas the electrophysiologist focuses more on waveform repetitiveness, fractionation, and activation pattern consistency. Furthermore, the VX1-software lacks the patient-specific approach of the electrophysiologist, which accounts for individual spatio-temporal characteristics, as it operates within a generalized framework that does not fully address the personalized complexity of IAE patterns and identification. These findings underscore the need for a more tailored approach that combines the standardization strengths of the VX1-software with the patient-specific insights of the electrophysiologist. Future research should integrate this patient-tailored capability into the VX1-software to improve its accuracy and adaptability across diverse patient populations, ultimately enhancing AF ablation strategies, reducing recurrence rates, and improving patient outcomes.

7. REFERENCES

[1] Tilz RR, Heeger CH, Wick A, Saguner AM, Metzner A, Rillig A, et al. Ten-Year Clinical Outcome after Circumferential Pulmonary Vein Isolation Utilizing the Hamburg Approach in Patients with Symptomatic Drug-Refractory Paroxysmal Atrial Fibrillation. Circ Arrhythmia Electrophysiol 2018;11. https://doi.org/10.1161/CIRCEP.117.005250/ASSET/8295C0F8-B835-4657-BB89-ECE53398F192/ASSETS/IMAGES/LARGE/E005250FIG04.JPG.

[2] Metzner A, Kuck KH, Chun JKR. What we have learned: is pulmonary vein isolation still the cornerstone of atrial fibrillation ablation? EP Eur 2022;24:ii8–13. https://doi.org/10.1093/EUROPACE/EUAB268.

[3] Chen PS, Chou CC, Tan AY, Zhou S, Fishbein MC, Hwang C, et al. The Mechanisms of Atrial Fibrillation. J Cardiovasc Electrophysiol 2006;17:S2–7. https://doi.org/10.1111/J.1540-8167.2006.00626.X.

[4] Khan R. Identifying and understanding the role of pulmonary vein activity in atrial fibrillation. Cardiovasc Res 2004;64:387–94. https://doi.org/10.1016/J.CARDIORES.2004.07.025/2/64-3-387-FIG4.GIF.

[5] Voskoboinik A, Moskovitch JT, Harel N, Sanders P, Kistler PM, Kalman JM. Revisiting pulmonary vein isolation alone for persistent atrial fibrillation: A systematic review and meta-analysis. Hear Rhythm 2017;14:661–7. https://doi.org/10.1016/J.HRTHM.2017.01.003.

[6] Benali K, Barré V, Hermida A, Galand V, Milhem A, Philibert S, et al. Recurrences of Atrial Fibrillation Despite Durable Pulmonary Vein Isolation: The PARTY-PVI Study. Circ Arrhythmia Electrophysiol 2023;16:E011354. https://doi.org/10.1161/CIRCEP.122.011354/SUPPL_FILE/CIRCAE_CIRCAE-2022-011354_SUPP1.PDF.

[7] Jilek C, Ullah W. Pulmonary vein reconnections or substrate in the left atrium: what is the reason for atrial fibrillation recurrences? A dialogue on a pressing clinical situation. EP Eur 2019;21:i12–20. https://doi.org/10.1093/EUROPACE/EUY289.

[8] Obergassel J, Nies M, Taraba S, Rottner L, Lemoine MD, My I, et al. Pulmonary vein reconnection and repeat ablation characteristics following cryoballoon-compared to radiofrequency-based pulmonary vein isolation. J Cardiovasc Electrophysiol 2024;35:1766–78. https://doi.org/10.1111/JCE.16343.

[9] Johner N, Namdar M, Shah DC. Individualised Approaches for Catheter Ablation of AF: Patient Selection and Procedural Endpoints. Arrhythmia Electrophysiol Rev 2019;8:184. https://doi.org/10.15420/AER.2019.33.2.

[10] Quintanilla JG, Shpun S, Jalife J, Filgueiras-Rama D. Novel approaches to mechanism-based atrial fibrillation ablation. Cardiovasc Res 2021;117:1662–81. https://doi.org/10.1093/CVR/CVAB108.

[11] Mody BP, Raza A, Jacobson J, Iwai S, Frenkel D, Rojas R, et al. Ablation of long-standing persistent atrial fibrillation. Ann Transl Med 2017;5:305. https://doi.org/10.21037/ATM.2017.05.21.

[12] Kistler PM, Chieng D, Sugumar H, Ling LH, Segan L, Azzopardi S, et al. Effect of Catheter Ablation Using Pulmonary Vein Isolation With vs Without Posterior Left Atrial Wall Isolation on Atrial Arrhythmia Recurrence in Patients With Persistent Atrial Fibrillation: The CAPLA Randomized Clinical Trial. JAMA 2023;329:127–35. https://doi.org/10.1001/JAMA.2022.23722.

[13] Verma A, Jiang C, Betts TR, Chen J, Deisenhofer I, Mantovan R, et al. Approaches to Catheter Ablation for Persistent Atrial Fibrillation. N Engl J Med 2015;372:1812–22. https://doi.org/10.1056/NEJMOA1408288/SUPPL_FILE/NEJMOA1408288_DISCLOSURES.PDF.

[14] Narayan SM, Baykaner T, Clopton P, Schricker A, Lalani GG, Krummen DE, et al. Ablation of rotor and focal sources reduces late recurrence of atrial fibrillation compared with trigger ablation alone: extended follow-up of the CONFIRM trial (Conventional Ablation for Atrial Fibrillation With or Without Focal Impulse and Rotor Modulation). J Am Coll Cardiol 2014;63:1761–8. https://doi.org/10.1016/J.JACC.2014.02.543.

[15] Steinberg JS, Shah Y, Bhatt A, Sichrovsky T, Arshad A, Hansinger E, et al. Focal impulse and rotor modulation: Acute procedural observations and extended clinical follow-up. Hear Rhythm 2017;14:192–7. https://doi.org/10.1016/J.HRTHM.2016.11.008.

[16] Buch E, Share M, Tung R, Benharash P, Sharma P, Koneru J, et al. Long-term clinical outcomes of focal impulse and rotor modulation for treatment of atrial fibrillation: A multicenter experience. Hear Rhythm 2016;13:636–41. https://doi.org/10.1016/J.HRTHM.2015.10.031.

[17] Huo Y, Gaspar T, Schönbauer R, Wójcik M, Fiedler L, Roithinger FX, et al. Low-Voltage Myocardium-Guided Ablation Trial of Persistent Atrial Fibrillation. NEJM Evid 2022;1. https://doi.org/10.1056/EVIDOA2200141.

[18] Valcher S, Villaschi A, Falasconi G, Chiarito M, Giunti F, Novelli L, et al. Low-Voltage Area Ablation in Addition to Pulmonary Vein Isolation in Patients with Atrial Fibrillation: A Systematic Review and Meta-Analysis. J Clin Med 2024;13:4541. https://doi.org/10.3390/JCM13154541/S1.

[19] Yang B, Jiang C, Lin Y, Yang G, Chu H, Cai H, et al. STABLE-SR (Electrophysiological Substrate Ablation in the Left Atrium During Sinus Rhythm) for the Treatment of Nonparoxysmal Atrial Fibrillation: A Prospective, Multicenter Randomized Clinical Trial. Circ Arrhythm Electrophysiol 2017;10. https://doi.org/10.1161/CIRCEP.117.005405.

[20] Seitz J, Bars C, Théodore G, Beurtheret S, Lellouche N, Bremondy M, et al. AF Ablation Guided by Spatiotemporal Electrogram Dispersion Without Pulmonary Vein Isolation: A Wholly Patient-Tailored Approach. J Am Coll Cardiol 2017;69:303–21. https://doi.org/10.1016/J.JACC.2016.10.065.

[21] Sakata K, Tanaka T, Yamashita S, Kobayashi M, Ito M, Yamashiro K. The spatiotemporal electrogram dispersion ablation targeting rotors is more effective for elderly patients than non-elderly population. J Arrhythmia 2023;39:315–26. https://doi.org/10.1002/JOA3.12860.

[22] Goldberger JJ, Zaatari G, Mitrani RD, Blandon C, Bohorquez J, Ng J, et al. Comparison of electrogram characteristics in persistent atrial fibrillation. J Cardiovasc Electrophysiol 2024;35:182–97. https://doi.org/10.1111/JCE.16133.

[23] Seitz J, Durdez TM, Albenque JP, Pisapia A, Gitenay E, Durand C, et al. Artificial intelligence software standardizes electrogram-based ablation outcome for persistent atrial fibrillation. J Cardiovasc Electrophysiol 2022;33:2250–60. https://doi.org/10.1111/JCE.15657.

[24] Deisenhofer I, Albenque J-P, Busch S, Gitenay E, Mountantonakis SE, Roux A, et al. Artificial intelligence for individualized treatment of persistent atrial fibrillation: a randomized controlled trial. Nat Med 2025 2025:1–8. https://doi.org/10.1038/s41591-025-03517-w.

[25] Bahlke F, Englert F, Popa M, Bourier F, Reents T, Lennerz C, et al. First clinical data on artificial intelligence-guided catheter ablation in long-standing persistent atrial fibrillation. J Cardiovasc Electrophysiol 2024;35:406–14. https://doi.org/10.1111/JCE.16184.

[26] Bars C, Gitenay E, Monteau J, Lotteau S, Reist M, Serdi M, et al. Artificial intelligence-enabled spatio-temporal dispersion mapping for persistent AF: Similarities and differences between pacing-induced or spontaneous AF. EP Eur 2023;25. https://doi.org/10.1093/EUROPACE/EUAD122.163.

[27] Manongi N, Kim J, Goldbarg S. Dispersion electrogram detection with an artificial intelligence software in redo paroxysmal atrial fibrillation ablation. Hear Case Reports 2023;9:948–53. https://doi.org/10.1016/J.HRCR.2023.10.006.

[28] Sousonis V, Voglimacci-Stephanopoli Q, Zeriouh S, Boveda S, Paul Albenque J. Pulsed field ablation of spatiotemporal electrogram dispersion following pulmonary vein isolation and left atrial linear lesions for persistent atrial fibrillation: a case report. Eur Hear Journal Case Reports 2024;8. https://doi.org/10.1093/EHJCR/YTAE085.

[29] Deisenhofer I, Albenque J-P, Busch S, Gitenay E, Mountantonakis S, Roux A, et al. LB-469805-01 TAILORED CARDIAC ABLATION PROCEDURE FOR PERSISTENT ATRIAL FIBRILLATION GUIDED BY ARTIFICIAL INTELLIGENCE: THE TAILORED-AF RANDOMIZED CLINICAL TRIAL. Hear Rhythm 2024;21:1199. https://doi.org/10.1016/J.HRTHM.2024.04.025.

[30] Nadim A, Li X, Chu G, Soriano D, Masè M, Ravelli F, et al. Unsupervised classification of dimension-reduced principal component scores from persistent atrial fibrillation electrograms 2021:28. https://doi.org/10.1117/12.2606172.

[31] Ghrissi A, Squara F, Montagnat J, Zarzoso V. Identification of spatiotemporal dispersion electrograms in persistent atrial fibrillation ablation using maximal voltage absolute values. Eur Signal Process Conf 2021;2021-Janua:1239–43. https://doi.org/10.23919/EUSIPCO47968.2020.9287681.

[32] Tzeis S, Gerstenfeld EP, Kalman J, Saad EB, Shamloo AS, Andrade JG, et al. 2024 European Heart Rhythm Association/Heart Rhythm Society/Asia Pacific Heart Rhythm Society/Latin American Heart Rhythm Society expert consensus statement on catheter and surgical ablation of atrial fibrillation. J Arrhythmia 2024;40:1217–354. https://doi.org/10.1002/JOA3.13082.

[33] Joglar JA, Chung MK, Armbruster AL, Benjamin EJ, Chyou JY, Cronin EM, et al. 2023 ACC/AHA/ACCP/HRS Guideline for the Diagnosis and Management of Atrial Fibrillation: A Report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. Circulation 2024;149:E1–156. https://doi.org/10.1161/CIR.0000000000001193/FORMAT/EPUB.

[34] Van Gelder IC, Rienstra M, Bunting K V, Casado-Arroyo R, Caso V, Crijns HJGM, et al. 2024 ESC Guidelines for the management of atrial fibrillation developed in collaboration with the European Association for Cardio-Thoracic Surgery (EACTS). Eur Heart J 2024. https://doi.org/10.1093/EURHEARTJ/EHAE176.

[35] Borlich M, Iden L, Kuhnhardt K, Paetsch I, Hindricks G, Sommer P. 3D Mapping for PVI-Geometry, Image Integration and Incorporation of Contact Force Into Work Flow. J Atr Fibrillation 2018;10:1795. https://doi.org/10.4022/JAFIB.1795.

[36] Calkins H, Hindricks G, Cappato R, Kim YH, Saad EB, Aguinaga L, et al. 2017 HRS/EHRA/ECAS/APHRS/SOLAECE expert consensus statement on catheter and surgical ablation of atrial fibrillation. Hear Rhythm 2017;14:e275–444. https://doi.org/10.1016/J.HRTHM.2017.05.012.

[37] Andronache M, Drca N, Viola G. High-resolution Mapping in Patients with Persistent AF. Arrhythmia Electrophysiol Rev 2019;8:111. https://doi.org/10.15420/AER.2018.57.1.

[38] De Ponti R, Verlato R, Bertaglia E, Del Greco M, Fusco A, Bottoni N, et al. Treatment of macro-re-entrant atrial tachycardia based on electroanatomic mapping: identification and ablation of the mid-diastolic isthmus. EP Eur 2007;9:449–57. https://doi.org/10.1093/EUROPACE/EUM055.

[39] Bhakta D, Miller JM. Principles of Electroanatomic Mapping. Indian Pacing Electrophysiol J 2008;8:32.

[40] Lankveld T, Zeemering S, Scherr D, Kuklik P, Hoffmann BA, Willems S, et al. Atrial Fibrillation Complexity Parameters Derived from Surface ECGs Predict Procedural Outcome and Long-Term Follow-Up of Stepwise Catheter Ablation for Atrial Fibrillation. Circ Arrhythmia Electrophysiol 2016;9. https://doi.org/10.1161/CIRCEP.115.003354/-/DC1.

[41] Electrophysiology (EP) Recording System – LABSYSTEM PRO - Boston Scientific n.d. https://www.bostonscientific.com/en-EU/medical-specialties/electrophysiology/arrhythmias/cardiac-mapping-system/electrophysiology-recording-system.html (accessed January 27, 2025).

[42] Williams SE, Roney CH, Connolly A, Sim I, Whitaker J, O'Hare D, et al. OpenEP: A Cross-Platform Electroanatomic Mapping Data Format and Analysis Platform for Electrophysiology Research. Front Physiol 2021;12:646023. https://doi.org/10.3389/FPHYS.2021.646023/BIBTEX.

[43] Starreveld R, Knops P, Roos-Serote M, Kik C, Bogers AJJC, Brundel BJJM, et al. The Impact of Filter Settings on Morphology of Unipolar Fibrillation Potentials. J Cardiovasc Transl Res 2020;13:953. https://doi.org/10.1007/S12265-020-10011-W.

[44] Frusone S, De Almeida RC, Squara F, Zarzoso V. Electrogram Fractionation Metrics in Spatio-Temporal Dispersion Based Catheter Ablation of Persistent Atrial Fibrillation n.d.

[45] Ravikumar V. Signal processing approaches for the spatiotemporal analysis of cardiac arrhythmias using intracardiac electrograms 2022.

[46] Baher A, Buck B, Fanarjian M, Paul Mounsey J, Gehi A, Chung E, et al. Recurrence quantification analysis of complex-fractionated electrograms differentiates active and passive sites during atrial fibrillation. J Cardiovasc Electrophysiol 2019;30:2229–38. https://doi.org/10.1111/JCE.14161.

[47] Almeida TP, Chu GS, Bell MJ, Li X, Salinet JL, Dastagir N, et al. The temporal behavior and consistency of bipolar atrial electrograms in human persistent atrial fibrillation. Med Biol Eng Comput 2018;56:71–83. https://doi.org/10.1007/S11517-017-1667-1.

[48] Horie T, Burioka N, Amisaki T, Shimizu E. Sample Entropy in Electrocardiogram During Atrial Fibrillation. Yonago Acta Med 2018;61:049–57. https://doi.org/10.33160/YAM.2018.03.007.

[49] Ganesan AN, Kuklik P, Lau DH, Brooks AG, Baumert M, Lim WW, et al. Bipolar electrogram Shannon entropy at sites of rotational activation implications for ablation of atrial fibrillation. Circ Arrhythmia Electrophysiol 2013;6:48–57. https://doi.org/10.1161/CIRCEP.112.976654/-/DC1.

[50] Ugarte JP, Tobón C, Orozco-Duque A. Entropy Mapping Approach for Functional Reentry Detection in Atrial Fibrillation: An In-Silico Study. Entropy (Basel) 2019;21. https://doi.org/10.3390/E21020194.

[51] Choi Y jung, Sohn JJ, Kwon S, Lee SR, Cha MJ, Choi EK, et al. Relationship between

dominant frequency, organization index, and left atrial size in patients with atrial fibrillation. J Cardiovasc Electrophysiol 2020;31:3159–65. https://doi.org/10.1111/JCE.14785.

[52] Halfar R, Lawson BAJ, dos Santos RW, Burrage K. Recurrence quantification analysis for fine-scale characterisation of arrhythmic patterns in cardiac tissue. Sci Reports 2023 131 2023;13:1–16. https://doi.org/10.1038/s41598-023-38256-w.

[53] Ravelli F, Masè M, Cristoforetti A, Del Greco M, Centonze M, Marini M, et al. Anatomic Localization of Rapid Repetitive Sources in Persistent Atrial Fibrillation: Fusion of Biatrial CT Images With Wave Similarity/Cycle Length Maps. JACC Cardiovasc Imaging 2012;5:1211–20. https://doi.org/10.1016/J.JCMG.2012.07.016.

[54] Podziemski P, Zeemering S, Kuklik P, van Hunnik A, Maesen B, Maessen J, et al. Rotors Detected by Phase Analysis of Filtered, Epicardial Atrial Fibrillation Electrograms Colocalize With Regions of Conduction Block. Circ Arrhythm Electrophysiol 2018;11:e005858. https://doi.org/10.1161/CIRCEP.117.005858/SUPPL_FILE/CIRCAE_CIRCAE-2017-005858_SUPP1.PDF.

[55] Lin YJ, Lo MT, Chang SL, Lo LW, Hu YF, Chao TF, et al. Benefits of Atrial Substrate Modification Guided by Electrogram Similarity and Phase Mapping Techniques to Eliminate Rotors and Focal Sources Versus Conventional Defragmentation in Persistent Atrial Fibrillation. JACC Clin Electrophysiol 2016;2:667–78. https://doi.org/10.1016/J.JACEP.2016.08.005.

[56] Lau DH, Zeemering S, Maesen B, Kuklik P, Verheule 1Sander, Schotten U. Catheter Ablation Targeting Complex Fractionated Atrial Electrogram in Atrial Fibrillation. J Atr Fibrillation 2013;6:907. https://doi.org/10.4022/JAFIB.907.

[57] Masuda M, Matsuda Y, Uematsu H, Sugino A, Ooka H, Kudo S, et al. Prognostic impact of atrial cardiomyopathy: Long-term follow-up of patients with and without low-voltage areas following atrial fibrillation ablation. Hear Rhythm 2024;21:378–86. https://doi.org/10.1016/J.HRTHM.2023.12.016.

[58] Ostertagová E, Ostertag O, Kováč J. Methodology and application of the Kruskal-Wallis test. Appl Mech Mater 2014;611:115–20. https://doi.org/10.4028/WWW.SCIENTIFIC.NET/AMM.611.115.

[59] Dinno A. Nonparametric pairwise multiple comparisons in independent groups using Dunn's test. Stata J 2015;15:292–300.

[60] Ghrissi A, Almonfrey D, Squara F, Montagnat J, Zarzoso V. Identification of Spatiotemporal Dispersion Electrograms in Atrial Fibrillation Ablation Using Machine Learning: A Comparative Study Identification of Spatiotemporal Dispersion Electrograms in Atrial Fibrillation Ablation Using Machine Learning: A Comparative Study Identification of Spatiotemporal Dispersion Electrograms in Atrial Fibrillation Ablation Using Machine Learning: A comparative Study. Biomed Signal Process Control n.d.:10. https://doi.org/10.1016/j.bspc.2021.103269ï.

[61] Almeida TP, Schlindwein FS, Salinet JL, Li X, Chu GS, Tuan JH, et al. Deterministic structures in fractionated atrial electrograms during human persistent atrial fibrillation. Comput Cardiol (2010) 2017;44:1–4. https://doi.org/10.22489/CINC.2017.196-321.

[62] Frusone S, Costa De Almeida R, Almonfrey D, Squara F, Zarzoso V. Identifying spatiotemporal dispersion in catheter ablation of persistent atrial fibrillation: a comparative study of machine learning techniques using both real and realistic synthetic multipolar electrograms n.d.

[63] Biviano AB, Ciaccio EJ, Knotts R, Fleitman J, Lawrence J, Iyer V, et al. Atrial Electrogram Discordance During Baseline versus Re-Induced Atrial Fibrillation: Potential Ramifications For Ablation Procedures. Heart Rhythm 2015;12:1448. https://doi.org/10.1016/J.HRTHM.2015.03.044.

[64] Alhusseini MI, Abuzaid F, Rogers AJ, Zaman JAB, Baykaner T, Clopton P, et al. Machine Learning to Classify Intracardiac Electrical Patterns During Atrial Fibrillation: Machine Learning of Atrial Fibrillation. Circ Arrhythmia Electrophysiol 2020;13:E008160. https://doi.org/10.1161/CIRCEP.119.008160/SUPPL_FILE/CIRCAE_CIRCAE-2019-008160_SUPP1.PDF.

[65] Frusone S, Almeida RC de, Squara F, Zarzoso V. Peak-based spatio-temporal dispersion classifier of multipolar intracardiac electrograms in persistent atrial fibrillation 2024;51. https://doi.org/10.13039/501100001665.

[66] Nicolet JJC, Restrepo JF, Schlotthauer G. Classification of intracavitary electrograms in atrial fibrillation using information and complexity measures. Biomed Signal Process Control 2020;57:101753. https://doi.org/10.1016/J.BSPC.2019.101753.

[67] Ma Y, Zhang D, Xu J, Pang H, Hu M, Li J, et al. Explainable machine learning model reveals its decision-making process in identifying patients with paroxysmal atrial fibrillation at high risk for recurrence after catheter ablation. BMC Cardiovasc Disord 2023;23:1–13. https://doi.org/10.1186/S12872-023-03087-0/FIGURES/6.

[68] Guo C-Y;, Chang K-H, Guo C-Y, Chang K-H. A Novel Algorithm to Estimate the Significance Level of a Feature Interaction Using the Extreme Gradient Boosting Machine. Int J Environ Res Public Heal 2022, Vol 19, Page 2338 2022;19:2338. https://doi.org/10.3390/IJERPH19042338.

[69] Wang H, Liang Q, Hancock JT, Khoshgoftaar TM. Feature selection strategies: a comparative analysis of SHAP-value and importance-based methods. J Big Data 2024;11:1–16. https://doi.org/10.1186/S40537-024-00905-W/TABLES/15.

[70] Hakkal S, Lahcen AA. XGBoost To Enhance Learner Performance Prediction. Comput Educ Artif Intell 2024;7:100254. https://doi.org/10.1016/J.CAEAI.2024.100254.

[71] Wade C. Hands-On Gradient Boosting with XGBoost and scikit-learn: Perform accessible ... - Corey Wade, Kevin Glynn - Google Boeken. Packt Publishing Ltd.; 2020.

[72] Introduction to Boosted Trees — xgboost 2.1.3 documentation n.d. https://xgboost.readthedocs.io/en/stable/tutorials/model.html (accessed February 17, 2025).

[73] XGBoost: A Comprehensive Guide, Model Overview, Analysis, and Code Demo using Paperspace GPUs n.d. https://blog.paperspace.com/xgboost-a-comprehensive-guide-to-model-overview-analysis-and-code-demo-using/ (accessed February 17, 2025).

[74] Sholeh M, Lestari U. Evaluation of Data Clustering Accuracy using K-Means Algorithm. Int J Multidiscip Approach Res Sci E n.d. https://doi.org/10.59653/ijmars.v2i01.504.

[75] Wang W. Clustering Task. Princ Mach Learn 2025:449–79. https://doi.org/10.1007/978-981-97-5333-8_14.

[76] Firman Ashari I, Dwi Nugroho E, Baraku R, Yanda IN, Liwardana R. Analysis of Elbow, Silhouette, Davies-Bouldin, Calinski-Harabasz, and Rand-Index Evaluation on K-Means Algorithm for Classifying Flood-Affected Areas in Jakarta. J Appl Informatics Comput 2023;7:2548–6861.

[77] Baiardi F, Kholidy H, Tekeoglu A, Chen M, Hilda M, Louk L, et al. Revisiting Gradient Boosting-Based Approaches for Learning Imbalanced Data: A Case of Anomaly

Detection on Power Grids. Big Data Cogn Comput 2022, Vol 6, Page 41 2022;6:41. https://doi.org/10.3390/BDCC6020041.

[78] Ma T, Wu L, Zhu S, Zhu H. Multiclassification Prediction of Clay Sensitivity Using Extreme Gradient Boosting Based on Imbalanced Dataset. Appl Sci 2022, Vol 12, Page 1143 2022;12:1143. https://doi.org/10.3390/APP12031143.

[79] Hancock J, Khoshgoftaar TM, Johnson JM. Informative Evaluation Metrics for Highly Imbalanced Big Data Classification. Proc - 21st IEEE Int Conf Mach Learn Appl ICMLA 2022 2022:1419–26. https://doi.org/10.1109/ICMLA55696.2022.00224.

[80] Rahman MM, Davis DN. Addressing the Class Imbalance Problem in Medical Datasets. Int J Mach Learn Comput 2013:224–8. https://doi.org/10.7763/IJMLC.2013.V3.307.

[81] Spelmen VS, Porkodi R. A Review on Handling Imbalanced Data. Proc 2018 Int Conf Curr Trends Towar Converging Technol ICCTCT 2018 2018. https://doi.org/10.1109/ICCTCT.2018.8551020.

[82] Michaud GF, Stevenson WG. Atrial Fibrillation. N Engl J Med 2021;384. https://doi.org/10.1056/NEJMCP2023658.

[83] Liu Y, Liu Q, Yang Y, Zhang C, Yin H, Wu J, et al. Effect of radiofrequency catheter ablation on left atrial structure and function in patients with different types of atrial fibrillation. Sci Reports 2022 121 2022;12:1–8. https://doi.org/10.1038/s41598-022-13725-w.

[84] Winkle RA, Mead RH, Engel G, Salcedo J, Brodt C, Barberini P, et al. Very long term outcomes of atrial fibrillation ablation. Hear Rhythm 2023;20:680–8. https://doi.org/10.1016/J.HRTHM.2023.02.002.

8. APPENDIX

The appendix in this thesis contain supplementary material that provide further context and clarification, allowing for a deeper understanding of the findings of the studies and the corresponding analysis.

8.1 LITERATURE REVIEW OF INTRA-ATRIAL ELECTROGRAM FEATURES

This section provides a detailed review of each category and its corresponding features, as outlined in *Chapter 2.3.4*, along with relevant formulas and supporting literature. The cited works offer foundational insights into the methodologies for analyzing the characteristics of both STD-based and non-STD-based IAEs.

ENTROPY ATTRIBUTES

The entropy attributes derived from each IAE, include the sample entropy (SampEn) and the Shannon entropy (ShEn), are used to measure the degree of regularity or irregularity in the intra-atrial signals [1–3]. These attributes provide valuable insights into cardiac dynamics, particularly in arrhythmias such as AF, by capturing the complexity and variability of intra-atrial electrical activity [4,5]. Furthermore, combining these entropy metrics enhances the ability to quantify arrhythmia severity and localize reentry circuits, or focal drivers, that contribute to persistent arrhythmias [4,6,7]. Identifying these reentry circuits or focal drivers provides a potential guiding ablation strategy, as they are hypothesized to play a leading role in sustaining AF [2,6].

The SampEn is a non-linear metric used to quantify unpredictability of a one-dimensional signal, such as the IAE, measuring the complexity and irregularity of a time series by evaluating the repetitiveness appearing of a pattern in the time series of the signal [3,7,8]. This approach allows the SampEn to characterize IAE fractionation, where a high SampEn indicates irregular and complex signals which are unpredictable [6,8,9]. SampEn was introduced by Richman et al [9]. as a measure of system complexity and has been utilized to analyze biomedical signals that are prone to noise [3]. It is derived by computing the conditional probability that two sequences, which are similar for m-points, remain similar for *(m+1)*-points at a dissimilarity level under a certain tolerance threshold r [2]. The following formula is used to calculate the SampEn [2,3]:

$$SampEn(m, r, N) = -\ln\left(\frac{\mathbb{1}(d(X_i^{m+1}, X_j^{m+1}) \leq r)}{\mathbb{1}(d(X_i^m, X_j^m) \leq r)}\right) \quad (1)$$

, where *m* is the length of the sequence being compared within the IAE (embedding dimension), *r* is the tolerance threshold, *N* is the total number of samples in the IAE, $d(X_i^m, X_j^m)$ is the distance function with the vectors $X_i^m = \{x(i),\ x(i+1), \ldots, x(i+m-1)\}$ and $X_j^m = \{x(j),\ x(j+1), \ldots, x(j+m-1)\}$, and $\mathbb{1}$ is the indicator function for $i \neq j$. The number of matches ($d(X_i^m, X_j^m) \leq r$) can be increased by either decreasing *m* or increasing *r*. However, this may affect the ability of SampEn to differentiate between classes [2]. Both parameters represent a trade-off between accuracy and discrimination capability, and there are currently no established guidelines for their optimal selection [2]. In the study, it is chosen to set $m = 2$ and $r = 0.2 \cdot \sigma_{X_i}$ [2,3].

The ShEn is a non-linear metric used to quantify the irregularity of a signal by measuring the uncertainty or randomness in the probability distribution in the time series of the signal [6,8]. By analyzing the amplitude distribution within the signal histogram, the ShEn reflects the unpredictability inherent in the time series of the signal [10]. This approach allows the ShEn to

provide an indicator of intra-atrial rotational activation patterns, such as reentry circuits, where a high ShEn is typically associated with the pivot zones of these circuits [6,10,11]. The ShEn is derived by calculating the expected amount of information required to describe the state of the signal, based on the probability distribution of the amplitude p_i within all possible signal states *i* [4]. The following formula is used to calculate the ShEn [4,8,11]:

$$ShEn\ (N) = -\sum_{i=1}^{N} p_i \cdot \log\ (p_i) \qquad (2)$$

, where *N* is the number of amplitude bins in the IAE, and p_i is the probability distribution of the IAE amplitude within a particular *i*-th bin. The probability p_i can be estimated by dividing the number of occurrences of amplitude values within each *i*-th bin by the total number of samples in the IAE. The choice of amplitude bins *N* affects the resolution of the ShEn calculation [12]. A higher number of bins *N* is advantageous for signals with significant amplitude variability, in case of arrhythmias [12]. Therefore, in this study, $N = 100$ bins are used to achieve a balance between capturing meaningful signal details and prevent overfitting or loss of important information [8,13].

In the analysis of STD-based IAEs, the SampEn and the ShEn provide metrics for quantifying the irregularity of fractionation patterns in the IAE [6]. The SampEn captures the timing of local electrical activations, identifying regions with high temporal dispersion [2,14]. In contrast, the ShEn assesses the amplitude distribution over time in the IAE, reflecting temporal complexity and amplitude dispersion [8,10,15]. Together, these entropy metrics capture the temporal complexities, enabling a detailed analysis of STDs. The hypothesis is that the STD-based IAEs show high SampEn and ShEn, compared to non-STD-based IAEs, because of the unpredictable and irregular electrical activations in complex, or fractionated, IAEs [10,15].

FREQUENCY ATTRIBUTES

The frequency attributes, include the dominant frequency (DF) and the organization index (OI), are used to measure the rate of local electrical activation in the atria along with spectral dispersion in the frequency domain [16,17]. These attributes provide valuable insights into the organization of the electrophysiological changes in the atria by assessing the complexity of AF conduction and spatial heterogeneity, with the aim to understand the pathophysiology of the electrical mechanism in AF [16]. Additionally, it is used to estimate the degree of atrial electrical remodeling by quantifying the level of organization [16–18]. Furthermore, by combining these metrics, critical, well-organized, and high frequency intra-atrial regions have been identified as areas for maintaining AF and are therefore considered potential ablation targets to improve success rates of AF termination [17,19,20]. Before calculating the DF and the OI, the IAEs are preprocessed using the following steps [18,19]: (1) bandpass filtering with cutoff frequencies of 30 Hz and 250 Hz, (2) rectification, (3) lowpass filtering with a 20 Hz cutoff, (4) detrending, (5) application of a Hanning window, and (6) zero-padding. These steps collectively optimize the IAE for frequency domain analysis. The bandpass and lowpass filters remove noise and extract significant signal components, while rectification and detrending ensure standardization for consistent analysis [18,19]. The Hanning window and zero-padding further refine the signal for accurate spectral representation, enhancing the reliability of DF and OI calculations [18].

The DF is a non-linear metric used to quantify the local activation rate of a signal, analogous to the AF cycle length, by measuring the most prominent frequency within the frequency spectrum of the signal [16,18]. This method allows the DF to characterize atrial regions with rapid electrical activation rates, reflected by a high DF, indicating primary drivers of electrical activity within the atrial tissue [19]. These regions are often associated with remodeled atrial substrate, signifying changes that contribute to the persistence of AF [16,17]. The DF is

determined by identifying the frequency with the highest amplitude in the frequency spectrum of the IAE [16]. This frequency spectrum is obtained by calculating the power spectral density using the *fast Fourier transform (FFT)* of the preprocessed IAE. The following formula is used to calculate the DF:

$$DF = max_{f\in(4,12)}(|FFT(IAE(t))|^2) \quad (3)$$

, where *f* represents the frequency, *IAE(t)* is the IAE in the time-domain, and *FFT* is the fast Fourier transform. The term $max_{f\in(4,12)}$ identifies the peak frequency within the power spectral density in the range of 4-12 Hz. This specific range is chosen based on literature indicating that DFs in IAEs during AF typically ranges between 4 and 12 Hz [17,18]. However, while DF is effective for analyzing stationary signals, it becomes less reliable when analyzing more chaotic signals, as these may be affected by noise, misleading phase information, and variable activation times that cause signals distortion [17,21]. Consequently, the OI is also calculated to complement this analysis [17].

The OI is a non-linear metric used to quantify the organization of the signal in relation to the DF by evaluating the spectral dispersion within the frequency domain of the IAE. This method allows the OI to assess both the periodicity of the signal and the variability of the frequency components within the spectrum [4,16]. The ability to measure how consistently periodic the atrial activity is, provide insight into how organized the rhythm is. More periodic and organized atrial activity, with high OI, is potentially more effective in driving the overall rhythm and may display a greater tendency to exhibit recurring spatial patterns across atrial regions. This capacity for identifying structured electrical behavior can be particularly valuable in understanding AF mechanisms [16,19,20]. The OI is determined by calculating the area under the spectral power of the DF peak and its three harmonics over a 0.75 Hz window, compared to the total power of the frequency spectrum in the range of 4 till 25 Hz [17]. The OI is bounded between zero and one and is calculated by the following formula:

$$OI = \frac{P_{DF,n}}{P_{total}} = \frac{\sum_{n=1}^{4}\sum_{f=(n\cdot DF)-0.75}^{(n\cdot DF)+0.75} PSD(f)}{\sum_{f=3}^{f=(4\cdot DF)+0.75} PSD(f)} \quad (4)$$

, where *P* is the power, *DF* is the dominant frequency, *n* is the number of harmonics of the DF, *f* is the frequency, and *PSD(f)* is the power spectrum density of the IAE.

Previous studies indicate that the DF and OI of STD-based IAEs provide significant insight into AF mechanisms [16,19,22]. Sustained AF leads to structural and electrical remodeling, creating substrates prone to rapid local activations that can perpetuate arrhythmias. Although high DF regions may indicate critical sites for sustaining AF, their inherent spatio-temporal instability poses challenges for consistent localization during ablation. However, the identification of recurrent high DF activity in these regions suggests a level of organized periodic behavior, which may improve understanding of AF dynamics and aid in targeting potential ablation sites [19]. Areas of high frequency and high degree of organization have been proposed as critical regions for maintaining AF [22]. These areas have been recently proposed as ablation targets to improve success rates in AF termination.

RECURRENCE ATTRIBUTES

The recurrence attributes are used to identify recurring patterns that reveal underlying non-linear and potentially unstable periodic dynamics of AF, derived through recurrence quantification analysis (RQA) [23]. The RQA assesses the relationship between the atrial rate and spectral center frequency, enabling the classification of fractionated or temporally unstable

IAEs [23]. Additionally, the RQA effectively identifies local intra-atrial dynamics, particularly regions of stable or repetitive driver activity that sustain AF through high recurrence patterns, distinguishing these critical atrial substrates from passive regions exhibiting chaotic or erratic intra-atrial signals [23,24]. The RQA characterizes these dynamics through state-space reconstruction in a recurrence plot (RP), a two-dimensional array composed of zeros and ones. A value of one in position *(i,j)* corresponding to the system being in the same place in phase-space at times t_i and t_j [23,25]. The RP is constructed using the following formula:

$$R_{i,j} = \theta \cdot \left(\varepsilon - \left\|x_i - x_j\right\|\right), \qquad with\ \theta(x) \coloneqq \begin{cases} 1, x > 0 \\ 0, x \leq 0 \end{cases}$$

In this equation, $\theta(x)$ is the Heaviside function, ε is the recurrence threshold, and $x_{i,j}$ denotes the state vector (*i,j = 1,…,N*) [23,25,26]. The recurrence threshold ε is set at 25% of the pairwise distances in the phase-space to optimize signal discrimination [26,27]. Prior to construction of the RP, the IAEs are preprocessed using a bandpass filter with cutoff frequencies of 30 Hz and 250 Hz to remove high-frequency noise and baseline drift [24]. Once the RP is constructed, various recurrence attributes are derived to quantify the general, diagonal line and vertical line structures, which further characterize the underlying intra-atrial dynamics [23,24].

The recurrence rate (RR) quantifies the density of recurring states in the RP, providing insights into the dynamics of atrial. The RR differentiates between regions of repetitive or structured activity, typically associated with stable AF drivers, and regions of chaotic or erratic signals. The RR is calculated using the following formula [25]:

$$RR = \frac{1}{N^2}\sum_{i,j} R_{i,j} \qquad (5)$$

In this equation, $R_{i,j}$ represents the recurrence matrix, equivalent to the RP, and *N* is the total number of analyzed points in time. A higher RR indicates more regularity and less chaos in the atrial electrical activity.

The phase-space trajectory of the IAEs often reveals irregular and complex patterns, which are represented as the diagonal line and vertical line structures in the RP. These diagonal line structures provide quantitative metrics, including the determinism (DET), the average diagonal line length (Davg), the maximum diagonal line length (Dmax) and the diagonal line entropy (Dent), that are essential for understanding the spatio-temporal, predictability and complexity of the system's dynamics. The DET is a non-linear metric used to quantify the extent to which recurring patterns in the IAEs remain consistent over time [25,28]. Despite the inherent irregularity of the IAEs in AF, periods of relative stability can still be identified, where closely evolving segments in the phase-space trajectory remain close for some time [27]. This is quantified as the probability that two closely evolving segments of the phase-space trajectory will remain near each other in subsequent time steps, and is calculated by the following formula [25]:

$$DET = \frac{\sum_{l=l_{min}}^{N} l \cdot P(l)}{\sum_{l=1}^{N} l \cdot P(l)} \qquad (6)$$

In this equation, $P(l)$ is the frequency distribution of the lengths *l* of the diagonal lines and l_{min} is the minimal length of a diagonal line necessary to be considered. In this analysis, a $l_{min} = 2$ is used. A lower DET indicates more chaotic or fragmented IAEs, while higher values indicate more predictable patterns, although it remain affected by the inherent irregularities of the system [7]. The D_{avg} is a non-linear metric providing an indication of the duration for which

consistent, recurring patterns in the IAE persist. This metric reflects the predictability time of the system, representing how long stable, complex conduction patterns remain intact before being disrupted, calculated with the following formula [27,28]:

$$D_{avg} = \frac{\sum_{l=l_{min}}^{N} l \cdot P(l)}{\sum_{l=l_{min}}^{N} P(l)} \quad (7)$$

Also in this equation, $P(l)$ is the frequency distribution of the lengths *l* of the diagonal lines and l_{min} is the minimal length of a diagonal line necessary to be considered. A lower D_{avg} indicate shorter, more frequent fragmentation and irregularity in both spatial and temporal domain, while higher values suggest more persistent and predictable patterns of recurrence [27,28]. The D_{max} is a non-linear metric that identifies the longest diagonal line in the RP, representing the maximum duration of recurrence for a dynamic state within the IAE [27]. This metric is associated with the predictability of the system, defined as [28]:

$$D_{max} = \max\{l_1, l_2, \dots, l_{N-1}, l_N\} \quad (8)$$

Here, $l_1, l_2, \dots, l_{N-1}, l_N$ are the lengths of the diagonal lines in the RP. Shorter D_{max} indicate limited predictability, often due to irregularities such as conduction block or re-entry circuits [23]. The D_{ent} is a non-linear metric measuring the complexity and variability of the diagonal line lengths in the RP, reflecting the heterogeneity and STD of the system's dynamics. This metric provides insight into the complexity of the spatio-temporal behavior in the IAE, calculated with the following formula [28]:

$$D_{ent} = -\sum_{l=l_{min}}^{N} \frac{P(l)}{\sum_{l=l_{min}}^{N} P(l)} \cdot \ln\left(\frac{P(l)}{\sum_{l=l_{min}}^{N} P(l)}\right) \quad (9)$$

In this equation, $P(l)$ is the frequency distribution of the lengths *l* of the diagonal lines, l_{min} is the minimal length of a diagonal line necessary to be considered, and *N* is the total number of analyzed diagonal lines in the RP. High D_{ent} exhibit more complex and irregular spatio-temporal patterns, whereas lower D_{ent} suggest more consistent and regular dynamics [23].

In contrast, the vertical line structures in the RP represent instances when the system revisits or remains near a specific state over time, indicating temporal stability rather than deterministic progression [23,27,29]. While diagonal line structures capture the predictability of the system's dynamics, vertical line structures highlight the duration the system stays in particular states or regions of phase-space, providing valuable insights into the stability and temporal recurrence [27]. These vertical line structures provide metrics, including the laminarity (LAM), the trapping time (TT), the maximal vertical line length (V_{max}) and the vertical line entropy (V_{ent}) , which quantify the persistence of states and the system's temporal behavior. The LAM is a non-linear metric used to quantify the persistence of vertical line structures in the RP, indicating periods when the system exhibits repeated stable behavior in specific regions of the phase-space trajectory [27]. This metric measures the probability that a state will not change for the next time step, reflecting the tendency for the system to remain a stable state for extended periods. The LAM is calculated by the following formula [27]:

$$LAM = \frac{\sum_{v=v_{min}}^{N} v \cdot P(v)}{\sum_{v=1}^{N} v \cdot P(v)} \quad (10)$$

In this equation, $P(v)$ is the frequency distribution of the lengths *v* of the vertical lines and v_{min} is the minimal length of a vertical line necessary to be considered. In this analysis, a $v_{min} = 2$ is used. A higher LAM suggests more periods of stable behavior, while lower values indicate more erratic and unpredictable dynamics. The TT is a non-linear metric that refers to the total time the system spends in recurrent states [27,28]. It quantifies how long the system remains in a particular state of phase-space before transitioning to another state [27]. This metric is an important indicator of the system's ability to sustain a given dynamic state over time, reflecting periods of stability or trapping within certain states. The TT is calculated with the following formula [25]:

$$TT = \frac{\sum_{v=v_{min}}^{N} v \cdot P(v)}{\sum_{v=v_{min}}^{N} P(v)} \quad (11)$$

Here also, $P(v)$ is the frequency distribution of the lengths *v* of the vertical lines and v_{min} is the minimal length of a vertical line necessary to be considered, and *N* is the total number of vertical lines. A higher TT indicates more frequent transitions and less temporal stability, suggesting that the system exhibits prolonged stability [25]. The V_{max} is a non-linear metric that measures the longest vertical line in the RP, representing the maximum duration the system remains in a particular state without transitioning [28]. This metric measures the maximum period of recurrence, reflecting the extend of temporal stability within the system, defined as [27]:

$$V_{max} = \max\{v_1, v_2, \dots, v_{N-1}, v_N\} \quad (12)$$

Here, $v_1, v_2, \dots, v_{N-1}, v_N$ are the lengths of the vertical lines in the RP. A shorter V_{max} indicates frequent state transitions and less temporal stability, while a longer V_{max} suggests sustained stability in certain regions of phase space. Thus, V_{max} is an important indicator of how long the system maintains stable dynamics before changing states. The V_{ent} is a non-linear metric that quantify the complexity and irregularity of the vertical line lengths in the RP,. It captures variations in recurrence durations, offering insights into the heterogeneity of the time the system spend in different states [30]. This metric provides important information into the temporal variability and irregularity of the IAE, reflecting the underlying dynamics and stability of the temporal behavior [29]. The Vent is calculated with the following formula [30]:

$$V_{ent} = -\sum_{v=v_{min}}^{N} \frac{P(v)}{\sum_{v=v_{min}}^{N} P(v)} \cdot \ln\left(\frac{P(v)}{\sum_{v=v_{min}}^{N} P(v)}\right) \quad (13)$$

In this equation, $P(v)$ is the frequency distribution of the lengths *v* of the vertical lines, v_{min} is the minimal length of a diagonal line necessary to be considered, and *N* is the total number of analyzed diagonal lines in the RP. High V_{ent} exhibit more complex and irregular spatio-temporal patterns, whereas lower D_{ent} suggest more consistent and regular dynamics. High V_{ent} indicate a more complex and less predictable system, characterized by greater irregularity in the durations of recurrences [23].

The RQA metrics provide valuable insights into the dynamic behaviors and underlying mechanisms of IAEs, particularly in the context of AF. These metrics assess both the predictability and complexity of the system's dynamics, highlighting regions of stability, irregularity, and transition in the phase-space trajectory of the IAEs [23,24]. They provide insights into the spatio-temporal structure of atrial electrical activity, allowing for the identification of stable regions, chaotic behavior, and recurrent dynamics. By quantifying state recurrence and persistence, RQA metrics reveal the stability and complexity of the rhythm

[23,24]. Higher values indicate more stable, predictable dynamics, while lower values point to more chaotic behavior. Overall, RQA provides valuable tools for analyzing electrogram properties, enhancing our understanding of AF mechanisms and aiding therapeutic approaches like ablation.

WAVE SIMILARITY ATTRIBUTES

The wave similarity attributes, including the intra-atrial cycle length (CL) with their standard deviation (CL_{std}), the peak-to-peak amplitude (PP) with their standard deviation (PP_{std}) and the wave similarity index (WSI), are used to measure waveform repetitiveness and temporal organization of consecutive atrial waveforms [31]. These attributes enable the identification of both repetitive and similar waveforms, such as those associated with rotors or focal drivers, providing insights into the temporal, amplitude, and waveform consistency of activation waves in IAEs, reflecting atrial conduction dynamics [32,33]. Combining these attributes is effective in distinguishing critical sites characterized with a fast activation rate and high organization from passive conduction regions [33]. Furthermore, these features enable the spatio-temporal characterization of AF, offering insights in supporting precise localization of arrhythmogenic drivers for targeted ablation and improved treatment outcomes [31,33,34].

The CL and CL_{std} are non-linear metrics for analyzing IAEs, offering insights into atrial conduction dynamics. The CL quantifies the average interval between successive activations, providing a temporal measure of atrial activation patterns. The CL_{std} complements this by capturing the variability in these intervals, offering a measure of rhythm regularity. Together, these features enable precise assessment of the temporal organization and stability of intra-atrial activation [32,35]. By calculating the CL and the CL_{std}, it is possible to characterize conduction properties, detect irregular activation patterns, and differentiate between regions of stable and unstable conduction. The following formulas are used for the CL and the CL_{std} [32]:

$$CL = \frac{1}{N}\sum_{i=1}^{N}(t_i - t_{i-1}) \quad (14)$$

, where N is the total number of cycle lengths, t_i is the time of the i-th peak in the IAE, and CL is the average cycle length. A prolonged CL indicates conduction delays due to fibrosis, ischemia, or scar tissue, which can predispose to reentrant arrhythmias and a high CL_{td} suggest irregular heart rhythm due to irregular atrial activation [34].

The PP and PP_{std} are non-linear metrics, providing insights into the amplitude and consistency of local intra-atrial electrical activity. The PP represents the average voltage difference between the highest and lowest points of activation waves, reflecting the strength of local atrial depolarization. The PP_{std} captures the variability in these amplitudes, offering a measure of stability and consistency in atrial conduction. These features are invaluable for assessing the electrophysiological state of the atrial tissue, identifying areas with stable or unstable electrical activity, and understanding the spatial dynamics of atrial activation [33,34]. The following formulas are used for the PP and PP_{std}:

$$PP = \frac{1}{N}\sum_{i=1}^{N}(V_{max,i} - V_{min,i}) \quad (15)$$

, where N is the total number of activation waves, $V_{max,i}$ is the maximal voltage of the i-th peak, $V_{min,i}$ is the minimal voltage of the i-th peak in the IAE, and PP is the average peak-to-peak amplitude. A high PP represents a strong local electrical activity, reflecting effective

depolarization and healthy atrial tissue and a high PP_{std} indicates variability in the PP suggesting unstable and unorganized electrical activity in that region in the atria [7].

The WSI is a quantitative non-linear metric used in IAE analysis to assess the repetitiveness and organization of activation waveforms. By evaluating the similarity between consecutive waveforms, the WSI provides a continuous grading scale that spans the spectrum of IAE morphologies, from highly disorganized complex signals to highly regular and repetitive waveforms. This metric is based on the rationale that regular sources of activation, such as rotors or focal drivers, produce consistent and repetitive waveforms, making WSI a valuable tool for identifying arrhythmogenic regions. Besides, the WSI also facilitates spatio-temporal characterization of AF dynamics, capturing temporal changes such as shortening cycle lengths and decreasing waveform regularity during the transition from organized to disorganized AF [33]. The WSI is calculated with the following formula:

$$WSI = \frac{1}{N}\sum_{i=1}^{N}\frac{cov(w_i, w_{i+1})}{\sqrt{var(w_i)\cdot var(w_{i+1})}} \quad (16)$$

, where w_i and w_{i+1} are the activation waveforms of two consecutive activation cycles, $cov(w_i, w_{i+1})$ represents the covariance between w_i and w_{i+1}, $var(w_i)$ and $var(w_{i+1})$ are the variances of w_i and w_{i+1}, and N is the total number of waveform pairs in the IAE. A high WSI indicates consistent and organized waveforms, typically associated with stable arrhythmogenic sources such as rotors or focal drivers, while a low WSI points to irregular and disorganized activation, characteristic of chaotic conduction patterns [31,33].

In STD-based IAEs, these features reveal significant variability in both the timing and amplitude of atrial activation, reflecting the heterogeneous and dynamic nature of conduction [7]. Specifically, higher CL_{std} and PP_{std} values indicate greater temporal and amplitude variability, pointing to regions with disorganized and irregular conduction, often associated with arrhythmogenic activity such as rotors or focal drivers in AF. The WSI in these regions tends to be lower, reflecting a lack of regularity and repetitive waveform patterns characteristic of chaotic atrial activation. Additionally, CL and PP further differentiate these two types of IAEs: STD-based IAEs typically exhibit more variability in both CL and PP, with shorter and more irregular cycle lengths and greater amplitude variability, indicative of faster and less stable local electrical activity. Non-STD-based IAEs, on the other hand, show more stable CL and PP, suggesting more regular depolarization and less arrhythmogenic potential [7,32,34]. Therefore, the hypothesize is that STD-based IAEs will exhibit greater variability in CL, CL_{std}, PP, PP_{std}, and WSI compared to non-STD-based IAEs. This variability is indicative of regions with irregular conduction patterns, which are more likely to be arrhythmogenic. Therefore, these features could serve as valuable markers for identifying critical sites of atrial fibrillation, facilitating targeted ablation strategies and improving treatment outcomes.

INTERVAL ATTRIBUTES

The interval attributes, including the interval confidence level (ICL), the shortest complex interval (SCI) and the average duration of complex intervals (ACI), are measuring temporal consistency of fractionation between successive voltage peaks [35]. These attributes provide valuable insights in temporal and electrical conduction properties for fractionation, activation rapidity, and conduction regularity, offering critical guidance for the design of targeted ablation strategies and improving the understanding of arrhythmic substrates [35–37].

The ICL quantifies the number of identified complex intervals within the IAE, indicating the fractionation present in the electrogram. Complex intervals are defined as time intervals between successive peaks and troughs occurring within the voltage window that are longer

than 50 milliseconds and shorter than 110 milliseconds [37]. The ICL is calculated with the following formula:

$$ICL = |\{x \, \epsilon \, (t_i - t_{i-1}) | 50 \leq x \leq 110\}|$$

(17)

, where t_i is the time of the *i*-th peak in the IAE where the signal is in the threshold of 0.05 mV and 0.15 mV, and *x* is the complex interval. A high ICL quantifies a large number of identified complex intervals within the IAE, indicating fractionation [35,37].

The SCI is a non-linear metric and provides insight into the fastest atrial conduction interval, which can be crucial for assessing the risk of rapid atrial rates and can be a marker for high-frequency atrial activity [36,37]. The SCI can be calculated using the following formula:

$$SCI = min(t_i - t_{i-1})$$

(18)

, where t_i is the time of the *i*-th peak in the IAE and *min()* is the shortest complex interval. A low SCI indicates that the atrial tissue is depolarizing at a faster rate, with very short intervals between activations. This often suggests rapid conduction within the atrium and can be a marker for high-frequency atrial activity.

The ACI is a non-linear metric and provides a measure of the average duration of complex intervals within the IAE. It reflects the overall consistency of activation across the atrial myocardium [36]. The ACI can be calculated using the following formula:

$$ACI = \frac{1}{N} \sum_{i=1}^{N} (t_i - t_{i-1})$$

(19)

where *N* is the total number of complex intervals, and t_i is the time of the *i*-th peak in the IAE. A high ACI indicates longer, more consistent complex intervals, suggesting stable conduction patterns, while a low ACI reflects more variable and disorganized activation across the atrial tissue [36].

These metrics offer valuable insights into the fractionation, rapidity, and regularity of atrial electrical conduction. High ICL values suggest stable, organized electrical activity, while low ICL values indicate more fractionated and irregular conduction. Low SCI values reflect faster atrial depolarization, often associated with regions exhibiting high-frequency activation, while lower ACI values point to increased variability in conduction, highlighting potential arrhythmogenic areas [36,37]. The hypothesis is that STD-based IAEs, characterized by more uniform and organized conduction, will show higher ICL values and lower SCI and ACI values compared to non-STD-based IAEs, which would exhibit more fractionated, rapid, and variable conduction. By analyzing these attributes, clinicians can refine their understanding of atrial arrhythmias, allowing for more precise targeting of arrhythmogenic regions, such as rotors or focal drivers, during ablation procedures.

VOLTAGE ATTRIBUTES

The voltage attributes, including the voltage absolute value (VAV), kurtosis (KUR), and skewness (SKE), measures the distribution of voltages over multiple bipoles at a certain time in different intra-atrial regions [38]. These attributes provide insights in the electrophysiological behavior of atrial tissue, revealing regions with varying electrical activity, abnormal conduction patterns, and potential arrhythmogenic drivers. This can refine ablation strategies, enabling more targeted and effective treatment for AF [38,39].

The VAV represents the average peak voltage observed over time, reflecting the strength of electrical activity at a given site [38]. The VAV is calculated using the following formula:

$$VAV = \frac{1}{N}\sum_{i=1}^{N}|V_i| \quad (20)$$

, where V_i represents the maximum voltage at the i-*th* time point and *N* is the total number of voltage bins. In this analysis it is chosen *N = 50*.

The KUR quantifies the tailedness of the voltage distribution, offering insight into the presence of extreme voltage values or outliers, which can indicate irregular or highly fragmented activation patterns [38]. The KUR is calculated using the following formula:

$$KUR = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{V_i - \mu}{\sigma}\right)^4 - 3 \quad (21)$$

, where V_i is the voltage at the *i-th* bin, μ is the mean of the voltages, σ is the standard deviation, and *N* is the total number of bins. A high positive KUR suggest the presence of extreme voltage values, while negative values indicate a more uniform distribution [38].

The SKE measures the asymmetry of the voltage distribution, indicating whether the voltage time series is skewed toward higher or lower values, potentially highlighting regions with uneven activation or abnormal electrical conduction [38]. The SKE is calculated with the following formula:

$$SKE = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{V_i - \mu}{\sigma}\right)^3 \quad (22)$$

, where V_i is the voltage at the *i-th* bin, μ is the mean of the voltages, σ is the standard deviation, and *N* is the total number of bins. A positive SKE indicates a distribution with a long tail to the right, while negative SKE indicates a tail on the left.

In STD-based IAEs, where there is significant variation in the activation pattern across different regions of the atrium, these attributes may show pronounced changes compared to non-STD-based IAEs, which exhibit more homogeneous and regular activation. A higher VAV may indicate increased voltage variation across regions, suggesting fragmented or disorganized conduction. Elevated KUR could reflect extreme voltage peaks, typical of localized arrhythmogenic sources like rotors. Positive skewness may signal rapid, high-voltage activations, while negative skewness could indicate areas with attenuated activity. In contrast, non-STD-based IAEs show more stable, uniform distributions, with lower VAV, near-zero kurtosis, and less skewness [38,39]. The hypothesis is that an STD-based IAE correlates with higher VAV, positive kurtosis, and greater skewness, pointing to irregular electrical behavior, while non-STD-based IAEs exhibit more regular, organized conduction.

8.1.1 REFERENCES

[1] Zhao L, Li J, Wan X, Wei S, Liu C. Determination of Parameters for an Entropy-Based Atrial Fibrillation Detector. Entropy 2021, Vol 23, Page 1199 2021;23:1199. https://doi.org/10.3390/E23091199.

[2] Cirugeda-Roldán EM, Molina Picó A, Novák D, Cuesta-Frau D, Kremen V. Sample Entropy Analysis of Noisy Atrial Electrograms during Atrial Fibrillation. Comput Math Methods Med 2018;2018. https://doi.org/10.1155/2018/1874651.

[3] Horie T, Burioka N, Amisaki T, Shimizu E. Sample Entropy in Electrocardiogram During Atrial Fibrillation. Yonago Acta Med 2018;61:49. https://doi.org/10.33160/YAM.2018.03.007.

[4] Goldberger JJ, Zaatari G, Mitrani RD, Blandon C, Bohorquez J, Ng J, et al. Comparison of electrogram characteristics in persistent atrial fibrillation. J Cardiovasc Electrophysiol 2024;35:182–97. https://doi.org/10.1111/JCE.16133.

[5] Nicolet JJC, Restrepo JF, Schlotthauer G. Classification of intracavitary electrograms in atrial fibrillation using information and complexity measures. Biomed Signal Process Control 2020;57:101753. https://doi.org/10.1016/J.BSPC.2019.101753.

[6] Ugarte JP, Tobón C, Orozco-Duque A. Entropy Mapping Approach for Functional Reentry Detection in Atrial Fibrillation: An In-Silico Study. Entropy 2019, Vol 21, Page 194 2019;21:194. https://doi.org/10.3390/E21020194.

[7] Nadim A, Li X, Chu G, Soriano D, Masè M, Ravelli F, et al. Unsupervised classification of dimension-reduced principal component scores from persistent atrial fibrillation electrograms 2021;12088:28. https://doi.org/10.1117/12.2606172.

[8] Frusone S, De Almeida RC, Squara F, Zarzoso V. Electrogram Fractionation Metrics in Spatio-Temporal Dispersion Based Catheter Ablation of Persistent Atrial Fibrillation n.d.

[9] Richman JS, Moorman JR. Physiological time-series analysis using approximate entropy and sample entropy. Am J Physiol Heart Circ Physiol 2000;278. https://doi.org/10.1152/AJPHEART.2000.278.6.H2039.

[10] Ganesan AN, Kuklik P, Lau DH, Brooks AG, Baumert M, Lim WW, et al. Bipolar electrogram Shannon entropy at sites of rotational activation implications for ablation of atrial fibrillation. Circ Arrhythmia Electrophysiol 2013;6:48–57. https://doi.org/10.1161/CIRCEP.112.976654/-/DC1.

[11] Ravikumar V. Signal processing approaches for the spatiotemporal analysis of cardiac arrhythmias using intracardiac electrograms 2022.

[12] Dharmaprani D, Dykes L, McGavigan AD, Kuklik P, Pope K, Ganesan AN. Information Theory and Atrial Fibrillation (AF): A Review. Front Physiol 2018;9:957. https://doi.org/10.3389/FPHYS.2018.00957.

[13] Purwani S, Nahar J, Twining C. Analyzing bin-width effect on the computed entropy. AIP Conf Proc 2017;1868. https://doi.org/10.1063/1.4995123/641266.

[14] Furutani N, Nariya Y, Takahashi T, Ito H, Yoshimura Y, Hiraishi H, et al. Neural Decoding of Multi-Modal Imagery Behavior Focusing on Temporal Complexity. Front Psychiatry 2020;11:746. https://doi.org/10.3389/FPSYT.2020.00746/FULL.

[15] Masè M, Faes L, Antolini R, Scaglione M, Ravelli F. Quantification of synchronization during atrial fibrillation by Shannon entropy: validation in patients and computer model of atrial arrhythmias. Physiol Meas 2005;26:911. https://doi.org/10.1088/0967-3334/26/6/003.

[16] Choi Y jung, Sohn JJ, Kwon S, Lee SR, Cha MJ, Choi EK, et al. Relationship between dominant frequency, organization index, and left atrial size in patients with atrial fibrillation. J Cardiovasc Electrophysiol 2020;31:3159–65. https://doi.org/10.1111/JCE.14785.

[17] Li X, Chu GS, Almeida TP, Vanheusden FJ, Salinet J, Dastagir N, et al. Automatic Extraction of Recurrent Patterns of High Dominant Frequency Mapping During Human Persistent Atrial Fibrillation. Front Physiol 2021;12:649486. https://doi.org/10.3389/FPHYS.2021.649486/BIBTEX.

[18] Johner N, Johner, Nicolas. Intracardiac electrogram analysis for an individualized approach to catheter ablation of atrial fibrillation 2021. https://doi.org/10.13097/ARCHIVE-OUVERTE/UNIGE:161102.

[19] Tobón C, Rodríguez JF, Ferrero JM, Hornero F, Saiz J. Dominant frequency and organization index maps in a realistic three-dimensional computational model of atrial fibrillation. EP Eur 2012;14:v25–32. https://doi.org/10.1093/EUROPACE/EUS268.

[20] Honarbakhsh S, Schilling RJ, Keating E, Finlay M, Hunter RJ. Drivers in AF colocate to sites of electrogram organization and rapidity: Potential synergy between spectral analysis and STAR mapping approaches in prioritizing drivers for ablation. J Cardiovasc Electrophysiol 2020;31:1340–9. https://doi.org/10.1111/JCE.14456.

[21] Kong X, Ravikumar V, Mulpuru SK, Roukoz H, Tolkacheva EG. A Data-Driven Preprocessing Framework for Atrial Fibrillation Intracardiac Electrocardiogram Analysis. Entropy (Basel) 2023;25. https://doi.org/10.3390/E25020332.

[22] Almeida TP, Li X, Sidhu B, Bezerra AS, Ehnesh M, Anton I, et al. Dominant Frequency and Organization Index for Substrate Identification of Persistent Atrial Fibrillation. Comput Cardiol (2010) 2021;2021-September. https://doi.org/10.23919/CINC53138.2021.9662648.

[23] Halfar R, Lawson BAJ, dos Santos RW, Burrage K. Recurrence quantification analysis for fine-scale characterisation of arrhythmic patterns in cardiac tissue. Sci Reports 2023 131 2023;13:1–16. https://doi.org/10.1038/s41598-023-38256-w.

[24] Baher A, Buck B, Fanarjian M, Paul Mounsey J, Gehi A, Chung E, et al. Recurrence quantification analysis of complex-fractionated electrograms differentiates active and passive sites during atrial fibrillation. J Cardiovasc Electrophysiol 2019;30:2229–38. https://doi.org/10.1111/JCE.14161.

[25] Navoret N, Jacquir S, Laurent G, Binczak S. Detection of complex fractionated atrial electrograms using recurrence quantification analysis. IEEE Trans Biomed Eng 2013;60:1975–82. https://doi.org/10.1109/TBME.2013.2247402.

[26] Mathunjwa BM, Lin YT, Lin CH, Abbod MF, Sadrawi M, Shieh JS. ECG Recurrence Plot-Based Arrhythmia Classification Using Two-Dimensional Deep Residual CNN Features. Sensors 2022, Vol 22, Page 1660 2022;22:1660. https://doi.org/10.3390/S22041660.

[27] Schinkel S, Dimigen O, Marwan N. Selection of recurrence threshold for signal detection. Eur Phys J Spec Top 2008 1641 2008;164:45–53. https://doi.org/10.1140/EPJST/E2008-00833-5.

[28] Ng J, Borodyanskiy AI, Chang ET, Villuendas R, Dibs S, Kadish AH, et al. Measuring the complexity of atrial fibrillation electrograms. J Cardiovasc Electrophysiol 2010;21:649–55. https://doi.org/10.1111/J.1540-8167.2009.01695.X.

[29] Zeemering S, van Hunnik A, van Rosmalen F, Bonizzi P, Scaf B, Delhaas T, et al. A Novel Tool for the Identification and Characterization of Repetitive Patterns in High-Density Contact Mapping of Atrial Fibrillation. Front Physiol 2020;11:570118. https://doi.org/10.3389/FPHYS.2020.570118/BIBTEX.

[30] Marwan N, Romano MC, Thiel M, Kurths J. Recurrence plots for the analysis of complex systems. Phys Rep 2007;438:237–329. https://doi.org/10.1016/j.physrep.2006.11.001.

[31] Ravelli F, Masè M. Computational mapping in atrial fibrillation: how the integration of signal-derived maps may guide the localization of critical sources. EP Eur 2014;16:714–23. https://doi.org/10.1093/EUROPACE/EUT376.

[32] Ravelli F, Masè M, Cristoforetti A, Del Greco M, Centonze M, Marini M, et al. Anatomic Localization of Rapid Repetitive Sources in Persistent Atrial Fibrillation: Fusion of Biatrial CT Images With Wave Similarity/Cycle Length Maps. JACC Cardiovasc Imaging 2012;5:1211–20. https://doi.org/10.1016/J.JCMG.2012.07.016.

[33] Lin YJ, Lo MT, Chang SL, Lo LW, Hu YF, Chao TF, et al. Benefits of Atrial Substrate Modification Guided by Electrogram Similarity and Phase Mapping Techniques to Eliminate Rotors and Focal Sources Versus Conventional Defragmentation in Persistent Atrial Fibrillation. JACC Clin Electrophysiol 2016;2:667–78. https://doi.org/10.1016/J.JACEP.2016.08.005.

[34] Podziemski P, Zeemering S, Kuklik P, van Hunnik A, Maesen B, Maessen J, et al. Rotors Detected by Phase Analysis of Filtered, Epicardial Atrial Fibrillation Electrograms Colocalize With Regions of Conduction Block. Circ Arrhythm Electrophysiol 2018;11:e005858. https://doi.org/10.1161/CIRCEP.117.005858/SUPPL_FILE/CIRCAE_CIRCAE-2017-005858_SUPP1.PDF.

[35] Almeida TP, Chu GS, Bell MJ, Li X, Salinet JL, Dastagir N, et al. The temporal behavior and consistency of bipolar atrial electrograms in human persistent atrial fibrillation. Med Biol Eng Comput 2018;56:71–83. https://doi.org/10.1007/S11517-017-1667-1.

[36] Almeida TP, Chu GS, Salinet JL, Vanheusden FJ, Li X, Tuan JH, et al. Minimizing discordances in automated classification of fractionated electrograms in human persistent atrial fibrillation. Med Biol Eng Comput 2016;54:1695–706. https://doi.org/10.1007/S11517-016-1456-2.

[37] Lau DH, Zeemering S, Maesen B, Kuklik P, Verheule 1Sander, Schotten U. Catheter Ablation Targeting Complex Fractionated Atrial Electrogram in Atrial Fibrillation. J Atr Fibrillation 2013;6:907. https://doi.org/10.4022/JAFIB.907.

[38] Ghrissi A, Squara F, Montagnat J, Zarzoso V. Identification of spatiotemporal dispersion electrograms in persistent atrial fibrillation ablation using maximal voltage absolute values. Eur Signal Process Conf 2021;2021-January:1239–43. https://doi.org/10.23919/EUSIPCO47968.2020.9287681.

[39] Masuda M, Matsuda Y, Uematsu H, Sugino A, Ooka H, Kudo S, et al. Prognostic impact of atrial cardiomyopathy: Long-term follow-up of patients with and without low-voltage areas following atrial fibrillation ablation. Hear Rhythm 2024;21:378–86. https://doi.org/10.1016/J.HRTHM.2023.12.016.

UNIVERSITY OF TWENTE. | isala