Chaos and Complexity Letters, 2007, Volume 3, Issue 3 ISSN 1555-3995
Editor: Franco F. Orsucci, pp.

NON-INVASIVE ASSESSMENT OF RISK FOR SEVERE TACHYARRHYTHMIAS BY MEANS OF NON-LINEAR ANALYSIS TECHNIQUES

***R. Pizzi*[*], *G. Inama*[**], *O. Durin*[*] and *C. Pedrinazzi*[*]**
[*]Department of Information Technologies, University of Milan
via Bramante 65 – 26013 Crema , Italy
[**]Cardiology Department, Hospital of Crema, L.go U.
Dossena 2 – 26013 Crema, Italy

ABSTRACT

Sudden death remains a phenomenon of disturbing proportions, displaying a mean incidence of 300,000-350,000 persons/year in the USA (0.1-0.2% of the general population). In Europe, the figures are very similar. In 90% of cases, sudden death has an arrhythmic cause.

Prevention of Sudden Cardiac Death (SCD) constitutes one of the most important challenges of modern cardiology. In order to make a real progress in this field it is crucial to identify precisely increased risk for serious ventricular tachyarrhythmias.

In this study the effectiveness of different methods of the non-linear analysis (NLA) of ECG in the risk stratification of patients with ventricular arrhythmias is evaluated, and these non-invasive parameters are correlated with the results of invasive electrophysiological study (EPS).

We evaluated 25 patients with history of cardiac arrest, syncope, sustained or nonsustained ventricular tachycardia (VT). The study group was compared with a control group of 25 healthy subjects. All patients underwent both electrophysiologic study (EPS) and non-linear analysis (NLA) of ECG.

Patients were classified through the application of a clustering procedure to the whole set of functions, and a comparison between the results of non-linear analysis of ECG and EPS was performed. Results are presented and discussed.

Keywords: ECG, Sudden Cardiac Death, Ventricular Tachyarrythmias, Non-linear Analysis, Chaos Theory, Clustering, Artificial Neural Networks, Attractors.

[*] E-mail: pizzi@dti.unimi.it

1. INTRODUCTION

Sudden cardiac death (SCD) represents one of the main problems of modern Cardiology (Myerburg 1993). A cardiac arrest can nowadays be recovered by means of Internal Cardiac Defibrillators (ICD), but these devices are expensive, complex and permanently invasive, and request periodic follow-ups for the remaining life (Moss 2001).

Nowadays our main problem in the strategy against SCD is not the lack of an effective therapy but the need of a test (or of some tests) able to stratify accurately the subjects suitable for the implant of an ICD. Today we have at our disposal many diagnostic tests to attain such stratification (Holter ECG, Signal Average ECG, T-wave alternans, induction of tachyarrhythmias during electrophysiologic endocavitary study an so on) but no one of these tests showed to be so valuable for prognostic evaluation, nor the combination of two or more of the showed a predictive value of future arrhythmic episodes greater than 30-40% (Buxton 1999, McLaughlin 2006, Furlanello 2004, Zipes 2006).

Therefore our purpose was to evaluate the possibility to assess the arrhythmic risk using a new non-invasive and low-cost method based on non-linear analysis (May 1976, Jackson 1989, Atmanspacher 1992, Kaplan 1995).

The study of heart rate variability using linear systems in both the time and frequency domains in order to analyse ECG signals is not proven to be able to solve the problem of the non-invasive stratification of subjects at risk: indeed, the variability of the ECG signal appears to be sustained by the complex interaction of several different systems, which would be better described by non-linear functions (Goldberger 1990,Signorini 1994, Bezarianos 1995).

The aim of our study was to assess the effectiveness of the non-linear analysis of ECG signals in the risk stratification of patients with ventricular arrhythmias, and to correlate these non-invasive parameters with the results of invasive electrophysiological study (EPS).

2. METHODS

2.1. Non-Linear Analysis of ECG Signals

Although research in this fascinating sector has been underway for some 15 years (Ravelli 1992, Goldberger 1996, Glass 1999, Braun 1998, Huikuri 2003), its results have unfortunately not found widespread application in clinical practice; indeed, the complexity of the calculations involved requires the availability of adequate computing capacity, in terms of both dedicated software and skilled human resources. This drawback has given rise to a certain fragmentation of the methods variously utilised (according to the means and experience of the various research centres). Moreover, it has also resulted in a shortage of literature in which the validity of such methods might be compared; indeed, the papers produced with the aim of analysing the previous and current situations are fairly recent (Lombardi 1996, Makikallio 1997, Owis 2002, Mohamed 2002, Huikuri 2003, Vigo 2004, Wu 2005, Stein 2005, Anderson 2005). Nevertheless, although the proliferation of numerous, and often very diverse, mathematical functions has sometimes produced apparently contradictory results, it has, at the same time, probably helped to explore the problem from

different standpoints, thereby providing new insights. For example, it has been possible to infer that the dynamics underlying the variability of a complicated biological system, such as that of the cardiac rhythm of a healthy individual, is probably characterised by a fairly organized degree of complexity (which can be assimilated to a system of deterministic chaos (Gleick 1987, Glass 1988, Stewart 1989, Lorenz 1993)); in a pathological state, however, this may tend either towards extreme simplification, or towards extreme, totally random disorder (Lombardi 2000, Otero 2003).

Healthy subjects are capable of satisfactory responses and adaptations to multiple, and often rapid, modifications of the internal and external environments.

The deterministic chaos that characterises a healthy biological system differs profoundly from completely haphazard variability (randomness), in that the extremely complex variability of the former is endowed with a self-similarity (fractal) which repeats in different timescales (days, hours, minutes) (Hausdorff 1919, Mandelbrot 1983, Otsuka 1997, Makikallio 1998, Makikallio 1999, Huikuri 2000, Otero 2003). Likewise, it differs from the simple periodic variability that characterises the ECG time series preceding serious cardiac arrhythmic events, which show a regular signal, dynamically characterised by limit cycles or punctual attractors. Indeed, a pathological substrate appears to be strictly bound to an increase in the order and regularity of the ECG dynamic system (or, by contrast, to the appearance of sharp randomness), displaying less flexibility and adaptability to great and sudden environmental changes.

The available data show that the most commonly used functions are: the Power Law Slope (Bezerianos 1995, Makikallio 1999, Huikuri 2000, Stein 2005), the Short-Term Fractal Scaling Exponent (or Alpha 1 exponent), calculated by means of Detrended Fluctuation Analysis (DFA1) (Lombardi 1996, Wu 2005), the quantification of Poincaré Plots (Makikallio1998, Anderson 2005) and the Correlation Dimension, or modified calculation of the Pointwise Correlation Dimension (Makikallio 1999, Kroll 1992), the results of which have been subjected to the usual statistical evaluations.

On the basis of the previous considerations we hypothesized the possibility to evaluate the risk of developing life-threatening cardiac arrhythmias by studying the non-linear features of ECG signals.

2.2. Recurrence Quantification Analysis

We treated the ECG time series on the basis of the Recurrence Quantification Analysis (Zbilut 1992, Kononov 1996).

A one-dimensional time series is extended to a higher dimensional space time series using the delay-time embedding technique (Grassberger 1983). In short, in order to extend a one-dimensional signal to an m-dimensional one, each observation in the original signal X(t) is substituted by a vector

$$y(i) = \{ x(i), x(i, d), x(i, 2d),..., x(i - (m-1) d \}$$

where i is the time index, m the embedding dimension, d the delay time.
As a consequence we obtain a series of vectors

$Y = \{ y(1), y(2), y(3), ., y(N - (m-1) d) \}$

where N is the length of the original series.

Once the dynamical system is reconstructed as above, it is possible to process different quantitative functions to evaluate the peculiar features of the dynamical system itself.

Recurrent values consecutive in time are graphically represented in a Recurrence Plot (RP) (Eckmann 1987) by lines parallel to the principal diagonal that are an important sign of deterministic structure. In fact it can be shown that the longest diagonal line corresponds to the value of the maximum Lyapounov exponent of the series. The Lyapounov exponent quantifies the mean rate of divergence of trajectories along the directions of the phase space. Chaotic systems have a positive maximum Lyapounov exponent.

A Recurrence Plot can highlight the hidden structure of the system and its internal structural changes. The basic idea of the RP is to form a color-coded matrix, where each [i][j]th entry is calculated as the Euclidean distance between all the pairs of vectors Y_i and Y_j in the reconstructed series and codified as colors.

For random signals, the uniform distribution of colors over the entire RP is expected. The more deterministic the signal, the more structured the RP. Hot colours (yellow, red, orange) are associated to short distances between vectors, cold colours (blue, black) represent long distances. Signals repeating fixed distances between vectors are organized, signals without repeating distances are not. In this way we obtain uniform colour distribution for random signals, but the more deterministic and self-similar is the signal, the more structured is the plot.

The RP yield several quantitative functions useful to evaluate the internal structure of the dynamical system underlying the EEG signals.

We considered the following functions:

Mutual Information

We can define Mutual Information as a quantity that measures the mutual dependence of two random variables.

Given the system $X(t)$, Mutual Information measures the number of bits predictable in $X(t+d)$, where d is the delay time. In case of independence between the two variables, Mutual Information is zero, otherwise is greater than zero.

Entropy

Entropy is calculated on the RP both in the space and in the time domain. This quantity compares the distribution of distances between all the pairs of vectors in the reconstructed space with the distribution of distances between the trajectories evolving in time.

The function compares the global distribution of colors inside the RP with the distribution of colors on each diagonal line. The more evident are the differences between the global distribution and the distributions in the single diagonal lines, the more the image is structured, the lower is entropy. Entropy is small when the length of the longest segment parallel to the diagonal is short. A high entropy is typical of periodic behavior, low entropy indicates chaotic behavior.

Recurrence

It measures the percentage of recurrent points: a point *(i,j)* is recurrent if the distance between the vectors *y(i)* and *y(j)* is less than a threshold, and is calculated as the ratio of the number of recurrent states measured with respect to all the possible states.

Determinism

It is the percentage of recurrent plots forming line segments parallel to the main diagonal. The presence of these lines reveals the existence of a deterministic structure.

Ratio

This function is the ratio between the value of Determinism and the value of Recurrence, therefore it is an index of self-organization, i.e. of a spatiotemporal structure that emerges spontaneously from the evolution of the system as a function of its dynamics.

After the calculation of these functions we classified the patients applying a clustering technique to the whole set of functions.

2.3. Hierarchical Clustering

Clustering is particularly suitable to find regularities inside large amounts of heterogeneous data (Anderberg 1973, Duran 1974, John 1975). The clustering method applied to a set of discriminat variables gives a global response, i.e. it does not give information of each single variable but of the interaction between all the used parameters.

The hierarchical clustering algorithms encountered a large popularity inside the community of biologists and physicians due to their effectiveness (Rohlf 1973, Day 1984, Olson 1993, Fraley 1996). Their purpose is the organization of data in a hierarchical structure, that collects similar observations into small cluster at a lower level, and more basically connected observations into larger clusters and so on, up to the whole set of data.

More precisely, we consider a sequence of h partitions of the N data, where n_h is the cardinality of the ith partition. The first partition of the sequence is represented by a unique set $C_1 = \{d_i \mid 1 < i < N\}$ including all the observations, the second partition forms $n_2 > 2$ disjointed subsets complementary of C_1 and so on.

The hierarchical methods use several techniques for the fusion of observations. We used the so-called Average Group Linkage, where each group is represented by the mean value of each variable, and the inter-group distance is defined as the distance between two mean vectors. Considering two hypothetical clusters r and s, they are grouped together into the cluster t in such a way as their mean distance *D(r,s)* is the lowest one. The distance between the two clusters *D(r,s)* can be defined as:

D (r,s) = Mean { d (i,j) : where i and j are in the cluster t, formed by the union of r and s }.

2.4. Artificial Neural Network Analysis

For comparison purposes, an alternative non-linear technique to classify arrhythmic patients was tested by means of a custom Artificial Neural Network, called Inductive Tracing Self-Organizing Map (ITSOM).

Artificial Neural Networks are intrinsically non-linear models able to classify complex patterns. In particular, the self-organizing networks as the Kohonen's Self Organizing Map (SOM) is well-known as a natural non-linear classifiers (Kohonen 1983, Kohonen 1993).

But two main reasons exist that limit the SOM's performances in case of strictly non-linear and time-variant input.

The first reason is that if the input topology is too tangled, the competitive layer is not able to unfold itself enough to simulate the input topology.

The second reason concerns the SOM's convergence conditions that are not easily verifiable. Due to the nature of the SOM's output (non-homologous to the input), it is not possible to settle either a network error for each epoch, or the number of epochs after that the network training has to be stopped.

Another problem of the SOM, typical of any clustering algorithm, is the lack of output explication. Once a classification is obtained, the user must analyze it, comparing it to the input values in order to extrapolate a significant output.

Thus we proposed ITSOM, a structural modification of the SOM The dynamical properties of the artificial neural networks and of the SOM in particular are well known (Ritter and Schulten, 1986, 1988; Ermentrout, 1992).

During simulations carried out with the SOM algorithm we observed that, even if the winning weights may vary at any presentation epoch, their temporal sequence tends to repeat itself.

A deeper analysis has shown that such a sequence, provided to keep the learning rates steady (instead of gradually decreasing them), constitutes chaotic attractors that repeat "nearly" exactly in time with the epochs succeeding, and that, once codified by the network, univocally characterize the input element that has determined them.

Actually the SOM learning rule, as explained above, makes it possible for the winning weight to represent an approximation of the input value. In this way at every epoch the new winning weight, together with the previous winner, constitutes a second-order approximation of the input value. At the n-th epoch, the set of n winning weights represents an nth-order approximation of the input value.

In this way, due to the countless variety of possible combinations among winning neurons, the configurations allow to determine finely the correct value, even in the case of tangled input topologies, despite of the small number of competitive neurons and their linear topology.

In the following step the network performs a real induction process, because after a many-to-few vector quantization from the input to the weight layer (to be precise, to the chaotic configurations of winning weights), a few-to-many procedure is performed from the chaotic configurations corresponding to the input set (Figure 1) codified by the network.

It should be stressed that the ITSOM crucial feature is that the network does not need to be brought to convergence, as the cyclic configurations stabilize their structure within a small number of epochs, then keep it steady through time.

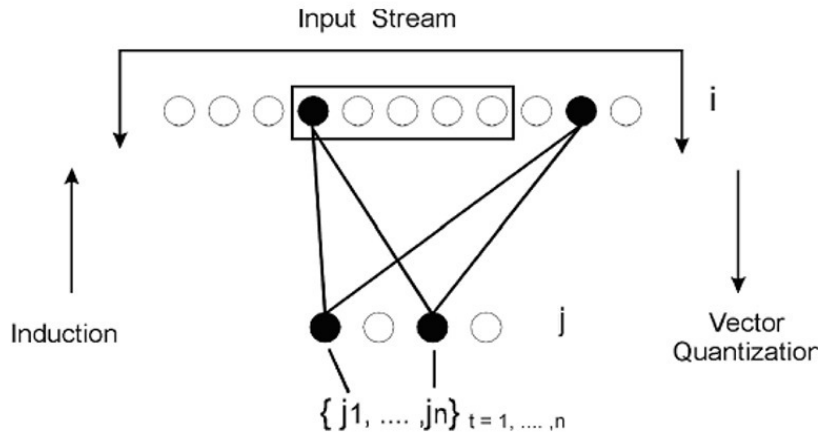

Figure 1. The ITSOM network.

After interrupting the network-processing phase, an algorithm is needed that codifies the obtained chaotic configurations into a small set of outputs.

The algorithm that has shown best performances and computational load among the tested pattern recognition algorithms is based on a *z*-score calculus.

The cumulative scores related to each input have been normalized following the distribution of the standardized variable *z* given by:

$$z = (x - \mu) / \sigma$$

where μ is the average of the scores on all the competitive layer weights and σ is the root mean squared deviation. Once fixed a threshold $0 < \tau \leq 1$, we have put

$$z = 1 \qquad \text{for } z > \tau,$$

$$z = 0 \qquad \text{for } z \leq \tau.$$

In this way every winning configuration is represented by a binary number with as many 1's and 0's as many the competitive layer weights.

Then the task of comparing these binary numbers is straightforward. It has been verified that the τ threshold size is not critical: fixing it to 0.5 we have obtained the best results with any input stream.

The *z*-score method has shown to be steady with regard of the performances, and computationally not expensive, being linear in the number of the competitive layer weights.

But it is worth emphasizing that the *z*-score algorithm allows the network to reach its best performances in a very small number of epochs (often less then 15).

This allows the network to complete its work within a negligible time.

The good performances of this network have been tested for equalization and demodulation of GSM signals (Favalli 1996), for image classification (Pizzi 1998) and for EEG analysis (Pizzi 2002, Pizzi 2007).

2.5. Chaotic Attractors

Another interesting discriminant method adopted in this study was the analysis of attractors in the time series generated by the winning weights of our Artificial Neural Network.

An attractor can be defined as a generalization of the steady state point, and represents the trajectory in a portion of state space where a dynamical system is attracted to (Gleick 1987, Lorenz 1993).

In line with the considerations above mentioned on the possible meaning of the ECG chaotic analysis, we tried to highlight the presence of dynamical attractors in the ECG signals using MATLAB and its SIMULINK module for the dynamical systems simulation.

The simulation allowed us to represent the dynamical trajectories generated by the Artificial Neural Network applied to the the ECG signals, obtaining chaotic attractors that represent the time series in the state space.

3. RESULTS

3.1. Study Population

We retrospectively evaluated 50 patients (19 males, 76%; mean age 50 + 16 years, range 23-76 years) admitted to our Electrophysiology Laboratory from 2000 to 2003 for EPS owing to a clinical history of aborted sudden cardiac death, sustained VT (more than 30 sec in duration), requiring emergency room admission, syncope of unknown origin or non-sustained VT evidenced by resting ECG or Holter monitoring. No significant differences were seen between the two groups in terms of age, gender, structural heart disease or ejection fraction.

The 25 patients were divided into two groups according to the inducibility (13 patients, 52%) or non-inducibility (48%) of VT during EPS. The study group was then compared with a control group of 25 healthy subjects of comparable age and gender (mean age 48 + 12 years; 18 males), who did not undergo EPS.

3.2. Electrophysiological Study

Electrophysiological study was always performed off antiarrhythmic drugs. A French tetrapolar recording and stimulating catheter was inserted through the right femoral vein and positioned in the right ventricular apex. Programmed Ventricular Stiulation (PVS) was performed. EPS was considered positive if sustained monomorphic VT was induced.

3.3. Non-Linear Analysis of ECG Signals

ECG signals were recorded in the basal state in all subjects by means of the Polygraph Lab System Duo EP Laboratory (Bard Electrophysiology Division, USA), before EPS was started. We chose 20” ECG strips, in sinus rhythm;

The files were recorded and converted into ASCII format by means of a suitable C code developed by the group.

These data were finally analyzed by non-linear mathematical functions, as shortly presented in (Pizzi 2006) and explained in detail below.

Among the selected patients, 12 developed ventricular tachyarrhythmia during the study, 13 did not.

For each patient we sampled ECG signals and calculated on them the above mentioned variables: Recurrence, Determinism, Entropy, Ratio, Recurrence, Mutual Information.

On these variables we applied the hierarchical clustering technique with the Average Group Linkage method.

The following tables report the results of the clustering procedures.

Table 1. 3 classes clustering

N°	CLASS	DETERMINISM	RATIO		ENTROPY	RECURRENCE	M.I.
		mean	min	max	media	min	
H25	1	84,52	6,18	10,74	3,34	7,46	96
H26	1	69,19	5,01	9,97	2,76	6,71	56
H27	1	77,57	3,70	12,26	2,64	7,19	55
H32	1	79,31	2,75	6,49	3,76	13,09	65
H34	1	81,41	3,55	6,26	3,78	12,62	59
H35	2	59,84	2,06	12,05	2,29	2,95	70
H38	1	76,59	3,79	7,82	3,08	10,32	57
H43	1	66,77	3,18	5,82	3,69	12,26	75
H51	1	80,70	2,98	7,25	3,70	4	94
H54	1	70,60	2,96	7,85	3,49	9,86	74
H61	1	65,86	2,60	6,88	3,17	7,17	63
H65	1	86,47	2,78	6,35	3,91	1,23	92
H71	1	73,98	3,76	8,08	3,40	8,10	63
I5	3	26,43	-1	44,23	0,90	0	4
I12	3	34,69	-1	22,27	0,91	0	16
I13	3	32,46	-1	24,71	0,88	0	11
I16	3	58,40	-1	24,47	1,50	0	18
I17	3	23,76	-1	21,55	0,72	0	11
I18	3	40,36	-1	31,45	1,15	0	16
I20	3	31,14	-1	50,00	0,23	0	11
I23	3	36,11	-1	16,69	1,20	0	6
I39	3	40,51	-1	26,33	1,47	0	7
I41	3	55,54	-1	20,46	1,96	0	4
I42	3	1,30	-1	26,72	0,02	0	6
I64	3	48,80	-1	24,63	0,87	0	6
N01	2	54,70	0	6,37	3,19	0,51	80
N02	2	43,59	0	10,94	2,48	0,26	54

Table 1. (Continued)

N°	CLASS	DETERMINISM	RATIO		ENTROPY	RECURRENCE	M.I.
		mean	min	max	media	min	
N03	2	44,95	0	8,01	2,60	2,87	78
N04	1	71,94	2	3,86	4,00	18,38	66
N05	2	53,27	2,16	10,81	2,74	4,55	62
N06	2	50,13	0	4,18	3,23	1,33	70
N07	2	63,49	0	5,13	3,53	1,52	88
N08	2	42,88	0	6,29	2,43	0,83	59
N09	1	60,72	2,78	5,45	3,41	11,92	69
N10	2	53,17	0	5,78	3,03	1,82	58
N12	1	52,86	2,77	8,66	2,85	7,21	61
N13	2	34,23	0	10,00	1,89	2,67	62
N15	2	49,32	0	6,61	2,61	2,14	57
N16	2	23,88	0	7,99	1,44	0,65	78
N17	2	55,62	0	7,59	3,13	1,17	84
N18	2	37,97	0	8,30	2,22	1,35	71
N19	2	60,25	0	7,26	3,12	2,10	63
N21	2	13,30	2,10	6,68	3,65	3,80	73
N24	1	70,82	2,48	6,59	3,88	12,38	79
N25	1	66,16	2,56	8,43	3,44	9,82	68
N26	2	59,93	0	8,04	3,21	0,79	73
N28	2	17,76	0	7,69	1,23	1,37	57
N29	2	61,81	2,82	8,21	3,18	2,44	65
N30	2	54,44	0	6,54	3,27	0,91	89
N31	2	60,86	0	5,19	3,62	0,61	66

Table 2. 2 classes clustering

N°	CLASS	DETERMINISM	RATIO		ENTROPY	RECURRENCE	M.I.
		mean	min	max	mean	min	
H25	1	84,52	6,18	10,74	3,34	7,46	96
H26	1	69,19	5,01	9,97	2,76	6,71	56
H27	1	77,57	3,70	12,26	2,64	7,19	55
H32	1	79,31	2,75	6,49	3,76	13,09	65
H34	1	81,41	3,55	6,26	3,78	12,62	59
H35	1	59,84	2,06	12,05	2,29	2,95	70
H38	1	76,59	3,79	7,82	3,08	10,32	57
H43	1	66,77	3,18	5,82	3,69	12,26	75
H51	1	80,70	2,98	7,25	3,70	4	94
H54	1	70,60	2,96	7,85	3,49	9,86	74
H61	1	65,86	2,60	6,88	3,17	7,17	63

N°	CLASS	DETERMINISM	RATIO		ENTROPY	RECURRENCE	M.I.
		mean	min	max	mean	min	
H65	1	86,47	2,78	6,35	3,91	1,23	92
H71	1	73,98	3,76	8,08	3,40	8,10	63
I5	2	26,43	-1	44,23	0,90	0	4
I12	2	34,69	-1	22,27	0,91	0	16
I13	2	32,46	-1	24,71	0,88	0	11
I16	2	58,40	-1	24,47	1,50	0	18
I17	2	23,76	-1	21,55	0,72	0	11
I18	2	40,36	-1	31,45	1,15	0	16
I20	2	31,14	-1	50,00	0,23	0	11
I23	2	36,11	-1	16,69	1,20	0	6
I39	2	40,51	-1	26,33	1,47	0	7
I41	2	55,54	-1	20,46	1,96	0	4
I42	2	1,30	-1	26,72	0,02	0	6
I64	2	48,80	-1	24,63	0,87	0	6
N01	1	54,70	0	6,37	3,19	0,51	80
N02	1	43,59	0	10,94	2,48	0,26	54
N03	1	44,95	0	8,01	2,60	2,87	78
N04	1	71,94	2	3,86	4,00	18,38	66
N05	1	53,27	2,16	10,81	2,74	4,55	62
N06	1	50,13	0	4,18	3,23	1,33	70
N07	1	63,49	0	5,13	3,53	1,52	88
N08	1	42,88	0	6,29	2,43	0,83	59
N09	1	60,72	2,78	5,45	3,41	11,92	69
N10	1	53,17	0	5,78	3,03	1,82	58
N11	1	52,86	2,77	8,66	2,85	7,21	61
N12	1	34,23	0	10,00	1,89	2,67	62
N13	1	49,32	0	6,61	2,61	2,14	57
N14	1	23,88	0	7,99	1,44	0,65	78
N15	1	55,62	0	7,59	3,13	1,17	84
N16	1	37,97	0	8,30	2,22	1,35	71
N17	1	60,25	0	7,26	3,12	2,10	63
N18	1	13,30	2,10	6,68	3,65	3,80	73
N19	1	70,82	2,48	6,59	3,88	12,38	79
N20	1	66,16	2,56	8,43	3,44	9,82	68
N21	1	59,93	0	8,04	3,21	0,79	73
N22	1	17,76	0	7,69	1,23	1,37	57
N23	1	61,81	2,82	8,21	3,18	2,44	65
N24	1	54,44	0	6,54	3,27	0,91	89
N25	1	60,86	0	5,19	3,62	0,61	66

The analysis shows that using the 3-classes clustering the procedure works fairly well but fails to classify some patients in the correct class.

A more interesting result is reached using a 2-classes procedure, where all the non-induced arrhythmic patients have been correctly included into the class of the healthy subjects, whereas the induced patients have been isolated into a separated cluster.

Another analysis, performed directly on the graphical appearence of the RPs, shows other interesting features of the ECG signals of the examined patients.

As described above, in the RPs random signals give rise to a uniform distribution of colors, whereas the more deterministic the signal, the more structured the plot.

Figure 2 shows the plots of healthy subjects. A massive presence of hot colors (red, yellow, orange) denotes small distances between vectors. The large bands of colors denote a chaotic behavior.

The plots of the non-induced patients (Figure 3) show a more uniform distribution of colors, index of the presence of random features inside the ECG signal , but some bands are still present under the background.

Finally (Figure 4), the plots of the induced patients are heavily characterized by a quite regular distribution of colors, synonymous of deterministic signals, i.e. signals that are more susceptible to be forecast than the first two groups.

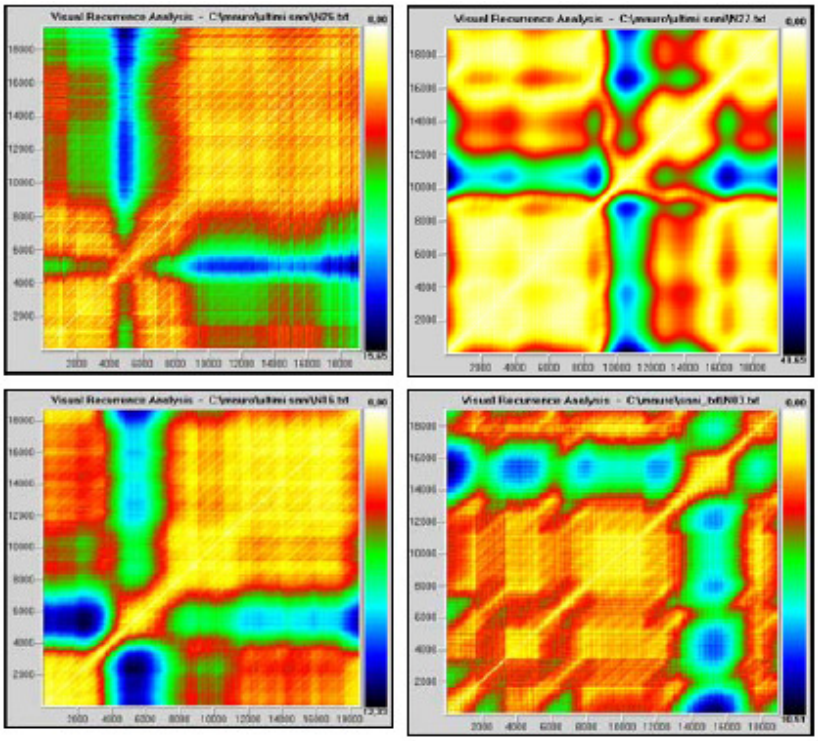

Figure 2. Recurrent Plots of healthy subjects.

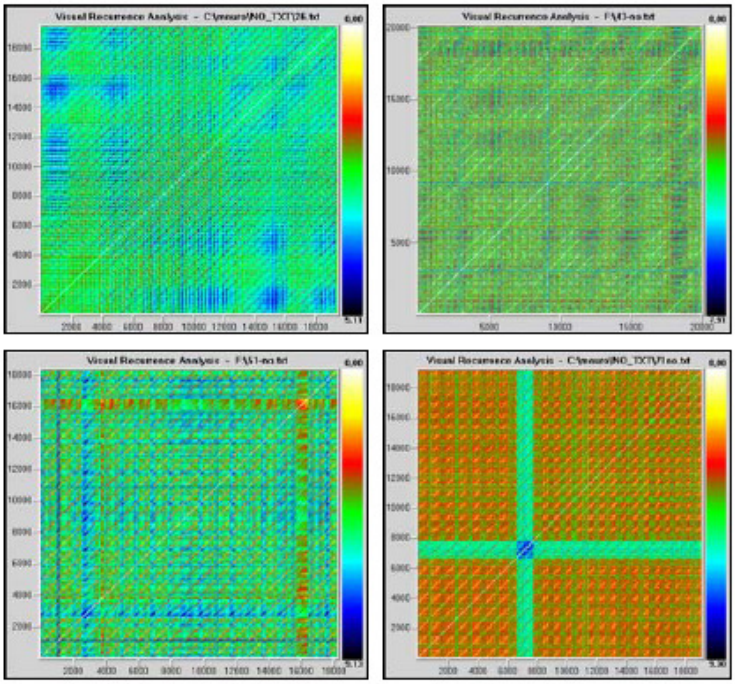

Figure 3. Recurrent Plots of non-induced patients.

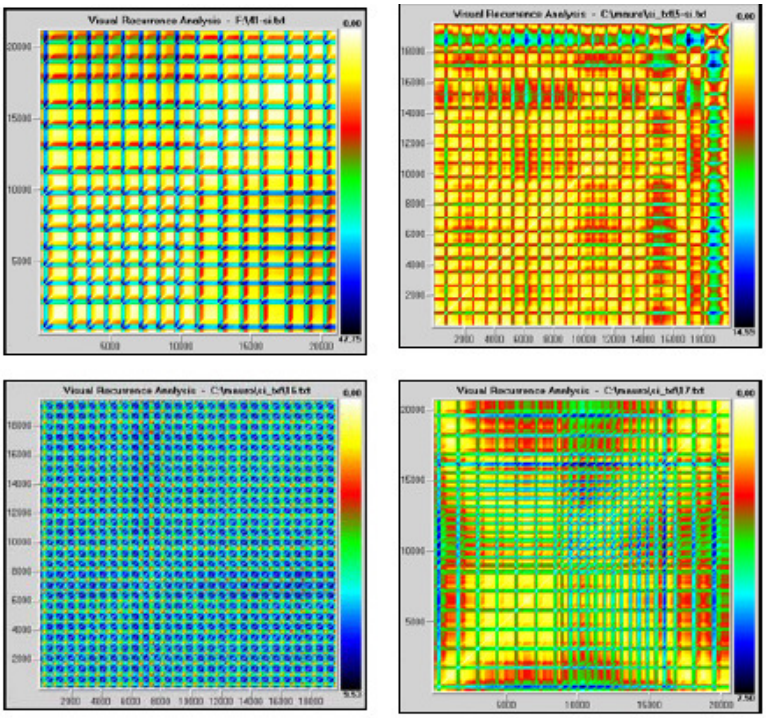

Figure 4. Recurrent Plots of induced patients.

3.4. Artificial Neural Networks Results

As the network yields an output sequence of winning neurons, we looked for the z-score codes generated by the sequences of winning neurons for different set of patients. We implemented an ITSOM network written in C that automatically changes its parameters. In particular we changed gradually the following parameters: number of input neurons, learning rate, number of epochs.

After a running test set we identified the best configuration as:

- number of input neurons: 500
- learning rate: 0.03
- number of epochs: 100.

We generated the binary z-scores corresponding to the three sets of patients with different competitive layer units: 12 and 15, obtaining the following results (Figures 5 and 6):

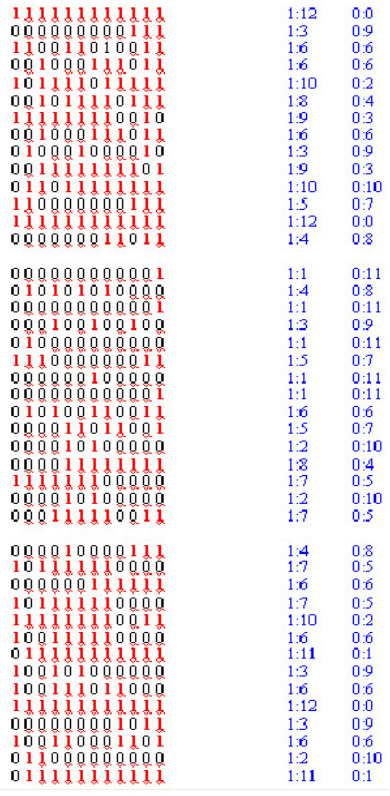

Figure 5. z-score codes with 12 competitive neurons (first group: healthy subjects; second group: induced subjects; third group: non-induced subjects).

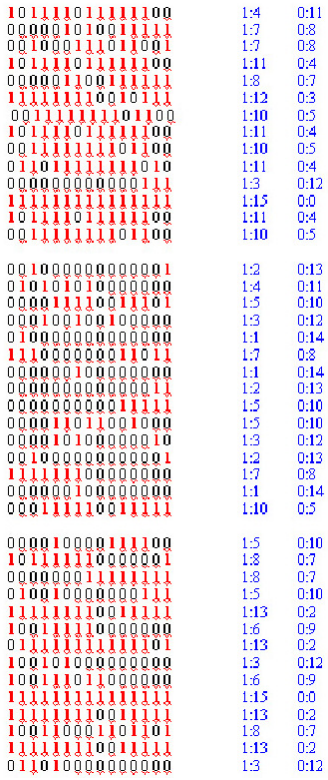

Figure 6. z-score codes with 15 competitive neurons (first group: healthy subjects; second group: induced subjects; third group: non-induced subjects).

Analysing these sequences we found out that the number of ones are definitely prevalent in the healthy and non-induced groups, indicating a tendency of their ECG signals to win in

the highest value neurons of the competitive layer, differently from the induced patients' signals , that tend to win on the lowest value neurons. On the basis of this observation, few lines of C code allowed us to automatically discriminate the patients by counting the zeros of the z-scores generated by their ECG signals.

This result is once again in agreement with the results drawn by the clustering and the recurrent plots analysis. The same agreement was reached by analysing the dynamic attractors generated by the ITSOM network.

3.5. Chaotic Attractors

In the following figures we show some examples of the chaotic attractors plots drawn from the three sets of winning weights of the Artificial Neural Networks applied to the ECG signals (induced, non induced, healthy). Our analysis pointed out that each set exhibits common features. The main differences between plots have been noted between healthy and induced patients. The subjects belonging to the set of non induced patients associate features common to both the induced and the healthy groups, but follow in general a behavior more similar to the healthy group (Figures 7,8,9).

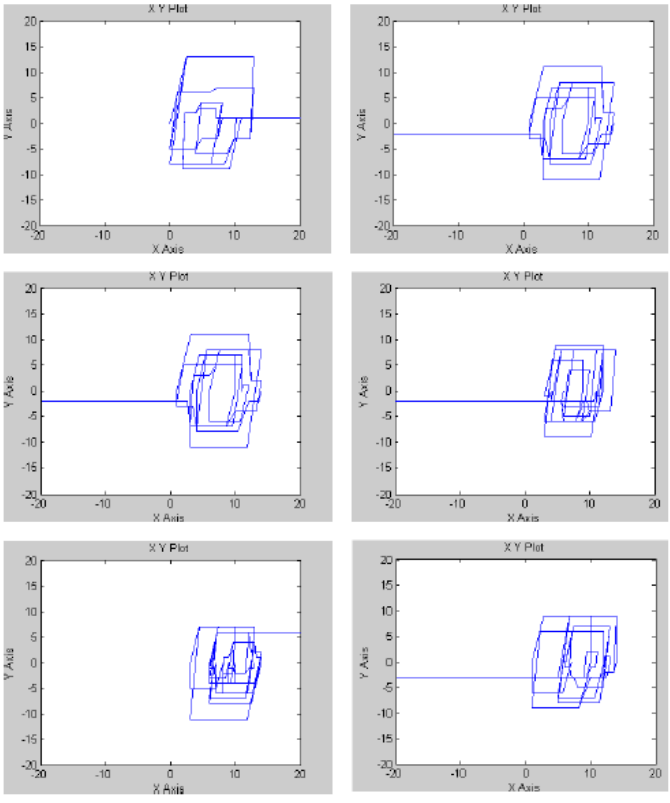

Figure 7. Attractors of healthy subjects.

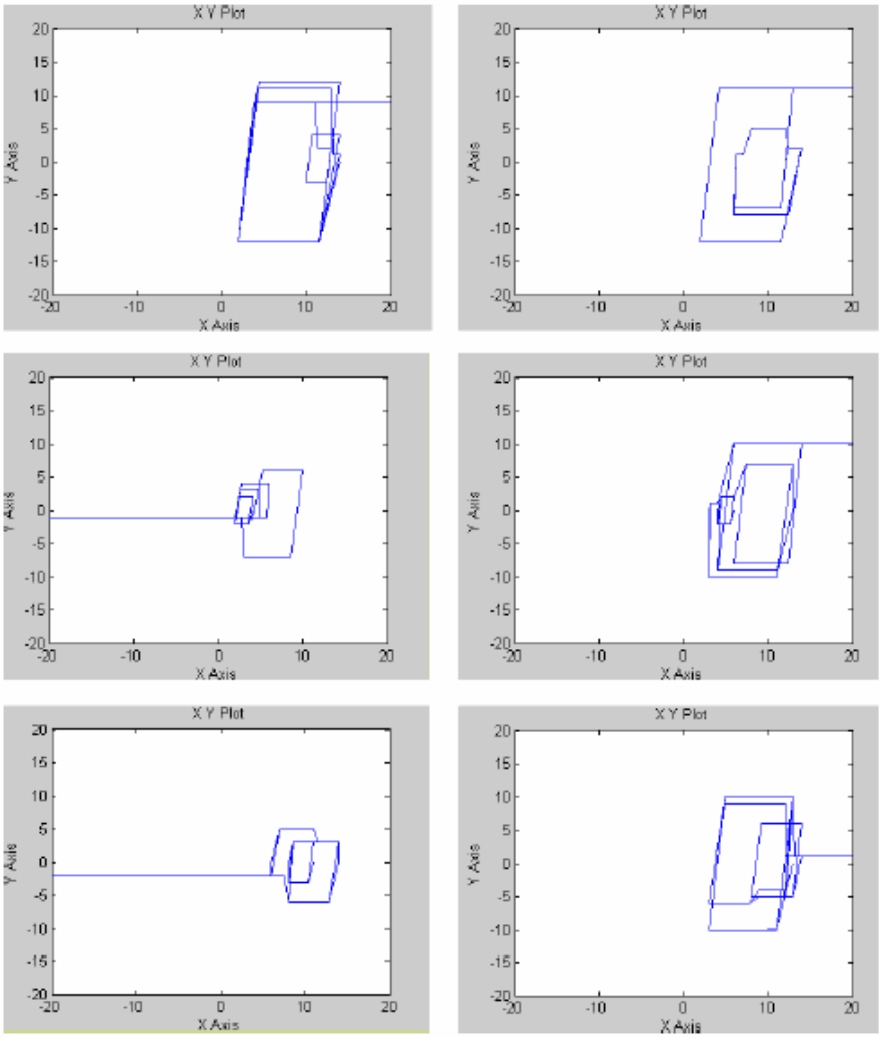

Figure 8, Attractors of induced subjects.

The healthy group attractors occupy a wider area compared with the induced group ones, with an evident chaotic behavior.

The plot of induced patients occupy an area smaller than the healthy subjects' plots, and but show a more homogeneous and regular graph.

As regards the non-induced group, we noted that they tend to occupy a wide area as the healthy one, but, differently from them, present a more homogeneous graph. We could also ascertain that the non-induced ECG signals share features similar to the healthy subjects' ones, with some features common to the induced patients.

These findings agree with the above mentioned researches that claim an increase of order and regularity in pathological ECG dynamical systems, and a more chaotic behavior in healthy ECG signals.

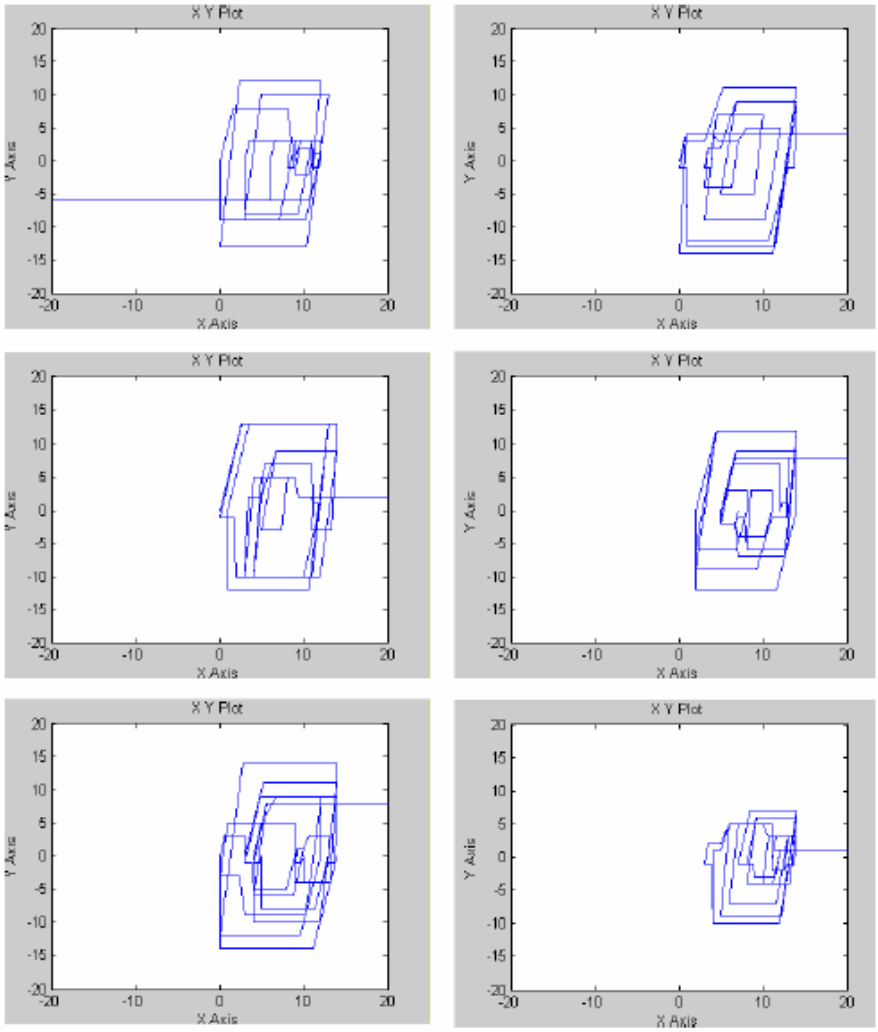

Figure 9. Attractors of non-induced subjects.

3.6. Statistical Analysis

Continuous variables were expressed as mean value + standard deviation and were compared by means of an unpaired t test. Categorical variables were compared by means of the Fisher exact test or Chisquare test with Yates' correction for continuity, where appropriate. Correlation between the results of EPS and those of non-linear analysis of ECG signals was analysed by means of logistic regression analysis. Furthermore, the negative and positive predictive value of non-linear analysis of ECG signals was calculated in comparison with the result of EPS.

When the Hierarchical Clustering Technique was applied to these numerical values in accordance with the Average Group Linkage Method, it showed that all patients with negative EPS had been assigned to the class of healthy subjects, whereas the patients in whom VT was

inducible had been correctly and clearly isolated into a separate cluster (Table 2). In our study, the result of non-linear analysis of ECG signals with application of the clustering technique was therefore significantly correlated to that of EPS ($p<0.001$), and was able to predict the result of EPS, with a negative predictive value of 100% and a positive predictive value of 100%.

4. DISCUSSION AND CONCLUSIONS

In EPS-positive patients, the functions produced clearly different results from those obtained in EPS-negative individuals and in normal control subjects; correspondence with the results of the invasive study was therefore excellent. A further interesting finding was that this correlation was totally independent of the underlying pathology. This raises intriguing questions with regard to the possible biological variables involved in the modification of these parameters.

It is also interesting that the results of all the functions applied were concordant, even though those of some functions (mutual information and entropy) were particularly marked. It therefore seems reasonable to claim that it is the entire set of functions which contributes to defining the dynamic features of the ECG signal in the subjects examined and which enables those with the greatest risk to be picked out from the rest of the group by means of the clustering procedure. These observations (in particular the sharply lower entropy values in the induced patients' signals with respect to the non-induces ones) seem to point to an increased regularity of the dynamic organisation of the electrocardiographic signal in those very patients who are most compromised, a finding which is in line the above mentioned studies carried out with different non-linear functions from those used in this study.

By contrast, the clustering procedure utilised in the present study appears to be totally different and innovative. Another undeniable advantage of the system of processing and calculation that we adopted is that it enables relatively small amounts of information to be used; indeed, using ECG strips of a duration of no more than 20 seconds improves the quality and control of the conditions in which the signal is recorded. This also has favourable implications for the feasibility of application in everyday clinical practice, an aspect that should not be overlooked in view of the fact that other methods, such as the calculation of Correlation Dimension, which requires a very broad range of data and therefore prolonged ECG recordings, are difficult to reconcile with the need to maintain acceptable stability of the system.

Our study applied a clustering procedure to the results obtained through the use of a set of non-linear calculation functions. This enabled a group of patients at particularly high arrhythmic risk according to both clinical and electrophysiological criteria to be clearly picked out from lower-risk patients and normal control subjects; moreover good agreement was seen between the results of the non-invasive evaluation and those of the invasive test.

The clustering procedure applied to a set of non-linear functions is also supported by the results obtained with the Recurrent Plots .

Besides, the use of a self-organizing artificial network, the ITSOM (another non-linear method) lead to a way to discriminate the patients on the basis of the chaotic attractors determinated by the sequences of winning neurons.

The artificial neural network splits the patients into 3 groups but maintains similarity between healthy and non-induced patients. The graphical representation of the dynamical attractors generated by the network shows an evident chaoticity in the healthy subjcet's time series and more regular patterns in induced patients. The non-induced patients share features both from the healthy and the induced patients.

Thus it is hypothesizable to use the inductive properties of the ITSOM network to recognize the new patient's clinical features comparing them with those collected by the present study. The ITSOM low computational load and the z-score comparison procedure make this task a real-time procedure that could be achieved during a daily clinical practice.

The results of this study are very encouraging and show that non-linear analysis can be furtherly studied as a tool to evaluate the complex dynamics of the heart by transforming qualitative diagnostic criteria into a quantitative problem.

The adopted model presents the evident advantage to avoid the risks and the uneasiness of the invasive methods, and moreover it can avoid the costs of the equipments needed for the current non-invasive prognostic procedures.

However, because it was retrospective and based on a somewhat small sample, the study cannot claim to be more than a preliminary experience.

Unfortunately, the gold standard in the prognostic stratification of patients at risk of life-threatening events remains follow-up alone, as no test, or combination of tests, can as yet provide a reliable point of reference. Nevertheless, we feel that our findings warrant particular attention, not least because they may pave the way both for prospective studies involving larger numbers of cases and for follow-up evaluations in the medium-long term.

ACKNOWLEDGEMENTS

This work could be realized thanks to the contribution of Dr. M. Panzeri, Dr. A. Bussi Roncalini, Dr. C. Patrini and Dr. V. Zitun, and to the encouragement of Prof. G. Degli Antoni (University of Milan).

REFERENCES

Anderberg M.R. (1973). *Cluster Analysis for Application*. New York: Academic Press.

Anderson J.L., Horne B.D. (2005). Non-linear Heart Rate Variability: a better ECG predictor of cardiovascular risk?, *J. Cardiovasc. Electrophysiol.* ,16 , 21-23.

Atmanspacher H., Kurths J. (1992). Complexity and Meaning in Nonlinear Dynamical Systems, in: *Open Systems and Information Dynamics*, 1, 269-289.

Bezerianos A, Bountis T, Papaioannou G, Polydoropoulos P. (1995) . Non linear time series analysis of electrocardiograms. *Chaos* 5, 95–101.

Braun C., Kowallik P., Freking A., Hadaler D., Kniefky K., Meesmann M. (1998). Demonstration of Non-Linear Components in Heart Rate Variability of Healthy Persons. *Am. J. Physiol*. 275, 1577-1584.

Buxton A.E., Leek L., Fisher J.D., et al. (1999). A randomized study of the prevention of sudden death in patients with coronary artery disease. Multicenter Unsustained Tachycardia Trial (MUSTT) Investigators. *N. Engl. J. Med*., 341, 1882-1890.

Fraley C. (1996). Algorithms for Model-Based Gaussian Hierarchical Clustering. Siam *Journal on Scientific Computing*, 20.

Clayton R.H., Murray A., Campbell R.W. (1995) . Objective features of the surface electrocardiogram during ventricular tachyarrhythmias. *Eur Heart J.* , 16(8), 1115-19.

Day W.H. E., Edelsbrunner, H. (1984).Efficient Algorithms for Agglomerative Hierarchical Clustering Methods. *Journal of Classification*, 1(1), 7-24.

Duran B.S. , Odell P.L. (1974). *Cluster Analysis: A Survey. Lecture Notes in Economics and Mathematical Systems*, 100, Berlin: Springer-Verlag Berlin .

Eckmann J.P., Kamphorst S.O., Ruelle D. (1987). Recurrence plots of dynamical systems. *Europhys. Lett.* 4:973-977.

Ermentrout B. (1992). Complex Dynamics in WTA Neural Networks with slow Inhibition. *Neural. Networks* 5, 415-431.

Favalli L., Pizzi R., Mecocci A. (1997). Non-linear mobiel radio channel estimation using neural networks, Crete: *Proc. DSP97, IEEE ed, Int. Conf. on Digital Signal Processing*, 287-290.

Furlanello F., Galanti G., Manetti P., et al. (2004). Microvolt T-wave alternans as predictor of electrophysiological testing results in professional competitive athletes. *A.N.E.*, 9, 1-6.

Glass L. (1999).Chaos and Heart Rate variability. *J. Cardiovasc Electrophysiol*, 10, 1358-1360.

Glass L., Mackey M.C. (1988). *From clocks to chaos: the rhythms of life*. Princeton University Press.

Gleick J. (1987). Chaos, New York:Viking .

Goldberger A. L. (1996). Non-linear dynamics for clinicians: chaos theory, fractals, and complexity at the bedside. *Lancet* , 347,1312-1314.

Goldberger A.L., Rigney D.R., West B.J. (1990). Chaos and fractals in human physiology. *Sci. Am*,. 262, 42-49.

Grassberger P. and Procaccia I. (1983). Measuring of Strangeness of Strange Attractors. *Physica* D, 9, 189.

Hausdorff, F. (1919). Dimension und ausseres. *Mass. Math. Ann*., 29.

Huikuri H.V. , Makikallio T.H. , Perkiomaki J. (2003). Measurement of Heart Rate Variability by Methods Based on Nonlinear Dynamics. *Journal of Electrocardiology*, *Suppl.* 2003, 36.

Huikuri H.V., Makikallio T.H., Peng C.K., Goldberger A.L., Hintze U., Moller M. (2000) . Fractal correlation properties of R-R internal dynamics and mortality in patients with depressed left ventricular function after an Acute Myocardial Infarction. *Circulation*,101, 47-53.

Huikuri H.V., Makikallio T.H., Perkiomaki J. (2003). Measurement of Heart Rate Variability by methods based on Non-Linear Dynamics. *J. Electrocardiol*., 36(Suppl), 95-99.

Jackson E. (1989). *Perspectives in Nonlinear Dynamics* I,II. Cambridge Un. Press..

John, A. Hartigan J.A. (1975). Clustering Algorithms. New York: John Wiley and Sons.

Kaplan D., Glass L. (1995). Understanding Nonlinear Dynamics. New York: Springer-Verlag.

Kohonen T (1983). *Self-Organisation and Association Memory*. New York: Springer Verlag.

Kohonen T. (1993). Physiological Interpretation of the Self-Organizing Map Algorithm. *Neural Networks*, 6,895-905.

Kononov E. (1996). URL: http://home.netcom.com/~eugenek/download.html.

Kroll M.W., Fulton K.W. (1992). Slope Filtered Pointwise Correlation Dimension algorithm and its evaluation with prefibrillation heart rate data. *J. Electrocardiol*., 24(Supp), 97-101.

Lombardi F. (2000). Chaos theory, heart rate variability, and arrhythmic mortality. *Circulation*, 101(11), 8-10.

Lombardi F., Sandrone G., Mortasa A., et al. Linear and non-linear dynamics of Heart rate (1996).Variability after Acute Myocardial Infarction with normal and reduced left ventricular Ejection Fraction. *Am. J. Cardiol*., 77,1283-1288.

Lorenz E. (1993). *The Essence of Chaos*. London :UCL Press.

Makikallio T.H., Hoiber S., Kober L., Torp-Pedersen C., Peng C.K., Goldberger A.L., Huikuri H.V. (1999). Fractal analysis of heart rate dynamics as a predictor of mortality in patients with depressed left ventricular function after Acute Myocardial Infarction. *Am. J. Cardiol*. , 83, 836-839.

Makikallio T.H., Ristimae T., Airaksinen K.E., Peng C.K., Goldberger A.L., Huikuri H.V. (1998) . Heart rate dynamics in patients with Stable Angina Pectoris and utility of Fractal and Complexity Measures. *Am. J. Cardiol*., 81, 27-31.

Makikallio T.H., Seppanen T., Airaksinen K.E., et al. (1997) . Dynamic analysis of heart rate may predict subsequent Ventricular Tachycardia after Myocardial Infarction. *Am. J. Cardiol*., 80, 779-783.

Mandelbrot B. (1983). *The Fractal Geometry of Nature*. New York: Freedman and Co.

May R.M. (1976). Simple mathematical models with very complicated dynamical behavior. *Nature,* 261, 459-467.

McLaughlin M.G., Zimetbaum P.J. (2006). Electrocardiographic Predictors of Arrhythmic Death. *A.N.E*., 11,327-337.

Mohamed I., Abou-Zied A.H., Abou-Bakr M. Y., Kadah Y.M. (2002). Study of Features Based on Nonlinear Dynamical Modeling in ECG Arrhythmia Detection and Classification. *IEEE Trans on Biomed. Eng.*, 49 (7),733 – 736.

Moss A.J., Fadl Y., Zareba W., Cannom D.S., Hall W.J. for the Multicenter Automatic Defibrillator Implantation Trial (MADIT) Research Group (2001). Survival benefit with an implanted defibrillator in relation to mortality risk in chronic coronary heart disease. *Am. J. Cardiol*.,88 (5), 516-520.

Myerburg R.J., Kessler K.M., Castellanos A. (1993). Sudden Cardiac Death: epidemiology, transient risk and intervention assessment. *Ann. Intern. Med*.,119, 1187-1197.

Olson C.F.(1993) . Parallel Algorithms for Hierarchical Clustering. Technical Report UCB//CSD-93-786, University of California at Berkeley.

Otero-Siliceo E., Arriada-Mendicoa N. (2003), Is it Healthy to Be Chaotic ?. *Medical. Hypotheses,* 60(2), 233-236.

Otsuka K., Cornelissen G., Halberg F. (1997). Circadian Rhythmic Fractal Scaling of Heart Rate Variability in Health and Coronary Artery Disease, *Clin. Cardiol*., 20, 631-638.

Owis M.I. , Abou-Zied A.H. , Youssef A.M., Kadah Y.M. (2002). Study of Features Based on Nonlinear Dynamical Modeling in ECG Arrhythmia Detection and Classification. *IEEE Trans. on Biomedical Engineering*, 49(7), 733-742.

Pizzi R., Cino G., Gelain F., Rossetti D. , Vescovi A.(2007). Learning in Human Neural Networks on Microelectrode Arrays. *Biosystems Journal*, 88(1-2) 1-15.

Pizzi R., Inama G., Durin O., Pedrinazzi C., (2006). Non-linear Analysis of ECG signals as a method for the prognosis of letal tachyarrhythmias. Cairo: *Proc. IV Int. Conf. on Information and Communication Technologies.*

Pizzi R., Sicurello F., Varini G. (1998). Development of an inductive self-organizing network for the real-time segmentation of diagnostic images. Pisa, Italy: *3th International Conference of Neural Networks and Expert Systems in Medicine and Healthcare* , 44-50.

Pizzi, R., de Curtis M., Dickson C. (2003). The Binding Problem: Evidence of Chaotic Attractors in Cortical Fast Oscillations Tested by an Artificial Neural Network. in J. Kacprzyk (Ed.), *Advances in Soft Computing*, Physica Springer Verlag.

Ravelli F., Antolini R. (1992). Complex dynamics underlying the human electrocardiogram. *Biol. Cybern*, 67(1), 57-65.

Ritter H., Schulten K. (1986). On the Stationary State of Kohonen's Self- Organizing Sensory Mapping, *Biological Cybernetics*, 54, 99-106.

Ritter H., Schulten K. (1988). Convergence properties of Kohonen's Topology Conserving Maps : Fluctuations, Stability, and Dimension Selection. *Biological Cybernetics* 60, 59-71.

Signorini M.G., Cerutti S., Guzzetti S., Parola R. (1994). Non-Linear Dynamics of Cardiovascular Variability Signals. *Methods of Information In Medicine*, 33, 81-84.

Rohlf F.J. (1973) . Hierarchical Clustering using the Minimum Spanning Tree. *Computer Journal* 16, 93-95.

Stein P.K, Reddy A. (2005). Non-linear heart rate variability and risk stratification in cardiovascular disease. *Ind. Pacing and Electrophysiol. J.* , 5, 210-220.

Stein P.K., Domitrovich P.P. (2005). Traditional and non-linear Heart Rate Variability are each independently associated with mortality after Myocardial Infarction. *J. Cardiovasc. Electrophysiol.*,16,13-20.

Stewart I.: (1989). Does God Play Dice? : The Mathematics of Chaos. London: B. Blackwell.

Vigo D.E., Siri L.N., Ladron de Guevara M.S., et al. (2004). Relation of depression to heart rate non-linear dynamics in Patients > 60 years of age with recent Unstable Angina Pectoris or Acute Myocardial Infarction. *Am. J. Cardiol*., 93, 756-760.

Wu Z.K., Vikman S., Laurikka J., Pehkonen E., Iivainen T., Huikuri H.V., Tarkka M.R. (2005) . Non-linear Heart Rate Variability in CABG patients and the preconditioning effect. *Eur. J. Cardiothorac. Surg*. 28,109-113.

Zbilut J.P. ,Webber C.L. (1992). Embeddings and delays as derived from quantification of recurrent plot. *Phys. Lett*., 171.

Zbilut J.P., Thomasson N., Webber C.L. Jr. (2002). Recurrence quantification analysis as a tool for nonlinear exploration of nonstationary cardiac signals. *Medical Engineering and Physics* 24 , 53-60.

Zipes D.P., Camm A.J., Borggrefe M., et al. (2006). ACC/AHA/ESC 2006 guidelines for management of patients with ventricular arrhythmias and the prevention of sudden cardiac death – executive summary. *Eur. Heart J.* , 27, 2099- 2140.