ISSN 0006-3509, Biophysics, 2006, Vol. 51, No. 1, pp. 115–119. © Pleiades Publishing, Inc., 2006.
Original Russian Text © O.Yu. Mironyuk, A.Yu. Loskutov, 2006, published in Biofizika, 2006, Vol. 51, No. 1, pp. 144–150.

COMPLEX SYSTEMS
BIOPHYSICS

Detection of Cardiac Pathologies Using Dimensional Characteristics of RR Intervals in Electrocardiograms

O. Yu. Mironyuk and A. Yu. Loskutov

Physical Faculty, Moscow State University, Vorob'evy gory, Moscow, 119992 Russia

Received April 26, 2005

Abstract—The possibility of detecting pathologies in patients with various types of heart failure by analyzing the correlation dimension and embedding dimension of RR intervals in electrocardiograms is estimated. Limitations of the proposed approach and methods of overcoming them are discussed. It is demonstrated that these methods are suitable for provisional diagnosis.

Key words: electrocardiogram, time series, correlation dimension

DOI: 10.1134/S0006350906010179

INTRODUCTION

Physiological characteristics of heart activity in various pathologies are intensely studied by biophysicists, biochemists, medical physicists, etc. The majority of theoretical approaches are based on the analysis of electrocardiograms (ECGs), because they have many advantages over other methods (they are simple, noninvasive, highly informative, etc.). In addition, any change in the myocardium immediately affects heart rate [1] and, hence, is reflected in an ECG. In turn, any heart disorder may lead, for the body as a whole, to severe consequences from which it is impossible to recover.

Myocardial disorders caused by various factors have been studied for decades. However, clinicians routinely use only elementary approaches to analyze patients' cardiac rhythms. Only recently have new methods based on the dynamic system theory attracted considerable attention [2–8]. The purpose of these studies is to determine the dependence of the dynamic characteristics of ECG time series on various physiological changes in the heart tissues and study the possibility of using them as clinically important parameters.

In medical practice, a normal heart rate is often referred to as a normal sinus rhythm. Despite this term, it is well known that intervals between heartbeats fluctuate. Even in healthy subjects at rest, heart rate shows considerable variations in frequency and amplitude, because the heart must permanently respond to all physiological processes in the body. Moreover, exactly periodic patterns of heart rate indicate a severe pathology of the heart [9].

In the late 1980s, a hypothesis was put forward [10, 11] that *deterministic chaos* was inherent in a normal human heart rate. Subsequent analysis demonstrated that affected and healthy subjects substantially differed from each other in the quantitative parameters of the chaos [6, 7] and that the original assumption that elements of dynamic chaos were present in the cardiac rhythm was acceptable. Therefore, the role of chaos in the pathogenesis of heart diseases is now the subject of numerous studies.

Here, we consider the relationship between the dimensions of the time series of ECG RR intervals and the complexity of cardiac rhythm. We analyze the correlation dimension of series derived from the ECGs of a *sufficiently large group* of patients with different types of heart failure, including angina pectoris, atrioventricular block, and consequences of myocardial infarction. Thus, the purpose of our study was to obtain evidence that patients with different pathologies differ from one another in ECG correlation dimensions. This, in turn, may help in provisional diagnosis by the methods proposed.

REPRESENTATION OF AN ECG AS A SERIES OF RR INTERVALS

An ECG is a curve showing the changes in the potential difference during consecutive myocardium contraction cycles. It can be used to detect various cardiac rhythm disorders. Figure 1 shows a typical ECG of a healthy human. Every peak (positive or negative) corresponds to the excitation or repolarization of different

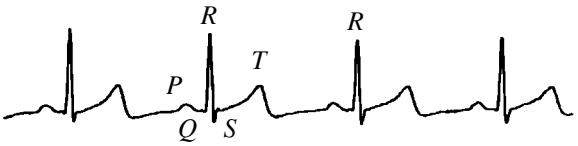

Fig. 1. A typical ECG.

Abbreviations: ECG, electrocardiogram; HRV, heart rate variability.

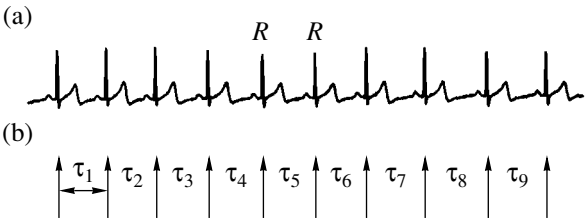

Fig. 2. Representation of ECG as a time series.

regions of the heart. The interval between the peaks of neighboring R waves of an ECG is equal to the duration of a complete cardiac cycle; it is referred to as the RR interval. Analysis of their sequence makes it possible to detect cardiovascular pathologies and diagnose individual diseases.

To study heart rate variability (HRV), an ECG is represented as a time series. In the general case, a time series is an array of N numbers that are values of a dynamic variable $x(t)$ at certain intervals of time. However, two methods of the construction of a time series are used for ECG analysis: the classic method [12–14], when the variable is fixed at equal intervals of time, and a series of RR intervals obtained by measuring distances between R peaks in the ECG [15–17]. In this study, we used the latter variant, consisting of the following steps.

Each R wave appearing at the moment of contraction t_i is substituted by a single pulse (Fig. 2), which is approximated by Dirac's δ function, $\delta(t - t_i)$. After that, the entire ECG is replaced by a series of RR intervals:

$$x(t) = \sum_i \delta(t - t_i). \quad (1)$$

Thus, the desired time series V_i is formed by the intervals between individual pulses, RR(i):

$$\mathrm{RR}_i = V(i), \quad V = (V(1), V(2), \ldots, V(N)), \quad (2)$$

where N is the total number of elements in the series. The formation of a time series as a sequence of the duration of intervals between peaks is widely used for analyzing biological systems where threshold values of variables are repeated according to a specific pattern. Although this representation is not a classical series of values recorded at equal time intervals, it still may serve as an actualization of the original nonlinear system (i.e., characterizing cardiac activity) [15].

This approach makes it possible not only to study the details of hidden consistent patterns, but also to reveal certain characteristics of chaotic processes inherent in the original series, e.g., to estimate the embedding dimension [17].

THE CORRELATION DIMENSION OF AN ECG

A heart tissue pathology causes a change in an ECG. These changes are clearly seen when quantitative parameters of the attractor reflecting the complexity of the behavior of the whole system, such as the correlation dimension and embedding dimension, have been calculated.

Correlation dimension. The correlation dimension as a characteristic of the attractor is easy to calculate on the basis of the time series composed of the ECG RR interval durations by the method suggested in [8]. According to this method, the reconstruction of the phase space and the restoration of the chaotic attractor of the system are reduced to the construction of a pseudoattractor, where the measured series itself, taken with a time delay, serves as the set of components of the vector.

Let m be the embedding dimension, i.e., the smallest dimension of the phase space containing the entire attractor of the dynamic system. Then, the original time series (2) can be used to obtain the "restored" attractor, which is constructed from the following vectors:

$$\begin{aligned} x_1 &= (V(1), V(1+\tau), \ldots, V(1+(m-1)\tau)), \\ x_2 &= (V(2), V(2+\tau), \ldots, V(2+(m-1)\tau)), \quad (3) \\ &\ldots \\ x_n &= (V(n), V(n+\tau), \ldots, V(n+(m-1)\tau)), \end{aligned}$$

where $n = N - (m-1)\tau$ is the number of vectors; τ is the delay time, i.e., the time interval at which the variable V is measured.

For our analysis of ECG, the components of these vectors should be taken in the form of a finite number of RR interval lengths. This means that the delay time τ in Eq. (3) is not constant; it varies for different components. As noted above, although this approach differs from the classic approach with a fixed sampling step, it permits characterizing the attractor of the original system and unambiguously identifying the state of this system. In addition, this representation is much better for analysis of most physiological processes.

Then the correlation integral for the series of correlation vectors (3) should be estimated:

$$C(r) = \lim_{N\to\infty} \frac{1}{N^2} \sum_{ij} H[r - |x_I - x_j|], \quad (4)$$

where $H(\alpha)$ is Heaviside's step function

$$H(\alpha) = \begin{cases} 0, & \text{if} \quad \alpha < 0 \\ 1, & \text{if} \quad \alpha \geq 0 \end{cases} \quad (5)$$

and $|x_i - x_j|$ is the distance between vector components in an m-dimensional phase space. Note that this distance can be determined by several methods [14].

Which of them will be used depends on the specific problem to be solved.

The correlation dimension is determined as the tangent of the slope of the linear segment of correlation integral on a logarithmic scale. Thus, the expression for the correlation dimension may be written as

$$d = \frac{\ln C(r)}{\ln r}. \quad (6)$$

As r increases, $C(r)$ reaches saturation, i.e., $C(r) \longrightarrow 1$ at the r values comparable with the attractor size. In this case, the distance between any two points of the attractor is obviously smaller than the parameter r. On the other hand, it is obvious that at a very small r the number of pairs of points $x_i x_j$ the distance between which is no larger than r will be small, because the number of points on the attractor is finite. Therefore, statistics become poor with a decrease in r, and, in practice, the correlation dimension (6) is found from the linear segment of the plot in a limited range of r values.

The embedding dimension m (i.e., the effective number of variables describing the system) is usually unknown a priori for the analyzed systems; hence, the common practice is to calculate correlation dimensions for different m values. As the embedding dimension is increased to a certain critical value, the correlation dimension also increases. However, when this critical value has been reached, d ceases to change: the slope of the linear segment of the $\ln C(r)$ curve as a function of $\ln r$ becomes constant. This critical value of m is the *minimal* embedding dimension and corresponds to the number of independent variables describing the system.

Restrictions of the methods for calculating the correlation dimension. The correlation dimension of the attractor of a dynamic system carries information on the complexity of its behavior. To date, the Grassberger–Procaccia method [18] is one of the most popular and informative algorithms for the procession of time series. However, the drawback to the advantages of this method is a large amount of calculations ($O(N^2)$ operations). Like many other methods of analyzing time series, this method is ineffective in the case of short series (less than 10^4 values) [19]. In addition, the Grassberger–Procaccia algorithm is practically inapplicable to nonstationary series.

Thus, some restrictions should be put on the time series to preclude errors when calculating the correlation dimension:

(1) the time series should be stationary [12, 20, 21] and

(2) the sample length should be at least $N_{\min} \approx 10^{d/2}$ to obtain a significant estimate of the dimension.

Because of the first restriction, the calculation of this HRV characteristic may prove to be incorrect. However, the precise correlation dimension is not so important for our analysis, because the main question is whether patients with different heart pathologies differ with respect to the correlation dimension (d) of the ECG. Determination of the d values for several fixed embedding dimensions m will suffice to answer this question. Indeed, at large values of the embedding dimension, curves of the dependence $d(m)$ almost reach a plateau, only slightly deviating from it. This method differs from the true one in that, generally speaking, the choice of m is strictly determined and the value cannot be set manually.

In this study, we compare the suggested HRV parameter for different ECGs at the same embedding dimension m. Hereinafter, we mean precisely this parameter when speaking of the correlation dimension of ECG RR parameters. Although this approach is somewhat incorrect, it makes it possible to attain the main goal, i.e., to divide the patients into groups with different pathologies.

In addition, since we were trying to obtain statistically significant results, we did not restrict our study to one ECG from the group selected. We used for analysis at least several (sometimes, very many) ECGs and then averaged the data.

Regarding the second restriction, we may note the following. The $N_{\min}$ ratio may be interpreted differently. A sample of a specified length N allows us to determine a dimension that does not exceed $d_{\max} \cong 2\log N$. Thus, for the values $N = 10^4$–10^5, which are the most typical in practice, we found $d_{\max} \cong 8$–10 [12, 22]. From the medical point of view, to obtain a time series of RR intervals of a patient's ECG consisting of 10^4–10^5 elements means that the ECG should be recorded for many hours (or even days) in a row. The longer the time series, the higher the significance of the resultant estimate of the correlation dimension. However, this difficulty is overcome with the use of modern mobile devices, e.g., a Halter monitor, making it possible to record ECGs continuously over several days.

ANALYSIS OF FACTUAL DATA

To obtain statistically significant data, we processed 159 series of RR intervals derived from the ECGs of patients with different cardiac rhythm disorders. All ECGs were obtained from the same source [23]. To exclude accidental outliers (which are inevitable in currently used mobile electrocardiographs), the records were preliminarily processed. Then, all ECGs were subdivided into groups according to individual heart diseases. After each ECG was processed, the data were averaged over each group. According to the first restriction of the method for calculating the correlation dimension, we tested whether the resultant series of RR intervals were stationary. As expected, all series studied proved to be nonstationary. Other researchers came to the same conclusion with the use of other criteria for the stationary/nonstationary state of time series [20, 24]. Therefore, we used the aforementioned procedure in subsequent analysis.

Patients grouped according to age and pathology

ECG group	d at $m = 3$	d at $m = 4$	d at $m = 5$
1	2.51 ± 0.04	3.11 ± 0.05	3.61 ± 0.1
2	2.13 ± 0.02	2.60 ± 0.03	2.78 ± 0.03
3	2.68 ± 0.04	3.29 ± 0.07	3.66 ± 0.08
4	2.74 ± 0.05	3.45 ± 0.06	3.89 ± 0.1
5	2.20 ± 0.02	2.47 ± 0.05	2.54 ± 0.03
6	2.57 ± 0.03	3.08 ± 0.08	3.78 ± 0.1
7	2.32 ± 0.03	2.81 ± 0.05	3.57 ± 0.1
8	2.51 ± 0.04	2.98 ± 0.05	3.75 ± 0.1
9	2.51 ± 0.04	1.36 ± 0.04	1.42 ± 0.06
10	2.51 ± 0.04	3.7 ± 0.043	4.44 ± 0.08
11	2.51 ± 0.04	3.19 ± 0.04	3.51 ± 0.08

Groups: 1, patients with coronary vessel disorders (age, 47–60 years; 21 ECGs); 2, patients that have had coronary artery shunt and myocardial infarction (age, 55 and 60 years; 12 ECGs); 3, patients with angina decubitis and coronary pathology (age, 40–54 years; 17 ECGs); 4, patients with angina decubitis that have had myocardial infarction (age, 58 years; 9 ECGs); 5, patients with angina decubitis and coronary pathology that have had myocardial infarction (age, 70 and 71 years; 7 ECGs); 6, patients with angina decubitis and pathology of the coronary artery (age, 45–51 years; 13 ECGs); 7, patients with angina of effort that have had myocardial infarction (age, 51–66 years; 16 ECGs); 8, patients with mixed angina pectoris and various coronary disorders (age, 48–63 years; 32 ECGs); 9, patients with atrioventricular block (age, 73–89 years; 6 ECGs); 10, patients with backward heart failure (8 ECGs); 11, patients with atrial flutter (18 ECGs).

The table shows the results of the calculation of the ECG correlation dimension corresponding to the embedding dimension values $m = 3$, $m = 4$, and $m = 5$. As evident from the table, group of patients with different heart pathologies differed from one another in correlation dimension. However, this was not true for all groups. For example, the ranges of d values sometimes overlapped within one embedding dimension ($m = 3$). This behavior of d was observed in the third (angina decubitis and pathology of coronary vessels) and fourth (angina decubitis and consequences of myocardial infarction) groups, as well as the first (pathologies of coronary vessels with different degrees of being affected) and eighth (mixed angina and pathology of coronary vessels) groups (Fig. 3). However, the *diagnoses* of these patients also *partly coincided.* This coincidence may have accounted for the overlapping of the correlation dimension ranges.

As the embedding dimension was increased ($m = 4$ and $m = 5$), the ranges of d values for the same groups of ECGs (the third and fourth and the first and eighth) no longer overlapped. This indicates that an increase in m allows the group of pathologies to be identified more precisely. Thus, when classifying the ECGs of patients with different types of heart failure, one should consider d values at several values of the embedding dimension. The results obtained at *only one* value of the embedding dimension *will not* unambiguously indicate a specific pathology.

Subsequent growth of the embedding dimension leads to an increase in the differences between individual groups. These results are shown in Fig. 4. For comparison, Fig. 4 also shows a straight line corresponding to a healthy patient.

These plots clearly demonstrate that

• in all cases studied here, the correlation dimensions of ECGs considerably differ from one another if the embedding dimension is sufficiently large;

• as m grows, the curve of the dependence of correlation integral on the embedding dimension almost reaches a plateau. In other words, d almost ceases to change at a certain value of m.

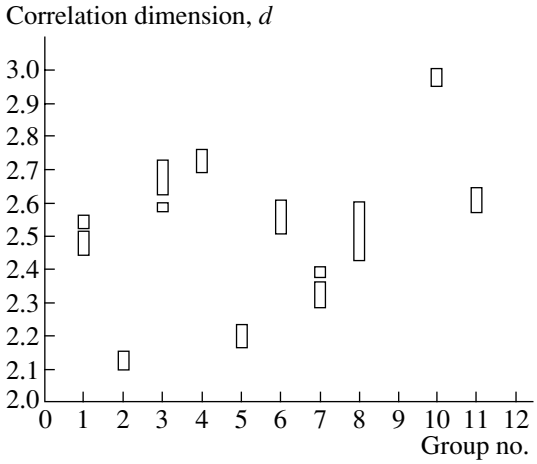

Fig. 3. The correlation dimensions of ECGs of patients with different heart diseases for $m = 3$.

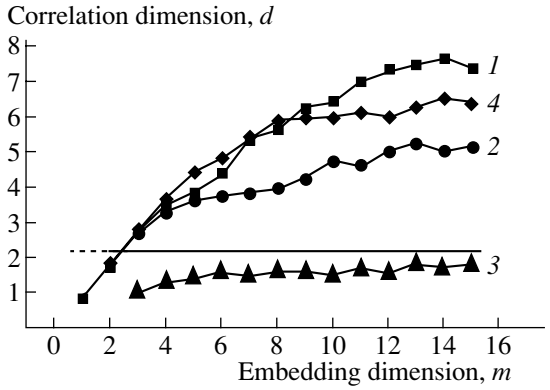

Fig. 4. Dependence of the correlation dimensions of the ECGs of patients with different heart diseases on the embedding dimension: *1*, angina pectoris and myocardial infarction; *2*, angina pectoris; *3*, Atrioventricular block; *4*, backwards heart failure. The straight line approximately corresponds to the normal state.

CONCLUSIONS

We have studied series of RR intervals of the ECGs of patients with certain heart pathologies, as well as time series corresponding to the normal cardiac rhythms of healthy people. Analysis of correlation dimension has demonstrated that this characteristic can be used to solve the reverse problem, i.e., to divide patients into groups according to the types of heart failure. The degree of chaos in ECG patterns increases or decreases depending on the type of heart pathology. Note, however, that this conclusion has been made after the treatment of sufficiently long series.

It is well known that heart pathologies may be caused not only by dysfunction of the cardiovascular system per se, but also by entirely different diseases. Therefore, of special interest is the study of so-called hidden heart pathologies, which may be caused, among other factors, by drugs used for treating diseases other than cardiovascular ones, but nevertheless affecting the heart. Standard methods often cannot detect these pathologies at an early stage. Therefore, the development of nonlinear dynamic methods, such as those described here, may help to make substantial progress in this field.

In addition, the use of the methods proposed here gives hope for solving the "direct" problem, i.e., a provisional diagnosis of a heart pathology and/or correction of it. We are now investigating this subject in collaboration with the Bakulev Research Center of Cardiovascular Surgery of the Russian Academy of Medical Sciences [25]. The solution of these and related processes would make it possible to determine the boundary beyond which chaotic processes that are often present in cardiac rhythm become incompatible with a healthy state and unambiguously indicate a disease.

REFERENCES

1. A. M. Sigal, *Cardiac Performance Rhythms and Their Disturbances* (Medgiz, Moscow, 1957) [in Russian].
2. E. A. Zhuchkova, A. Yu. Loskutov, and S. D. Rybalko, Biofizika **47** (5), 892–901 (2002).
3. N. B. Janson and A. G. Balanov, et al., Phys. Rev. E **65**, 036 212 (2002).
4. N. Wessel, Ch. Ziebmann, et al., Phys. Rev. E **61** (1), 733–739 (2000).
5. V. Schulte-Frohlinde, Yo. Ashkenazy, et al., Phys. Rev. Lett. **87** (6), 068104–1 (2001).
6. C. K. Peng, Sh. Halvin, H. E. Stanley, and A. L. Goldberger, Chaos **5** (1), 82–87 (1995).
7. J. Kurths, A. Voss, et al., Chaos **5** (1), 88–94 (1995).
8. R. B. Govindar, K. Narayanan, and M. S. Gopinathan, Chaos **8** (2), 495–502 (1998).
9. A. L. Guldberger and P. R. Rigney, *Dynamic Patterns in Complex Systems* (1988), pp. 248–261.
10. A. L. Goldberger, P. R. Rigney, and B. J. West, Sci. Am., No. 262, 42–49 (1990).
11. A. L. Goldberger, Ann. Biomed. Eng, No. 18, 195–198 (1990).
12. H. G. Shuster, *Deterministic Chaos* (Physik-Verlag, Weinheim, 1984; Mir, Moscow, 1988).
13. A. Yu. Loskutov and A. S. Mikhailov, *Introduction to Synergetics* (Nauka, Moscow, 1990) [in Russian].
14. G. G. Malinetskii and A. B. Potapov, *Current Problems of Nonlinear Dynamics* (URSS, Moscow, 2002) [in Russian].
15. T. Sauer, Phys. Rev. Lett. **72** (24), 3811–3814 (1994).
16. T. Sauer, *Nonlinear Dynamics and Time Series*, Ed. by C. Cutler and D. Kaplan **11**, 63–75 (1997).
17. D. M. Racicot and A. Longtin, Phys. D (Amsterdam) **104** (2), 184–204 (1997).
18. P. Grassberger and I. Procaccia, Phys. D (Amsterdam) **9** (1), 189–208 (1983).
19. A. Yu. Loskutov, O. L. Kotlyarov, I. A. Istomin, and D. I. Zhuravlev, Vestn. Mosk. Univ., Ser. 3. Phys., Astron. **6**, 3–21 (2002).
20. P. Ch. Ivanov, P. Bernaola-Galvan, et al., 7th Int. Conf. "Fractal-2002" (2002), pp. 55–63.
21. V. V. Sychev, Extended Abstract of Candidate's Dissertation in Mathematics and Physics (Pushchino, 1999).
22. A. B. Potapov, Extended Abstract of Doctoral Dissertation in Mathematics and Physics (Moscow, 1999).
23. *PhysioNet* (http://physionet.org/). A Public Source of Research Resources of Complex Physiological Data PhysioNet.
24. M. B. Kennel, Phys. Rev. E **56** (1), 316–321 (1997).
25. I. P. Polyakova, O. Yu. Mironyuk, and A. Yu. Loskutov (in press).