Using High-Resolution Voltage Maps to Predict "redo" in the Treatment of Atrial Fibrillation (AF)

Jean Bragard[1], Leire Moriones[1], Blas Echebarria[2], Susana Ravassa[3], Javier Ibero[4] & Ignacio García-Bolao[4]

[1] School of Sciences, Universidad de Navarra, Pamplona, Spain
[2] Universitat Politècnica de Catalunya, Barcelona, Spain
[3] CIMA, Universidad de Navarra, Pamplona, Spain
[4] CUN, Universidad de Navarra, Pamplona, Spain.

Abstract

The aim of this work is to use biomarkers extracted from high-resolution voltage maps of atrial fibrillation (AF) patients in order to make predictions about future "redo" procedures. We collected maps of the left atrium of 122 patients, prior of being treated for AF The bipolar voltage maps were extracted with the Rythmia system from Boston Scientific and subsequently analyzed in the MATLAB environment. The present study focuses on three biomarkers extracted from those maps. Two are associated with the bipolar voltage measurements on the map, i.e., the mean voltage and the voltage dispersion on the map. The third indicator is the area of the atrium evaluated from the map. The data are used for feeding a supervised classification algorithm. The output variable is a binary variable that is set to 1 if the patient will need a "redo" procedure in the twelve months following the cardiac intervention and 0 otherwise. We show that the biomarkers have some statistical power in predicting future outcomes. Especially the mean voltage on the map is the best predictor of the future outcome. We determine the cutoff value for the mean voltage based on the best prediction accuracy of Vm=0.542 mV in agreement with previous studies. We discuss some extensions of this study that could allow improvements in predictive power.

1. Introduction

Cardiovascular diseases have been identified as a research priority in the European Union (EU). The time has come to efficiently combine the Information Technology resources in the applied medical field. This synergy will optimize medical treatment and reach the long sought-after "personalized medicine ." Although diagnostic approaches and therapies have drastically improved over the last decade, effective, validated, and auditable tools for the integrated assessment of cardiac function in clinical practice have yet to be developed. Atrial fibrillation (AF) is one of the most prevalent cardiac pathologies, considered an epidemic by the World Health Organization, affecting 7.6 million people over 65 years of age in Europe [1]. AF incidence will progressively increase in the coming years because it is closely associated with aging (projected to be 14.4 million by 2060). AF is associated with a high risk of morbidity and mortality and a high social and healthcare cost, mainly because 10 to 40% of patients with AF require hospitalization. AF is a progressive disease: most patients begin with paroxysmal AF and progress to persistent AF over time, although some of them have lifelong AF crises without persistent episodes, while in others, persistent AF is the first arrhythmic event. Some studies suggest that the progression from paroxysmal to persistent AF is associated with an increase in the morbidity and mortality of AF. However, to date, it is unknown how to predict this evolution. To assess this point, we have access to a cohort of 122 patients. Each of them has been treated for AF. Before the pulmonary vein PV ablation therapy, a very high-definition voltage map of the left atrium was acquired (as shown in Fig.1). In this paper, we use the tools of machine learning ML to build a classifier to predict the patient's state in the year following the cardiac intervention. Statistical learning has emerged recently as a subfield of Statistics, focused on supervised and unsupervised modeling and predictions. In recent years, progress in statistical learning has been accompanied by increased availability and user-friendly tools. In this paper, we will extensively use MATLAB's Statistics and ML Toolbox.

2. Methods

We included 122 consecutive patients with paroxysmal or persistent AF for pulmonary vein isolation (PVI) using an ultra-high definition voltage mapping (uHDM) sys-

tem in the University Hospital (CUN). The study was conducted following the ethical principles of the Declaration of Helsinki. All patients gave their informed consent.

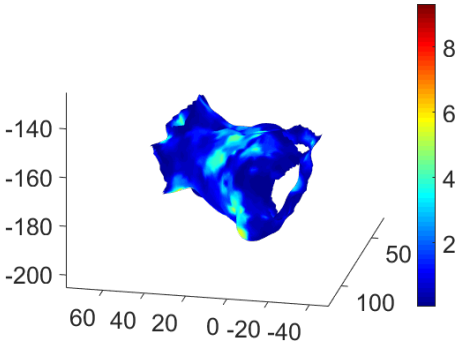

Figure 1. Bipolar voltage map, color scale indicates voltage in mV units. Axes indicate relative positions of the left atrium in mm units. The PVs have been removed from the map (see text for details).

2.1. Data acquisition

Mapping of the LA was conducted with a uHDM system (Rhythmia; Boston Scientific Corporation, Marlborough, MA) and a 64−electrode basket−type catheter (IntellaMap Orion, Boston Scientific Corporation) during paced atrial rhythm. Bipolar electrogram recordings were filtered at 40 to 400Hz and were saved in a file format suitable for further analysis in the MATLAB environment. Only points located within 2 mm of the external surface of the map were considered for analysis. After mapping, PVI was performed in all cases following a previously described protocol. The follow-up of each patient includes a status binary variable "redo" that corresponds to a recurrence of the AF (1 for TRUE and 0 for FALSE) inside the year following the surgical procedure.

2.2. Data analysis

All the 122 uHD maps were analyzed to obtain several indicators for the subsequent statistical analysis. All the data analysis was done using the commercial mathematical software *MATLAB* [MATLAB and Statistics Toolbox Release R2012a, The MathWorks Inc., Natick, Masachusetts, United States] in several stages:

Preprocessing of the signal

Bipolar electrogram recordings were filtered at 40 to 400Hz and were saved in MATLAB file format. The values of the bipolar potentials are expressed in millivolts (mV) at each vertex site. The PV of the acquired maps were extracted semi-manually, i.e., a first extraction was automatically done by the Rhythmia software, and a second, supervised extraction (correction) was done manually for each map. The resulting map is shown in Fig. 1. The number of points defining the map is in the range 12,567 $\pm$ 5,486 points per map.

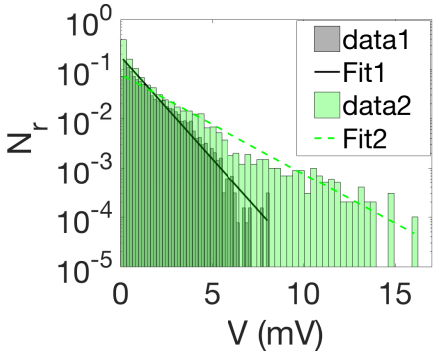

Figure 2. Spatial bipolar voltage distributions for 2 selected individuals. The fits indicate the best linear fits for these log-lin scale histograms.

Features extractions

For each map, the following parameters were evaluated. (a) The spatial average value of all bipolar potentials on the LA surface (Vm). (b) The slope of the voltage histogram (Vs), where in the histogram, we represent the relative frequency in a logarithmic scale on the vertical axis and the bipolar potential on a linear scale on the horizontal axis. Since the spatial voltage distribution typically follows an exponential distribution, we choose the slope of the scatterplot in these log-linear scales as the simplest way to characterize it. In each case, we verified that the corresponding adjusted R^2 was appropriate and that the 95% confidence interval for the slope estimation was sufficiently narrow. We used the MATLAB command "hist(y,M)" for the binning method, with M number of bins set to 50 for all the analyzed data sets. From this scatterplot, we used linear interpolation to calculate the slope. By measuring the histogram slope, we obtained a characteristic voltage decay for each patient (expressed in mV^{-1}) (see details in Fig. 2). (c) a third "geometrical" indicator was extracted from each map. It corresponds to the area of the left atrium (in mm^2) after the removal of the PV.

In the two illustrative examples shown in Fig. 2, we computed that for $data_1$ VS= -0.9511 (units mV^{-1}), with a standard error SE=0.0332 and R^2=0.95, we also evaluated that Vm=1.2137 mV and A=95.4 mm^2. Likewise, for $data_2$, we computed that VS= -0.4625 (units mV^{-1}), with a standard error SE=0.0208 and R^2=0.92, we also evaluated that Vm=1.1941 mV and A=97.2 mm^2 [2].

All the data were collected in a data table of 122 rows and 4 columns. The first three columns correspond to the three indicators defined above and the last column corre-

sponds to the binary "redo" variable.

Correlations between the 3 predictors: Vm, VS, Area

We evaluate the possible relation between the three covariates variables Vm, VS, and Area. We do that by computing the correlation matrix and the dispersion plots of the indicators taken two-by-two. The MATLAB commands are *corrcoef* and *corrplot*.

Logistic regression between "redo" and the 3 predictors

If the outcome variable is binary, a standard model is a logistic model when the probability of the outcome variable is modeled with the covariates X following the logistic function:

$$p(X) = \frac{e^{\beta_0+\beta_i X_i}}{1+e^{\beta_0+\beta_i X_i}}\,. \qquad (1)$$

If one uses more than one covariate, the model is called multiple logistic regression. The logistic regression is computed with the MATLAB command *fitglm*. The goodness of fit is evaluated with the deviance tests and the standard errors (SE) and p-values of the best coefficient estimates of the logistic equation Eq. (1).

Comparison of different ML classifiers

The advent of Machine Learning (ML) algorithms has brought forward many of the classifier tools as an additional statistical method [3]. There exist many classification techniques, or classifiers, that one can use to predict a qualitative response We can cite without being exhaustive: logistic regression, linear discriminant analysis, quadratic discriminant analysis, naive Bayes, random forest, and K-nearest neighbors. MATLAB conveniently integrates a tool called *classificationLearner* that allows one to probe several different classifiers and rank them according to their performances. The ranking between classifiers is usually done by comparing the overall accuracy of the classifiers, the sensitivity (also called true positive rate TPR), and the specificity (1-FPR), where FPR is the false positive rate, and the area under the curve (AUC) for the receiver operating characteristic (ROC) curve.

3. Results and discussions

We compute and discuss the several aspects that were defined in the Method section.

Evidence of correlations between the 3 predictors

We note that the two electrical biomarkers, Vm and VS, show a significant level of association. The Pearson coefficient of correlation is given by r=0.534, with a corresponding p-value$< 10^{-9}$. The geometrical indicator (Area) has no significant linear correlation with the two

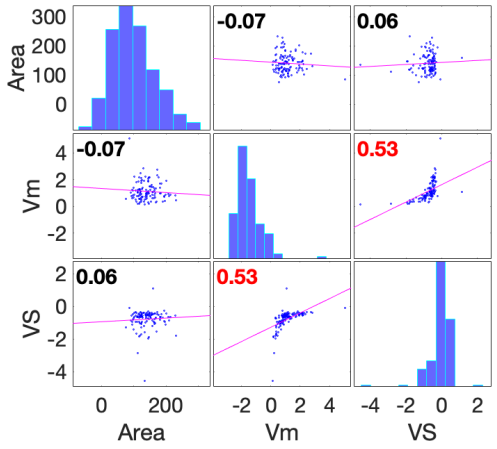

Figure 3. Scatterplots between the 3 predictor variables. A positive correlation shows up between Vm and VS.

electrical biomarkers. Note that the scatterplot of Vm versus VS shows clearly that the existing association between the two variables is not linear.

Results for the logistic regression fits

Table 1. Values for the logit coefficients, their SE and p-values, the deviance and the deviance test p-value.

β_i	SE	p-value	Deviance	Dt. p-val.
only Vm			119.6	$< 10^{-6}$
$\beta_0 = 1.201$	0.524	0.022		
$\beta_1 = -2.287$	0.578	$< 10^{-4}$		
only VS			127.1	$< 10^{-4}$
$\beta_0 = -2.470$	0.493	$< 10^{-6}$		
$\beta_1 = -1.865$	0.540	$5.6\,10^{-4}$		
only Area			142.8	0.211
$\beta_0 = -2.080$	0.935	0.026		
$\beta_1 = 0.0079$	0.0063	0.211		
All three			116.5	$3.8\,10^{-6}$
$\beta_0 = -1.212$	1.501	0.419		
$\beta_{1\mathrm{Vm}} = -1.767$	0.700	0.012		
$\beta_{1\mathrm{VS}} = -0.667$	0.603	0.269		
$\beta_{1\mathrm{Ar.}} = 0.0098$	0.0069	0.162		

We have performed three univariate logistic regressions for the three different predictors and a multivariate logistic fit where we have taken the three variables together. The results for the coefficients of the fit and the goodness of the fits are summarized in Table 1.

We used the standard deviance test to rank the different logistic regressions. Unsurprisingly, we observe that the best predictor for the outcome is the mean voltage Vm. When the three predictors are considered, we see that the

VS is no longer significant at the $\alpha = 0.05$ level and could be removed from the predictors. The strong correlation between Vm and VS makes VS redundant.

As an application of the logistic regression, we have computed the confusion matrices associated with a variable cutoff probability π in the case of the fit with the only Vm predictor corresponding to the first case in Table 1. We obtain that the optimum cutoff value for maximizing the overall accuracy is $\pi = 0.49$ with a corresponding value of Vm= 0.542 mV. For this value of the optimum cutoff, the overall accuracy is 82 %, with a specificity of 95 % and a sensitivity of 47 %. The confusion table is given below:

Table 2. Confusion matrix for the "redo" predicted values for the logistic regression associated with Vm and a selected Vm cutoff equals to 0.542 mV.

		Predicted logit	
		Yes	No
Actual	Yes	16	18
	No	4	84

Results for different classifiers

Finally, we were interested in testing the performance of other ML classifiers. MATLAB command *classificationLearner* allows testing on a dataset with several classifiers at once. We have used our data to test the classifiers and indicated the five best classifiers according to the overall accuracy in Table 3. The different classifiers are trained through the standard 5-fold cross-validation algorithm.

Table 3. Results from the five best ML classifiers ranked by their overall validation accuracy.

Method	Accuracy	AUC	Specific.	Sensitiv.
Logit	0.77	0.75	0.91	0.41
Coarse tree	0.762	0.63	0.90	0.41
Bagged trees	0.762	0.75	0.91	0.38
Naive Bayes	0.754	0.76	0.85	0.50
Lin. SVM	0.746	0.72	0.94	0.24

The ROC curves are shown in Fig. 4, where the classifiers have very similar specificity and sensitivity, as indicated by the corresponding colored symbols.

Looking at Table 3, we see that while the overall accuracy is not great, the specificity is relatively high for all the classifiers. Type 1 error (false positive) is maintained at small values. This agrees with what is usually recommended in the medical literature when the prevailing assumption is the status quo or default assumption.

The present study indicates that we obtain some predictive power (somewhat limited) by analyzing the patients' high-definition voltage maps of the atrium. Further studies are needed to estimate the limit of prediction that we can reach with those maps.

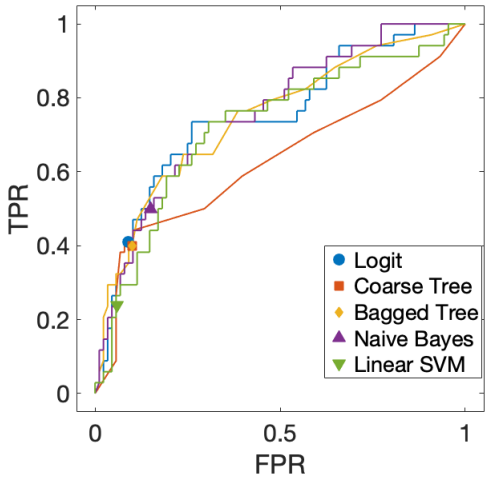

Figure 4. Roc curves for several ML classifiers. The symbols indicate the selected classifiers for each method (see Table 3 for corresponding specificity and sensitivity).

Acknowledgements

This work was supported by the reference project PID2020-116927RB-C22 of the Ministry of Economy and Competitivity (Spain). L.M. acknowledges the financial support of the "Asociación de Amigos de la Universidad de Navarra".

References

[1] SS Chugh, et al. "Worldwide epidemiology of atrial fibrillation: a Global Burden of Disease 2010 Study", Circulation 2014; 129(8): 837–847.

[2] G. Ballesteros, et al. "Association of left atrium voltage amplitude and distribution with the risk of atrial fibrillation recurrence and evolution after pulmonary vein isolation: An ultrahigh-density mapping study", J. Cardiov. Electro. 2019; 30(8): 1231–40.

[3] G. James et al. "An introduction to Statistical Learning" (2nd Edition), Springer (New York) 2021.

Address for correspondence:

Jean Bragard
c/ Irunlarrea, 1
Universidad de Navarra,
31008 Pamplona
Email: jbragard@unav.es