in: Proc. Cardiotechnix-2014, 2nd International Congress on Cardiovascular Technologies, Rome, Italy, Oct. 25-26, 2014.

Predicting Catheter Ablation Outcome in Persistent Atrial Fibrillation via Multivariate Analyses of ECG Fibrillatory Wave Amplitude

A. R. Hidalgo-Muñoz[1], V. Zarzoso[1], M. Meo[2], O. Meste[1], D. G. Latcu[3], I. Popescu[3] and N. Saoudi[3]

[1] *Laboratoire I3S, UMR 7271, Université Nice Sophia Antipolis, CNRS, France*

[2] *Department of Anesthesiology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA*

[3] *Service de Cardiologie, Centre Hospitalier Princesse Grace, Monaco*

{hidalgo, zarzoso, meste}@i3s.unice.fr, mmeo@partners.org, {dg.latcu, ipopescu, nsaoudi}@chpg.mc

1 INTRODUCTION

Atrial fibrillation (AF) is the most common sustained arrhythmia encountered in clinical practice. Among the possible strategies to face AF, catheter ablation (CA) is a well established therapy with proven efficacy to maintain sinus rhythm during follow-up. The highest rate of AF recurrence is reported to occur during the first 6 months following ablation (Calkins et al., 2012). Hence, an accurate selection of patients who can really benefit from this intervention at long-term is essential.

In this work, two multivariate approaches for selecting the most relevant features linked to CA outcome are proposed. The first one is known as *logistic regression (LR)* (Indrayan, 2012). The second one is based on a machine learning technique, named *support vector machine (SVM)* (Burges, 1998). The goal of both analyses is to distinguish the AF patients more prone to a successful CA outcome from those who are more sensitive to arrhythmia recurrence. More specifically, patients who suffer from persistent AF are considered in this study.

There are many possibilities to tackle the analysis of the fibrillatory waves (f-waves) such as time or frequency domains (Matsuo et al., 2009; Jones et al., 2013). One of the most direct and easily interpretable measurements obtained from the f-waves is the fibrillatory wave amplitude (Nault et al., 2009; Meo et al., 2013), which is the focus of the present work.

2 METHODS

2.1 Database and ECG Acquisition

Sixty-two patients (52 male, age = 61.5±10.4 years) having persistent and long-lasting persistent AF (Calkins et al., 2012) were recruited at the Cardiology Department, Princess Grace Hospital, Monaco, after giving their informed consent. One minute standard 12-lead ECG was acquired at a sampling rate of 977 Hz at the beginning of the CA procedure. ECG signals were filtered by a 4th-order zero-phase bandpass Chebyshev filter with cut-off frequencies of 0.5 Hz and 30 Hz. For every lead of each recorded ECG, the fibrillatory wave amplitude was computed as in (Meo et al., 2013). Also as in that reference, this work focused exclusively on the 8 linearly independent leads I, II, V1-V6.

All patients underwent stepwise CA (Calkins et al., 2012), including lasso-guided circumferential pulmonary vein (PV) disconnection, fractionated potentials, non-PV triggers, roof line and mitral isthmus line right atrial ablation. Outcome was followed up during at least 6 months after the CA procedure. Finally, 47 patients were free from AF (*success*) while 15 patients had documented AF recurrences (*failure*).

2.2 Statistical Analysis

Normal distribution was first checked by the Kolmogorov-Smirnov test. A U-Mann Whitney test was performed to verify any statistically significant differences between the groups of interest. Both the statistical analysis and the subsequent logistic regression model were performed by using the software SPSS version 13.0.

LR consists of a probabilistic statistical model aimed at predicting the outcome of a categorical dependent variable (success and failure after CA) based on one or more predictor variables (features). The next equation is an example for M predictors:

$$LR = log\left(\frac{\theta}{1-\theta}\right) = b_0 + b_1 x_1 + ... + b_M x_M \quad (1)$$

where θ is the probability of belonging to the success CA class, $\{b_0, b_1, ..., b_M\}$ are the coefficients of the regression model and $\{x_1, ..., x_M\}$ correspond to the numerical values of the different features. In our case, x_m corresponds to the fibrillatory amplitude mean value for the mth lead selected by the model from the complete set of 8 leads. A backward elimination method that discards the least significant variable at each iteration according to the Wald index was employed (Indrayan, 2012) to keep only the most relevant ECG leads.

On the other hand, SVM separates the given set of binary labeled training data with a hyperplane that is maximally distant (Euclidean distance) from the two classes success and failure CA outcome). Linear discriminant functions define a decision hyperplane in a multidimensional feature space as:

$$g(\mathbf{x}) = \mathbf{w}^T\mathbf{x} + w_0 = 0 \tag{2}$$

where $\mathbf{w}$ is known as the weight vector and $\mathbf{x}$ is the feature vector (set of lead amplitude values). Symbol w_0 denotes the threshold and $(\cdot)^T$ the transpose operator. Note that elements in vector $\mathbf{w}$ close to zero have little influence on $g(\mathbf{x})$ and hardly contribute to the decision. In (Guyon et al., 2002) it was suggested a recursive feature elimination technique (SVM-RFE) that discards iteratively the features corresponding to the lowest values of $|w_i|$. After feature elimination, the classifier is evaluated using a leave-one-out cross-validation strategy. In our case, the initial feature vector $\mathbf{x}$ is made up of the mean fibrillatory amplitude values computed in each of the 8 leads considered in this study. The SVM-RFE algorithm was implemented using MATLAB version R2011a.

3 RESULTS

No significant inter-class differences were found for any of the single-lead amplitude parameters considered separately.

After applying LR by using the backward elimination method, the remaining leads were I, V1, V2 and V5. Table 1 summarizes the best results in terms of AUC by applying univariate and multivariate approaches. The best cut-off point was $\theta = 0.7$ yielding 83% sensitivity, 73.3% specificity and 80.6% total accuracy.

The SVM-RFE algorithm ranked the leads as follows: I, V5, V2, V1, II, V4, V6 and V3 (ordered from the most to the least relevant). The best result in terms of accuracy is obtained by using only the three most relevant leads (I, V5 and V2) reaching up to 82.3% accuracy, 70% specificity and 84.6% sensitivity.

4 CONCLUSION

This work has shown that analyzing the fibrillatory amplitudes measured in multiple ECG leads improves the accuracy of previous CA outcome predictors based on a single lead (Nault et al., 2009). The multivariate amplitude analysis model reaches an AUC of up to 0.854 with four leads. Both LR and SVM-RFE agree in pointing to I, V1, V2 and V5 as the most discriminant leads to determine mid- and long-term CA outcome. Along the lines of (Meo et al., 2013), these results support the suitability of taking into consideration the information of several leads that are often neglected in CA outcome prediction and AF analysis (Nault et al., 2009; Matsuo et al., 2009).

Table 1: Best accuracy results by the univariate contrast with the highest AUC and the multivariate analysis after LR backward elimination. AUC: area under the curve; IC: AUC's interval of confidence; p: AUC's statistical significance.

Analysis	AUC	IC	p
Univariate (lead I)	0.678	0.50 to 0.85	0.039
Multivariate (θ)	0.854	0.75 to 0.96	<0.001

Acknowledgements. Work supported in part by the French National Research Agency under contract ANR-2010-JCJC-0303-01 "PERSIST" and the Centre Scientifique de Monaco. A. R. Hidalgo-Muñoz holds a postdoctoral research fellowship awarded by Univ. Nice Sophia Antipolis.

REFERENCES

Burges, C. (1998). A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2(2):121–167.

Calkins, H., Kuck, K. H., Cappato, R., Brugada, J., Camm, A. J., et al. (2012). 2012 HRS/EHRA/ECAS expert consensus statement on catheter and surgical ablation of atrial fibrillation: recommendations for patient selection, procedural techniques, patient management and follow-up, definitions, endpoints, and research trial design. *Europace*, 14:528–606.

Guyon, I., Weston, J., Barnhill, S., and Vapnik, V. (2002). Gene selection for cancer classification using support vector machines. *Machine Learning*, 46:389–422.

Indrayan, A. (2012). *Medical Biostatistics*. CRC Press.

Jones, A. R., Krummen, D. E., and Narayan, S. M. (2013). Non-invasive identification of stable rotors and focal sources for human atrial fibrillation: mechanistic classification of atrial fibrillation from the electrocardiogram. *Europace*, 15(9):1249–1258.

Matsuo, S., Lellouche, N., Wright, M., Bevilacqua, M., Knecht, S., et al. (2009). Clinical predictors of termination and clinical outcome of catheter ablation for persistent atrial fibrillation. *Journal of the American College of Cardiology*, 54(9):788–795.

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

Nault, I., Lellouche, N., Matsuo, S., Knecht, S., Wright, M., et al. (2009). Clinical value of fibrillatory wave amplitude on surface ECG in patients with persistent atrial fibrillation. *Journal of Interventional Cardiac Electrophysiology*, 26(1):11–19.