ENTROPY MEASURES OF HEART RATE VARIABILITY FOR SHORT ECG DATASETS IN PATIENTS WITH CONGESTIVE HEART FAILURE*

Beata Graff

Hypertension Unit, Department of Hypertension and Diabetology
Medical University of Gdańsk, Dębinki 7C, 80-952 Gdańsk, Poland

Grzegorz Graff†, Agnieszka Kaczkowska‡

Faculty of Applied Physics and Mathematics, Gdańsk University of Technology
Narutowicza 11/12, 80-233 Gdańsk, Poland

(Received December 5, 2011)

We investigated the usefulness of entropy measures calculated for short ECG series in distinguishing healthy subjects from patients with congestive heart failure (CHF). Four entropy measures were tested: Approximate Entropy (ApEn), Sample Entropy (SampEn), Fuzzy Entropy (FuzzyEn) and Permutation Entropy (PE), each computed for ECG series of 1000, 500, 250 and 100 RR intervals. We found that with a reduction of the data set length up to 250 RR intervals, values of ApEn, SampEn, FuzzyEn and PE can remain significantly different in patients with CHF compared to healthy individuals. SampEn and FuzzyEn differentiated considered groups even for data sets of 100 RR intervals.

DOI:10.5506/APhysPolBSupp.5.153
PACS numbers: 87.19.Hh, 89.70.Cf

1. Introduction

An important challenge in contemporary cardiology is rapid and accurate identification of patients in pathological conditions. Entropy-based measures are powerful methods used to analyze the degree of irregularity of short time series, so they can be considered for the estimation of heart rate complexity.

* Presented at the 8th European Conference on Mathematical and Theoretical Biology, Mini-symposium "Heart Rate Dynamics: Models and Measures of Complexity", Kraków, Poland, June 28–July 2, 2011.

† graff@mif.pg.gda.pl
‡ akaczkowska@mif.pg.gda.pl

Although usually the entropy-based measures were applied to the series of at least 750 data points [1], they seem to provide valuable information even for shorter sequences. For example, recently Batchinsky *et al.* observed that strong data set reduction up to 100 RR for SampEn and 200 RR for ApEn, does not diminish their usefulness to predict mortality in trauma patients [2].

In 2006 Signorini *et al.* successfully applied ApEn and SampEn to find out the difference between patients with congestive heart failure (CHF) and healthy subjects in 24 hours, 5 hours and 15 minutes recordings [3]. The aim of this paper is to examine whether ApEn, SampEn, FuzzEn and PE can provide an effective separation of pathological condition of CHF and the physiological state despite the recording length reduction from 1000 till 100 RR intervals.

2. Methods

2.1. Clinical data

The ECG data analyzed in the paper were downloaded from the Physionet Congestive Heart Failure RR Interval Database and Normal Sinus Rhythm RR Interval Database [4]. We analyzed the RR intervals in the segment $[1000, 2000]$ of each recordings. The CHF group consisted of 23 patients (7 men, aged 39–68, 2 women, aged 38 and 59, and 14 subjects, sex not known, aged 34–66) and the control group consisted of 26 healthy subjects (9 men, aged 59–72, and 14 women, aged 58–71). The sampling frequency was 128 Hz. Recordings with number of artifacts and ectopic beats larger than 5% of all beats were not included in the study. In order to remove the artifacts we applied the method recommended in [3], *i.e.* each RR-interval that differs by more than 30% from the mean of the 6 previous values was replaced by the mean of the previous 6 samples.

2.2. Entropy-based measures

In all studied subjects ApEn, SampEn, FuzzyEn and PE values were obtained. The algorithms for computing each kind of entropy are described in the following papers: ApEn [5], SampEn [6], FuzzyEn [7], PE [8].

In order to compute ApEn, SampEn or FuzzyEn one needs three input parameters have to be fixed: m — the length of compared runs; r — tolerance distance expressed as a percentage of standard deviation of the data sets; N — the length of time series. In our investigations the most recommended values of m and r, *i.e.* $m = 2$ and $r \in \{0, 1\mathrm{SD}, 0, 2\mathrm{SD}\}$ were chosen.

PE depends on two parameters: L — the number of considered patterns and N — the length of time series. Following the suggestions in [8] that optimal values of L should satisfy $L! \leq N - L + 1$ we considered the values of $L \in \{3, 4, 5\}$.

We studied the changes of each kind of entropy as a result of the decrease of N: $N \in \{1000, 500, 250, 100\}$.

3. Results and discussion

The entropy values are presented in Table I and Table II and expressed as means ± SD. Distributions of variables were determined by the Kolmogorov–Smirnov test. Group differences were assessed by unpaired t-test or Mann–Whitney 2-sample test, as appropriate. Data analysis was performed using STATISTICA (Version 9, StatSoft) software. A value of $p < 0.05$ was considered statistically significant.

TABLE I

ApEn, SampEn and FuzzyEn for $r = 0.2$ SD and $r = 0.1$ SD and for data length $N = 1000, 500, 250$ and 100.

	$r = 0.2$ SD			$r = 0.1$ SD		
N	CHF	Healthy	p	CHF	Healthy	p
			ApEn(2,r,N)			
1000	1.238 ± 0.23	0.981 ± 0.27	**<0.01**	1.269 ± 0.17	1.068 ± 0.20	**<0.001**
500	1.128 ± 0.18	0.917 ± 0.21	**<0.001**	1.142 ± 0.14	0.880 ± 0.14	**<0.001**
250	0.958 ± 0.20	0.824 ± 0.16	**0.01**	0.967 ± 0.17	0.744 ± 0.16	**<0.001**
100	0.707 ± 0.23	0.673 ± 0.12	ns	0.724 ± 0.20	0.512 ± 0.18	**<0.001**
			SampEn(2,r,N)			
1000	1.375 ± 0.39	0.971 ± 0.39	**<0.01**	1.438 ± 0.32	1.616 ± 0.45	ns
500	1.372 ± 0.40	0.967 ± 0.38	**<0.001**	1.431 ± 0.34	1.633 ± 0.50	ns
250	1.362 ± 0.45	0.963 ± 0.38	**<0.01**	1.419 ± 0.38	1.695 ± 0.63	ns
100	1.306 ± 0.44	1.005 ± 0.38	**0.02**	1.351 ± 0.37	1.733 ± 0.73	**0.03**
			FuzzyEn(2,r,N)			
1000	1.278 ± 0.39	0.798 ± 0.27	**<0.0001**	1.548 ± 0.38	1.386 ± 0.34	ns
500	1.285 ± 0.38	0.805 ± 0.25	**<0.001**	1.557 ± 0.37	1.394 ± 0.33	ns
250	1.267 ± 0.41	0.809 ± 0.26	**<0.001**	1.533 ± 0.39	1.398 ± 0.31	ns
100	1.250 ± 0.43	0.828 ± 0.27	**<0.001**	1.499 ± 0.41	1.421 ± 0.33	ns

TABLE II

Permutation entropy for $L = 3, 4, 5$ and for data length $N = 1000, 500, 250$ and 100.

	PE(L, N)		
N	CHF	Healthy	p
	$L = 3$		
1000	1.207 ± 0.04	1.235 ± 0.03	**0.01**
500	1.207 ± 0.04	1.233 ± 0.04	**0.03**
250	1.198 ± 0.05	1.228 ± 0.03	**0.02**
100	1.180 ± 0.05	1.209 ± 0.05	ns
	$L = 4$		
1000	1.388 ± 0.06	1.427 ± 0.04	**0.02**
500	1.385 ± 0.07	1.416 ± 0.05	ns
250	1.367 ± 0.07	1.400 ± 0.04	ns
100	1.322 ± 0.07	1.350 ± 0.06	ns
	$L = 5$		
1000	1.510 ± 0.08	1.561 ± 0.06	**0.02**
500	1.492 ± 0.09	1.529 ± 0.07	ns
250	1.439 ± 0.08	1.474 ± 0.06	ns
100	1.323 ± 0.08	1.347 ± 0.07	ns

The results proved that entropy-based measures are able to distinguish healthy subjects from patients with CHF even for very short ECG recordings. The difference between compared groups depended on the kind of entropy measure and the choice of input parameters (*i.e.* m, r, L). The specially good discriminative properties were shown by ApEn(2,0.1,N) and FuzzyEn(2,0.2,N) although in different ways: ApEn(2,0.1,N) values in both groups gradually decreased with a reduction of the data set length (figure 1) while in case of FuzzyEn(2,0.2,N) values in each group were almost constant (figure 2).

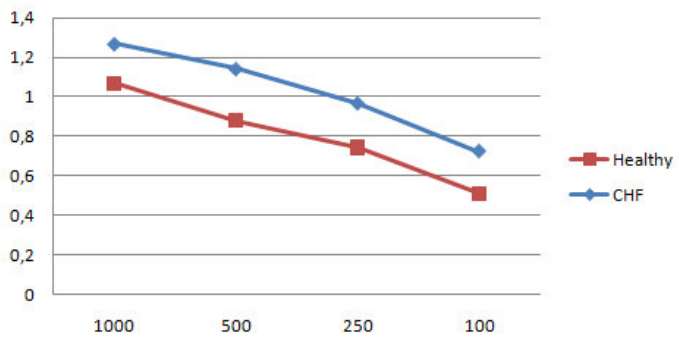

Fig. 1. Values of ApEn(2,0.1,N) for decreasing N.

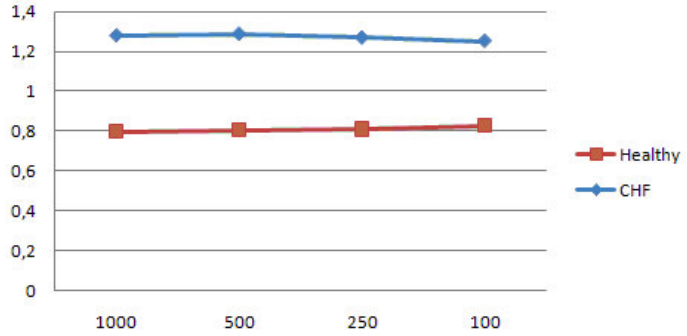

Fig. 2. Values of FuzzyEn(2,0.2,N) for decreasing N.

Higher entropy values were obtained in patients with CHF compared to healthy subjects. In the study by Beckers *et al.* [9] values of ApEn were also higher in CHF group but the difference was significant only during the night. Our findings are fully consistent with those obtained by Signorini *et al.* who reported higher ApEn and SampEn values in short data sets of 15 minutes [3]. The increase of heart rate complexity in patients with heart failure was also found by Struzik *et al.* [10].

4. Conclusion

Entropy-based measures obtained from short ECG data sets up to 100 RR intervals can separate effectively healthy subjects and patients with congestive heart failure.

This work was supported by the Foundation for Polish Science TEAM/2008-2/5 (B.G.) and the Polish Ministry of Science and Higher Education Grant no. N N201 373236 (G.G.).

REFERENCES

[1] M. Costa, C.-K. Peng, A.L. Goldberger, J.M. Hausdorff, *Physica A* **330**, 53 (2003).

[2] A.I. Batchinsky *et al.*, *Shock* **32**, 565 (2009).

[3] M.G. Signorini, M. Ferrario, M. Marchetti, A. Marseglia, *Conf. Proc. IEEE Eng. Med. Biol. Soc.* **1**, 3431 (2006).

[4] A.L. Goldberger *et al.*, *Circulation* **101**, e215 (2000).

[5] S.M. Pincus, *Proc. Nati. Acad. Sci. USA* **88**, 2297 (1991).

[6] J.S. Richman, J.R. Moorman, *Am. J. Physiol. Heart Circ. Physiol.* **278**, H2039 (2000).

[7] W. Chen, J. Zhuang, W. Yu, Z. Wang, *Med. Eng. Phys.* **31**, 61 (2009).

[8] J. M. Amigó, *Permutation Complexity in Dynamical Systems*, Springer-Verlag, 2010.

[9] F. Beckers, D. Ramaekers, A.E. Aubert, *Cardiovasc. Eng.* **1**, 177 (2001).

[10] Z.R. Struzik *et al.*, *Europhys. Lett* **82**, 28005-p1 (2008).