Noninvasive Prediction of Catheter Ablation Acute Outcome in Persistent Atrial Fibrillation Based on Logistic Regression of ECG Fibrillatory Wave Amplitude and Spatio-temporal Variability

Marianna Meo*, *Student Member, IEEE*, Vicente Zarzoso, *Senior Member, IEEE*, Olivier Meste, *Member, IEEE*, Decebal G. Latcu, and Nadir Saoudi

Abstract— **Catheter ablation (CA) is increasingly employed to treat persistent atrial fibrillation (AF), yet assessment of procedural AF termination is still a subject of debate in the medical community. This has motivated the development of different criteria based on the standard electrocardiogram (ECG) to characterize ablation immediate effectiveness. However, most of conventional descriptors are merely computed in one ECG lead, thus neglecting significant information provided by the other leads. The present study proposes a novel predictor of CA outcome by exploiting a subset of the 12 leads in the standard ECG. Our method predicts the need for electrical cardioversion subsequent to CA by suitably combining two sets of multilead features, namely, a measure of fibrillatory wave amplitude and an index of AF spatio-temporal variability per lead. These features are obtained on a reduced-rank approximation determined by principal component analysis emphasizing the highest-variance components in the multilead atrial activity signal, and are then combined by logistic regression. On a database of over 50 persistent AF patients, our method provides reliable predictive measures and proves more robust and informative than classical AF descriptors.**

I. INTRODUCTION

Nowadays, atrial fibrillation (AF) is one of the most common cardiac diseases. Catheter ablation (CA) is currently considered as a first-line therapy [1]. Nevertheless, its action on heart substrate has not been completely clarified yet, thus its effectiveness is not always guaranteed. Indeed, due to uncertainty about its outcome, criteria defining the main steps of the procedure as well as its endpoint are quite disparate, leading to different medical protocols. All approaches generally aim at either sinus rhythm restoration or AF transformation into an intermediate arrhythmia, for instance, atrial flutter. In this framework , AF termination is sometimes not exclusively obtained by CA. Different theories about therapy combination have been put forward. In [2] it is stated that most patients undergoing electrical cardioversion (CEE) after CA have recurrences. Conversely, the study

led in [3] claims that acute AF termination by CA has no influence over the long-term outcome, thus justifying the use of complementary curative strategies. The main procedural endpoints may depend on the type of AF and include completion of a predetermined lesion set, depending on technical choices. Some groups aim at AF termination during ablation [4], while other rather address noninducibility of AF following ablation [5]. In any case, there is still debate surrounding the predictive value of such endpoints, and various approaches are currently adopted in medical practice. One of the most adopted strategies to restore sinus rhythm is combining CA with CEE immediately applied after the ablation procedure. In this context, our research puts forward a tool able to individuate candidates to CEE after CA, i.e., those for whom CA did not terminate AF during the procedure. This tool is based on the heart electrical content present in standard electrocardiogram (ECG). More specifically, it takes into account information about fibrillatory wave (f-wave) amplitude and AF spatio-temporal variability (STV) extracted from the rank-1 approximation to the multilead atrial activity (AA) signal and subsequently processed by logistic regression (LR). The proposed method enables a more efficient prediction of acute CA success, as well as a more detailed design of patient's treatment protocol.

II. METHODS

A. ECG Database and Preprocessing

Fifty-one patients underwent ablation, and 3 of them experienced a double procedure. In 37 out of the 54 procedures, AF was not terminated by CA, and CEE was applied immediately after ablation. In the remaining cases, AF was terminated either by CA exclusively or accompanied by pharmacological treatment in the follow-up. Stepwise CA was performed with the aid of Prucka Cardiolab and Biosense CARTO electrophysiology measurement systems at the Cardiology Department, Princess Grace Hospital, Monaco. One-minute 12-lead ECG was acquired at a sampling rate of 1 kHz at the beginning of each CA procedure. Accordingly, 54 ECG recordings have been processed in our analysis. An example on lead V_1 is illustrated in Fig. 1.

As explained in [6], [7], TQ intervals are automatically segmented, mean-centered and concatenated so as to yield the $(L \times N)$ matrix of the multilead AA signal:

$$\mathbf{Y}_{\text{AA}} = [\mathbf{y}_{\text{AA}}(1), \mathbf{y}_{\text{AA}}(2), \ldots, \mathbf{y}_{\text{AA}}(N)] \in \mathbb{R}^{L \times N} \quad (1)$$
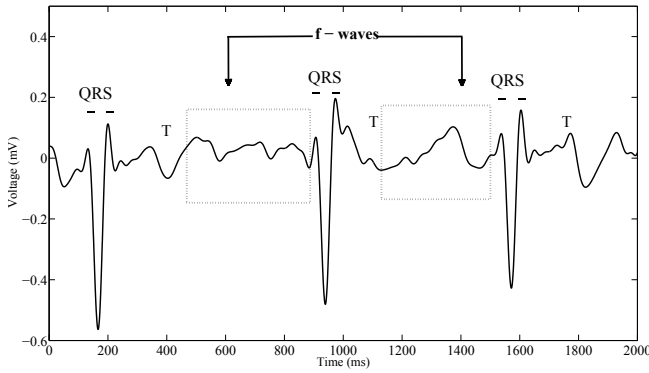
Fig. 1. Example of ECG recording during AF and its fiducial points on lead V_1. Dotted boxes highlight TQ intervals subsequently concatenated to yield the AA signal $\mathbf{Y}_{AA}$ in eqn. (1).

where vector $\mathbf{y}_{AA}(n) = [y_1(n), y_2(n), \dots, y_L(n)]^T$ represents the AA signal activity recorded at a certain temporal sample n for each ECG lead $\ell = 1, \dots, L$. Furthermore, a reduced set of 8 linearly independent ECG leads is analyzed, namely, Einthoven's leads I and II and the precordial ones V_1-V_6.

B. Rank-1 Approximation by Principal Component Analysis

As mentioned in previous works [6], [7], the multilead ECG signal can be decomposed by principal component analysis (PCA) as a linear combination of uncorrelated sources, the so-called principal components (PCs), linked to heart electrical phenomena. Knowing that the first source $\mathbf{x}_1$ retains the highest percentage of the AA signal energy, as in [7], the multilead AA signal can be effectively approximated using this source only:

$$\hat{\mathbf{y}}_{AA}(n) = \mathbf{m}_1 x_1(n). \tag{2}$$

The dominant PC $x_1(n)$ is derived from the observed signals by a spatial filter defined by vector $\mathbf{m}_1$, the maximum-variance direction of the multivariate observation. The rank-1 truncation enhances the most descriptive component in terms of variance, besides compression of input data. The signal approximation in eqn. (2) is further processed to compute some features as explained in the following sections.

C. Multivariate Amplitude and Spatio-temporal Variability

F-wave amplitude is automatically computed on each ECG lead of the aforementioned ensemble of $\hat{\mathbf{Y}}_{AA}$ by the piece-wise cubic interpolation Hermite interpolating polynomial (PCHIP) algorithm proposed in [7], and its mean value d_ℓ is then obtained. F-wave amplitude is thus described by an L-component vector $\mathbf{d} = [d_1, d_2, \dots, d_L]$ for each CA procedure.

As in the study presented in [6], AF variability is assessed on each electrode (thus rendering ECG spatial distribution) and on several temporal segments (so as to quantify the degree of temporal repetitiveness) on the same subset of surface ECG. The same tuning parameter values are set for computing another L-component vector $\boldsymbol{\mu} = [\mu_1, \mu_2, \dots, \mu_L]$, where the ℓth component μ_ℓ stands for the normalized mean square error (NMSE) of the PCA approximation to the

AA signal computed on the ℓth lead and averaged over all segments.

D. Introduction to Logistic Regression

Logistic regression (LR) determines the impact of multiple independent variables presented simultaneously to predict membership to one of the two categories of the dependent or response variable, with mutually exclusive levels θ and $(1-\theta)$. LR forms a best fitting equation using maximum likelihood, which aims at maximizing the probability of assigning observations $\mathbf{z}$ to the correct category given the fitted regression coefficients $\mathbf{b} = [b_1, b_2, \dots, b_L]$. In our application, $\mathbf{z}$ is the vector containing features computed in each ECG lead. Such coefficients weight the contribution of each independent variable (i.e., ECG leads $\ell = 1, \dots, L$) to the response variable (i.e., CA outcome prediction). The LR model is defined as:

$$LR = \log \frac{\theta}{1-\theta} = \mathbf{b}^T \mathbf{z} \tag{3}$$

where $(\cdot)^T$ denotes transposition. The LR score is defined as the ratio of the probability of occurrence of an event θ to that of nonoccurrence with level $(1-\theta)$. In our problem, $(1-\theta)$ represents the probability of acute AF termination. It follows that the higher the score (3), the more likely CA procedural failure, hence the need for a complementary CEE treatment. Unlike linear regression, there is no closed-form expression for the estimates of model coefficients, but an iterative procedure is needed for their computation.

E. LR of Multivariate ECG Descriptors

Application of the LR model to each of the aforementioned multivariate predictors yields the scores subsequently exploited as classifiers to discriminate between the classes under examination. Therefore, the model is first applied to the vector $\mathbf{z}_A = \mathbf{d}$, thus yielding the LR score:

$$LR_A = \mathbf{b}_A^T \mathbf{z}_A. \tag{4}$$

Similarly, regression of AF STV data $\mathbf{z}_{STV} = \boldsymbol{\mu}$ results in:

$$LR_{STV} = \mathbf{b}_{STV}^T \mathbf{z}_{STV}. \tag{5}$$

A further application of regression is performed on the two multivariate predictors concatenated in a unique variable $\mathbf{z}_{(A;ST)}$, giving the score $LR_{(A;ST)}$ as output. In this case, the sample vector is characterized by a $2L$-component descriptor.

$$LR_{(A;STV)} = \mathbf{b}_{(A;STV)}^T \mathbf{z}_{(A;STV)}. \tag{6}$$

F. Statistical Analysis and Prediction Quality Assessment

The categories of interest are referred to as "CEE" and "No CEE", and LR scores related to each class are expressed as mean $\pm$ standard deviation. Interclass differences are assessed by the p-value output by the unpaired Student's t-test under a confidence level $\alpha = 0.05$. Prediction accuracy is quantified by the area under curve (AUC) criterion. The optimal cutoff value guaranteeing the maximum discrimination level between the two classes is reported as well. In addition, the related maximum rates of positive and negative

detections, namely, sensitivity and specificity, are indicated. Leave-one-out cross validation (LOOCV) technique is carried out so as to validate classification results. Indeed, AUC values are computed several times by keeping a subset of 53 procedures out of 54 and thus discarding one procedure at each iteration, computing their average value at the final step. The same computation is repeated for the optimal cutoff associated with the AUC index. All the aforementioned values are reported in Table I; values corresponding to best prediction performance are shown in bold. Finally, the benefits of multilead analysis are demonstrated in Fig. 2. More specifically, AUC index is determined for each value of lead-subset size L ranging from 1 up to 8, once computed the proposed multilead LR predictor on all $8!/((8-L)!L!)$ possible lead combinations for a fixed L value. LOOCV validation is applied as well. The minimum, maximum, and mean AUC values over all L-lead subset combinations have been obtained as a function of the subset dimension L. Such analysis is performed on all scores LR_A, LR_{STV} and $LR_{(A;STV)}$.

III. RESULTS

A. Comparison with Conventional AF Descriptors

The proposed LR-based scores are compared with some classical AF descriptors in relation with the medical therapies examined. All numerical results are reported in Table I. First, mean peak-to-peak amplitude on lead V_1 $D(V_1)$ is taken into account, as several studies demonstrated that it is predictive of CA outcome [8]. In addition, sample entropy $SampEn(L_s, r_s)$, widely regarded as a predictor of AF termination by CEE [9], is computed for all procedures. Tuning parameters, namely, length of the sequences compared L_s and similarity threshold r_s, are set to the values proposed in [10] for sake of comparison. Consequently, L_s is equal to 2, whereas the second parameter is a fraction of σ_{V_1}, the AA signal standard deviation on lead V_1, i.e., $r_1 = 0.1\sigma_{V_1}$ and $r_2 = 0.2\sigma_{V_1}$. A normalized version of the sample entropy, i.e., the squared sample entropy (QSE), recently proposed in [11], is also tested here under the same conditions. The computational time T_C of each algorithm is also examined on the whole database in Table I.

B. LR Technique Analysis

The proposed method required on average no more than 3.4 seconds for processing 2 leads on an Intel® Core™ 2 Quad 2.66 GHz Processor running MATLAB2012a (The MathWorks Inc.) when combining ECG features. Its execution on the whole 8-lead ensemble provides a satisfactory trade-off between computational load and classification accuracy compared with other methods. Influence of weighting coefficients over prediction has been investigated as well. Accordingly, for each procedure, f-wave amplitude computed on $\hat{\mathbf{Y}}_{AA}$ has been averaged over the ECG leads examined, thus giving $\overline{D}_8$ as output [7]. The same procedure has been repeated for STV indices obtained on each ECG lead, whose mean valued is represented by $\overline{\mu}_8$. This is equal to attributing the same importance to all leads, according to

a uniform weighting scheme ($\mathbf{b} = \mathbf{1}$). Furthermore, the regression coefficients' weighting action has been investigated. More precisely, values of LR coefficients $\mathbf{b}_A$ computed on f-wave amplitude data only are compared with those obtained in the first L entries of vector $\mathbf{b}_{(A;STV)}$ when amplitude and AF complexity are processed together. The same analysis is repeated for coefficients $\mathbf{b}_{STV}$ and the last L entries of $\mathbf{b}_{(A;STV)}$ to verify how STV information provided by every ECG lead influences prediction scores in both cases. Results of our examination are plotted in Fig. 3, providing spatial distribution of regression coefficients over ECG leads.
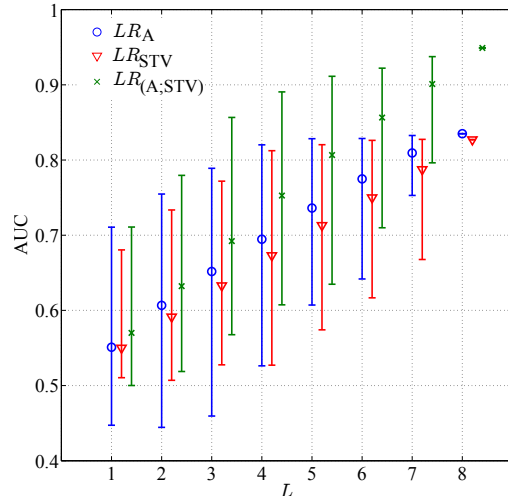

Fig. 2. Evolution of the AUC index of the proposed LR scores LR_A, LR_{STV} and $LR_{(A;STV)}$ as function of the number of ECG leads L.
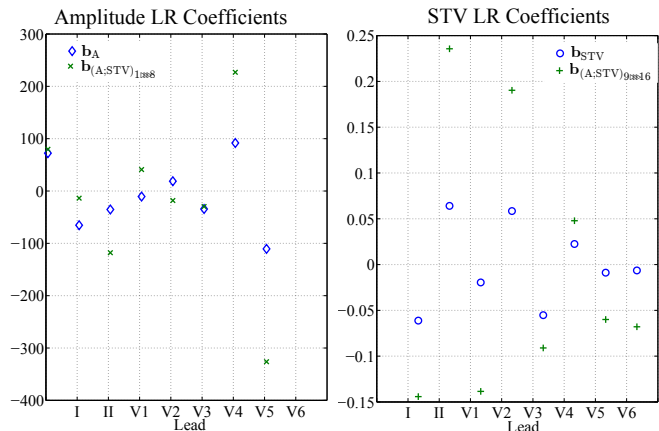

Fig. 3. LR regression coefficients of ECG multivariate features, both in separate and combined analysis. Left: multivariate f-wave amplitude ($\mathbf{b}_A$ and first 8 coefficients of $\mathbf{b}_{(A;STV)}$). Right: multivariate AF STV ($\mathbf{b}_{STV}$ and last 8 coefficients of $\mathbf{b}_{(A;STV)}$).

IV. DISCUSSION

Our results show the superiority of LR-based scores over conventional AF descriptors in acute CA outcome prediction. Indeed, amplitude and STV features seem to be enhanced when contributions computed on each lead are combined into the LR linear combination and properly weighted by LR coefficients. This approach remarkably improves CA outcome prediction and helps discriminating between successful CA procedures from failing ones, thus followed by CEE.

TABLE I

INTERCLASS STATISTICAL DIFFERENCE ASSESSMENT

	CEE	No CEE	p-value	AUC	Sensitivity	Specificity	Best cutoff	T_C
LR_A	0.78 ± 0.20	0.48 ± 0.27	$1.98 \cdot 10^{-5}$	0.83	0.78	0.82	0.71	5.88
LR_{STV}	0.78 ± 0.20	0.48 ± 0.25	$2.32 \cdot 10^{-5}$	0.83	0.81	0.71	0.59	22.17
$LR_{(A;STV)}$	0.88 ± 0.20	0.26 ± 0.26	$\mathbf{9.19 \cdot 10^{-13}}$	**0.95**	**0.95**	**0.88**	0.46	27.36
$D(V_1)$	0.06 ± 0.02	0.07 ± 0.03	0.18	0.61	0.62	0.65	0.06	1.20
$SampEn(L_s, r_1)$	2.78 ± 0.18	2.77 ± 0.26	0.82	0.49	0.35	0.82	2.87	$2.96 \cdot 10^3$
$SampEn(L_s, r_2)$	2.09 ± 0.18	2.08 ± 0.25	0.83	0.49	0.35	0.76	2.18	$3.89 \cdot 10^3$
$QSE(L_s, r_1)$	5.21 ± 0.38	5.19 ± 0.46	0.88	0.51	0.22	0.76	5.00	$2.21 \cdot 10^3$
$QSE(L_s, r_2)$	4.37 ± 0.36	4.34 ± 0.42	0.77	0.52	0.46	0.53	4.36	$2.22 \cdot 10^3$
$\overline{D}_8$	0.04 ± 0.02	0.03 ± 0.02	0.41	0.62	0.51	0.71	0.03	6.63
$\overline{\mu}_8$	51.45 ± 12.18	49.36 ± 12.31	0.56	0.53	0.40	0.71	49.46	22.13

Such a multilead strategy provides a more accurate prediction than classical single-lead parameters, for instance, f-wave mean amplitude $D(V_1)$, which is not able to underline significant interclass differences. Similar remarks can be made for the sample entropy and the QSE index, as its discriminative power is considerably weak. What is more, regardless of tuning parameters' values, results obtained do not match basic assumptions made in previous studies [9], namely, correlation between low entropy values and CEE success due to a higher degree of AF organization.

Concerning mean values of f-wave amplitude and NMSE, ($\overline{D}_8$ and $\overline{\mu}_8$, respectively), note that attributing the same weight to all ECG leads by simply averaging their contributions does not improve prediction, as no significant differences can be observed between the "CEE" and "No CEE" classes. It turns out that CA outcome prediction takes benefit from multivariate data processing and classification, provided that independent variables (i.e., ECG leads) are properly combined, as when LR is applied.

LR coefficients are graphically interpreted in Fig. 3. We can note that combining information does not only selectively enhance contributions provided by certain leads rather than others, but also that LR coefficient dispersion is considerably lower when analyzing each set of features separately, as quantified by standard deviation ($\sigma_{\mathbf{b}_A} = 68.00$ vs. $\sigma_{\mathbf{b}_{(A;STV)_{1...8}}} = 159.51$, and $\sigma_{\mathbf{b}_{STV}} = 0.05$ vs. $\sigma_{\mathbf{b}_{(A;STV)_{9...16}}} = 0.15$). Benefits from feature combination are confirmed by Fig. 2 as well. First, it shows that CA outcome prediction is improved by multilead processing, as the higher the number of leads L examined, the higher the AUC, as previously observed in [7]. Secondly, regardless of the number of electrodes L exploited, examining such ECG properties together probably provides a more complete overview of AF activity, which enhances classification quality. On one hand, f-wave amplitude is predictive of CA outcome and it is strictly correlated with AA signal energy; on the other hand, STV quantifies the degree of regularity and temporal repetitiveness of AF patterns, besides their spatial distribution over ECG leads. These characteristics are quite complementary with each other. Indeed, f-wave amplitude can effectively depict very regular signals; conversely, sharp patterns can hamper signal interpolation and do not render AF temporal evolution. In contrast, NMSE-based parameters can easily capture signal diversity and are more robust to spurious peaks. However, their descriptive power is reduced when dealing with very regular waveforms, where it is harder to extract information about their variability; hence the advantage of merging knowledge about these two aspects, so as to enrich AF characterization and improve prediction accuracy. Prediction benefits from combining features, as also shown by the high AUC values in Table I.

V. CONCLUSIONS

This work has put forward a novel tool for predicting the need for CEE application subsequent to ablation, regarded as a criterion of acute CA outcome in persistent AF. This technique properly combines amplitude and STV measures derived from different ECG leads and effectively predicts AF termination by CA. The interpretation of LR coefficients requires further investigation. In addition, the algorithm should be tested on other features determined on standard ECG (e.g., frequency domain parameters). Despite these limitations, the present multiple-feature framework notably improves CA outcome prediction and helps AF characterization.

REFERENCES

[1] H. Calkins, J. Brugada, D. L. Packer, and et al. HRS/EHRA/ECAS expert consensus statement on catheter and surgical ablation of atrial fibrillation: recommendations for personnel, policy, procedures and follow-up. *Heart Rhythm*, 4(6):816–861, 2007.

[2] M. Haïssaguerre, P. Jaïs, D. C. Shah, and et al. Electrophysiological end point for catheter ablation of atrial fibrillation initiated from multiple pulmonary venous foci. *Circulation*, 101:1409–1417, 2000.

[3] C.S. Elayi, L. Di Biase, C. Barrett, and et al. Atrial fibrillation termination as a procedural endpoint during ablation in long-standing persistent atrial fibrillation. *Heart Rhythm*, 7(9), 2010.

[4] M. Haïssaguerre, P. Sanders, M. Hocini, and et al. Catheter ablation of long-lasting persistent atrial fibrillation: critical structures for termination. *J Cardiovasc Electrophysiol*, 1125–1137, 2005.

[5] P. Jaïs, M. Hocini, P. Sanders, and et al. Long-term evaluation of atrial fibrillation ablation guided by noninducibility. *Heart Rhythm*, 3:140–145, 2006.

[6] M. Meo, V. Zarzoso, O. Meste, D. G. Latcu, and N. Saoudi. Catheter ablation outcome prediction in persistent atrial fibrillation based on spatio-temporal complexity measures of the surface ECG. *in: Computing in Cardiology*, 38:261–265, Sept. 18-21, 2011.

[7] M. Meo, V. Zarzoso, O. Meste, D. G. Latcu, and N. Saoudi. Spatial variability of the 12-lead surface ECG as a tool for noninvasive prediction of catheter ablation outcome in persistent atrial fibrillation. *IEEE Trans. Biomed. Eng.*, 60(1):20–27, 2013.

[8] I. Nault, N. Lellouche, S. Matsuo, and et al. Clinical value of fibrillatory wave amplitude on surface ECG in patients with persistent atrial fibrillation. *J. Interv. Card. Electrophysiol.*, 26(1):11–19, 2009.

[9] R. Alcaraz and J.J. Rieta. A novel application of sample entropy to the electrocardiogram of atrial fibrillation. *Nonlinear Analysis: Real World Applications*, 11:1026–1035, 2010.

[10] D.E. Lake, J.S. Richman, M.P. Griffin, and et al. Sample entropy analysis of neonatal heart rate variability. *Am. J. Physiol. Heart Circ Physiol.*, 283:789–797, 2002.

[11] D. E. Lake and J. R. Moorman. Accurate estimation of entropy in very short physiological time series. *Am J Physiol Heart Circ Physiol*, 300(1):319–325, 2011.