Combined Analysis of Time and Frequency Series Regularity Applied to the Study of Atrial Fibrillation

C Vayá[1], JJ Rieta[2]

[1]Innovation in Bioengineering, Universidad de Castilla-La Mancha, Spain
[2]Biomedical Synergy, Universidad Politécnica de Valencia, Spain

Abstract

Atrial Fibrillation (AF) is the most common arrhythmia encountered in the clinical practice. In this work, a new method based on electrocardiogram (ECG) signal processing is carried out in order to distinguish between AF episodes that will terminate immediately and those that will sustain. This new method is based on a combined analysis of the atrial activity (AA) series regularity in both time and frequency domains. The regularity is measured by using the non-linear regularity estimator Sample Entropy ($SampEn$). The $SampEn$ of these series of spectral parameters and the $SampEn$ of the AA in the time-domain are studied jointly in a discriminant analysis. In global percentages, forty eight of fifty recordings (96%) were classified correctly by the discriminant analysis, which provided an improvement of 6% with respect to the univariate analysis of $SampEn$.

1. Introduction

Atrial Fibrillation (AF) is the most common arrhythmia encountered at advanced age. The prevalence of AF remains lower than 1% among the general population, but it increases considerably from sixty years old [1]. The spontaneously terminating AF, frequently refereed as paroxysmal AF (PAF), is often the precursor to sustained AF [2]. Given that sustained AF increases the likelihood of suffering myocardial infarctions and strokes [1], its accurate recognition by means of noninvasive techniques is of great interest to the regular clinical practice. The prediction of PAF mantainance can help to choose the appropriate intervention that may terminate the arrythmia. Otherwise, the prediction of the spontaneous termination of PAF episodes could avoid unnecessay therapies and their associated clinical costs [3].

In this work, an analysis the of AA spectral parameters organization is carried out with the aim to classify between terminating and non-terminating AF episodes. The analysis of these parameters is made in terms of mathematical regularity of their series. The organization is measured by using the entropy estimator Sample Entropy ($SampEn$) [4, 5]. Entropy estimators have already been used in the characterization of biomedical signals different from ECG [6]. The ECG recordings analysis is completed in five main steps: extraction of the AA, computation of the spectrogram, curve fitting, construction of spectral parameter series, and $SampEn$ computation. The $SampEn$ values are then studied by using univariate and discriminant analyses.

2. Materials

For the present work we have used the database of PhysioNet/Computers in Cardiology Challenge 2004. The database is divided into a learning set and a test set. Each record in the database is a one minute segment of AF that has been extracted from a long term ECG recording. The learning set contains 10 segments of non-terminating AF (group N) and 10 segments of AF that terminates within a second after the end of the record (group T). The test set contains 30 records, approximately half of them belonging to group N, and the rest to group T.

Butterworth filtering of eight order and pass-band from 1 to 50 Hz is applied to each recording. The original sampling rate (f_s) of the Holter system was 128 samples per second, but the ECG recordings were interpolated by a factor of 8 so that a f_s equal to 1024 resulted. The resultant time-domain higher resolution allowed us to obtain a better cancellation of the QRS complex and a higher length of parameter sequences.

3. Methods

In order to use the ECG as a suitable tool for the analysis of AF, we need to separate the atrial activity from other cardioelectric signals. The extraction of the AA during AF requires nonlinear signal processing since spectra of atrial and ventricular activities (VA) overlap and, in consequence, they cannot be separated by simple linear filtering [7]. Given that the challenge database consisted of

two-lead ECG recordings, we opted to use ABS to separate the AA from the rest of the cardioelectric signal as a previous step to the rest of the analysis [8, 7].

After the obtention of the AA signal, its spectrogram [9] is computed using Hamming windows of 1024 samples in length and 75% overlap. In order to facilitate the spectral parameters extraction, cubic spline fitting is applied to each of the Fourier transforms that constitute the spectrogram. The cubic spline fitting curve from the original data is interpolated so that the resulting frequency increment is 0.01 Hz. Finally, the $SampEn$ of all aforementioned series is obtained as an estimation of their mathematical regularity.

The size of series is around 600 elements, what is large enough since the $SampEn$ is meaningfully applied to more than 100 data points [4]. The $SampEn$ values are evaluated by the *t*-test and by discriminant analysis, which results are exposed in section 4. The objective pursued by the discriminant analysis is to know if there exists any combination of the aforementioned parameters which improves the decision reliability. The discriminant analysis is carried out in two stages. First, the discriminant function is obtained by considering the learning data set. Then this discriminant line is applied to the test set in order to validate the results.

The first spectral parameter we consider is the main peak frequency (f_{p1}), which is known to be highly relevant in the characterization of AF [10]. The second parameter is the main peak magnitude, A_1. The third and fourth parameters are the second largest frequency peak (f_{p2}) and its related peak magnitude, A_2. The fifth parameter is the Spectral Concentration (SC). The SC can be defined as [11]:

$$SC = \frac{\sum_{f=0.82f_{p1}}^{1.17f_{p1}} P_{AA}(f)}{\sum_{f=0}^{0.5f_s} P_{AA}(f)} \quad (1)$$

where P_{AA} is the power spectral density of the AA signal, f is the frequencies vector, f_s is the sampling rate (1024 Hz), and f_{p1} is the main peak frequency of the AA. Other two parameters related to the width of the spectrum main lobe have been used: the 3-dB width of the peak, $w3dB$, and the power in the 3-dB band, $pb3dB$. This two last parameters have been used in [12] to characterize AF. Two derived parameters, Δf_p and $\bar{A}_2$ are referred to the spectral shape of AA. Similar parameters are used in [13]. The first derived parameter is the normalized distance between f_{p1} and f_{p2}, which is expressed as:

$$\Delta f_p = \frac{(f_{p1} - f_{p2})}{f_{p1}} \quad (2)$$

The second derived parameters is the normalized amplitude of the second largest peak, which is defined as:

$$\bar{A}_2 = \frac{A_2}{A_1} \quad (3)$$

The deviation of the main and second peak magnitudes from their respective mean values are also computed as a dispersion measurement:

$$d_1 = f_{p1} - \boldsymbol{E}(f_{p1}) \quad (4)$$

$$d_2 = f_{p2} - \boldsymbol{E}(f_{p2}) \quad (5)$$

where $\boldsymbol{E}(\cdot)$ represents the average value over the set of periodogram. Finally, the Median Frequency (MF) is obtained as the center of mass of the spectrum:

$$MF = \frac{\sum_{f=0}^{0.5f_s} |FT_{AA}(f)| \cdot f}{\sum_{f=0}^{0.5f_s} f} \quad (6)$$

where FT_{AA} is the Fourier Transform of the AA. This parameter was previously used in other works to characterize the ventricular fibrillation [14].

4. Results

The results of the *t*-test applied to the $SampEn$ of the numerical series for the learning set are summarized in Figure 1. These results reveal that it is possible to distinguish between terminating and non-terminating AF in six of the twelve parameters, considering a parameter to be relevant when its bilateral significance is less than 0.05. These six relevant parameters are f_{p1}, f_{p2}, Δf_p, A_1, d_1 and SC, which bilateral significances are, respectively, 0.001, 0.005, 0.003, 0.004, 0.015, and 0.001. The mean $SampEn$ in type N recordings is higher than in type T recordings for all these relevant parameters. An optimal decision threshold of 0.1173 has been chosen for f_{p1}. By considering this value of threshold, 19 out of 20 learning recordings have been classified correctly. Taking the same threshold for the test set, 26 out of 30 recordings have been classified correctly. This resulted in a percentage of recordings properly classified equal to 95% for the learning set and equal to 86.67% for the test set. The results obtained by this classification are presented in Figure 2 for every recording.

The previous univariate analysis revealed that the $SampEn$ of the spectral parameters f_{p1}, f_{p2}, Δf_p, A_1, d_1 and SC have a bilateral significance lower than 0.05 and, in consequence, all of them are suitable to be used

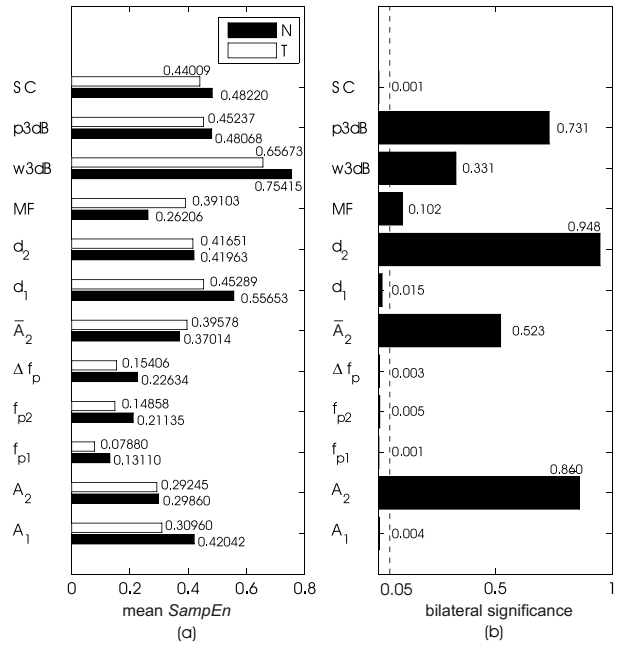

Figure 1. Results of the *t*-test for the $SampEn$ of all the spectral parameters. a) Mean $SampEn$ for groups N and T, b) $SampEn$ bilateral significance between groups.

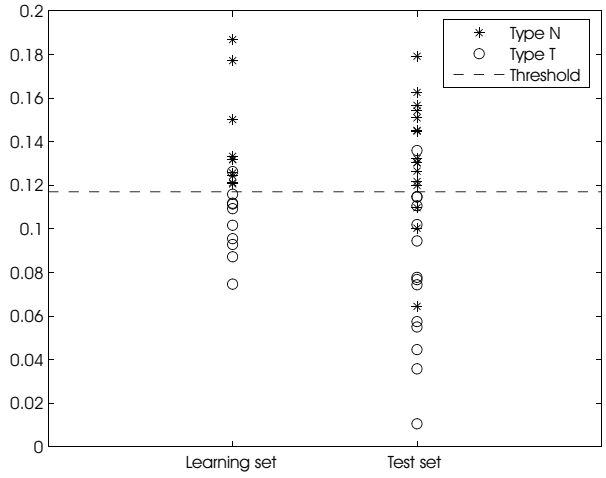

Figure 2. Classification of type N and T episodes using a threshold value for the $SampEn$ of f_{p1} equal to 0.1173.

	Canonical coefficient
$SampEn$ of f_{p1}	1,880
$SampEn$ of AA	1,662
$SampEn$ of Δf_p	0,816

Table 1. Standardized canonical discriminant function coefficients from the stepwise analysis.

in discriminant analysis, which is the next step that has been taken. Furthermore, the $SampEn$ of the AA was also computed and a mean difference of 0.25081, with the greatest mean value for the N group, and a bilateral significance equal to 0.004 were figured out by the *t*-test. This fact suggested to include this parameter in the discriminant analysis along with the spectral parameters. This made possible, in addition, to combine the information obtained from both time and frequency domains.

The obtained discriminant function is a plane given by the equation $x_3 = 0.0355 \cdot x_1 + 1.6 \cdot x_2 + 0.4653$, where x_1, x_2 and x_3 represent the $SampEn$ of f_{p1}, Δf_p and the AA, respectively. The standardized canonical coefficients of the discriminant function are presented in Table 1. These coefficients are ordered in the table by their importance in the discriminant function. A small value of Wilk's lambda test [15] significance ($p < 0.001$) was obtained, which indicates the great discriminatory ability of the function. All of the cases used to create the model, i.e. the learning set, were correctly classified. Regarding the test set, 15 out of 16 type N cases and 13 out of 14 type T cases were correctly classified (see Table 2). Expressing this results in percentages, 100 % cases of the learning set and 93.75 % cases of the test set were classified properly.

5. Conclusions

The discriminant analysis has provided an improvement of the results with respect to the classification by threshold (5% for the learning set and 7.08% for the test set). For that reason, it is worth considering the discriminant analysis in predicting the evolution of AF because this improvement in the classification of AF could be of great importance in the regular clinical practice.

To sum up, a new method based on the mathematical regularity of spectral parameters has been introduced as an original and improved way to predict the evolution of paroxysmal AF episodes. From the results we can deduce that the the future evolution of AF affects not only to the values of spectral parameters but also to their variability in time. The $SampEn$ of the spectral parameters is higher for the non-terminating than for the terminating episodes. The spectral parameters mathematical regularity might be associated with the physiological organization of the atrial activation This good results make this new method a useful tool that can help clinicians in the management of AF.

			Predicted Group		Total
			N	T	
Learning cases	Count	N	10	0	10
		T	0	10	10
	%	N	100	0	100
		T	0	100	100
Test cases	Count	N	15	1	16
		T	1	13	14
	%	N	93,75	6,25	100
		T	7,14	92,86	100

Table 2. Type N and T correctly classified recordings for both learning and test sets by using the discriminant analysis.

Acknowledgements

This work was partly supported by the projects PAID-04-08 and TEC2007–64884 from the Spanish Ministry of Education and Science.

References

[1] Fuster V, Ryden LE, Asinger RW, et al. CC/AHA/ESC guidelines for the management of patients with atrial fibrillation. Journal of the American College of Cardiology 2001;38(4):1266/I–1266/LXX.

[2] Kerr CR, Humphries KH, et al. Progression to chronic atrial fibrillation after the initial diagnosis of paroxysmal atrial fibrillation: Results from the canadian registry of atrial fibrillation. American Heart Journal 2005;149(3):489–496.

[3] Blomström-Lundqvist C, Scheinman MM, Aliot EM, Alpert JS, et al. ACC/AHA/ESC guidelines for the management of patients with supraventricular arrhythmias-executive summary. Circulation 2003;108(15):1871–1909.

[4] Pincus SM. Approximate entropy (ApEn) as complexity measure. Chaos 1995;5(1):110–117.

[5] Richman JS, Moorman JR. Physiological time-series analysis using approximate entropy and sample entropy. Am J Physiol Heart Circ Physiol 2000;(278):H2039–H2049.

[6] Hornero R, Abasolo D, Jimeno N, I SC, Poza J, Aboy M. Variability, regularity and complexity of time series generated by schizophrenic patients and control subjects. EEE Transactions on Biomedical Engineering 2006;53(2):210–218.

[7] Bollmann A, Husser D, Mainardi L, Lombardi F, Langley P, Murray A, Rieta JJ, Millet J, Olsson SB, Stridh M, L S. Analysis of surface electrocardiograms in atrial fibrillation: Techniques, research and clinical applications. Europace 2006;8(11):911–926.

[8] Rieta JJ, Castells F, Sánchez C, Zarzoso V. Atrial activity extraction for atrial fibrillation analysis using blind source separation. IEEE Transactions on Biomedical Engineering 2004;51(7):1176–1186.

[9] Auger F, Flandrin P, Gonçalvès P, Lemoine O. Time-Frequency Toolbox for use with MATLAB. Centre National de la Recherche Scientifique, 1996.

[10] Bollmann A, Lombardi F. Electrocardiology of atrial fibrillation. IEEE engineering in medicine and biology 2006; 25(6):15–23.

[11] Castells F, Rieta JJ, Millet J, Zarzoso V. Spatiotemporal blind source separation approach to atrial activity estimation in atrial tachyarrhythmias. IEEE Transactions on Biomedical Engineering 2005;52(2):258–267.

[12] Chiarugi F, Varanini M, Cantini F, Conforti F, Vrouchos G. Noninvasive ECG as a tool for predicting termination of paroxysmal atrial fibrillation. IEEE Transactions on Biomedical Engineering 2007;54(8):1399–1406.

[13] Stridh M, Bollmann A, Olsson SB, Sörnmo L. Detection and feature extraction of atrial tachyarrhythmias. IEEE engineering in medicine and biology 2006;25(6):31–39.

[14] Clayton R H MA. Comparison of techniques for time-frequency analysis of the ECG during human ventricular fibrillation. IEE Proc Sei Meas Technol 1998;145(6):301–306.

[15] Huberty CJ. Applied Discriminant Analysis. John Wiley and sons, 1994.

Address for correspondence:

Carlos Vayá
Innovation in Bioengineering.
Universidad de Castilla-La Mancha.
Camino del Pozuelo, s/n, 16071, Cuenca.
carlos.vaya@uclm.es