Study of Sample Entropy Ideal Computational Parameters in the Estimation of Atrial Fibrillation Organization from the ECG

R Alcaraz[1], D Abásolo[2], R Hornero[2], JJ Rieta[3]

[1]Innovation in Bioengineering Research Group. University of Castilla-La Mancha, Cuenca, Spain
[2]Biomedical Engineering Group, University of Valladolid, Valladolid, Spain
[3]Biomedical Synergy, Electronic Engineering Dept., Universidad Politécnica de Valencia, Spain

Abstract

Sample Entropy (SampEn) is a nonlinear regularity index that requires the a priori *selection of three parameters: the length of the sequences to be compared, m, the patterns similarity tolerance, r, and the number of samples under analysis, N. Appropriate values for m, r and N have been recommended in some cases, such as heart rate, hormonal data, etc., but no guidelines exist for the selection of that values. Hence, an optimal parameters study should be required for the application of SampEn to not previously analyzed biomedical signals. In this work, a thorough analysis on the optimal SampEn parameter values within two different scenarios of AF organization estimation, such as the prediction of paroxysmal AF termination and the electrical cardioversion outcome in persistent AF, is presented. Results indicated that, (i) the proportion between N and the sampling rate (f_s) should be higher than one second and $f_s \geq 256$ Hz, (ii) overlapping between adjacent N-length windows does not improve organization estimation and (iii) values of m and r maximizing classification should be considered within a range wider than the proposed in the literature for heart rate analysis.*

1. Introduction

The application of nonlinear regularity metrics to physiological signals is a valuable tool because "hidden information" related to underlying mechanisms can be obtained [1, 2]. To this respect, the employment of sample entropy (SampEn) to estimate non-invasively atrial fibrillation (AF) organization has revealed clinically useful information, which could be used for a better treatment of the arrhythmia [3].

Given a time series with N data points, the a priori selection of two unknown parameters, m and r, is required to compute SampEn [2]. The parameter m determines the length of the sequences to be compared and r is the tolerance for accepting similar patterns between two segments. Although these parameters are critical in determining the outcome of SampEn, no guidelines exist for optimizing their values. Typically recommended m and r values are $m = 1$ and $m = 2$ and r between 0.1 and 0.25 times the standard deviation (SD) of the data [4]. This recommendation is largely based on the application of approximate entropy (ApEn) to relatively slow dynamic signals such as heart rate [1, 4] and hormone secretion data [5]. However, a recent work has demonstrated that the typically recommended values for ApEn are not always appropriate for fast dynamic signals [6]. As a consequence, since only few values, within the range suggested in the literature [4], were tested in the previous works where SampEn was applied to AF organization estimation [3], the main goal of the present study is to carry out an in depth analysis on SampEn parameters able to achieve optimized classification of AF events which are directly dependent on AF organization.

2. Materials

Two different scenarios, such as the prediction of paroxysmal AF termination the electrical cardioversion (ECV) outcome in persistent AF, in which organization plays an important role, were analyzed.

Regarding paroxysmal AF, 50 Holter recordings of 30 seconds in length and two leads (II and V1) available in Physionet [8] were analyzed. The database included non-terminating AF episodes (group N), which were observed to continue in AF for, at least, one hour following the end of the excerpt, and AF episodes terminating immediately after the end of the extracted segment (group T). These signals were digitized at a sampling rate of 128 Hz and 16-bit resolution.

With regard to persistent AF, 63 patients with persistent AF lasting more than 30 days, undergoing ECV were followed during four weeks. A standard 12-lead ECG was acquired for each patient during the whole procedure and a segment of 30 seconds in length was extracted from each recording for the analysis. All the signals were digi-

tized at a sampling rate of 1024 Hz and 16-bit resolution. After ECV, 22 patients (34.93%) maintained normal sinus rhythm (NSR) during the first month. In 31 patients (49.20%), NSR duration was below one month and the remaining 10 (15.87%), relapsed to AF immediately after ECV. These 41 patients constituted the group of AF recurrence. All the patients were under drug treatment with amiodarone.

3. Methods

3.1. AF organization estimation

In both databases, lead V_1 was chosen for the analysis because previous works have shown that AA is prevalent in this lead [7]. To estimate AF organization, the application of SampEn to the surface ECG requires the fulfillment of several steps. Firstly, the ventricular activity has to be removed making use of a cancellation technique. Next, the main atrial wave (MAW) has to be extracted from the AA by applying a selective filtering centered on the dominant atrial frequency (DAF), i.e. the highest amplitude frequency within the 3–9 Hz range. Finally, SampEn computation can be applied to this wave. This approach has been described in detail in previous works [3].

3.2. Selection of N and f_s

Previous works, in which heart rate and hormonal data regularity were analyzed [1, 4, 5, 9], have shown that for $m = 2$, values of r from 0.1 to 0.25 times the SD of the data and values of N between 100 and 5000 samples produce good statistical validity of ApEn. Additionally, Pincus and Goldberger [1] suggested that N should be at least 10^m and, preferably, at least 30^m when heart rate regularity is analyzed. However, this recommendation is not applicable to the case studied in this work, because the MAW length depends on the ECG sampling rate. To this respect, the selection of a specific number of samples N will stretch out more or less the MAW information as a function of the sampling rate (f_s). Hence, for high f_s, a low MAW time interval would be analyzed and the opposite also holds.

To obtain the optimal values of these parameters, the MAW sampling rate was varied and, for each case, SampEn was computed with different N values. More precisely, the selected N values were 30, 60, 120, 240, 480, 960, 1920, ..., L samples, where L is the closest series length lower than the analyzed MAW segment. Given that the minimum considered sampling rate was 64 Hz and that the ECG lengths were 30 seconds, their combination yields 1920 samples. As a consequence, all the selected N values took that number as reference, divided or multiplied by a power of two. On the other hand, the MAW was downsampled or upsampled to 64, 128, 256, 512, 1024 and 2048 Hz. Since the original sampling rates were 128 Hz for paroxysmal AF and 1024 Hz for persistent AF, a method based on cubic splines was applied when interpolation was necessary.

The discriminative abilities for both predictions of paroxysmal AF termination and ECV result were calculated making use of the receiver operating characteristic (ROC) curves. The total number of paroxysmal AF patients and ECVs precisely classified was considered as the diagnostic accuracy corresponding to each prediction. A leave-one-out cross-validation scheme was used to validate the statistical robustness of obtained results. Finally, statistically significant differences between groups were evaluated making use of Student's t–test.

3.3. Selection of m and r

Once the optimal combinations of N and f_s were obtained, the most adequate selection of m and r was investigated. The accuracy and confidence of SampEn estimate improves as the number of length m matches increases. The number of matches can be increased by choosing small m (short templates) and large r (wide tolerance). However, penalties appear when too relaxed criteria are used [9]. For small r values, poor conditional probability estimates are achieved, while for large r values, too much detailed system information is lost and SampEn tends to 0 for all the processes. To avoid a significant noise contribution on SampEn computation, one must choose r larger than most of the noise [9]. Overall, to get optimal m and r values, an approach similar to the developed by Lake et al [10, 11] was used. SampEn was computed using a 10×20 matrix of combinations of $m = 1, 2, \ldots, 10$ and $r = 0.05, 0.1, 0.15, \ldots, 1$ times the SD of the analyzed segment.

4. Results

4.1. Selection of N and f_s

Results showed that for a given N, when f_s increases, SampEn decreases. On the other hand, for a given f_s of 256 Hz and above, that SampEn increases when N also increases up to a saturation point. Thus, for higher N, very similar SampEn values were obtained with variations lower than $\pm 5\%$. On the contrary, for f_s of 128 Hz and below, no clear increasing tendency was observed. Moreover, the value of N from which the limited increasing behavior was noticed was dependent on f_s. Thus, for rates of 256, 512, 1024 and 2048 Hz, the reduced variation N values for SampEn were 240, 480, 960 and 1920, respectively.

The discriminative differences between subsets of both databases were maintained for f_s of 256 Hz and above. Indeed, a high predictive accuracy was achieved, such as

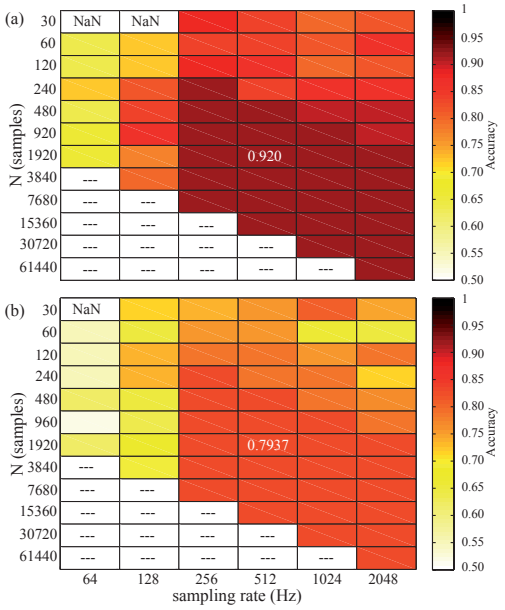

Figure 1. Diagnostic accuracy color maps obtained with different combinations of N and f_s for (a) paroxysmal AF termination and (b) cardioversion of persistent AF. NaN in some cells stands for Not a Number because SampEn did not report similarity in such a short time series.

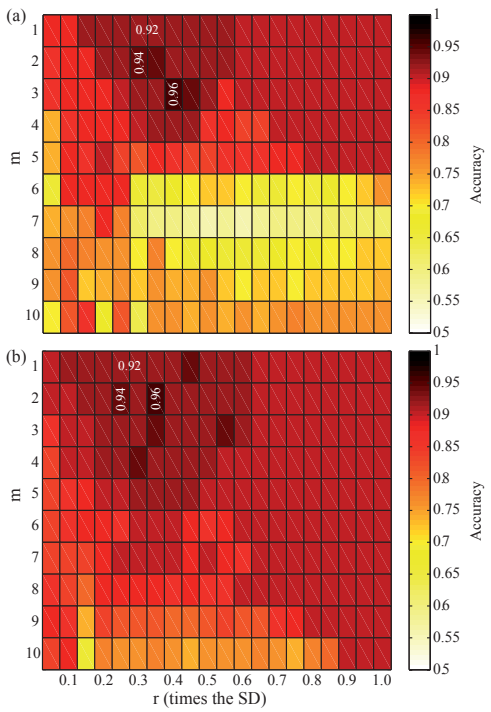

Figure 2. Diagnostic accuracy color maps for paroxysmal AF termination prediction obtained as a function of $m = 1, 2, \ldots, 10$ and $r = 0.05, 0.1, 0.15, \ldots, 1$ times the SD of the data, being N and f_s (a) 240 samples—256 Hz and (b) 960 samples—1024 Hz, respectively.

Fig. 1 shows. Precisely, for $f_s \geq 256$ Hz the lowest accuracy was 86% for paroxysmal AF and 73.02% for persistent AF, being 90% and 74.60% of cross-validated grouped cases correctly classified, respectively. In addition, differences between groups were noticed, given that $p < 0.01$ for all the cases. On the contrary, for $f_s \leq 128$ Hz a limited predictive ability was obtained for most pairs of N and f_s. Moreover, in most of the cases, $p > 0.01$.

It is also noteworthy that for N equal or higher than the saturation values for SampEn, the diagnostic accuracy presented a constant result of 92% for paroxysmal AF and 79.37% for persistent AF, which was higher than the obtained with lower N values for a given f_s (see Fig. 1). In addition, a higher statistical significance ($p < 0.0001$) together with an identification accuracy of 96% and 82.54% of cross-validated grouped cases, respectively, were also yielded.

4.2. Selection of m and r

Results from the previous section showed that several combinations of N and f_s can provide good classification for both AF databases. Thereby, two different combinations were selected for the optimal selection of m and r. In this case, the outcomes related to the diagnostic accuracy are presented in Figs. 2 and 3. A remarkable region in which a high accuracy for the two AF databases was reached can be appreciated. In addition, a very good statistical significance ($p < 0.00001$) was also obtained. As can be seen, the region was slightly larger for the pair $N = 960$ samples and $f_s = 1024$ Hz. This area can be approximately delimited, for both AF databases, by $m = 1$ to $m = 5$ and r between 0.1 and 0.6 times the SD. Additionally, in all the cases, the optimal combination of m and r providing the highest diagnostic accuracy values was found within the indicated area.

5. Discussion and conclusions

Results showed that, for f_s of 256 Hz and above, the differences between each subset, both for paroxysmal and persistent AF, were maintained independently of N. On the other hand, for a specific f_s higher than 128 Hz, when N increased, SampEn also increased up to a saturation point. However, this behavior was not observed for $f_s = 128$ Hz and below. Overall, a first recommendation is that the MAW sampling rate should be equal or higher than 256 Hz for an appropriate AF organization evaluation with SampEn.

On the other hand, considering the saturation point of N, where a constant accuracy was reached for both AF

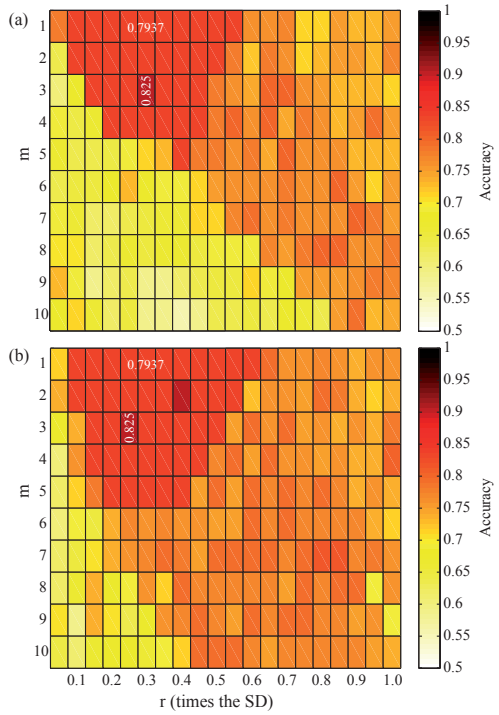

Figure 3. Diagnostic accuracy color maps for ECV result prediction obtained as a function of $m = 1, 2, \ldots, 10$ and $r = 0.05, 0.1, 0.15, \ldots, 1$ times the SD of the data, being N and f_s (a) 240 samples—256 Hz and (b) 960 samples—1024 Hz, respectively.

types, a clear relationship between N and f_s can be elucidated for all of these situations. As the time period analyzed with SampEn, T, is the proportion between N and f_s, for all these cases we have that $T = \frac{240}{256} = \frac{480}{512} = \frac{960}{1024} = \frac{1920}{2048} = 0.9375$ s. As a consequence, any combination such that the analyzed time period is longer than one second would be appropriate to evaluate AF organization with SampEn.

With regard to optimal values for m and r, note that the region outlined by the combinations with highest accuracy was considerably larger than the one typically recommended in the literature [4]. Nevertheless, these results prove that previous works used adequate values for the SampEn parameters in the estimation of AF organization [3]. However, a better diagnostic precision could be reached for each prediction with a combination of m and r placed within the wider indicated region. An interesting observation is that optimal values reporting the highest accuracy were dissimilar depending on the combination of N and f_s and on each type of AF. As a consequence, it can be suggested that the optimal combination of m and r has to be searched, within the proposed range, before the first application of SampEn to a non-previously analyzed database.

Acknowledgments

This work was supported by the projects TEC2007–64884 and TEC2010–20633 from the Spanish Ministry of Science and Innovation and PII2C09-0224-5983 and PII1C09-0036-3237 from Junta de Comunidades de Castilla La Mancha.

References

[1] Pincus SM, Goldberger AL. Physiological time-series analysis: what does regularity quantify? Am J Physiol Apr 1994;266(4 Pt 2):H1643–H1656.

[2] Richman JS, Moorman JR. Physiological time-series analysis using approximate entropy and sample entropy. Am J Physiol Heart Circ Physiol Jun 2000;278(6):H2039–H2049.

[3] Alcaraz R, Rieta JJ. A review on sample entropy applications for the non-invasive analysis of atrial fibrillation electrocardiograms. Biomed Signal Proc Control 2010;5:1–14.

[4] Pincus SM. Assessing serial irregularity and its implications for health. Ann N Y Acad Sci Dec 2001;954:245–267.

[5] Pincus SM, Keefe DL. Quantification of hormone pulsatility via an approximate entropy algorithm. Am J Physiol May 1992;262(5 Pt 1):E741–E754.

[6] Lu S, Chen X, Kanters JK, Solomon IC, Chon KH. Automatic selection of the threshold value r for approximate entropy. IEEE Trans Biomed Eng 2008;55(8):1966–1972.

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

[8] Goldberger AL, Amaral LA, Glass L, Hausdorff JM, Ivanov PC, Mark RG, Mietus JE, Moody GB, Peng CK, Stanley HE. Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals. Circulation Jun 2000;101(23):E215–E220.

[9] Pincus SM. Approximate entropy as a measure of system complexity. P Natl Acad Sci USA 1991;88(6):2297–2301.

[10] Lake DE, Richman JS, Griffin MP, Moorman JR. Sample entropy analysis of neonatal heart rate variability. Am J Physiol Regul Integr Comp Physiol Sep 2002; 283(3):R789–R797.

[11] Lewis MJ, Short AL. Sample entropy of electrocardiographic RR and QT time-series data during rest and exercise. Physiol Meas Jun 2007;28(6):731–744.

Address for correspondence:

Raúl Alcaraz Martínez
E. U. Politécnica de Cuenca
Campus Universitario
16071 Cuenca (Spain)
raul.alcaraz@uclm.es