Non-linear properties of R–R distributions as a measure of heart rate variability

Article in Chaos Solitons & Fractals · June 2003
DOI: 10.1016/S0960-0779(02)00403-4

6 authors, including:

Isabel Iruzun
Universidad Nacional de La Plata
61 PUBLICATIONS 270 CITATIONS

Paula E Bergero
Universidad Nacional de La Plata
47 PUBLICATIONS 130 CITATIONS

J. L. Vicente
Universidad Nacional de La Plata
79 PUBLICATIONS 519 CITATIONS

Some of the authors of this publication are also working on these related projects:

- PHYSICO-CHEMICAL ADSORPTION ON HETEROGENEOUS SOLIDS View project
- Adsorption on carbonaceous materials View project

All content following this page was uploaded by Paula E Bergero on 11 July 2019.
The user has requested enhancement of the downloaded file.
Non-linear properties of $R–R$ distributions as a measure of heart rate variability

I.M. Irurzun $^{a,b}$, P. Bergero $^a$, M.C. Cordero $^a$, M.M. Defeo $^c$, J.L. Vicente $^a$, E.E. Mola $^{a,*}$

$^a$ INIFTA, Instituto de Investigaciones Fisicoquímicas Teóricas y Aplicadas (UNLP, CONICET, CICPBA), Sue. 4, Casilla de Correo 16, La Plata 1900, Argentina

$^b$ Facultad de Ciencias Exactas y Naturales, Departamento de Física, Universidad de Buenos Aires, Buenos Aires 1428, Argentina

$^c$ Servicio de Cardiología, Hospital “Prof. Rodolfo Rossi”, La Plata 1900, Argentina

Accepted 3 September 2002

Abstract

We analyze the dynamic quality of the $R–R$ interbeat intervals of electrocardiographic signals from healthy people and from patients with premature ventricular contractions (PVCs) by applying different measure algorithms to standardised public domain data sets of heart rate variability. Our aim is to assess the utility of these algorithms for the above mentioned purposes.

Long and short time series, 24 and 0.50 h respectively, of interbeat intervals of healthy and PVC subjects were compared with the aim of developing a fast method to investigate their temporal organization.

Two different methods were used: power spectral analysis and the integral correlation method.

Power spectral analysis has proven to be a powerful tool for detecting long-range correlations. If it is applied in a short time series, power spectra of healthy and PVC subjects show a similar behavior, which disqualifies power spectral analysis as a fast method to distinguish healthy from PVC subjects.

The integral correlation method allows us to study the fractal properties of interbeat intervals of electrocardiographic signals.

The cardiac activity of healthy and PVC people stems from dynamics of chaotic nature characterized by correlation dimensions $d_c$ equal to $3.40 \pm 0.50$ and $5.00 \pm 0.80$ for healthy and PVC subjects respectively.

The methodology presented in this article bridges the gap between theoretical and experimental studies of non-linear phenomena. From our results we conclude that the minimum number of coupled differential equations to describe cardiac activity must be six and seven for healthy and PVC individuals respectively.

From the present analysis we conclude that the correlation integral method is particularly suitable, in comparison with the power spectral analysis, for the early detection of arrhythmias on short time (0.5 h) series.

$^*$ Corresponding author. Fax: +54-221-4254642.

E-mail address: eemola@inifta.unlp.edu.ar (E.E. Mola).

2002 Elsevier Science Ltd. All rights reserved.

1. Introduction

Sudden cardiac death remains one of the continuing challenges to the modern clinician. Currently, it accounts for the death of the majority of people with cardiovascular disease [1]. Abnormalities can appear as a consequence of several
pathologies, such as coronary artery disease, left ventricular hypertrophy, cardiomyopathy, etc. [2]. These pathologies will finally be observed as abnormalities in the rhythm of the heart. Premature ventricular contractions (PVCs) are a typical example (see Fig 1).

Isolated PVCs frequently appear in healthy subjects without any clinical symptom of illness. In such cases medical treatment is not advised. However, isolated PVCs can get worse in more frequent and abnormal premature beats. Thus the early detection of arrhythmias can be difficult and the 24-h Holter electrocardiogram is recommended to evaluate the seriousness of the arrhythmia.

Based on the hypothesis that the pattern of PVCs in patients with heart disease may provide clues to prognosis, Stein et al. [3,4] determined the fractal dimension of these events (PVC) embedded along a one-dimensional time axis. This analysis yields a fractal dimension, which can range from zero to one and is inversely related to clustering of PVCs. Neither of these studies determined the statistical distribution nor the correlations between the interbeat intervals.

It might be more convenient to evaluate the correlation dimension $d_t$ from a time series of finite length using the algorithms developed by Grassberger and Procaccia [5]. This approach was employed by Babloyantz and Destexhe [6], who evaluated $d_t$ from four (4) healthy resting individuals. They found values ranging from $d_t = 3.6 \pm 0.1$ to $5.2 \pm 0.1$ by employing time series 40 min long and a time delay $\tau$ ranging from 120 to 240 ms. By choosing the interbeat sequence as the time series, the same authors have evaluated the correlation dimension of the $R$–$R$ intervals. They found a rather high dimensional chaos characterized by $d_t = 5.9 \pm 0.4$. In Section 3 we perform a thorough investigation of the correlation dimension of $R$–$R$ intervals using the databases of PhysioNet.

To confirm the coherent nature of the $R$–$R$ interval variability, Babloyantz and Destexhe [6] constructed a, map of the interval variation:

![Fig. 1. (a) Normal ECG signal for a healthy heart of a 22-year-old man. (b) ECG signal from a subject with PVCs. The tracing demonstrate a late or end-diastolic premature ventricular beat.](image-url)
$\Delta(R-R)_{n+1} = \{(R-R)_{n+2} - (R-R)_{n+1}\}/(R-R)$ \hspace{1cm} (1)

vs

$\Delta(R-R)_{n} = \{(R-R)_{n+1} - (R-R)_{n}\}/(R-R)$ \hspace{1cm} (2)

where $(R-R)_{n}$ is the $n$th $R-R$ time interval and $(R-R)$ is the average $R-R$ time interval. This representation reveals the correlation between the variabilities of three consecutive $R-R$ intervals and they claim that from the set of four healthy individuals that they studied it is possible to distinguish between a pseudo-random fluctuation around a periodic oscillator and the behavior of a normal heart.

We constructed the map of the interval variation using Eqs. (1) and (2) for every healthy individual of the MIT-BIH Database [7,8], and we determined that there is a great variability among healthy subjects, which makes it difficult to differentiate between the two behaviors studied by Babloyantz and Destexhe in Ref. [6]. The reason of this uncertainty will be clarified in Section 3.

Peng et al. [9,10] studied the scale behavior of heartbeat series using a power spectral method and showed that healthy and ill people present different power behavior. We have applied this method to PVC subjects and found that they cannot be distinguished from healthy people.

Recent studies have detected unstable periodic orbits in human cardiac rhythms for healthy and pathological cases [11–13]. Although this analysis would allow distinguishing among different states of cardiac systems, its complexity becomes useless as a diagnostic method.

To summarize we compare two different methods to study the dynamics of the cardiac rhythms of healthy and PVC people: the power spectral analysis and the correlation integral method. Analyses were made over the time series of $R-R$ interbeat intervals, using series of almost 30 min. By comparing both methods we conclude that the correlation integral method is more adequate than the power spectral analysis to provide clues to distinguish between healthy and PVC subjects. This work is organized as follows: in Sections 2 and 3 the power spectral analysis and the correlation integral methods respectively are presented. Time series of healthy people are analyzed and similar studies are performed over time series of PVC subjects. Results are discussed in Section 4 with our conclusions.

2. Power spectral analysis

2.1. Healthy adults

Peng et al. [9,10] studied scale-invariant properties of the human heartbeat time series. The analysis was based on very long time series of $R-R$ interbeat intervals (up to 24 h $\approx 10^5$ beats). These time series obtained by the sequential intervals between beat $n$ and beat $n+1$, and denoted by $B(n)$ or $(R-R)_{n}$, typically reveals a complex type of variability. In Fig. 2(a) $B(n)$ is plotted for the healthy subject 16273 taken from the MIT-BIH Database Distribution [7,8]. To study the correlation properties of this time series we investigated the heartbeat intervals $I(n) \equiv B(n+1) - B(n)$ using the method proposed by Peng et al. [9,10]. Fig. 2(b) shows a plot of $I(n)$ vs beat numbers $(n)$. Since $I(n)$ is stationary, standard spectral analysis techniques can be applied. Fig. 3(a) shows the power spectra $S_{I}(f)$, the square of the Fourier transform amplitudes for $I(n)$, derived from the same data set as that used in Fig. 2(a). The spectral data in Fig. 3(a) are smoothed by averaging over 50 values [9,10]. The fact that the log–log plot of $S_{I}(f)$ vs $f$ is linear implies

$S_{I}(f) \sim f^{-\beta}$ \hspace{1cm} (3)

We found $\beta \approx -1$ in agreement with the results obtained by Peng et al. [9,10] in a non-specified set of healthy people. These findings are not simply attributable to different levels of daily activities [9].

To analyze the validity of the scaling when a reduced number of beats are employed, we took sets of different length from the original data (Fig. 3(a)) and their power spectra are also shown in Fig. 3(b) and (c).

We found that the scaling behavior with $\beta \approx -1$ found in Fig. 3(a) on three decades using approximately $6 \times 10^4$ beats is reduced to almost one decade employing $\sim 3 \times 10^3$ beats. Smoothing effects were also investigated, Table 1 summarizes this effect for the subject 16273 from the MIT-BIH Database Distribution of Fig. 2(a). We learn that while statistical dispersion grows by reducing the number of values employed to obtain the average, $\beta$ values remain almost constant.

Therefore the log–log linear relationship found by Peng et al. [9,10] on healthy subjects remains valid, over a decade, using $\sim 3 \times 10^3$ beats.
2.2. PVC adult subjects

ECG’s from twenty-two (22) different pathological PVC cases were taken from the MIT-BIH Database Distribution [7,8]. We took time series made up by approximately $3 \times 10^3$ beats, which is roughly the information included for every PVC subject.

To study the correlation properties of this short PVC time series we investigated the heartbeat intervals $I(n) \equiv B(n+1) - B(n)$ using the method proposed by Peng et al. [9,10].

Fig. 4 shows the power spectra $S(f)$, the square of the Fourier transform amplitudes for $I(n)$, derived from the PVC subject 107 of the MIT-BIH Database. The time series is made up by approximately $2 \times 10^3$ beats. For comparison purposes the power spectra derived from the short time series of the healthy subject in Fig. 2(a) (see Fig. 3(c)), as well a straight line with slope $\beta = -1$ are included in Fig. 4.

From Fig. 4 we learn that it is difficult to establish a difference between healthy and non-healthy subjects, due to the reduced frequency interval ($10^{-2}$–$10^{-1}$ beats$^{-1}$) where linear scaling of Eq. (3) holds, because of the short time series employed. This feature disqualifies power spectral analysis as a fast method to distinguish between healthy and non-healthy subjects. Even more, at that particular frequency interval it is possible to observe the emergence of intermittent relatively low frequency heart rate oscillations associated with the well recognized, and common to different pathologies, syndrome of periodic (Cheyne-Stokes) respiration.

3. Correlation integral method

3.1. Time series: $R–R$ interbeat intervals

For physiological systems such as the cardiac system, it is generally difficult to confirm the presence of deterministic chaos from the usual dimensional analysis [14]. Although clinicians describe the normal electrical activity of the heart as regular sinus rhythm, cardiac interbeat intervals fluctuate in a complex, apparently erratic manner, in healthy subjects even at rest. We wish to characterize whether there is any strange attractor underlying its dynamics. Dynamical information will be extracted by the analysis of the time series of $R–R$ interbeat intervals.
We used a modification of the correlation integral method [5] to analyze the R–R interbeat intervals of an ECG.

Given a time series

$$\frac{1}{C_{138}} \frac{R_R}{C_1} = 1, 2, \ldots, N$$

where each term defines the \(i\)th interbeat interval, and the correlation integral over a given length scale, \(l\), is defined as

$$C(l) = \lim_{N \to \infty} \frac{\text{Number} \{ (i, j) : |(R_R)_i - (R_R)_j | < l \}}{N^2}$$

When the correlation integral is calculated over varying length scales, it obeys the scaling relationship

$$C(l) \sim l^\nu$$

for small length scales. In this relationship, \(\nu\) is the correlation exponent. The correlation exponent characterizes temporal correlations among time series elements and, thus, describes the homogeneity of the series in phase space. When the dimensionality of \((R_R)_i\) in Eq. (4) is finite, the series can be embedded in a higher dimensional space to

---

**Table 1**

<table>
<thead>
<tr>
<th>Number of the adjacent points averaged</th>
<th>(\beta) (Eq. (3)), healthy subject (16273)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>(-1.07 \pm 0.03)</td>
</tr>
<tr>
<td>30</td>
<td>(-1.05 \pm 0.02)</td>
</tr>
<tr>
<td>50</td>
<td>(-1.06 \pm 0.01)</td>
</tr>
</tbody>
</table>

---

Fig. 3. Power spectra \(S_j(f)\), the square of the Fourier transform amplitudes for \(I(n)\), for sets of interbeat intervals increase sequences for the same subject as that of Fig. 2(a): (a) \(n = 65536\), (b) \(n = 20000\), (c) \(n = 2500\). Spectral data are smoothed by averaging over 50 values [7,8]. The heart rate spectrum is plotted as a function of inverse beat number (beat\(^{-1}\)) rather than frequency (time\(^{-1}\)). For comparison purposes a straight line with slope \(\beta = -1\), see Eq. (3), is included.
eliminate error in the measurement of $\nu$. This embedding is accomplished by creating $d$-dimensional vectors $\xi_i$, from the time series:

$$\xi_i = \{(R-R)_i, (R-R)_{i+1}, \ldots, (R-R)_{i+(d-1)}\}$$  \quad (7)

and substituting $\xi_i$ for $(R-R)_i$ in Eq. (5). When we measure the correlation exponent from Eqs. (5) and (6) as the series of Eq. (4) is embedded in spaces of increasing dimensions, its value will increase with the spatial dimension if the series is random. If the series is governed by a strange attractor of finite fractal dimension ($d_f$), the correlation exponent will assume a constant value at spaces of dimension greater than or equal to $d_f$. For a chaotic series, the correlation exponent follows the relationship

$$\nu \leq d_f$$ \quad (8)

The fractal dimension is necessarily less than or equal to the dimensionality of the space where the strange attractor is completely embedded.

3.2. Normal healthy people

ECG’s from 18 different normal healthy people were taken from the MIT-BIH Database Distribution [7,8]. In Fig. 5(a) we show the calculations performed using a particular $R$–$R$ time series of a healthy subject (subject 16273).

Fig. 5(a) shows several plots from which we calculated the correlation exponent $\nu$ from a log–log plot of Eq. (6) as we embed the time series given by Eq. (4) in increasing dimensional space ($d = 1, 2, \ldots, 7$). The correlation exponent $\nu$, which is the slope of successive lines, increases from a value of $0.84 \pm 0.01$ at $d = 1$ up to a value of $3.77 \pm 0.03$ at $d = 6$, where a saturation value is reached. The time series of $R$–$R$ interbeat intervals employed to perform these calculations consists of $6 \times 10^4$ points.

To investigate the validity of the scaling when a reduced number of beats is employed from the original data (Fig. 5(a)) we took a time series of 3500 points, which is roughly of the same length as the time series of PVC subjects.

Log–log plots of $C(l)$ vs $l$ for different dimensional spaces ($d = 1, \ldots, 6$) are shown in Fig. 5(b). The correlation exponent $\nu$, which is the slope of successive lines, increases from a value of $0.80 \pm 0.01$ at $d = 1$ up to a value of $3.71 \pm 0.04$ at $d = 6$, where a saturation value is reached. By comparing Fig. 5(a) and (b) we learn that the scaling behavior can be found even with rather short time series of 3500 points.

Fig. 4. Power spectra $S_\nu(f)$, the square of the Fourier transform amplitudes for $I(n)$, for the PVC subject 107 of the MIT-BIH Database. The time series is made up by approximately $2.1 \times 10^3$ beats. For comparison purposes a straight line with slope $\beta = -1$, see Eq. (3), is included. The spectral peak at about $10^{-1}$ beat$^{-1}$ in the power spectrum of the PVC subject is associated with Cheyne-Stokes respiration.
As was pointed out by Schreiber and Kostelich [15,16], correlation exponent calculations of chaotic time series are sensitive to the presence of small amounts of noise or random contributions. Noise places a lower bound on the values of $l$ (see Eq. (6)) that can be used to estimate $\nu$, and could be affecting the scaling relationship of our data at higher scales.

As was pointed out by Schreiber and Kostelich [15,16], correlation exponent calculations of chaotic time series are sensitive to the presence of small amounts of noise or random contributions. Noise places a lower bound on the values of $l$ (see Eq. (6)) that can be used to estimate $\nu$, and could be affecting the scaling relationship of our data at higher scales.

Fig. 5. Log–log plot of the correlation integral, Eq. (6), for the healthy subject of Fig. 2(a) for various scale values $l$ (s). We embed the time series in increasing dimensional space ($d = 1, 2, \ldots, 7$). The correlation exponent ($\nu$) is the slope of the linear plot. The time series employed is (a) 6 x $10^4$ points and (b) 3500 points.

Fig. 6. Log–log plot of the correlation integral, Eq. (6), for the 107 PVC subject of Fig. 4 for various length values $l$ (s). We embed the time series in increasing dimensional space ($d = 1, 2, \ldots, 7$). The correlation exponent ($\nu$) is the slope of the linear plots.
dimensions (see Fig. 5(a)). Even though several noise-reduction methods exist, in our opinion, they have similar advantages and shortcomings. Hence, we postpone an exhaustive analysis to a subsequent work [17].

The presence of a non-linear structure in the time series of $R-R$ interbeat intervals can be established by performing surrogate data analysis [11]. This analysis has been applied to our healthy subjects time series of 3500 beats. These surrogates will destroy the non-linear structure, if any, present in the time series. Since the original and the surrogate data sets will have the same linear properties, any difference in the discriminating metric between the original and the surrogate must be only because of the non-linear structure present in the original time series. The significance $S$ [11], defined as $S = (v_{sur} - v_{ori})/\sigma$ (where $v_{sur}$ and $v_{ori}$ are the values for surrogate and original data, using the correlation exponent at $d = 6$ as the discriminating metric, and $\sigma$ is the standard deviation of the $v_{sur}$), was calculated, and as an example we found $S = 11.4 \pm 0.2$, for the healthy 16273 subject in Fig. 2(a). The saturation of the correlation exponent as a function of the dimension of space and the presence of a non-linear structure inferred from the surrogate data analysis suggest that the time series of $R-R$ interbeat intervals is deterministically chaotic.

3.3. Premature ventricular contraction

ECG’s from 22 different pathological PVC cases were taken from the MIT-BIH Database Distribution [7,8]. We took time series made up by approximately $\sim 3 \times 10^3$ points, which is roughly the information included for every PVC subject. A surrogate analysis like those for healthy people was performed but the correlation exponent at $d = 7$ was used as discriminating metric. For the subject 107 in Fig. 4 we found a significance $S = 9.4 \pm 0.2$.

Fig. 6 shows several plots corresponding to the PVC subject 107 of Fig. 4, from which we calculated the correlation exponent $v$ by the same procedure employed with healthy people. As we embed the time series in increasing dimensional space, the correlation exponent $v$ increases from a value of $0.777 \pm 0.006$ at $d = 1$ up to a value of $4.16 \pm 0.04$ at $d = 7$, where a saturation value is reached.

Finally, Fig. 7 shows the average correlation exponent ($v$) as a function of spatial dimension for healthy and PVC subjects. Average correlation exponents were obtained using the 18 healthy subjects and the 22 PVC cases. Error bars

---

Fig. 7. Average correlation exponent ($v$) as a function of embedding dimension for healthy and PVC subjects. Average correlation exponents were obtained using the 18 healthy subjects and the 22 PVC cases quoted in the MIT-BIH Database. A saturation is observed at a dimension $d = 6$ and 7 for healthy and PVC subjects respectively.
were calculated as the standard deviation. In both cases a linear relationship is observed, but there is a significant difference between them.

1. Within the statistical uncertainty (given by the standard deviation), the two series have approximately the same correlation exponent when the embedding dimension is equal to one \((d = 1)\).
2. There is a clear-cut difference, beyond statistical uncertainties, between correlation exponents at any embedding dimension greater than \(d = 1\).
3. Correlation exponents of healthy people saturate \((\langle \nu \rangle = 3.40 \pm 0.50)\) at spatial dimensions \(d \geq 6\).
4. Correlation exponents of PVC people saturate \((\langle \nu \rangle = 5.00 \pm 0.80)\) at dimensions \(d \geq 7\).

For comparison purposes we also indicated the behavior of the average correlation exponent for 10 different time series of 3500 random numbers uniformly distributed. The correlation exponent describes the homogeneity of the series in phase space. For a random number series, it shows an increase with the spatial dimension without reaching a saturation value (see Fig. 7). Since the series should be distributed homogeneously in spaces of arbitrary dimension, a linear relationship ideality with slope \(b = 1\) and origin coordinate \(a = 0\) is expected. With our random series, we have obtained \(b = 0.91 \pm 0.02\), and \(a = 0.11 \pm 0.09\).

### 4. Conclusions

We have analyzed two different methods to study the dynamics of the cardiac rhythms of healthy and PVC subjects: the power spectral analysis and the correlation integral method. Analyses were made over the time series of \(R-R\) interbeat intervals, using short series of almost 30 min.

Power spectral analysis has proven to be a powerful tool for detecting long-range correlations. However, from Fig. 4 we learn that it is difficult to establish a difference between healthy and PVC subjects employing short time series, due to the reduced frequency interval \((10^{-2} - 10^{-1} \text{ beat}^{-1})\) where the linear scaling of Eq. (3) holds. This feature disqualifies power spectral analysis as a fast method to distinguish between healthy and PVC subjects. Even more, at that particular frequency interval it is possible to observe the emergence of intermittent relatively low frequency heart rate oscillations associated with the well recognized, and common to different pathologies, syndrome of periodic (Cheyne-Stokes) respiration.

Fractal properties were studied using the integral correlation method. We found that both the normal human heart and those with PVC follow deterministic dynamics of chaotic nature. Time series are governed by strange attractors of fractal dimension \(d_f = 3.40 \pm 0.50\) and \(d_f = 5.00 \pm 0.80\) respectively. The strange attractors can be completely embedded in spaces of dimensions 6 and 7 respectively. Hence, we conclude that the minimum number of coupled differential equations to describe cardiac activity must be six and seven for healthy and PVC individuals respectively. A basic prerequisite is to know the number of independent variables to set up a model to understand the dynamics of heart activity.

Our results on healthy subjects are in agreement with those obtained by Babloyantz and Destexhe \([6]\). Analyzing four healthy individuals these authors found an embedding dimension \(d = 7\) and sometimes \(d = 6\). They also constructed a map of the interval variation (see Eqs. (1) and (2)). That representation reveals the correlation between the variabilities of three consecutive \(R-R\) intervals and they claim that it is possible to distinguish between a pseudo-random fluctuation around a periodic oscillator and the behavior of a normal heart.

But we determined that it is difficult to differentiate between the two behaviors studied by Babloyantz and Destexhe. The origin of such uncertainty comes from the fact that maps of the interval variation, as are stated by Eqs. (1) and (2), partially explore the correlation properties at a space of dimension \(d = 2\), a dimension value significantly lower than the saturation value.

The analysis developed in Section 3 allows to assess the importance of the correlation integral method in comparison with the power spectral analysis for the early detection of arrhythmias without using the 24-h Holter electrocardiogram as is actually recommended at the present.

Work is in progress in La Plata to apply the present analysis to endemic pathologies that modify heart activity.

### Acknowledgements

The authors are indebted to the valuable information provided by MIT-BIH Database, Harvard-MIT Division of Health Sciences and Technology. This research project was financially supported by the Consejo Nacional de
References