Discrimination Between CFAEs of Paroxysmal and Persistent Atrial Fibrillation with Simple Classification Models of Reduced Features

Emanuela Finotti[1], Edward J Ciaccio[2], Hasan Garan[2],
Vicente Bertomeu-González[3], Raúl Alcaraz[4], José J Rieta[1*]

[1] BioMIT.org, Electronic Engineering Department, Universitat Politecnica de Valencia, Spain.
[2] Department of Medicine - Division of Cardiology, Columbia University, New York, USA.
[3] Clinical Medicine Department, Miguel Hernández University, Elche Spain;
[4] Research Group in Electronic, Biomed. and Telecomm. Eng., Univ. of Castilla-La Mancha, Spain.

Abstract

A significant number of variables to discriminate between paroxysmal and persistent atrial fibrillation (ParAF vs. PerAF) has been widely exploited, mostly assessed with statistical tests aimed to suggest adequate approaches for catheter ablation (CA) of AF. However, in practice, it would be desirable to utilize simple classification models readily understandable. In this work dominant frequency (DF), AF cycle length (AFCL), sample entropy (SE) and determinism (DET) of recurrent quantification analysis were applied to recordings of complex fractionated atrial electrograms (CFAEs) of AF patients, aimed to create simple models to discriminate between ParAF and PerAF. Correlation matrix filters removed redundant information and Random Forests ranked the variables by relevance. Next, coarse tree models were built, optimally combining high-ranking indexes, and tested with leave-one-out cross-validation. The best classification performance combined SE and DF with an Accuracy (Acc) of 88.2% to discriminate ParAF from PerAF, while the highest single Acc was provided by DET reaching 82.4%. Hence, careful selection of reduced sets of features feeding simple classification models is able to discriminate accurately between CFAEs of ParAF and PerAF.

1. Introduction

Atrial fibrillation (AF) requires therapeutic interventions even in patients which suffer no subjective discomfort [1]. Its treatment includes electrical [2] and pharmacological [3] cardioversion, but catheter ablation (CA) has become the first line therapy [4] due to its superiority to other therapies in improving all-cause mortality [5].

In clinical practice, one of the goals for the straightforward characterization of atrial substrate aims at discriminating between patients into paroxysmal (ParAF) and persistent (PerAF) AF, as statistics have shown that the success rate of CA is strictly dependent on the area of ablation and the AF type [6]. Accordingly, this defines a challenge for a precise characterization of the atrial substrate aimed at guiding optimally CA, where, methodologies to distinguish between complex fractionated atrial electrograms (CFAE) of ParAF and PerAF would be very interesting and useful in fast and efficient atrial substrate mapping [7].

Previous works proposed several discriminatory strategies based on non-linear variables assessed with statistical tests, aimed at classifying ParAF and PerAF from CFAEs. To this respect, Ciaccio *et al.* measured CFAEs repetitiveness [7] and quantified the degree of morphological heterogeneity in CFAE deflections [8]; Acharya *et al.* adopted recurrence plots, recurrence quantification analysis (RQA) and entropy measures [9]; Ndrepepa *et al.* and Ravi *et al.* used the atrial fibrillation cycle length [10,11] and Sanders *et al.* employed spectral analysis and the dominant frequency mapping [12].

However, during clinical practice, it would be desirable to utilize straightforward classification models which are readily understandable, fed with features easily applicable to CFAEs. This work exploits non-linear strategies and simple models to discriminate between ParAF and PerAF from CFAEs of patients undergoing CA in AF. We studied the dominant frequency (DF), AF cycle length (AFCL), sample entropy (SE) and determinism (DET) of RQA. The indexes extracted from CFAEs were processed and selected prior to be converted into features for coarse tree classification models.

The main hypothesis is that a careful selection of reduced sets of features feeding simple classification models is able to discriminate accurately between CFAEs of ParAF and PerAF, thus being able to provide straightforward insights into atrial substrate evaluation and improved therapeutical decisions on AF management.

2.   Database and preprocessing

A total of 204 electrograms of 16 seconds were sampled at 977 Hz from ParAF and long-standing PerAF patients who underwent CA at the cardiac electrophysiological laboratory of Columbia University Medical Center. The Internal Review Board approved its acquisition and analysis. For ParAF patients in sinus rhythm, AF was induced by rapid pacing at the coronary sinus or at the lateral wall of the right atrium. 90 and 114 CFAEs were recorded for ParAF and PerAF, respectively, which were identified observing the published criteria in [7] from the four pulmonary veins (PV): left superior (LSPV), left inferior (LIPV), right superior (RSPV) and right inferior (RIPV); and the anterior (ANT) and posterior (POS) free wall of the left atrium.

The CFAEs were segmented into 1, 2 and 4 second-length sequences with no overlapping, thus creating three datasets. The first one consisted of 3264 one second-length sequences; the second one had 1632 two second-length sequences and the third dataset accounted on 816 four second-length sequences.

The signals were band-pass filtered by the acquisition system (0.5-450 Hz), which removed baseline drift and high frequency noise (CardioLab, GE Healthcare, Waukesha, WI), then resampled at 1 KHz and filtered with a wavelet-based method for noise reduction [13].

3.   Methods

3.1.   Sample entropy

SE has been chosen due to its ability to outline the dynamics of atrial activity in AF patients as well as the severity of the atrial remodeling [14]. SE assigned a non-negative value reflecting the complexity of each sequence, with lower values corresponding to more self-similarity in the time series. The maximum template length m and tolerance r were set at 2 and at 0.35 times the standard deviation of the segment [14]. SE was applied to 1, 2 and 4 s-length sequences from the corresponding datasets.

3.2.   Recurrence Plots and RQA measures

In order to produce recurrence plots (RPs), the CFAE segments were considered as time series and their trajectories were projected into a d-dimensional phase space (PS). The PS was reconstructed with Takens' theorem, using an embedding dimension (d) and a time delay (τ) found with Nearest Neighbors and Mutual Information strategies, respectively. The RPs were produced representing each pair of trajectory states as black dots if their distance was below the threshold, which was set at 10% of the mean of the PS diameter [15], and with white dots otherwise.

RPs were used to visualize recurrent patterns within 1, 2 and 4 s-length sequences. The RQA measure selected was the determinism (DET), which quantifies the sequence predictability by measuring the percentage of recurrence points that belong to diagonal lines of a minimum length, set at 50.

3.3.   DF and AF Cycle Length

The DF was measured by identifying the highest spectral peak in the electrophysiological range of interest of 3-12 Hz. The DF has been proven to be generally greater in PerAF and lower in ParAF, as revealed by previous studies [8]. The DF was computed on the 16 s-length CFAEs at the different recording sites using the Welch Periodogram.

The AFCL was measured by detecting local activation waves within the CFAEs [16], inverting the distances between consecutive activations and averaging them. The AFCL was computed on the 16 s-length CFAEs at all the recording sites.

3.4.   Statistical tests and feature selection

The Mann-Whitney test verified the null hypothesis that the indexes were similar for ParAF and PerAF, with significance value set at 0.05. The test was performed for the indexes at each atrial recording site. For the three datasets, the p-values returned were averaged for channels and the resulted values are reported in section 4.

Feature selection was used to remove data containing redundant or irrelevant variables [17]. Two main techniques were combined: the *filters* method, which selected the features independently from the predictor choice in a preprocessing step, and the *wrapper* method, which selected the features with machine learning techniques for classification. For the filters method we applied a filter to the correlation matrix, removing the most redundant variables with respect to the others, with a cutoff value of 0.60. After the correlation matrix filtering, the wrapper method was applied via the Random Forest (RF) algorithm for variable importance, which provides a ranking of the features considered as the most relevant for classification. The RF assigned to each variable a score (Gini impurity) from 0 to 100. The higher the value the more important the feature. In this study, just the variables having scores higher than 40 were selected.

3.5.   Classification models

The classification of ParAF and PerAF was performed using coarse tree models together with leave-one-out cross-validation. Each coarse tree model was built using one possible combination of the variables that passed the features selection process. In particular, for all segment

lengths, all the possible subsets of features were used to find out the coarse tree providing the highest accuracy. Aimed at simplifying even more the classification, a coarse tree model for each variable was created having just one input feature. Accuracy values obtained with the classification process were reported as percentages.

4.   Results

The mean and the standard deviation values resulted by applying SE and DET to 1, 2 and 4 s-length datasets are reported in Table 1 organized by AF type. The SE mean values decreased with the segment length, with greater values in PerAF, as compared to ParAF, thus reflecting a higher degree of disorganization in PerAF. Contrarily, DET values increased with segment length, showing greater values in ParAF than in PerAF, providing ParAF as more predictable than PerAF.

The AFCL and the DF applied to the 16 s-length CFAEs revealed values of AFCL with average and standard deviation of 7.01±1.46 in ParAF, and 7.75±1.56 in PerAF, while DF had respectively average and standard deviation of 5.59±1.36 in ParAF and 6.20±1.08 in PerAF.

As shown in Table 2 for SE and DET, the Mann-Whitney test rejected the null hypothesis in most of the channels, with particularly low p-values in the RSPV, while in the RIPV the null hypothesis was accepted, revealing a similar atrial remodeling in both AF types. For the AFCL, the p-values resulted always higher than the significance level of 0.05, but in the RSPV. Finally, the statistical tests run for DF showed no differences between ParAF versus PerAF just in the LSPV and in the POS.

The correlation matrix obtained by averaging correlation values of the 1, 2 and 4 s-length datasets provided a strong negative correlation between the indexes pairs of SE-DET and AFCL-DET recorded at the same measurement sites, while a strong positive correlation appeared between the pair SE-AFCL measured at the same channel.

Application of the correlation matrix filter to 1 and 2 s-length datasets removed the indexes DF_{LIPV} and entirely the indexes of DET and AFCL, due to their strong correlation with SE. For the 4 s-length dataset, the subset of indexes kept were SE_{LSPV}, SE_{LIPV}, SE_{RSPV}, SE_{RIPV}, SE_{POS}, $AFLC_{ANT}$, DF_{LSPV}, DF_{RSPV}, DF_{RIPV}, DF_{ANT}, DF_{POS}.

For the three datasets, the variables scoring provided by the Random Forest is presented in Figure 1. For the 1 and 2 s-length datasets, the variables ranked with a score higher than 40 were DF_{LSPV}, DF_{RSPV}, DF_{RIPV}, SE_{LIPV} and SE_{pos}, while in the 4 s-length dataset also the $AFLC_{ANT}$ was included as it surpassed the threshold value.

After testing all the possible sets of highly ranked features, the group SE_{POS}, DF_{LSPV}, DF_{RSPV} and DF_{RIPV} provided the best classification performance with

Table 1. Sample entropy and Determinism mean and standard deviation values for datasets of 1, 2 and 4 s-length segments. Values obtained by averaging among CFAEs.

	SE		DET	
	ParAF	PerAF	ParAF	PerAF
1s	0.123 ± 0.042	0.148 ± 0.057	0.608 ± 0.135	0.524 ± 0.160
2s	0.120 ± 0.043	0.145 ± 0.056	0.656 ± 0.140	0.556 ± 0.177
4s	0.117 ± 0.043	0.143 ± 0.057	0.688 ± 0.135	0.585 ± 0.179

Table 2. Mann-Whitney test p-values obtained, for each measurement site, by averaging the p-values returned in the test performed for the three segment length datasets.

channel	SE	DET	AFCL	DF
LSPV	0.0006	0.0001	> 0.05	> 0.05
LIPV	0.0253	0.0080	> 0.05	0.0287
RSPV	< 0.0001	< 0.0001	0.0192	0.0007
RIPV	> 0.05	> 0.05	> 0.05	0.0035
ANT	0.0035	0.0016	> 0.05	0.0387
POS	0.0002	0.0006	> 0.05	> 0.05

an accuracy of 88.2% for all segment lengths. The same accuracy value was reached adding to the group also SE_{LIPV}. However, in order to simplify as much as possible the model, we decided to consider as optimal the minimal set of features providing the highest accuracy. Accuracy values of the models built with the other possible combinations of the highest-ranked features (score > 40) had mean and standard deviation of 70.7±8.7% for 1 s-length, 70.8±8.3% for 2 s-length and 69.2±8.8% for 4 s-length. Which is a significant reduction in accuracy.

The highest accuracy achieved by a single index was provided by DET_{LIPV} with 82.4% for any segment length as well, while the averaged single accuracies reached by the other indexes were 60.2±11% for 1 s-length, 54.7±13.7% for 2 s-length and 57.3±11.4% for 4 s-length.
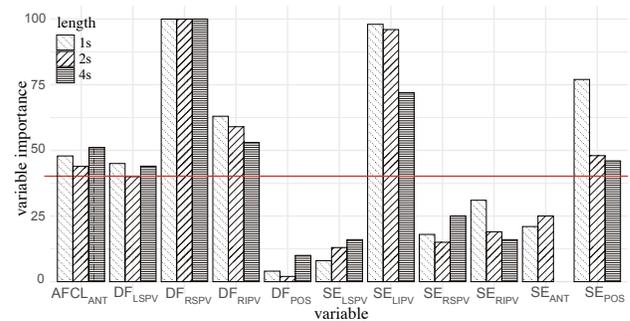

Figure 1.   Variable importance ranked with the Random Forest algorithm for 1, 2 and 4 s-length analyses. The cut-off value of 40 is shown in red.

5. Discussion and conclusion

The study of the discriminatory power of DF_{RIPV} using statistical test led to the wrong conclusion that the index is not discriminative for ParAF and PerAF, as the null hypothesis of no differences between ParAF and PerAF was accepted. Contrarily, the feature selection ranked the DF_{RIPV}, together with SE_{POS}, DF_{LSPV}, DF_{RSPV}, as the set of most important variables to classify the AF type with the patients analyzed, reaching the highest accuracy (88.2%). We can conclude that a careful selection of reduced sets of features feeding simple classification models are able to discriminate accurately between CFAEs of ParAF and PerAF, thus providing simplified insights into atrial substrate evaluation and improved therapeutical decisions on AF management.

Acknowledgements

References

[1] January CT, Wann LS, Calkins H, Chen LY, Cigarroa JE, Cleveland Jr JC, Ellinor PT, Ezekowitz MD, Field ME, Furie KL, Heidenreich PA, Murray KT, Shea JB, Tracy CM, Yancy CW. 2019 AHA/ACC/HRS focused update of the 2014 AHA/ACC/HRS guideline for the management of patients with atrial fibrillation: A report of the american college of cardiology/american heart association task force on clinical practice guidelines and the heart rhythm society in collaboration with the society of thoracic surgeons. Circulation 07 2019;140(2):e125–e151.

[2] Kwon CH. External electrical cardioversion is an easy and safe intervention for rhythm control in persistent atrial fibrillation. Korean Circ J Jun 2020;50(6):524–526.

[3] Capucci A, Cipolletta L, Guerra F, Giannini I. Emerging pharmacotherapies for the treatment of atrial fibrillation. Expert Opin Emerg Drugs 03 2018;23(1):25–36.

[4] Hesselson AB. Catheter ablation in the treatment of atrial fibrillation. Int J Angiol Jun 2020;29(2):108–112.

[5] Turagam MK, Garg J, Whang W, Sartori S, Koruth JS, Miller MA, Langan N, Sofi A, Gomes A, Choudry S, Dukkipati SR, Reddy VY. Catheter ablation of atrial fibrillation in patients with heart failure: A meta-analysis of randomized controlled trials. Ann Intern Med 01 2019; 170(1):41–50.

[6] Latchamsetty R, Morady F. Catheter ablation of atrial fibrillation. Heart Fail Clin Apr 2016;12(2):223–33.

[7] Ciaccio EJ, Biviano AB, Whang W, Vest JA, Gambhir A, Einstein AJ, Garan H. Differences in repeating patterns of complex fractionated left atrial electrograms in longstanding persistent atrial fibrillation as compared with paroxysmal atrial fibrillation. Circ Arrhythm Electrophysiol Aug 2011;4(4):470–7.

[8] Ciaccio EJ, Biviano AB, Whang W, Gambhir A, Garan H. Different characteristics of complex fractionated atrial electrograms in acute paroxysmal versus long-standing persistent atrial fibrillation. Heart Rhythm Sep 2010;7(9):1207–15.

[9] Acharya UR, Faust O, Ciaccio EJ, Koh JEW, Oh SL, Tan RS, Garan H. Application of nonlinear methods to discriminate fractionated electrograms in paroxysmal versus persistent atrial fibrillation. Comput Methods Programs Biomed Jul 2019;175:163–178.

[10] Ndrepepa G, Karch MR, Schneider MAE, Weyerbrock S, Schreieck J, Deisenhofer I, Zrenner B, Schömig A, Schmitt C. Characterization of paroxysmal and persistent atrial fibrillation in the human left atrium during initiation and sustained episodes. J Cardiovasc Electrophysiol Jun 2002; 13(6):525–32.

[11] Ravi KC, Krummen DE, Tran AJ, Bullinga JR, Narayan SM. Electrocardiographic measurements of regional atrial fibrillation cycle length. Pacing Clin Electrophysiol Mar 2009;32 Suppl 1:S66–71.

[12] Sanders P, Berenfeld O, Hocini M, Jaïs P, Vaidyanathan R, Hsu LF, Garrigue S, Takahashi Y, Rotter M, Sacher F, Scavée C, Ploutz-Snyder R, Jalife J, Haïssaguerre M. Spectral analysis identifies sites of high-frequency activity maintaining atrial fibrillation in humans. Circulation Aug 2005; 112(6):789–97.

[13] Martínez-Iniesta M, Ródenas J, Rieta JJ, Alcaraz R. The stationary wavelet transform as an efficient reductor of powerline interference for atrial bipolar electrograms in cardiac electrophysiology. Physiol Meas 07 2019;40(7):075003.

[14] Alcaraz R, Abásolo D, Hornero R, Rieta JJ. Optimal parameters study for sample entropy-based atrial fibrillation organization analysis. Comput Methods Programs Biomed Jul 2010;99(1):124–32.

[15] Schinkel S, Dimigen O, Marwan N. Selection of recurrence threshold for signal detection. The European Physical Journal Special Topics 10 2008;164:15–53.

[16] Osorio D, Alcaraz R, Rieta JJ. A fractionation-based local activation wave detector for atrial electrograms of atrial fibrillation. In 2017 Computing in Cardiology (CinC). IEEE, 2017; 1–4.

[17] Guyon I, Elisseeff A. An introduction of variable and feature selection. J Machine Learning Research Special Issue on Variable and Feature Selection 2003;3.

Address for correspondence:

José J. Rieta
BioMIT.org, Electronic Engineering Department, Building 7F-5
Universitat Politecnica de Valencia
Camino de Vera, s/n, 46022, Valencia, Spain
E-mail: jjrieta@upv.es