

Multiscale Analysis of Heart Rate Dynamics: Entropy and Time Irreversibility Measures

Madalena D. Costa · Chung-Kang Peng ·
Ary L. Goldberger

© Springer Science+Business Media, LLC 2007

Abstract Cardiovascular signals are largely analyzed using traditional time and frequency domain measures. However, such measures fail to account for important properties related to multiscale organization and non-equilibrium dynamics. The complementary role of conventional signal analysis methods and emerging multiscale techniques, is, therefore, an important frontier area of investigation. The key finding of this presentation is that two recently developed multiscale computational tools—multiscale entropy and multiscale time irreversibility—are able to extract information from cardiac interbeat interval time series not contained in traditional methods based on mean, variance or Fourier spectrum (two-point correlation) techniques. These new methods, with careful attention to their limitations, may be useful in diagnostics, risk stratification and detection of toxicity of cardiac drugs.

Keywords Nonlinear dynamics · Heart rate variability · Information theory · Nonequilibrium systems

Introduction

Physiologic control mechanisms exist from subcellular to systemic levels and operate over multiple time scales. Continuous interplay among these different regulatory systems ensures that information is constantly exchanged across all levels of organization, even at rest, and enables

an organism to adjust to an ever-changing environment and to perform a variety of activities necessary for survival. These dynamic processes, especially under healthy conditions, are evident in the complex fluctuations of physiologic output signals, such as heart rate, blood pressure, brain electrical activity, and hormone levels. In contrast, aging and pathologic systems, which have degraded control mechanisms, are likely to generate less complex outputs [1, 2]. This loss of complexity is manifest as increased randomness, (e.g., the heart rate fluctuations of a patient with atrial fibrillation), or greater periodicity (e.g., cardiopulmonary oscillations in Cheyne–Stokes syndrome with heart failure). Both classes of pathology are characterized by a “collapse of complexity,” which is associated with a breakdown of the long-range correlations, multiscale variability and time irreversibility properties of the signals observed in healthy subjects.

There is no unifying consensus definition of complexity. In information theory, complexity measures the degree of compressibility of a string of characters. Therefore, random uncorrelated strings, which are virtually incompressible, are considered the most complex [3]. In contrast, from a complex systems approach [4], which we adopt here, random uncorrelated strings of characters are among the least complex signals, and those with long-range correlations are among the most complex. Complex signals typically exhibit one (and usually) more of the following properties: (i) *non-linearity*—the relationships among components are not additive; therefore small perturbations can cause large effects; (ii) *non-stationarity*—the statistical properties of the system’s output change with time; (iii) *time irreversibility or asymmetry*—systems dissipate energy as they operate far-from-equilibrium and display an “arrow of time” signature; (iv) *multiscale variability*—systems exhibit spatio-temporal patterns over a range of scales.

M. D. Costa (✉) · C.-K. Peng · A. L. Goldberger
Division of Interdisciplinary Medicine and Biotechnology, Beth Israel Deaconess Medical Center/Harvard Medical School, 330 Brookline Avenue, GZ-431, Boston, MA 02215, USA
e-mail: mcosta3@bidmc.harvard.edu

The purpose of this paper is fourfold: (i) to briefly review two multiscale methods that quantify different properties of complex cardiovascular signals: multiscale entropy [5, 6] and multiscale time irreversibility (asymmetry) [7], (ii) to present a simplified version of the time irreversibility measure, (iii) to apply the two methods to the challenge of characterizing and distinguishing physiologic and surrogate time series not separable using conventional techniques, and (iv) to test models of heart rate variability previously proposed as part of an international competition.

Our underlying conceptual construct is that multiscale complexity is a marker of healthy dynamics, which have the highest functionality and adaptability, and these properties degrade with aging and disease. Multiscale entropy quantifies the information content of a signal over multiple time scales. Time irreversibility quantifies its degree of temporal asymmetry. Of note, both measures probe aspects of cardiovascular signals that are independent of traditional time and frequency domain measures such as mean, variance and Fourier spectrum, and are also independent of each other as demonstrated below. For continuous heart rate time series obtained, for example from 24 h electrocardiograms (Holter recordings), the analyzed time scales typically range from milliseconds to minutes.

Methods

Multiscale Entropy

The multiscale entropy method was developed to quantify a central aspect of complex signals, namely their multiscale variability over a range of scales. At each level of resolution, the multiscale entropy algorithm yields a value that reflects the mean rate of creation of information. The overall degree of complexity of a signal is calculated by integrating the values obtained for a pre-defined range of scales.

In practice the algorithm comprises two steps: (1) a coarse-graining procedure that allows us to look at representations of the system's dynamics at different time scales, and (2) the quantification of the degree of irregularity of each coarse-grained time series, which can be accomplished using sample entropy (SampEn), a statistic introduced by Moorman et al. [8].

Sample entropy is the negative natural logarithm of the conditional probability that two patterns of length m , $x_m(i) = \{x_i, \dots, x_{i+m-1}\}$ and $x_m(j) = \{x_j, \dots, x_{j+m-1}\}$, which are similar to each other within a tolerance r (meaning that $d[x_m(i), x_m(j)] \leq r$, where d is a function that measures the distance between vectors), will still be considered similar to each other when points, x_{i+m} and x_{j+m} are added to patterns $x_m(i)$ and $x_m(j)$, respectively.

Sample entropy is a measure of irregularity. For regular signals, SampEn is very low. For uncorrelated random signals, SampEn is the highest.

Although complex signals are irregular not all irregular signals are complex. Consider, for example, a physiologic time series, which is the output of a system regulated by multiple control mechanisms, and the time series obtained by shuffling these data points. The surrogate shuffled time series is less complex than the physiologic time series. However, an entropy measure, such as SampEn, assigns a higher entropy value to the surrogate shuffled time series than to the physiologic one because the former is more random than the latter. Therefore, the use of single scale-based measures of entropy to assess complexity may lead to misleading results.

By distinction, our [5] objective is to compute entropy over multiple time scales. The underlying hypothesis is that complex systems, in particular physiologic systems controlled by regulatory mechanisms that operate on different time scales, generate time series that exhibit highly variable fluctuations at multiple levels of resolution. To quantify irregularity at different time scales we coarse-grain the original time series. Given a signal with N data points sampled at Δ Hz, the coarse-grained time series are constructed as follows. For scale 1, the coarse-grained time series is the same as the original signal. In this case the time interval between consecutive data points is $1/\Delta$ s. For scale n (Fig. 1), we divide the data into consecutive non-overlapping blocks with n data points each and calculate the mean inside each block. The sequence of average values is the coarse-grained time series for scale n that corresponds to n/Δ s. This time series is shorter than the original one. It contains N/n data points, which should be taken into consideration when calculating the error bars for the entropy values. Of note, the choice of calculating the mean value to summarize the dynamics inside each block of data was motivated by Zhang's work [9].

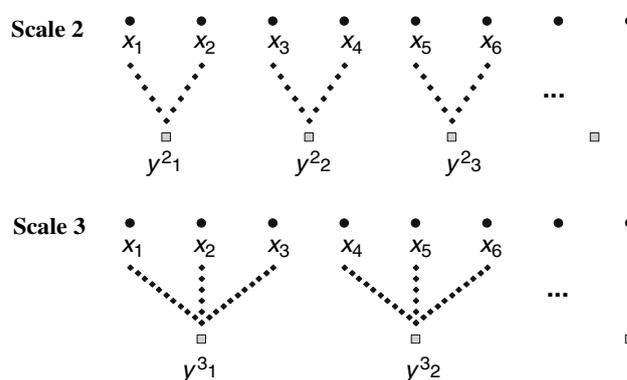


Fig. 1 Schematic illustration of the coarse graining procedure for scales 2 and 3 used for the multiscale entropy method. (Adapted from ref. 6.)

The next step is to calculate the sample entropy for each coarse-grained time series and to plot the results as a function of scale, which yields the so-called multiscale entropy curve. The multiscale entropy curve [6] for correlated fractal ($1/f$) noise is a straight line (the entropy is 1.8 for all scales) since at all levels of resolution new patterns of variability are revealed and, therefore, new information is created. In contrast, the multiscale entropy curve for white noise is a function that monotonically decreases with scale. This profile is characteristic of single scale signals, which at low resolution levels (higher scales) contain less information than at high resolution levels (lower scales.) Of note, the multiscale entropy does not distinguish between a given time series $X = \{x_1, x_2, \dots, x_N\}$ and that obtained by time reversing the series $X' = \{x_N, x_{N-1}, \dots, x_1\}$ because the degree of irregularity/unpredictability is the same for both signals.

Multiscale Time Irreversibility

Time irreversibility (“the arrow of time”), a fundamental property of far-from-equilibrium systems, is related to the unidirectionality of the energy flow across the boundaries of the system [10]. Living beings are paradigmatic examples of systems operating far-from equilibrium. They utilize energy to evolve to and maintain ordered structural configurations, through inherently time irreversible processes. Perhaps counterintuitively, death is a state of maximum equilibrium, since there are no driving forces or consumption of energy. To the extent that all processes occurring under equilibrium conditions are time reversible, states approaching death are expected to be more time reversible than those representing far-from-equilibrium healthy physiology.

In time series analysis, time irreversibility refers to the lack of invariance of the statistical properties of a signal under the operation of time reversal [11]. Figure 2 shows the heart rate time series of a healthy young subject and of a patient with congestive heart failure, both in the forward and backward (reversed) time directions. The pathologic signal is more symmetric than the healthy one, indicating a loss of temporal irreversibility with pathology.

A method [7] that we recently developed to quantify the degree of time irreversibility comprises three steps: a coarse-graining procedure, the computation of the degree of time irreversibility for each coarse-grained time series, and the integration of the results obtained for a pre-defined range of scales. We made the simplifying assumptions that transitions between consecutive values (an increase or decrease in heart rate) are independent and require a certain amount of “energy.” Based on a statistical physics approach [12] we considered that the relationship between

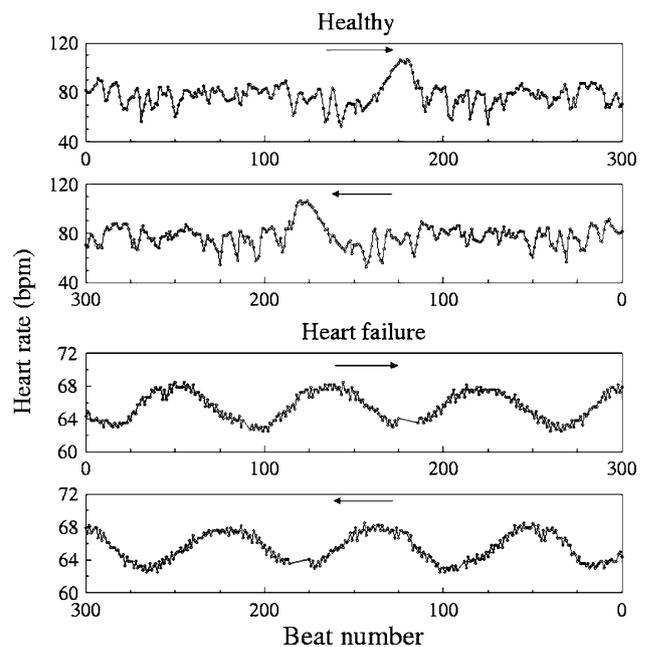


Fig. 2 Heartbeat time series from a healthy subject (top panels) and patient with severe congestive heart failure (bottom panels) shown in both the forward (first and third panels) and backward time directions (second and fourth panels). Note that the time series from the congestive heart failure patient “reads the same” in both forward and backward time directions, in contrast to the more asymmetric time series from the healthy subject

energy E for each transition and the probability p of observing that transition was: $E = p \ln p$. Then, we calculated the difference between the average energy for the activation ($p_+ \ln p_+$) and the relaxation ($p_- \ln p_-$) of the heart, over a range of time scales (where p_+ is the probability that the value of the recorded variable increases at any time instant and p_- is the probability that it decreases). Of note, the relationship between $p \ln p$ and the energy E is only valid if the transitions between states are independent. If they are not independent, we still can use the differences between $p_+ \ln p_+$ and $p_- \ln p_-$ to quantify the degree of temporal asymmetry of a signal but there is no simple physical interpretation for these terms.

In this paper we present a simplified version of this algorithm that yields comparable results and is easier to implement. The simplified version is based on the observation that the number of increments (number of times $x_{i+1} - x_i > 0$) is equal to the number of decrements (number of times $x_{i+1} - x_i < 0$) for a symmetric function. We use this finding for calculating the asymmetry of the original time series and for the coarse-grained time series. Consider a time series $X = \{x_i\}$, $1 \leq i \leq N$. For scale 1, we construct the time series $Y_1 = \{y_i\}$, $y_i = x_{i+1} - x_i$, $1 \leq i \leq N - 1$. Then, we calculate the difference A_1 between the percentage of increments and decrements according to

$$A_1 = \frac{\sum H[-y_i] - \sum H[y_i]}{N-1}$$

where H is the Heaviside function ($H(a) = 0$ if $a < 0$ and $H(a) = 1$ if $a \geq 0$ and $1 \leq i \leq N-1$.) For scale j , we construct the time series $Y_j = \{y_i\}$, $y_i = x_{i+j} - x_i$, $1 \leq i \leq N-j$. Then, we calculate the difference A_j between the percentage of increments and decrements according to

$$A_j = \frac{\sum H[-y_i] - \sum H[y_i]}{N-j}, \text{ where } 1 \leq i \leq N-j.$$

The time asymmetry (irreversibility) index is defined as ΣA_j for a pre-defined range of scales.

The time irreversibility and multiscale entropy methods quantify different aspects of complex systems and are independent of each other. Consider, for example, two time series X_1 , an asymmetric triangle function of period 10 ($X_1 = \{x_i\}$, where $x_i = 3$ if i is a multiple of 10 and $x_i = 0.5(1+i)$ if i is not a multiple of 10, and X_2 , the time series obtained from X_1 by time reversing the sequence of data points. The two time series X_1 and X_2 have the same multiscale entropy values because for all coarse-grained time scales the conditional probability that $d[x_{m+1}(i), x_{m+1}(j)] \leq r$ given that $d[x_m(i), x_m(j)] \leq r$ is the same that the conditional probability that $d[x'_{m+1}(i), x'_{m+1}(j)] \leq r$ given that $d[x'_m(i), x'_m(j)] \leq r$, where $x_m(i) = \{x_i, \dots, x_{i+m-1}\}$ and $x'_m(i) = \{x_{i+m-1}, \dots, x_i\}$. However, X_1 and X_2 have different degrees of temporal irreversibility.

For time series X_1 , the degree of time irreversibility is $8/10 - 2/10 = 3/5$ because, in this example, the probability that the value of x_i increases is $8/10$ while the probability that x_i decreases is $2/10$. For the temporally reversed series, the asymmetry index is $-3/5$.

Results

To illustrate the application of these two techniques, we apply them to a physiologic and surrogate data pair and to data from a time series modeling competition sponsored by the NIH PhysioNet Resource.

Physiologic/Surrogate Data Test

In Fig. 3 we show two time series with the same means and standard deviations and identical power spectra. One is the heart rate time series of a healthy young subject and the other is a surrogate time series generated by a computational algorithm [13] that degrades the information content of the original signal through the process of phase randomization. Although both signals have the same statistical properties as measured by conventional biostatistical methods including Fourier spectral analysis, their underlying mechanisms are very different. The surrogate time series is less complex than the physiologic time series.

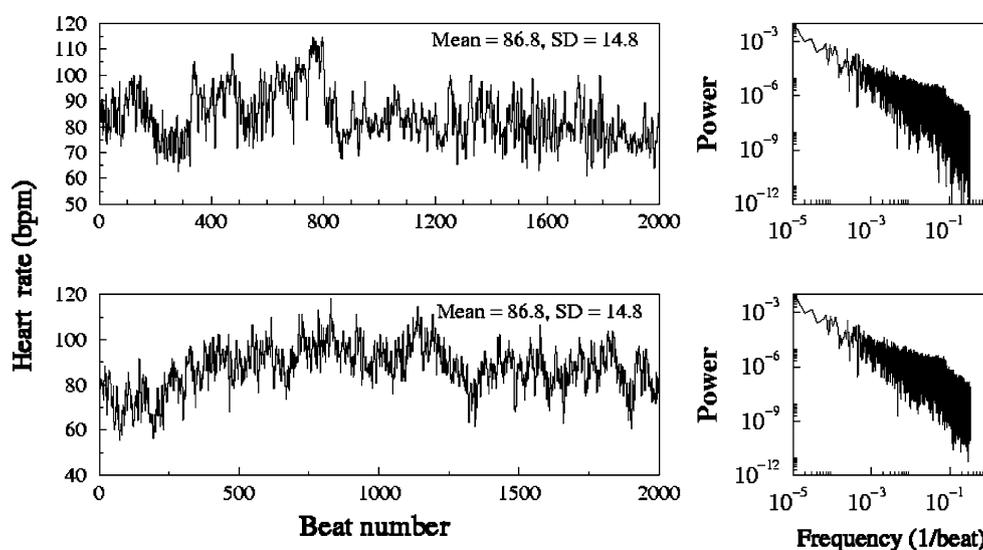


Fig. 3 Two time series with fundamentally distinct dynamics but the same mean values and standard deviations, and identical power spectra. *Top panels:* the heart rate time series (day time) from a young healthy subject and its Fourier power spectrum. *Bottom panels:* the surrogate, phase randomized series of the physiologic signal shown in the top panel and its Fourier power spectrum. The average value of entropy per scale (arbitrary units) is consistently higher for the physiologic (1.55) than for the surrogate time series (1.46). Similarly,

the asymmetry index is also higher for the physiologic time series (0.37) than for the surrogate series (0.04). Thus, in contrast to traditional time and frequency measures, both the multiscale entropy and time irreversibility algorithms capture the differences between the physiologic and the surrogate time series. Furthermore the results are consistent with a degradation of complexity (nonlinearity) with the phase randomization process

However, conventional time and frequency domain measures fail to fully quantify the information content of these signals, which is subtly encoded in the temporal sequence of the values. In contrast, both the multiscale entropy and time irreversibility metrics distinguish the two time series. Of interest, multiscale entropy and time irreversibility are consistently higher for the physiologic than for the surrogate time series.

Tests on Physiologic Models of Heart Rate Variability

Models of physiologic control should account for the complex structure of physiologic time series, but in general they fail to do so. For example, in 2002, the NIH Research Resource for Complex Physiologic Signals (<http://www.physionet.org>), in conjunction with Computers in Cardiology (<http://www.cinc.org>), sponsored an international competition on cardiac interbeat interval time series modeling. This challenge had two parts. Part I called for the development of models that could generate dynamics simulating those observed in heartbeat time series from healthy subjects. Participants entering Part II were challenged to distinguish between physiologic and model-generated time series. Using only the multiscale entropy method, we were able to correctly identify the origin of 48 out of 50 synthetic time series [14]. Of note, most of the models proposed failed to account for the multiscale, fractal properties of the cardiac interbeat interval time series.

In Fig. 4, we show the time irreversibility analysis for the physiologic and synthetic time series. None of the models proposed generated time irreversible signals. Of particular interest, the time series numbered 14 and 16, which have the most marked negative time asymmetry values, were not generated by any model. Instead, they were obtained by time-reversing two physiologic series [15]. This result shows that all proposed models fail to generate time irreversible signals across a wide range of scales, a dynamical signature of real-world systems with multiscale, nonequilibrium properties. Our algorithm indicates that datasets #23 and #49 have a similar degree of asymmetry. The relatively lower asymmetry value computed for the physiologic time series (#23) could be due to unknown factors such as age, level of physical activity, drugs effects, etc. In any case, we do not expect the time irreversibility algorithm to provide absolute discrimination between different classes of time series.

Discussion and Conclusions

The key finding of this presentation is that two recently developed multiscale techniques are able to extract

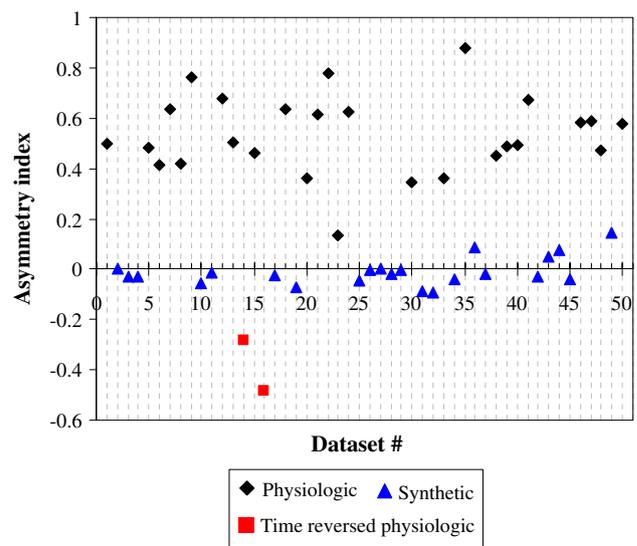


Fig. 4 Time irreversibility analysis of 50 cardiac interbeat intervals time series from the PhysioNet/CinC2002 Challenge database. Note that the physiologic time series are more time asymmetric (asymmetry index = 0.53 ± 0.16) than those generated by computational models (asymmetry index = -0.01 ± 0.06 .) Time series numbered 14 and 16 (as given in <http://www.physionet.org/challenge/2002/>) were generated by time-reversing two physiologic time series. The time irreversibility algorithm identifies these time series. Although the absolute values of the degree of irreversibility for the time reversed signals were within the physiologic range, their sign is negative indicating that time does not run in the “correct” direction

information from cardiac interbeat interval time series not contained in traditional time and frequency domain techniques. The complementary role of conventional and newer multiscale analytical methods in cardiovascular signal analysis is a frontier area of investigation. Another technique that may be useful in this regard is multifractal analysis, which is described elsewhere [16]. Quantifying multiscale variability and time irreversibility may have important applications, including risk stratification, aging effects and assessment of therapeutic interventions and development of models of biologic control. These two measures quantify different properties of complex systems and are independent of each other. In addition they are independent of other conventional heart rate time and frequency domain measures.

Limitations

Both the multiscale entropy and time irreversibility require relatively long time series (a couple of thousand data points) and the data need to be stationary. These two requirements are often inter-related. For example, what appears to be a slow drift of the baseline, i.e., a trend without apparent physiologic meaning when we follow a process for a shorter period, may, in fact, be a dynamical

pattern that is only revealed upon tracking the process over a more extended time interval.

Determining whether a “real-world” time series is stationary or not, requires the definition of a time scale of interest. In general, a process is stationary only if the characteristic scale of the dynamical structures it generates is much smaller than the recording time. For “relatively” stationary time series, the number of data points required to calculate sample entropy is of the order of a couple of hundred [4, 6]. One way of working around the nonstationarity problem is to detrend the data to eliminate those structures whose characteristic scales are not substantially much shorter (at least two orders of magnitude) than the recording length [17]. Further, defining the utility and limitations of newer multiscale complexity measures requires testing on open-access databases. Finally, we note that efforts to develop real world models of short and longer-term cardiovascular regulation need to account for the observed multiscale properties.

Acknowledgements We gratefully acknowledge support from the NIH Research Resource for Complex Physiologic Signals (NIBIB and NIGMS), the G. Harold and Leila Y. Mathers Charitable Foundation, the James S. McDonnell Foundation, the Ellison Medical Foundation, and the Defense Advanced Research Projects Agency (HR0011-05-1-0057).

References

- Goldberger AL, Amaral LA, Hausdorff JM, Ivanov PCh, Peng CK, Stanley HE. Fractal dynamics in physiology: alterations with disease and aging. *Proc Natl Acad Sci U S A* 2002;19(99 Suppl 1): 2466–72.
- Buchman TG. The community of the self. *Nature* 2002;420: 246–51.
- Cover TM, Thomas JA. Elements of information theory. 2nd ed. New York: John Wiley & Sons, Inc; 1991.
- Grassberger P. In: Atmanspacher H, Scheingraber H, editors. Information dynamics. New York: Plenum; 1991.
- Costa M, Goldberger AL, Peng C-K. Multiscale entropy analysis of complex physiologic time series. *Phys Rev Lett* 2002;89:068102-1-4.
- Costa M, Goldberger AL, Peng C-K. Multiscale entropy analysis of biological signals. *Phys Rev E* 2005;71:021906-1-18.
- Costa M, Goldberger AL, Peng C-K. Broken asymmetry of the human heartbeat: loss of time irreversibility in aging and disease. *Phys Rev Lett* 2005;95:198102-1-4.
- Richman JS, Moorman JR. Physiological time-series analysis using approximate entropy and sample entropy. *Am J Physiol* 2000;278:H2039-49.
- Zhang Y-C. Complexity and 1/f noise. A phase space approach. *J Phys I* 1991;1:971–7.
- Prigogine I, Antoniou I. Laws of nature and time symmetry breaking. *Ann NY Acad Sci* 1999;879:8–28.
- Weiss G. Time-reversibility of linear stochastic processes. *J Appl Probab* 1975;12:831–6.
- Jou D, Casas-Vazquez J, Lebon G. Extended irreversible thermodynamics. Berlin: Springer; 2001.
- Schreiber T, Schmitz A. Improved surrogate data for nonlinearity tests. *Phys Rev Lett* 1996;77:635–8.
- Costa M, Goldberger AL, Peng C-K. Multiscale entropy to distinguish physiologic and synthetic RR time series *Comput Cardiol* 2002;29:137–40.
- Moody GB. RR interval time series modeling: the PhysioNet/Computers in Cardiology Challenge 2002. *Comput Cardiol* 2002;29:125–8.
- Ivanov PCh, Amaral LAN, Goldberger AL, Havlin S, Rosenblum MG, Struzik Z, Stanley HE. Multifractality in human heartbeat dynamics. *Nature* 1999;399:461–5.
- Costa M, Priplata AA, Lipsitz LA, Wu Z, Huang NE, Goldberger AL, Peng C-K. Noise and poise: Enhancement of postural complexity in the elderly with a stochastic resonance-based therapy. *Europhys Lett* 2007;77:68008-1-5.