

Chapter 6

HEART RATE VARIABILITY: MEASURES AND MODELS

Malvin C. Teich, Steven B. Lowen, Bradley M. Jost,
Karin Vibe-Rheymer, and Conor Heneghan

1. INTRODUCTION

The human heart generates the quintessential biological signal: the heartbeat. A recording of the cardiac-induced skin potentials at the body's surface, an electrocardiogram (ECG), reveals information about atrial and ventricular electrical activity. Abnormalities in the temporal durations of the segments between deflections, or of the intervals between waves in the ECG, as well as their relative heights, serve to expose and distinguish cardiac dysfunction. Because the electrical activity of the human heart is influenced by many physiological mechanisms, electrocardiography has become an invaluable tool for the diagnosis of a variety of pathologies that affect the cardiovascular system [1]. Electrocardiologists have come to excel at visually interpreting the detailed form of the ECG wave pattern and have become adept at differential diagnoses.

Readily recognizable features of the ECG wave pattern are designated by the letters P-QRS-T; the wave itself is often referred to as the *QRS complex*. Aside from the significance of various features of the QRS complex, the timing of the *sequence* of QRS complexes over tens, hundreds, and thousands of heartbeats is also significant. These inter-complex times are readily measured by recording the occurrences of the peaks of the large R waves, which are, perhaps, the most distinctive feature of the normal ECG.

In this chapter we focus on various measures of the fluctuations of this sequence of interbeat intervals and how such fluctuations can be used to assess the presence or likelihood of cardiovascular disease [2]. This approach has come to be called *heart rate variability* (HRV) analysis [3, 4] even when it is the time *intervals* whose fluctuations are studied (heart rate has units of inverse time rather than time). HRV analysis serves as a marker for cardiovascular disease because cardiac dysfunction is often manifested by systematic changes in the variability of the RR-interval sequence relative to that of normal controls [1, 3, 5, 6]. A whole host of HRV measures, some scale dependent and others scale independent, have been developed and examined over the years in an effort to develop readily available, inexpensive, and noninvasive measures of cardiovascular function.

We examine 16 HRV measures and their suitability for correctly classifying ECG records of various lengths as normal or revealing the presence of cardiac dysfunction. Particular attention is devoted to HRV measures that are useful for discriminating patients with congestive heart failure from normal subjects. Using receiver operating

characteristic (ROC) analysis, we demonstrate that scale-dependent HRV measures (e.g., wavelet and spectral measures) are substantially superior to scale-independent measures (such as wavelet and spectral fractal exponents) for discriminating these two classes of data over a broad range of record lengths. The wavelet-transform standard deviation at a scale near 32 heartbeat intervals and its spectral counterpart near $1/32$ cycles/interval turn out to provide reliable results using ECG records just minutes long.

A long-standing issue of importance in cardiac physiology is the determination of whether the normal RR sequence arises from a chaotic attractor or has an underlying stochastic origin [6]. We present a phase-space analysis in which *differences* between adjacent RR intervals are embedded. This has the salutary effect of removing most of the correlation in the time series, which is well known to be deleterious to the detection of underlying deterministic dynamics. We demonstrate that RR sequences, from normal subjects and from patients with cardiac dysfunction alike, have stochastic rather than deterministic origins, in accord with our earlier conclusions [7, 8].

Finally, we develop a mathematical point process that emulates the human heartbeat time series for both normal subjects and patients with heart failure. Using simulations, we show that a jittered integrate-and-fire model built around a fractal-Gaussian-noise kernel provides a realistic, although not perfect, simulation of real heartbeat sequences. A construct of this kind may well be useful in a number of venues, including pacemaker excitation.

2. METHODS AND MEASURES

2.1. The Heartbeat Sequence as a Point Process

The statistical behavior of the sequence of heartbeats can be studied by replacing the complex waveform of an individual heartbeat recorded in the ECG (an entire QRS complex) with the time of occurrence of the contraction (the time of the peak of the R phase), which is a single number [8, 9]. In mathematical terms, the heartbeat sequence is then modeled as an unmarked point process. This simplification greatly reduces the computational complexity of the problem and permits us to use the substantial methodology that exists for point processes [10–12].

The occurrence of a contraction at time t_i is therefore simply represented by an impulse $\delta(t - t_i)$ at that time, where δ is the Dirac delta function, so that the sequence of heartbeats is represented by

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

A realization of a point process is specified by the set of occurrence times $\{t_i\}$ of the events. A single realization of the data is often all that is available to the observer, so the identification of the point process, and the elucidation of the mechanisms that underlie it, must be gleaned from this one realization.

One way in which the information in an experimental point process can be made more digestible is to reduce the data into a statistic that emphasizes a particular aspect of the data (at the expense of other features). These statistics fall into two broad classes

that derive from the sequence of interevent intervals and the sequence of counts, as illustrated in Figure 1 [10, 13].

Figure 1 illustrates how an ECG may be analyzed to obtain the sequence of inter-beat intervals as well as the sequence of counts. Figure 1a illustrates an ECG (sequence of QRS complexes) recorded from a patient. The R waves are schematically represented by a sequence of vertical lines, as shown in Figure 1b. The time between the first two R waves is τ_1 , the first RR (or interbeat) interval as indicated by the horizontal arrows in this figure. The time between the second and third R waves is τ_2 , and so forth. In Figure 1c, the time axis is divided into equally spaced, contiguous time windows, each of duration T seconds, and the (integer) number of R waves that fall in the i th window is counted and denoted N_i . This sequence $\{N_i\}$ forms a discrete-time random counting process of nonnegative integers. Varying the duration T yields a family of sequences $\{N_i\}(T)$. The RR intervals $\{\tau_i\}$ themselves also form a sequence of positive real-valued random numbers, which is shown schematically in Figure 1d. Here the abscissa is the interval number, which is not a simple function of time.

In this section we examine several statistical measures (including some that are novel) to characterize these stochastic processes; the development is assisted by an understanding of point processes.

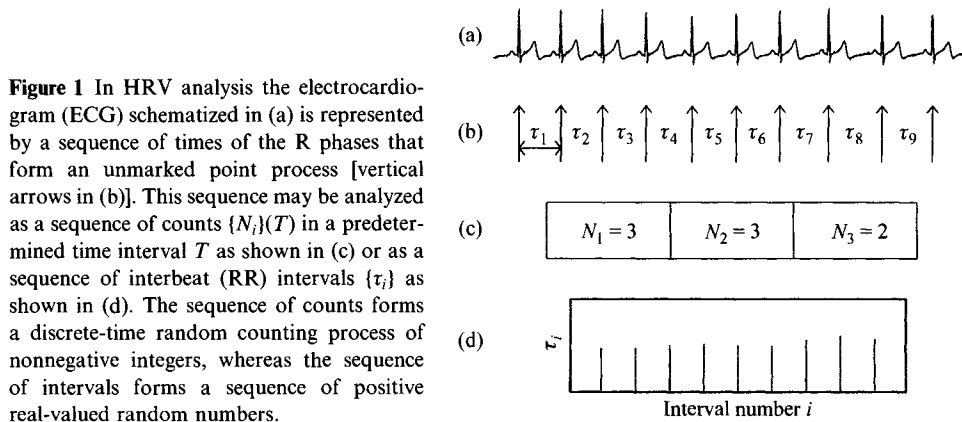


Figure 1 In HRV analysis the electrocardiogram (ECG) schematized in (a) is represented by a sequence of times of the R phases that form an unmarked point process [vertical arrows in (b)]. This sequence may be analyzed as a sequence of counts $\{N_i\}(T)$ in a predetermined time interval T as shown in (c) or as a sequence of interbeat (RR) intervals $\{\tau_i\}$ as shown in (d). The sequence of counts forms a discrete-time random counting process of nonnegative integers, whereas the sequence of intervals forms a sequence of positive real-valued random numbers.

2.1.1. Conventional Point Processes

The homogeneous Poisson point process, perhaps the simplest of all stochastic point processes, is described by a single parameter, the rate λ . This point process is memoryless: the occurrence of an event at any time t_0 is independent of the presence (or absence) of events at other times $t \neq t_0$. Because of this property, both the intervals $\{\tau_i\}$ and counts $\{N_i\}$ form sequences of independent, identically distributed random variables. The homogeneous Poisson point process is therefore completely characterized by the interevent-interval distribution (also referred to as the interbeat-interval histogram), which is exponential, or the event-number distribution (also referred to as the counting distribution), which is Poisson, together with the property of being independent. This process serves as a benchmark against which other point processes are measured; it therefore plays the role that the white Gaussian process enjoys in the realm of continuous-time stochastic processes.

A related point process is the nonparalyzable fixed-dead-time-modified Poisson point process, a close cousin of the homogeneous Poisson point process that differs only by the imposition of a dead-time (refractory) interval after the occurrence of each event, during which other events are prohibited from occurring [10, 14]. Another cousin is the gamma- r renewal process which, for integer r , is generated from a homogeneous Poisson point process by permitting every r th event to survive while deleting all intermediate events [10, 15]. Both the dead-time-modified Poisson point process and the gamma- r renewal process require two parameters for their description.

Some point process exhibit no dependences among their interevent intervals at the outset, in which case the sequence of interevent intervals forms a sequence of identically distributed random variables and the point process is completely specified by its interevent-interval histogram, i.e., its first-order statistic. Such a process is called a renewal process [10], a definition motivated by the replacement of failed parts, each replacement of which forms a renewal of the point process. Both examples of point processes presented above belong to the class of renewal point processes.

The interevent-interval histogram is, perhaps, the most commonly used of all statistical measures of point processes in the life sciences. The interevent-interval histogram estimates the interevent-interval probability density function $p_{\tau}(\tau)$ by computing the relative frequency of occurrence of interevent intervals as a function of interval size. Its construction involves the loss of interval ordering and therefore of information about dependences among intervals; a reordering of the sequence does not alter the interevent-interval histogram because the order plays no role in the relative frequency of occurrence.

The interevent-interval probability density function for the homogeneous Poisson point process assumes the exponential form

$$p_{\tau}(\tau) = \lambda \exp(-\lambda \tau) \quad (2)$$

where λ is the mean number of events per unit time. The interevent-interval mean and variance are readily calculated to be $E[\tau] = \int_0^{\infty} \tau p_{\tau}(\tau) d\tau = 1/\lambda$ and $\text{Var}(\tau) = E[\tau^2] - E^2[\tau] = 1/\lambda^2$, respectively, where $E[\cdot]$ represents expectation over the quantity inside the brackets. The interevent-interval probability density function for the dead-time-modified Poisson point process exhibits the same exponential form as for the homogeneous Poisson point process but is truncated at short interevent intervals as a result of the dead time [10]:

$$p_{\tau}(\tau) = \begin{cases} 0 & \tau < \tau_d \\ \lambda \exp[-\lambda(\tau - \tau_d)] & \tau \geq \tau_d \end{cases} \quad (3)$$

Here τ_d is the dead time and λ is the rate of the process before dead time is imposed.

If a process is nonrenewal, so that dependencies exist among its interevent intervals, then the interevent-interval histogram does not completely characterize the process [13]. In this case, measures that reveal the nature of the dependencies provide information that is complementary to that contained in the interevent-interval histogram. The heartbeat time series is such a nonrenewal process.

2.1.2. Fractal and Fractal-Rate Point Processes

The complete characterization of a stochastic process involves a description of all possible joint probabilities of the various events occurring in the process. Different statistics provide complementary views of the process; no single statistic can in general describe a stochastic process completely. Fractal stochastic processes exhibit scaling in their statistics. Such scaling leads naturally to power-law behavior, as demonstrated in the following. Consider a statistic w , such as the Allan factor for long counting times (see Section 2.5.1), which depends continuously on the scale x over which measurements are taken [16, 17]. Suppose changing the scale by any factor a effectively scales the statistic by some other factor $g(a)$, related to the factor but independent of the original scale:

$$w(ax) = g(a)w(x) \quad (4)$$

The only nontrivial solution of this scaling equation, for real functions and arguments, that is independent of a and x is

$$w(x) = bx^c \quad \text{with} \quad g(x) = x^c \quad (5)$$

for some constants b and c [16–18]. Thus, statistics with power-law forms are closely related to the concept of a fractal [19–21]. The particular case of fixed a admits a more general solution [22]:

$$g(x; a) = x^c \cos[2\pi \ln(x)/\ln(a)] \quad (6)$$

Consider once again, for example, the interevent-interval histogram. This statistic highlights the behavior of the times between adjacent events but reveals none of the information contained in the relationships among these times, such as correlation between adjacent time intervals. If the interevent-interval probability density function follows the form of Eq. 5 so that $p(\tau) \sim \tau^c$ over a certain range of τ where $c < -1$, the process is known as a fractal renewal point process [17, 19], a form of fractal stochastic process.

A number of statistics may be used to describe a fractal stochastic point process, and each statistic that scales will in general have a different scaling exponent c . Each of these exponents can be simply related to a more general parameter α , the fractal exponent, where the exact relation between these two exponents will depend upon the statistic in question. For example, the exponent c of the interevent-interval probability density function defined above is related to the fractal exponent α by $c = -(1 + \alpha)$. As the fractal exponent is a constant that describes the overall scaling behavior of a statistic, it does not depend on the particular scale and is therefore *scale independent*. Scale-independent measures are discussed in Sections 2.6 and 3.6.1.

Sample functions of the fractal renewal point process are true fractals; the expected value of their generalized dimensions assumes a nonintegral value between the topological dimension (zero) and the Euclidean dimension (unity) [19].

The sequence of unitary events observed in many biological and physical systems, such as the heartbeat sequence, do not exhibit power-law distributed interevent-interval histograms but nevertheless exhibit scaling in other statistics. These processes therefore have integral generalized dimensions and are consequently not true fractals. They may

nevertheless be endowed with rate functions that are either fractals or their increments: fractal Brownian motion, fractal Gaussian noise, or other related processes. Therefore, such point processes are more properly termed *fractal-rate* stochastic point processes [17]. It can be shown by surrogate data methods, e.g., shuffling the order of the intervals (see Section 5.1.3), that it is the ordering and not the relative interval sizes that distinguish these point processes [8].

2.2. Standard Frequency-Domain Measures

A number of HRV measures have been used as standards in cardiology, both for purposes of physiological interpretation and for clinical diagnostic applications [3]. We briefly describe some of the more commonly used measures that we include in this chapter for comparison with several novel measures that have been recently developed.

Fourier transform techniques provide a method for quantifying the correlation properties of a stochastic process through spectral analysis. Two definitions of power spectral density have been used in the analysis of HRV [9]. A *rate*-based power spectral density $S_\lambda(f)$ is obtained by deriving an underlying random continuous process $\lambda(t)$, the heart rate, based on a transformation of the observed RR interbeat intervals. The power spectral density of this random process is well defined, and standard techniques may be used for estimating the power spectral density from a single observation of $\lambda(t)$. An advantage of this technique is that the power spectral density thus calculated has temporal frequency as the independent variable, so that spectral components can be interpreted in terms of underlying physiological processes with known time scales. The power spectral density itself is usually expressed in units of sec^{-1} . However, the choice of how to calculate $\lambda(t)$, the underlying rate function, may influence the calculated power spectral density.

The second spectral measure that is more widely used is an *interval*-based power spectral density $S_\tau(f)$ that is directly calculated from measured RR interbeat intervals without transformation [9]. In this case, the intervals are treated as discrete-index samples of an underlying random process, and there is no intermediate calculation of an underlying rate function. The power spectral density in this case has cycles/interval as the independent variable and therefore has units of $\text{sec}^2/\text{interval}$.

The two types of power spectral densities are easily confused and care must be taken in their interpretation. For example, one could mistakenly interpret the abscissa of an interval-based power spectral density plot as being equivalent to temporal frequency (e.g., cycles/s). Although this is generally incorrect, for point processes whose interevent-interval coefficient of variation is relatively small [9], the interval-based and rate-based power spectral density plots can be made approximately equivalent by converting the interval-based frequency f_{int} (in cycles/interval) to the time-based frequency f_{time} (in cycles/s) using

$$f_{\text{time}} = f_{\text{int}}/E[\tau] \quad (7)$$

For typical interbeat-interval sequences, the coefficient of variation is indeed relatively small and this conversion can be carried out without the introduction of significant error [9]. In the remainder of this chapter we work principally with the interval-based power-spectral density. We use the notation $f \equiv f_{\text{int}}$ for the interval-based frequency (cycles/interval) and retain the notation f_{time} for temporal frequency (cycles/s).

We make use of nonparametric technique for estimating the spectral density. A simple reliable method for estimating the power spectral density of a process from a set of discrete samples $\{\tau_i\}$ is to calculate the averaged periodogram [23–25]. The data are first divided into K nonoverlapping blocks of L samples. After the optional use of a Hanning window, the discrete Fourier transform of each block is calculated and squared. The results are then averaged to form the estimate

$$\hat{S}_\tau(f) \equiv \frac{1}{K} \sum_{k=1}^K |\tilde{\tau}_k(f)|^2 \quad (8)$$

Here $\tilde{\tau}_k(f)$ is the discrete Fourier transform of the k th block of data and the hat explicitly indicates that we are dealing with an *estimate* of $S_\tau(f)$, which is called an averaged periodogram.

The periodogram covers a broad range of frequencies that can be divided into bands that are relevant to the presence of various cardiac pathologies. The power within a band is calculated by integrating the power spectral density over the associated frequency range. Some commonly used measures in HRV are [3]:

VLF. The power in the very-low-frequency range: 0.003–0.04 cycles/interval. Physiological correlates of the VLF band have not been specifically identified [3].

LF. The power in the low-frequency range: 0.04–0.15 cycles/interval. The LF band may reflect both sympathetic and vagal activity but its interpretation is controversial [3].

HF. The power in the high-frequency range: 0.15–0.4 cycles/interval. Efferent vagal activity is a major contributor to the HF band [26–28].

LF/HF. The ratio of the power in the low-frequency range to that in the high-frequency range. This ratio may either mirror sympathovagal balance or reflect sympathetic modulations [3].

2.3. Standard Time-Domain Measures

We consider three time-domain measures commonly used in HRV analysis. The first and last are highly correlated with each other inasmuch as they estimate the high-frequency variations in the heart rate [3]. They are:

pNN50. The relative proportion of successive NN intervals (normal-to-normal intervals, i.e., all intervals between adjacent QRS complexes resulting from sinus node depolarizations [3]) with interval differences greater than 50 ms.

SDANN. The Standard Deviation of the Average NN interval calculated in 5-minute segments. It is often calculated over a 24-hour period. This measure estimates fluctuations over frequencies smaller than 0.003 cycles/sec.

SDNN (σ_{int}). The Standard Deviation of the NN interval set $\{\tau_i\}$ specified in units of seconds. This measure is one of the more venerable among the many scale-dependent measures that have long been used for HRV analysis [3, 5, 29, 30].

2.4. Other Standard Measures

There are several other well-known measures that have been considered for HRV analysis. For completeness, we briefly mention two of them here: the event-number

histogram and the Fano factor [7, 8]. Just as the interevent-interval histogram provides an estimate of the probability density function of interevent-interval magnitude, the event-number histogram provides an estimate of the probability mass function of the number of events. Construction of the event-number histogram, like that of the interevent-interval histogram, involves loss of information, in this case the ordering of the counts. However, whereas the time scale of information contained in the interevent-interval histogram is the mean interevent interval, which is intrinsic to the process under consideration, the event-number histogram reflects behavior occurring on the adjustable time scale of the counting window T . The Fano factor, which is the variance of the number of events in a specified counting time T divided by the mean number of events in that counting time, is a measure of correlation over different time scales T . This measure is sometimes called the index of dispersion of counts [31]. In terms of the sequence of counts illustrated in Figure 1, the Fano factor is simply the variance of $\{N_i\}$ divided by the mean of $\{N_i\}$.

2.5. Novel Scale-Dependent Measures

The previous standard measures are all well-established scale-dependent measures. We now describe a set of recently devised scale-dependent measures whose performance we evaluate. Throughout this chapter, when referring to intervals, we denote the fixed scale as m ; when referring to time, we employ T .

2.5.1. Allan Factor $A(T)$

In this section we present a measure we first defined in 1996 [32] and called the Allan factor. We quickly found that this quantity was a useful measure of HRV [8]. The Allan factor is the ratio of the event-number Allan variance to twice the mean:

$$A(T) \equiv \frac{E\{[N_{i+1}(T) - N_i(T)]^2\}}{2E\{N_{i+1}(T)\}} \quad (9)$$

The Allan variance, as opposed to the ordinary variance, is defined in terms of the variability of *successive* counts [17, 33, 34]. As such, it is a measure based on the Haar wavelet. The Allan variance was first introduced in connection with the stability of atomic-based clocks [33]. Because the Allan factor functions as a derivative, it has the salutary effect of mitigating against linear nonstationarities.

The Allan factor of a point process generally varies as a function of the counting time T ; the exception is the homogeneous Poisson point process. For a homogeneous Poisson point process, $A(T) = 1$ for any counting time T . Any deviation from unity in the value of $A(T)$ therefore indicates that the point process in question is not Poisson in nature. An excess above unity reveals that a sequence is less ordered than a homogeneous Poisson point process, whereas values below unity signify sequences which are more ordered. For a point process without overlapping events, the Allan factor approaches unity as T approaches zero.

A more complex wavelet Allan factor can be constructed to eliminate polynomial trends [35–37]. The Allan variance, $E[(N_{i+1} - N_i)^2]$, may be recast as the variance of the integral of the point process under study multiplied by the following function:

$$\psi_{\text{Haar}}(t) = \begin{cases} -1 & \text{for } -T < t < 0 \\ +1 & \text{for } 0 < t < T \\ 0 & \text{otherwise} \end{cases} \quad (10)$$

Equation 10 defines a scaled wavelet function, specifically the Haar wavelet. This can be generalized to any admissible wavelet $\psi(t)$; when suitably normalized the result is a wavelet Allan factor [36, 38].

2.5.2. Wavelet-Transform Standard Deviation [$\sigma_{\text{wav}}(m)$]

Wavelet analysis has proved to be a useful technique for analyzing signals at multiple scales [39–44]. It permits the time and frequency characteristics of a signal to be simultaneously examined and has the advantage of naturally removing polynomial nonstationarities [36, 37, 45]. The Allan factor served in this capacity for the counting process $\{N_i\}$, as discussed earlier. Wavelets similarly find use in the analysis of RR-interval series. They are attractive because they mitigate against the nonstationarities and slow variations inherent in the interbeat-interval sequence. These arise, in part, from the changing activity level of the subject during the course of a 24-hour period.

Wavelet analysis simultaneously gives rise to both scale-dependent and scale-independent measures [46], affording the experimenter an opportunity to compare the two approaches. In this latter capacity wavelet analysis provides an estimate of the wavelet-transform fractal (scaling) exponent α_w [46, 47], as discussed in the context of HRV in Sections 2.6.2 and 3.6.1. As a result of these salutary properties, we devote particular attention to the wavelet analysis of HRV in this chapter.

A dyadic discrete wavelet transform for the RR-interval sequence $\{\tau_i\}$ may be defined as [41–43]

$$W_{m,n}(m) = \frac{1}{\sqrt{m}} \sum_{i=0}^{L-1} \tau_i \psi(i/m - n) \quad (11)$$

The quantity ψ is the wavelet basis function, and L is the number of RR intervals in the set $\{\tau_i\}$. The scale m is related to the scale index j by $m = 2^j$. Both j and the translation variable n are nonnegative integers. The term dyadic refers to the use of scales that are integer powers of 2. This is an arbitrary choice; the wavelet transform could be calculated at arbitrary scale values, although the dyadic scale enjoys a number of convenient mathematical properties [41, 42].

The dyadic discrete wavelet transform calculated according to this prescription generates a three-dimensional space from a two-dimensional signal graph. One axis is time or, in our case, the RR-interval number i ; the second axis is the scale m ; and the third axis is the strength of the wavelet component. Pictorially speaking, the transform gives rise to a landscape whose longitude and latitude are RR-interval number and scale of observation, while the altitude is the value of the discrete wavelet transform at the interval i and the scale m .

Figure 2 provides an example of such a wavelet transform, where $\psi(x)$ is the simple Haar wavelet. Figure 2a illustrates the original wavelet, a function that is by definition $\psi(x) = 1$ for x between 0 and 0.5; $\psi(x) = -1$ for x between 0.5 and 1; and $\psi(x) = 0$ elsewhere. Figure 2b illustrates the wavelet scaled by the factor $m = 16$, which causes it to last for 16 samples rather than 1; and delayed by a factor of $n = 3$ times the length of the wavelet, so that it begins at $x = nm = 48$. Figure 2c shows a sequence of interbeat-interval values multiplied by the scaled and shifted wavelet (the summand in Eq. 11). The abscissa is labeled i rather than x to indicate that we have a discrete-time process composed of the sequence $\{\tau_i\}$. In this particular example, only values of τ_i between $i = 48$ and 63 survive. Adding them (with the appropriate sign) provides the wavelet transform beginning at interval number $i = 48$ at a scale of $m = 16$.

For the Haar wavelet the calculation of the wavelet transform is therefore tantamount to adding the eight RR intervals between intervals 48 and 55 inclusive and then subtracting the eight subsequent RR intervals between intervals 56 and 63 inclusive, as illustrated in Figure 2c. Moving this window across interval number allows us to see how the wavelet transform evolves with interval number, whereas varying the scale of the window permits this variation to be observed over a range of resolutions, from fine to coarse (smaller scales allow the observation of more rapid variations, i.e., higher frequencies).

A simple measure that can be devised from the wavelet transformation is the standard deviation of the wavelet transform as a function of scale [46–49]:

$$\sigma_{\text{wav}}(m) = \left[E \left\{ |W_{m,n}(m) - E[W_{m,n}(m)]|^2 \right\} \right]^{1/2} \quad (12)$$

where the expectation is taken over the process of RR intervals and is independent of n . It is readily shown that $E[W_{m,n}(m)] = 0$ for all values of m so that a simplified form for the wavelet-transform standard deviation emerges:

$$\sigma_{\text{wav}}(m) = \left\{ E \left[|W_{m,n}(m)|^2 \right] \right\}^{1/2} \quad (13)$$

This quantity has been shown to be quite valuable for HRV analysis [46–50]. The special case obtained by using the Haar-wavelet basis and evaluating Eq. 13 at $m = 1$ yields the standard deviation of the difference between pairs of consecutive interbeat intervals. This special case is therefore identical to the well-known HRV measure

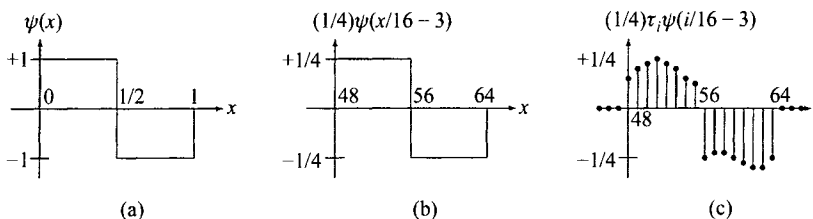


Figure 2 Estimating the wavelet transform using the Haar wavelet: (a) original Haar wavelet, (b) delayed and scaled version of the wavelet ($m = 16$, $n = 3$), and (c) time series multiplied by this wavelet.

referred to as **RMSSD** [3], an abbreviation for **Root-Mean-Square of Successive-interval Differences**.

Figure 3 provides an example in which the discrete wavelet transform is calculated using an RR-interval data set. In Figure 3a, the original RR interbeat-interval series is shown, and Figure 3b shows the dyadic discrete wavelet transform at three different scales as a function of RR-interval number. It is important and interesting to note that the trends and baseline variations present in the original time series have been removed by the transform. As illustrated in Figure 3b, the wavelet-transform standard deviation σ_{wav} typically increases with the scale m . When plotted versus scale, this quantity provides information about the behavior of the signal at all scales. In Section 3 we show how this measure can be effectively used to separate patients with heart failure from normal subjects.

2.5.3. Relationship of Wavelet [$\sigma_{\text{wav}}(m)$] and Spectral Measures [$S_{\tau}(f)$]

Is there a spectral measure equivalent to the wavelet-transform standard deviation? We proceed to show that the wavelet-transform standard deviation $\sigma_{\text{wav}}(m)$ and the interval-based power spectral density $S_{\tau}(f)$ are isomorphic [49], so the answer is yes under conditions of stationarity. Although their equivalence is most easily analyzed in the continuous domain, the results are readily translated to the discrete domain by interpreting the discrete wavelet transform as a discretized version of a continuous wavelet transform.

The continuous wavelet transform of a signal $\tau(t)$ is defined as

$$W_{\tau}(s, r) = \frac{1}{\sqrt{s}} \int_{-\infty}^{\infty} \tau(t) \psi^* \left(\frac{t-r}{s} \right) dt \quad (14)$$

where s and r are continuous-valued scale and translation parameters, respectively, ψ is a wavelet basis function, and $*$ denotes complex conjugation. Since $E[W_{\tau}] = 0$, the variance of W_{τ} at scale s is

$$D(s) = \sigma_{\text{wav}}^2(s) = E \left[|W_{\tau}(s, r)|^2 \right] \quad (15)$$

which can be written explicitly as

$$D(s) = E \left[\frac{1}{\sqrt{s}} \int_{-\infty}^{\infty} \tau(t) \psi^* \left(\frac{t-r}{s} \right) dt \frac{1}{\sqrt{s}} \int_{-\infty}^{\infty} \tau^*(t') \psi \left(\frac{t'-r}{s} \right) dt' \right] \quad (16)$$

For a wide-sense stationary signal the variance can be written as

$$D(s) = \frac{1}{s} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R(t-t') \psi^* \left(\frac{t-r}{s} \right) \psi \left(\frac{t'-r}{s} \right) dt dt' \quad (17)$$

where R is the autocorrelation function of τ . Routine algebraic manipulation then leads to

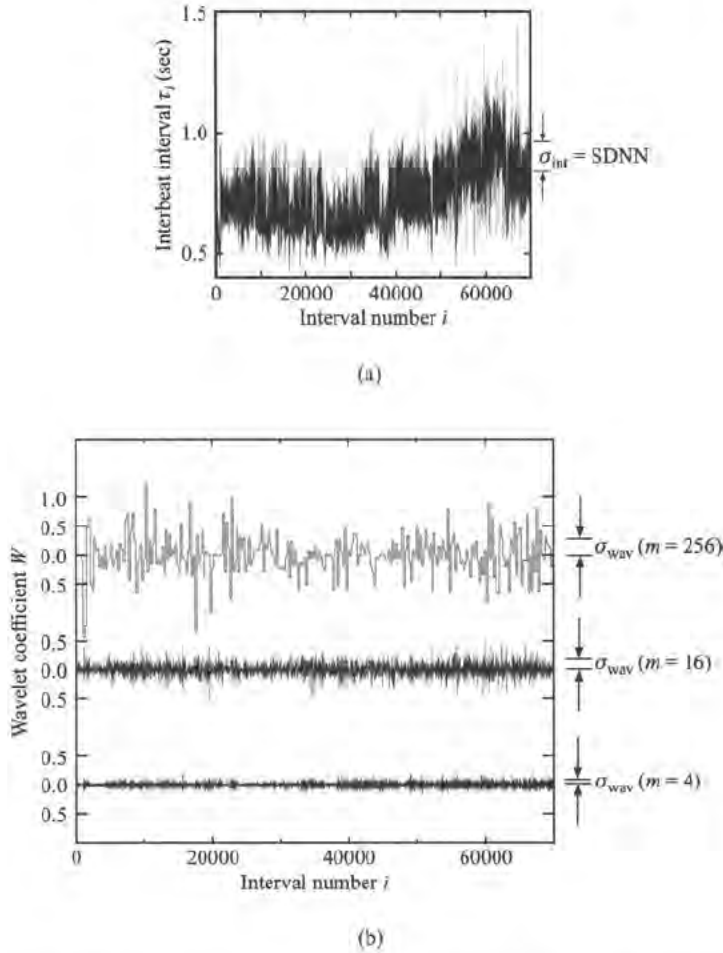


Figure 3 (a) Series of interbeat intervals τ_i versus interval number i for a typical normal patient (data set n16265). (Adjacent values of the interbeat interval are connected by straight lines to facilitate viewing.) Substantial trends are evident. The interbeat-interval standard deviation $\sigma_{int} \equiv \text{SDNN}$ is indicated. (b) Wavelet transform $W_{m,n}(m)$ (calculated using a Daubechies 10-tap analyzing wavelet) at three scales ($m = 4, 16, 256$) versus interval number $i (= mn)$ for the RR-interval sequence shown in (a). The trends in the original interbeat-interval time series are removed by the wavelet transformation. The wavelet-transform standard deviation $\sigma_{wav}(m)$ for this data set is seen to increase with the scale m .

$$D(s) = s \int_{-\infty}^{\infty} R(sy) W_{\psi}(1, y) dy \quad (18)$$

or, alternatively,

$$D(s) = s \int_{f=-\infty}^{\infty} S_{\tau}(f) \left[\int_{y=-\infty}^{\infty} W_{\psi}(1, y) \exp(j2\pi fsy) dy \right] df \quad (19)$$

where $W_\psi(1, y)$ is the wavelet transform of the wavelet itself (termed the wavelet kernel), and $S_\tau(f)$ is the power spectral density of the signal.

For the dyadic discrete wavelet transform that we have used, Eq. 19 becomes

$$\begin{aligned} D(m) &= \sigma_{\text{wav}}^2(m) = m \int_{f=-\infty}^{\infty} S_\tau(f) \left[\int_{y=-\infty}^{\infty} W_\psi(1, y) \exp(j2\pi f m y) dy \right] df \\ &= m \int_{-\infty}^{\infty} S_\tau(f) H(mf) df \end{aligned} \quad (20)$$

We conclude that for stationary signals the interval-based power spectral density $S_\tau(f)$ is directly related to the wavelet-transform standard deviation $\sigma_{\text{wav}}(m)$ through an integral transform. This important result has a simple interpretation: the factor in square brackets in Eq. 20 represents a bandpass filter $H(mf)$ that passes only spectral components in a bandwidth surrounding the frequency f_m that corresponds to the scale m . This is because the Fourier transform of a wavelet kernel is constrained to be bandpass in nature. For a discrete-index sequence, the sampling "time" can be arbitrarily set to unity so that a frequency f_m corresponds to $1/m$. We conclude that information obtained from a $D(m)$ -based statistic is also accessible through interval-based power spectral density measures. In Section 3 we explicitly show that the two measures are comparable in their abilities to discriminate patients with heart failure from normal subjects.

2.5.4. Detrended Fluctuation Analysis [DFA(m)]

Detrended fluctuation analysis (DFA) was originally proposed as a technique for quantifying the nature of long-range correlations in a time series [51–53]. As implied by its name, it was conceived as a method for detrending variability in a sequence of events. The DFA computation involves the calculation of the summed series

$$y(k) = \sum_{i=1}^k \{\tau_i - E[\tau]\} \quad (21)$$

where $y(k)$ is the k th value of the summed series and $E[\tau]$ denotes the average over the set $\{\tau_i\}$. The summed series is then divided into segments of length m and a least-squares fit is performed on each of the data segments, providing the trends for the individual segments. Detrending is carried out by subtracting the local trend $y_m(k)$ in each segment. The root-mean-square fluctuation of the resulting series is then

$$F(m) = \left\{ \frac{1}{L} \sum_{k=1}^L [y(k) - y_m(k)]^2 \right\}^{1/2} \quad (22)$$

The functional dependence of $F(m)$ is obtained by evaluations over all segment sizes m .

Although detrended fluctuation analysis was originally proposed as a method for estimating the scale-independent fractal exponent of a time series [51], as discussed in Section 2.6.1, we consider its merits as a scale-dependent measure. As will be demonstrated in Section 3, a plot of $F(m)$ versus m reveals a window of separation between

patients with congestive heart failure and normal subjects over a limited range of scales, much as that provided by the other scale-dependent measures discussed in this section. Because DFA is an ad hoc measure that involves nonlinear computations, it is difficult to relate it to other scale-dependent measures in the spirit of Eq. 20. Furthermore, as will become clear in Section 3.6.2, relative to other measures DFA is highly time intensive from a computational point of view.

2.6. Scale-Independent Measures

Scale-independent measures are designed to estimate fractal exponents that characterize scaling behavior in one or more statistics of a sequence of events, as discussed in Section 2.1.2. The canonical example of a scale-independent measure in HRV is the fractal exponent α_S of the interbeat-interval power spectrum, associated with the decreasing power-law form of the spectrum at sufficiently low frequencies f : $S_\tau(f) \propto f^{-\alpha_S}$ [3, 30, 54]. Other scale-independent measures have been examined by us [7, 8, 16, 17, 46] and by others [6, 51, 55, 56] in connection with HRV analysis. For exponent values encountered in HRV and infinite data length, all measures should in principle lead to a unique fractal exponent. In practice, however, finite data length and other factors introduce bias and variance, so that different measures give rise to different results. The performance of scale-independent measures has been compared with that of scale-dependent measures for assessing cardiac dysfunction [46, 47].

2.6.1. Detrended-Fluctuation-Analysis Power-Law Exponent (α_D)

The DFA technique, and its use as a scale-dependent measure, has been described in Section 2.5.4. A number of studies [51, 53, 56] have considered the extraction of power-law exponents from DFA and their use in HRV. As originally proposed [51], $\log[F(m)]$ is plotted against $\log(m)$ and scaling exponents are obtained by fitting straight lines to sections of the resulting curve—the exponents are simply the slopes of the linearly fitted segments on this doubly logarithmic plot. The relationship between the scaling exponents has been proposed as a means of differentiating normal from pathological subjects [51, 55, 56].

2.6.2. Wavelet-Transform Power-Law Exponent (α_W)

The use of the wavelet transform as a scale-dependent measure was considered in Section 2.5.2. It was pointed out that a scale-independent measure also emerges from the wavelet-transform standard deviation. The wavelet-transform fractal exponent α_W is estimated directly from the wavelet transform as twice the slope of the curve $\log[\sigma_{\text{wav}}(m)]$ versus $\log(m)$, measured at large values of m [46]. The factor of 2 is present because the fractal exponent is related to variance rather than to standard deviation.

2.6.3. Periodogram Power-Law Exponent (α_S)

The description of the periodogram as a scale-dependent measure was provided in Section 2.2. The periodogram fractal exponent α_S [17, 54] is obtained as the least-

square-fit slope of the spectrum when plotted on doubly logarithmic coordinates. The range of low frequencies over which the slope is estimated stretches between $10/L$ and $100/L$ where L is the length of the data set [17].

2.6.4. Allan-Factor Power-Law Exponent (α_A)

The use of the Allan factor as a scale-dependent measure was considered in Section 2.5.1. The Allan factor fractal exponent α_A [8, 17] is obtained by determining the slope of the best-fitting straight line, at large values of T , to the Allan factor curve (Eq. 9) plotted on doubly logarithmic coordinates. Estimates of α obtained from the Allan factor can range up to a value of 3 [32]. The use of wavelets more complex than the Haar enables an increased range of fractal exponents to be accessed, at the cost of a reduction in the range of counting time over which the wavelet Allan factor varies as T^{α_A} . In general, for a particular wavelet with regularity (number of vanishing moments) R , fractal exponents $\alpha < 2R + 1$ can be reliably estimated [36, 38]. For the Haar basis, $R = 1$, whereas all other wavelet bases have $R > 1$. A wavelet Allan factor making use of bases other than the Haar is therefore required for fractal-rate stochastic point processes for which $\alpha \geq 3$. For processes with $\alpha < 3$, however, the Allan factor appears to be the best choice [36, 38].

2.6.5. Rescaled-Range-Analysis Power-Law Exponent (α_R)

Rescaled range analysis [19, 57–59] provides information about correlations among blocks of interevent intervals. For a block of k interevent intervals, the difference between each interval and the mean interevent interval is obtained and successively added to a cumulative sum. The normalized range $R(k)$ is the difference between the maximum and minimum values that the cumulative sum attains, divided by the standard deviation of the interval size. $R(k)$ is plotted against k . Information about the nature and the degree of correlation in the process is obtained by fitting $R(k)$ to the function k^H , where H is the so-called Hurst exponent [57]. For $H > 0.5$ positive correlation exists among the intervals, whereas $H < 0.5$ indicates the presence of negative correlation; $H = 0.5$ obtains for intervals with no correlation. Renewal processes yield $H = 0.5$. For negatively correlated intervals, an interval that is larger than the mean tends, on average, to be preceded or followed by one smaller than the mean.

The Hurst exponent H is generally assumed to be well suited to processes that exhibit long-term correlation or have a large variance [19, 57–59], but there are limits to its robustness because it exhibits large systematic errors and highly variable estimates for some fractal sequences [6, 60, 61]. Nevertheless, it provides a useful indication of correlation in a sequence of events arising from the ordering of the interevent intervals alone.

The exponent α_R is ambiguously related to the Hurst exponent H , because some authors have used the quantity H to index fractal Gaussian noise whereas others have used the same value of H to index the integral of fractal Gaussian noise (which is fractional Brownian motion). The relationship between the quantities is $\alpha_R = 2H - 1$ for fractal Gaussian noise and $\alpha_R = 2H + 1$ for fractal Brownian motion. In the context of this work, the former relationship holds.

2.7. Estimating the Performance of a Measure

We have, to this point, outlined a variety of candidate measures for use in HRV analysis. The task now is to determine the relative value of these measures from a clinical perspective. We achieve this by turning to estimation theory [62].

A statistical measure obtained from a finite set of actual data is characterized by an estimator. The fidelity with which the estimator can approximate the true value of the measure is determined by its bias and variance. The bias is the deviation of the expected value of the estimator from its true underlying value (assuming that this exists), whereas the variance indicates the expected deviation from the mean. An ideal estimator has zero bias and zero variance, but this is not achievable with a finite set of data. For any unbiased estimator the Cramér–Rao bound provides a lower bound for the estimator variance; measures that achieve the Cramér–Rao bound are called efficient estimators. The estimator bias and variance play a role in establishing the overall statistical significance of conclusions based on the value returned by the estimator.

2.7.1. Statistical Significance: p , d' , h , and d

The concept of statistical significance extends the basic properties of bias and variance [63]. It provides a probabilistic interpretation of how likely it is that a particular value of the estimator might occur by chance alone, arising from both random fluctuations in the data and the inherent properties of the estimator itself.

A frequently used standard of statistical significance is the p value, the calculation of which almost always implicitly assumes a Gaussian-distributed dependent variable. A lower p value indicates greater statistical significance, and a measure is said to be statistically significant to a value of p_0 when $p < p_0$. The distributions obtained from HRV measurements are generally not Gaussian, however, so the usual method for estimating the p value cannot be used with confidence. Because other methods for estimating the p value require more data than is available, we do not consider this quantity further.

Another often-used distribution-dependent standard is the d' value. It serves to indicate the degree of separation between two distributions and has been widely used in signal detection theory and psychophysics, where the two distributions represent noise and signal plus noise [64]. The most common definition of d' is the difference in the means of two Gaussian distributions divided by their common standard deviation. Two closely related distribution-dependent cousins of d' are the detection distance h , defined as the difference in the means of the two Gaussian distributions divided by the square root of the sum of their variances, and the detection distance d , defined as the difference in the means of the two Gaussian distributions divided by the sum of their standard deviations. Larger values of d' , h , and d indicate improved separation between the two distributions and therefore reduced error in assigning an outcome to one or the other of the hypotheses.

Because HRV measures are intended to provide diagnostic information in a clinical setting and do not return Gaussian statistics, the evaluation of their performance using distribution-independent means is preferred. Two techniques for achieving this, positive and negative predictive values and receiver operating characteristic (ROC) analysis, are described next. Neither requires knowledge of the statistical distribution of the measured quantities and both are useful.

2.7.2. Positive and Negative Predictive Values

The performance of the various HRV measures discussed previously can be effectively compared using positive predictive values and negative predictive values, the proportion of correct positive and negative identifications, respectively. When there is no false positive (or negative) detection, the predictive value is equal to unity and there is perfect assignment. Furthermore, when the individual values of a measure for normal subjects and patients do not overlap, the predictive value curves are typically monotonic, either increasing or decreasing, with the threshold. A detailed discussion of positive and negative predictive values is provided in Section 3.4.

2.7.3. Receiver Operating Characteristic (ROC) Analysis

ROC analysis [62, 64] is an objective and highly effective technique for assessing the performance of a measure when it is used in binary hypothesis testing. This format provides that a data sample be assigned to one of two hypotheses or classes (e.g., pathologic or normal) depending on the value of some measured statistic relative to a threshold value. The efficacy of a measure is then judged on the basis of its sensitivity (the proportion of pathologic patients correctly identified) and its specificity (the proportion of normal subjects correctly identified). The ROC curve is a graphical presentation of sensitivity versus $1 - \text{specificity}$ as a threshold parameter is swept. Note that sensitivity and specificity relate to the status of the patients (pathologic and normal), whereas predictive values relate to the status of the identifications (positive and negative).

The area under the ROC curve serves as a well-established index of diagnostic accuracy [64]; the maximum value of 1.0 corresponds to perfect assignment (unity sensitivity for all values of specificity), whereas a value of 0.5 arises from assignment to a class by pure chance (areas < 0.5 arise when the sense of comparison is reversed). ROC analysis can be used to choose the best of a host of different candidate diagnostic measures by comparing their ROC areas, or to establish for a single measure the trade-off between data length and misidentifications (misses and false positives) by examining ROC area as a function of record length. A minimum record length can then be specified to achieve acceptable classification accuracy.

As pointed out before, ROC analysis relies on no implicit assumptions about the statistical nature of the data set [62], so it is generally more suitable [47] for analyzing non-Gaussian time series than are measures of statistical significance such as p value, h , and d . Another important feature of ROC curves is that they are insensitive to the units employed (e.g., spectral magnitude, magnitude squared, or log magnitude); ROC curves for a measure M are identical to those for any monotonic transformation thereof such as M^x or $\log(M)$. In contrast, the values of d' , h , and d are generally modified by such transformation, as demonstrated in Section 3.5.1.

3. DISCRIMINATING HEART-FAILURE PATIENTS FROM NORMAL SUBJECTS

We now proceed to examine the relative merits of various HRV measures for discriminating patients with congestive heart failure (CHF) from normal subjects.

Specifically, we contrast and compare the performance of the 16 measures set forth in Section 2: VLF, LF, HF, LF/HF, pNN50, SDANN, SDNN (σ_{int}), $A(T)$, $\sigma_{\text{wav}}(m)$, $S_t(f)$, DFA(m), α_D , α_W , α_S , α_A , and α_R .

After discussing the selection of an appropriate scale m , we use predictive value plots and ROC curves to select a particular subset of HRV markers that appears to be promising for discerning the presence of heart failure in a patient population.

3.1. Database

The RR recordings analyzed in this section were drawn from the Beth-Israel Hospital (Boston, MA) Heart-Failure Database, which includes 12 records from normal subjects (age 29–64 years, mean 44 years) and 12 records from patients with severe CHF (age 22–71 years, mean 56 years). The recordings were made with a Holter monitor digitized at a fixed value of 250 samples/s. Also included in this database are 3 RR records for CHF patients who also suffered from atrial fibrillation (AF); these records are analyzed as a separate class. All records contain both diurnal and nocturnal segments. The data were originally provided to us in 1992 by D. Rigney and A. L. Goldberger.

A detailed characterization of each of the records is presented in Table 1 of Ref. 8; some statistical details are provided in Table A1. Of the 27 recordings, the shortest contained $L_{\text{max}} = 75,821$ RR intervals; the remaining 26 recordings were truncated to this length before calculating the 16 HRV measures.

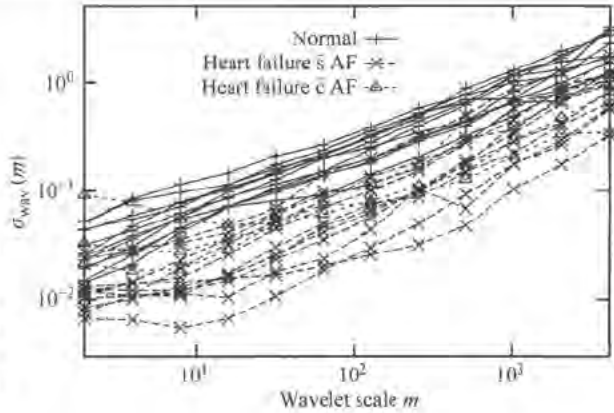
3.2. Selecting a Scale

A value for the scale m that suitably discriminates heart-failure patients from normal subjects can be inferred from our recent wavelet studies of the CHF and normal records from the same database as discussed in Section 3.1 [46, 47]. With the help of the wavelet-transform standard deviation $\sigma_{\text{wav}}(m)$ discussed in detail in Section 2.5.2, we discovered a critical scale window near $m = 32$ interbeat intervals over which the normal subjects exhibited greater fluctuations than those afflicted with heart failure. For these particular long data sets, we found that it was possible to discriminate perfectly between the two groups [46, 47, 48].

The results are displayed in Figure 4, where $\sigma_{\text{wav}}(m)$ is plotted against wavelet scale m for the 12 normal subjects (+), the 12 CHF patients without atrial fibrillation (\times), and the 3 CHF patients with atrial fibrillation (Δ), using Haar-wavelet analysis. The AF patients (Δ) typically fell near the high end of the non-AF patients (\times), indicating greater RR fluctuations, particularly at small scales. This results from the presence of nonsinus beats. Nevertheless it is evident from Figure 4 that the wavelet measure σ_{wav} completely separates the normal subjects from the heart-failure patients (both without and with AF) at scales of 16 and 32 heartbeat intervals, as reported in Ref. 46. One can do no better. This conclusion persists for a broad range of analyzing wavelets, from Daubechies 2-tap (Haar) to Daubechies 20-tap [46].

The importance of this scale window has been confirmed in an Israeli–Danish study of diabetic patients who had not yet developed clinical signs of cardiovascular disease [50]. The reduction in the value of the wavelet-transform standard deviation $\sigma_{\text{wav}}(32)$ that leads to the scale window occurs not only for CHF (both with and without AF) and diabetic patients but also for heart-transplant patients [48, 50] and also in records preceding sudden cardiac death [46, 48]. The depression of $\sigma_{\text{wav}}(32)$ at these scales is prob-

Figure 4 Haar-wavelet-transform standard deviation $\sigma_{\text{wav}}(m)$ vs scale m for the 12 normal subjects (+), 12 CHF patients without (̄) atrial fibrillation (×), and 3 CHF patients with (̄) atrial fibrillation (Δ). Each data set comprises the first 75,821 RR intervals of a recording drawn from the Beth-Israel Hospital heart-failure database. Complete separation of the normal subjects and heart-failure patients is achieved at scales $m = 16$ and 32 interbeat intervals.



ably associated with the impairment of autonomic nervous system function. Baroreflex modulations of the sympathetic or parasympathetic tone typically lie in the range 0.04–0.09 cycles/s (11–25 s), which corresponds to the scale where $\sigma_{\text{wav}}(m)$ is reduced.

These studies, in conjunction with our earlier investigations that revealed a similar critical scale window in the *counting* statistics of the heartbeat [8, 35] (as opposed to the *time-interval* statistics under discussion), lead to the recognition that scales in the vicinity of $m = 32$ enjoy a special status. The measures that depend on a particular scale are therefore evaluated at $m = 32$ and $f = 1/32$ in the expectation that these values maximize discriminability in the more usual situation in which the two classes of data cannot be fully separated.

3.3. Individual Value Plots

Having devised a suitable scale value m , we now proceed to evaluate the 16 measures for all 27 normal and CHF data sets, each comprising 75821 RR intervals. The results are presented in Figure 5, where each of the 16 panels represents a different measure. For each measure the individual values for normal subjects (+), CHF patients without AF (×), and CHF patients with AF (Δ) comprise the left three columns, respectively. Values in the right four columns correspond to other cardiovascular pathologies and will be discussed in Section 4.

To illustrate how particular measures succeed (or fail to succeed) in distinguishing between CHF patients and normal subjects, we focus in detail on two measures: VLF power and pNN50. For this particular collection of patients and record lengths, the normal subjects all exhibit larger values of VLF power than the CHF patients; indeed a horizontal line drawn at $\text{VLF} = 0.000600$ completely separates the two classes. On the other hand, for pNN50, although the normals still have larger values on average, there is a region of overlap of CHF patients and normal subjects near 0.05, indicating that the two classes of patients cannot be entirely separated using this measure. Thus, for the full data set, comprising 75,821 RR intervals, VLF succeeds in completely distinguishing CHF patients and normal subjects whereas pNN50 does not.

Examining all 16 panels, we find that 6 measures manage to completely separate the normal subjects (first column) from the heart-failure patients (second and third columns) while the remaining 10 fail to do so. The six successful measures are highlighted by boldface font in Figure 5: **VLF**, **LF**, **$A(10)$** , **$\sigma_{\text{wav}}(32)$** , **$S_c(1/32)$** , and **DFA(32)**.

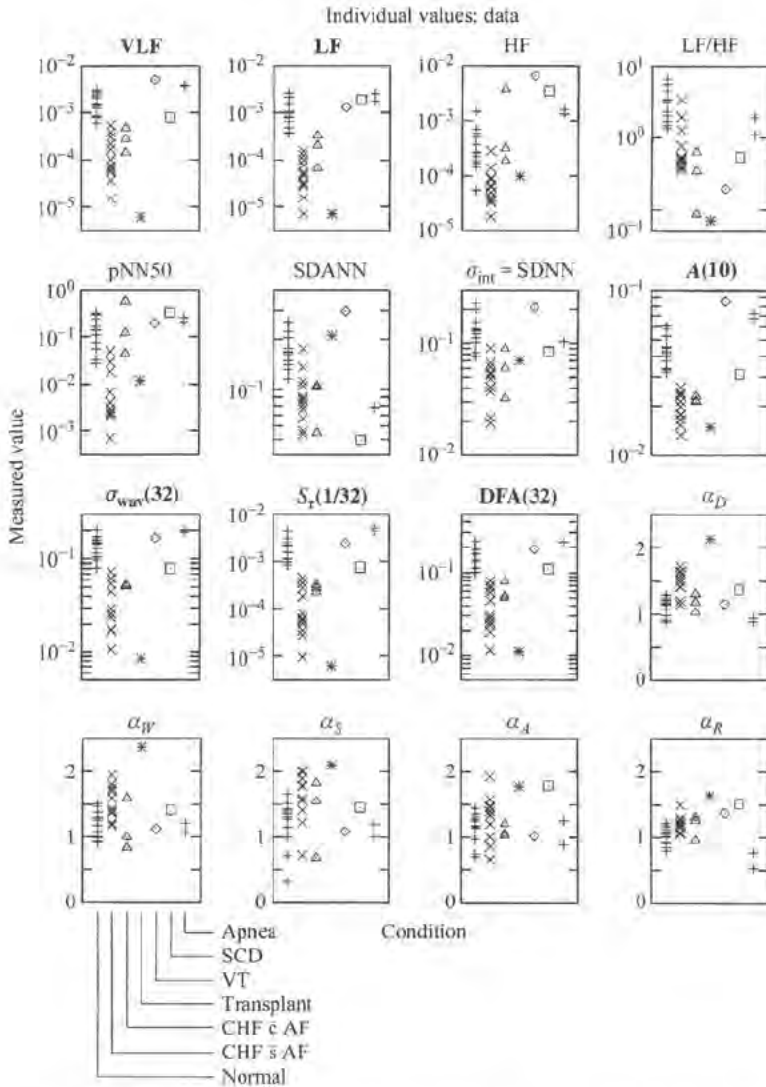


Figure 5 Individual value plots (data) for the 16 measures. Each panel corresponds to a different HRV measure. The seven columns in each panel, from left to right, comprise data for (1) 12 normal subjects (+), (2) 12 CHF patients \bar{s} AF (\times), (3) 3 CHF patients \bar{c} AF (Δ), (4) 1 heart-transplant patient ($*$), (5) 1 ventricular-tachycardia patient (\square), (6) 1 sudden-cardiac-death patient (\diamond), and (7) 2 sleep-apnea patients (+). Each data set comprises 75,821 RR intervals except for the two sleep-apnea data sets, which comprise 16,874 and 15,751 RR intervals, respectively. The six measures highlighted in boldface font succeed in completely separating normal subjects and CHF patients (\bar{s} and \bar{c} atrial fibrillation) **VLF**, **LF**, **A(10)**, **$\sigma_{wav}(32)$** , **$S_1(1/32)$** , and **DFA(32)**.

3.4. Predictive Value Plots

How can the ability of a measure to separate two classes of subjects be quantified? Returning to the VLF panel in Figure 5, we place a threshold level θ at an arbitrary position on the ordinate and consider only the leftmost two columns: normal subjects and heart-failure patients who do not suffer from atrial fibrillation. We then classify all subjects for whom the VLF values are $< \theta$ as CHF patients (positive) and for those for whom the VLF values are $> \theta$ as normal (negative). (Measures that yield smaller results for normal patients, on average, obey a reversed decision criterion.)

If a subject labeled as a CHF patient is indeed so afflicted, then this situation is referred to as a true positive (P_T); a normal subject erroneously labeled as a CHF patient is referred to as a false positive (P_F). We define negative outcomes that are true (N_T) and false (N_F) in an analogous manner. As pointed out in Section 2.7.2, the positive predictive value $V_P = P_T/(P_T + P_F)$ and negative predictive value $V_N = N_T/(N_T + N_F)$ represent the proportion of positives and negatives, respectively, that are correctly identified. This determination is carried out for many values of the threshold θ .

Figure 6 shows the positive (solid curves) and negative (dotted curves) predictive values for all 16 measures, plotted against the threshold θ , each in its own panel. These curves are constructed using the 12 normal and 12 heart-failure (without AF) records that constitute the CHF database discussed in Section 3.1. For the VLF measure, both predictive values are simultaneously unity in the immediate vicinity of $\theta = 0.000600$. This occurs because P_F and N_F are both zero at this particular value of θ and reconfirms that the two classes of data separate perfectly in the VLF panel of Figure 5 at this threshold.

For threshold values outside the range $0.000544 < \theta < 0.000603$, some of the patients will be incorrectly identified by the VLF measure. If we set $\theta = 0.000100$ for example, 6 of the 12 CHF patients will be incorrectly identified as normal subjects, which is confirmed by examining the VLF panel in Figure 5. This yields $V_N = N_T/(N_T + N_F) = 12/(12 + 6) \doteq 0.67 < 1$, which is the magnitude of the negative predictive value (dotted curve) in the VLF panel in Figure 6. At this value of the threshold ($\theta = 0.000100$), the positive predictive value remains unity because P_F remains zero.

The pNN50 panel in Figure 5, in contrast, reveals a range of overlap in the individual values of the normal subjects and CHF patients. Consequently, as θ increases into the overlap region, V_P decreases below unity and this happens before V_N attains unity value. Thus, there is no threshold value, or range of threshold values, for which the positive and negative predictive values in Figure 6 are both unity. The best threshold for this measure lies in the range $0.026 < \theta < 0.050$, with the choice depending on the relative benefit of being able to predict accurately the presence or absence of CHF in a patient.

There are six measures in Figure 6 (indicated in boldface font) for which the positive and negative predictive values are both unity over the same range of threshold values. These measures are, of course, the same six measures for which the normal subjects and heart-failure patients fall into disjoint sets in Figure 5.

3.5. ROC Curves

Two other important clinically relevant quantities that depend on the threshold θ are the sensitivity, the proportion of heart-failure patients that are properly identified [$P_T/(P_T + N_F)$], and the specificity, the proportion of normal subjects that are properly

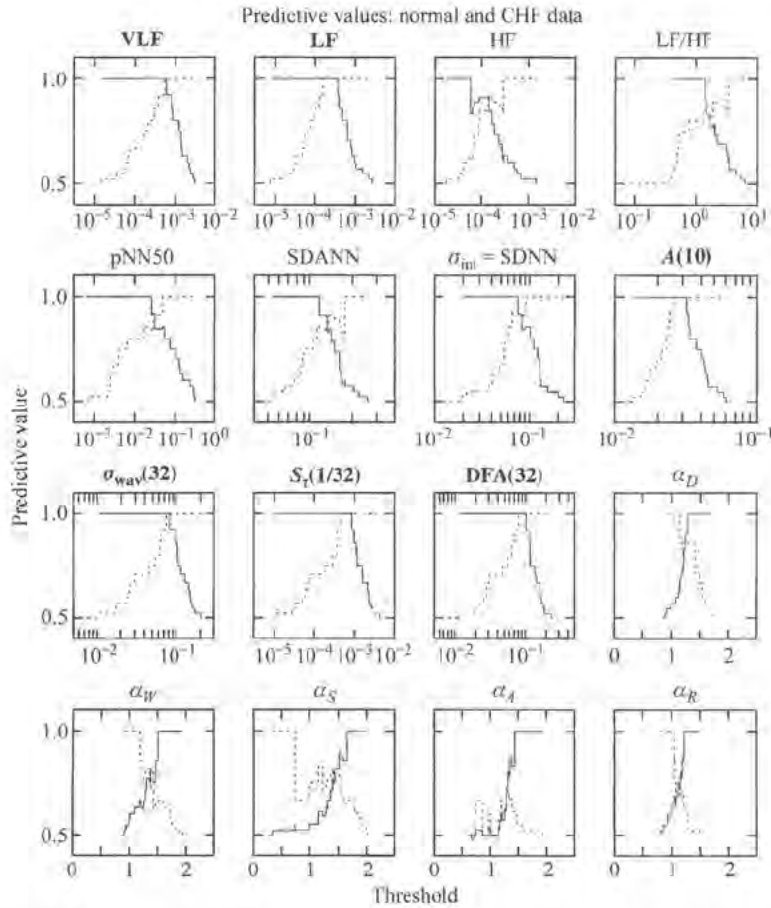


Figure 6 Positive (solid curves) and negative (dotted curves) predictive values for all 16 HRV measures plotted against the threshold θ , each in its own panel. These curves are constructed using the 12 normal and 12 heart-failure (\bar{s} AF) records drawn from the CHF database, each of which has been truncated to 75,821 RR intervals. The six measures highlighted in boldface font exhibit threshold regions for which both the positive and negative predictive values are unity: **VLF**, **LF**, **A(10)**, **$\sigma_{wav}(32)$** , **$S_T(1/32)$** , and **DFA(32)**. This indicates that the normal subjects and CHF (\bar{s} AF) patients can be completely distinguished by these six measures, in accordance with the results established in Figure 5.

identified $[N_T/(N_T + P_F)]$. As pointed out in Sections 2.7.2 and 2.7.3, sensitivity and specificity are related to patient status (pathologic and normal, respectively), whereas the positive and negative predictive values are related to identification status (positive and negative, respectively). Sensitivity and specificity are both monotonic functions of the threshold, but this is not generally true for the predictive values. The monotonicity property is salutary in that it facilitates the use of a parametric plot which permits these quantities to be represented in compact form. A plot of sensitivity versus $1 - \text{specificity}$, traced out various values of the threshold θ , forms the ROC curve (see Section 2.7.3).

ROC curves are presented in Figure 7 for all 16 measures, again using the same 12 normal and 12 heart-failure (without AF) records that constitute the CHF database discussed in Section 3.1. Because of the complete separation between the two classes of patients (leftmost two columns of the VLF panel in Figure 5) near $\theta = 0.000600$, the VLF ROC curve in Figure 7 simultaneously achieves unity (100%) sensitivity and unity (100%) specificity (the point at upper left corner of the ROC curve). For the pNN50 statistic, in contrast, the overlap evident in Figure 5 prevents this, so that the upper left corner of the pNN50 ROC curve in Figure 7 instead reveals smaller simultaneous values of sensitivity and specificity.

Six measures in Figure 7 simultaneously exhibit unity sensitivity and specificity; these are indicated by boldface font and have ROC curves that are perfectly square. They are clearly the same measures for which the normal subjects and heart-failure

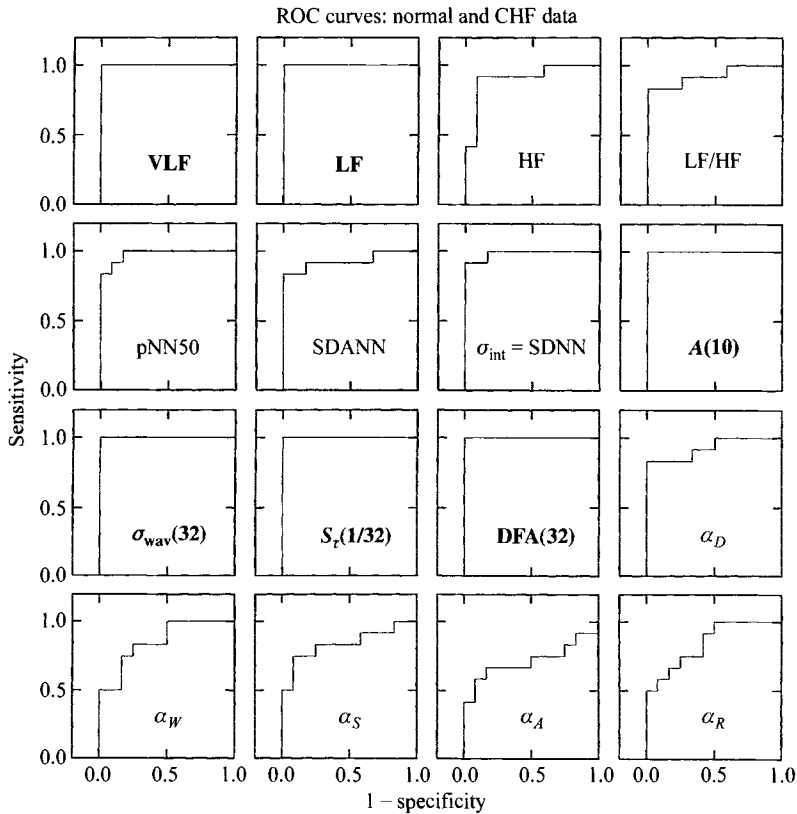


Figure 7 ROC curves (sensitivity vs. $1 - \text{specificity}$) for all 16 HRV measures, each in its own panel. These curves are constructed using the 12 normal and 12 heart-failure (\bar{s} AF) records drawn from the CHF database, each of which has been truncated to 75,821 RR intervals. The six measures highlighted in boldface font simultaneously achieve 100% sensitivity and 100% specificity so that the ROC curve is perfectly square: **VLF**, **LF**, **A(10)**, **$\sigma_{\text{wav}}(32)$** , **$S_{\tau}(1/32)$** , and **DFA(32)**. This indicates that the normal subjects and CHF (\bar{s} AF) patients can be completely distinguished by these six measures, in accordance with the results established in Figures 5 and 6.

patients fall into disjoint sets in Figure 5 and for which simultaneous positive and negative predictive values of unity are observed in Figure 6.

3.5.1. Comparison with Detection-Distance Measures

For didactic purposes we compare the ROC results presented immediately above with those obtained using detection-distance analysis. As we indicated in Section 2.7.1, care must be exercised when using these techniques for anything other than Gaussian-distributed quantities. The calculations were carried out using the same 12 normal-subject and 12 CHF-patient records, each comprising $L_{\max} = 75,821$ intervals. In Table 1 we provide the detection distances h and d , in order of descending value of h , for all 16 measures. Large values are best because they indicate that the two distributions are well separated.

Five of the six measures that entirely separate the CHF patients and normal subjects using ROC analysis fall in the top five positions in Table 1. The sixth measure, LF, falls in the ninth position. This confirms that detection-distance analysis applied to these long recordings provides results that qualitatively agree with those obtained using ROC analysis. However, detection-distance analysis does not provide any indication of how many (or indeed whether any) of the measures at the top of the list completely separate the two classes of patients, nor does it provide estimates of sensitivity and specificity. Moreover, the rankings according to d differ from those according to h .

Finally, the detection distance for a particular measure, as well as the relative ranking of the measure, depends on what appear to be insignificant details about the

TABLE 1 Detection Distances h and d for All 16 HRV Measures Applied to the 12 Normal and 12 CHF Records of Length L_{\max} ^a

Measure	h	d
DFA(32)	2.48253	1.81831
$\sigma_{\text{wav}}(32)$	2.33614	1.70153
$A(10)$	2.32522	1.77482
VLF	1.84285	1.56551
$S_1(1/32)$	1.77422	1.55200
σ_{int}	1.74750	1.32475
$\sigma_{\text{wav}}^2(32)$	1.71343	1.47165
α_D	1.64883	1.17679
SDANN	1.46943	1.04079
LF	1.36580	1.28686
pNN50	1.36476	1.20896
LF/HF	1.24507	0.91444
α_W	1.09916	0.77800
α_R	1.02367	0.72463
HF	0.85361	0.73077
α_S	0.82125	0.58071
α_A	0.38778	0.27895

^aThe measures are listed in order of descending value of h . The rankings according to d differ from those according to h . The wavelet-transform standard deviation $\sigma_{\text{wav}}(32)$ and variance $\sigma_{\text{wav}}^2(32)$, although related by a simple monotonic transformation, yield different values of h and have different rankings.

specific form in which the measure is cast. For example the h values for $\sigma_{\text{wav}}(32)$ and its square $\sigma_{\text{wav}}^2(32)$ are substantially different, and so are their rankings in Table 1. As discussed in Section 2.7.3, ROC analysis is invariant to monotonic transformations of the measure and therefore does not suffer from this disadvantage.

3.5.2. ROC-Area Curves

Perfectly square ROC curves, associated with the group of six boldface-labeled measures in Figures 5–8, exhibit unit area. These ROCs represent 100% sensitivity for all values of specificity, indicating that every patient is properly assigned to the appropriate status: heart failure or normal. Although the perfect separation achieved by these six measures endorses them as useful diagnostic statistics, the results of most studies are seldom so clear-cut. ROC area will surely decrease as increasing numbers of out-of-sample records are added to the database, because increased population size means increased variability [46].

ROC area also decreases with diminishing data length; shorter records yield less information about patient condition and these patients are therefore more likely to be misclassified [47, 48]. The ROC curves in Figure 7 have been constructed from Holter-monitor records that contain many hours of data (75,821 RR intervals). It would be useful in a clinical setting to be able to draw inferences from HRV measures recorded over shorter times, say minutes rather than hours. It is therefore important to examine the performance of the 16 HRV measures as data length is decreased. As indicated in Section 2.7.3, ROC analysis provides an ideal method for carrying out this task [47, 48].

In Figure 8 we present ROC-area curves as a function of the number of RR intervals (data length) L analyzed ($64 \leq L \leq 75,821$). The ROC areas for the full-length records ($L_{\text{max}} = 75,821$), which correspond to the areas under the ROC curves presented in Figure 7, are the rightmost points in the ROC-area curves shown in Figure 8. Results for shorter records were obtained by dividing the 12 normal and 12 heart-failure (without AF) records that comprise the CHF database (Section 3.1) into smaller segments of length L . The area under the ROC curve for that data length L was computed for the first such segment for all 16 measures, and then for the second segment, and so on, for all segments of length L (remainders of length $< L$ of the original files were not used for the ROC-area calculations). From the L_{max}/L values of the ROC area, the mean and standard deviation were computed and are plotted in Figure 8. The lengths L examined ranged from $L = 2^6 = 64$ to $L = 2^{16} = 65,536$ RR intervals, in powers of two, in addition to the entire record of $L_{\text{max}} = 75,821$ intervals.

To illustrate the information provided by these curves, we direct our attention to the VLF and pNN50 panels in Figure 8. For the full-length records the rightmost point in the VLF panel reveals unit area while that for pNN50 lies somewhat lower, as expected from the corresponding ROC curves in Figure 7. VLF clearly outperforms pNN50. As the data length analyzed decreases, so too do the ROC areas for both measures while their variances increase, also as expected. However, when the data length dips to 256 or fewer RR intervals, the performance of the two measures reverses so that pNN50 outperforms VLF. There is an important point to be drawn from this example. Not only does the performance of a measure depend on data length, but so too does the relative performance of different measures.

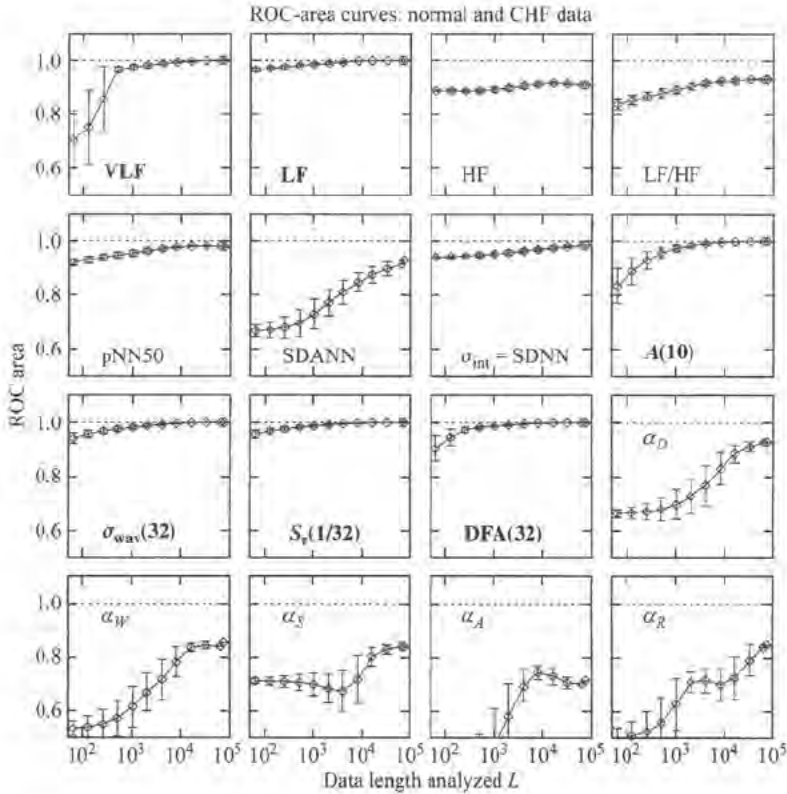


Figure 8 Diagnostic accuracy (area under ROC curve) as a function of record length analyzed L (number of RR intervals) for the 16 HRV measures (mean \pm 1SD). An area of unity corresponds to perfect separability of the two classes of patients. The six measures highlighted in boldface font—VLF, LF, $A(10)$, $\sigma_{wav}(32)$, $S_r(1/32)$, and $DFA(32)$ —provide such perfect separability at the longest record lengths, in accordance with the results in Figure 7. As the record length decreases, performance degrades at a different rate for each measure. The five scale-independent measures, α_D , α_W , α_S , α_A , and α_R , perform poorly at all data lengths.

3.6. Comparing the Measures for Various Data Lengths

Based on their overall ability to distinguish between CHF patients and normal subjects over a range of data lengths, the 16 measures shown in Figure 8 divide roughly into three classes. The six measures that fall in the first class, VLF, LF, $A(10)$, $\sigma_{wav}(32)$, $S_r(1/32)$, and $DFA(32)$, succeed in completely separating the two classes of patients for data lengths down to $L = 2^{15} = 32,768$ RR intervals. These six measures share a dependence on a single scale, or small range of scales, near 32 heartbeat intervals. For this collection of data sets, this scale appears to yield the best performance. Members of this class outperform the other 10 measures at nearly all data lengths. Apparently, the scale value itself is far more important than the measure used to evaluate it.

The second class, consisting of HF, the ratio LF/HF, pNN50, and σ_{int} , fail to achieve complete separation for any data size examined. Nevertheless, the members

of this class are not devoid of value in separating CHF patients from normal subjects. Interestingly, all but LF/HF provide better results than $A(10)$, a member of the first class, for the shortest data lengths. Results for these four measures varied relatively little with data size, thus exhibiting a form of robustness.

Members of the third class, consisting of SDANN and the five scale-independent measures $\alpha_{[.]}$, exhibit poor performance at all data lengths. These six measures require long sequences of RR intervals to make available the long-term fluctuations required for accurate estimation of the fractal exponent. Data lengths $L < 5000$ RR intervals lead to large variance and (negative) bias and are not likely to be meaningful. As an example of the kind of peculiarity that can emerge when attempting to apply scale-independent measures to short records, the α_A ROC area decreases below 0.5 when the data size falls below 2048 intervals (reversing the sense of the comparison only for these data sizes increases the ROC area, although not above 0.7; however this clearly violates the spirit of the method). SDANN requires several 5-minute segments to determine the standard deviation accurately.

3.6.1. Scale-Independent versus Scale-Dependent Measures

As indicated in the previous subsection, all five scale-independent measures (α_D , α_W , α_S , α_A , and α_R) perform poorly at all data lengths. These fractal-exponent estimators return widely differing results as is plainly evident in Figure 5. This suggests that there is little merit in the concept of a single exponent for characterizing the human heartbeat sequence, no less a "universal" one as some have proposed [51, 55, 56].

A variation on this theme is the possibility that pairs of fractal exponents can provide a useful HRV measure. At small scales m , Figure 4 reveals that heart-failure patients exhibit smaller values of the wavelet-transform standard-deviation slope than do normal subjects. Following Peng et al. [51], who constructed a measure based on differences of DFA scaling exponents in different scaling regions in an attempt to discriminate CHF patients from normal subjects, Thurner et al. [46] constructed a measure based on differences in the wavelet-transform standard-deviation slope at different scales. However, the outcome was found to be unsatisfactory when compared with other available measures; we concluded the same about the results obtained by Peng et al. [51]. Using ROC analysis, as described in Section 2.7.3, we determined that the ROC area for the measure described by Thurner et al. [46] was sufficiently small (0.917 for $m = 4, 16$, and 256) that we abandoned this construct.

Four of the techniques we have discussed in this chapter (spectral, wavelet, detrended fluctuation analysis, and Allan factor) yield both scale-independent and scale-dependent measures and therefore afford us the opportunity of directly comparing these two classes of measures in individual calculations: $\alpha_W \leftrightarrow \sigma_{\text{wav}}(32)$; $\alpha_S \leftrightarrow S_\tau(1/32)$; $\alpha_D \leftrightarrow \text{DFA}(32)$; $\alpha_A \leftrightarrow A(10)$. In each of these cases the fixed-scale measure is found to greatly outperform the fractal-exponent measure for all data sizes examined, as we previously illustrated for the pairs $\alpha_W \leftrightarrow \sigma_{\text{wav}}(32)$ and $\alpha_S \leftrightarrow S_\tau(1/32)$ [47]. These results were recently confirmed in a follow-up Israeli–Danish study [65]. Moreover, in contrast to the substantial variability returned in fractal-exponent estimates, results for the different scale-dependent measures at $m = 32$ intervals bear reasonable similarity to each other.

Nunes Amaral et al. [56] concluded exactly the opposite, namely that scaling exponents provide superior performance to scale-dependent measures. This may be because they relied exclusively on the distribution-dependent measures $\eta \equiv h^2$ and d^2 (See Sections 2.7.1. and 3.5.1) rather than on distribution-independent ROC analysis. The same authors [56] also purport to glean information from higher moments of the wavelet coefficients, but the reliability of such information is questionable because estimator variance increases with moment order [17].

3.6.2. Computation Times of the Various Measures

The computation times for the 16 measures considered in this chapter are provided in Table 2. All measures were run 10 times and averaged except for the two DFA measures, which, because of their long execution time, were run only once. These long execution times are associated with the suggested method for computing DFA [66], which is an N^2 process. DFA computation times therefore increase as the square of the number of intervals, whereas all 14 other methods, in contrast, are either of order N or $N \log(N)$.

Based on computation time, we can rank order the six scale-dependent measures that fall into the first class from fastest to slowest: $\sigma_{\text{wav}}(32)$, $S_r(1/32)$, $A(10)$, LF, VLF, and DFA(32). Because of its computational simplicity, the wavelet-transform standard deviation $\sigma_{\text{wav}}(32)$ computes more rapidly than any of the other measures. It is 3 times faster than its nearest competitor $S_r(1/32)$, 16.5 times faster than LF, and 32,500 times faster than DFA(32).

3.6.3. Comparing the Most Effective Measures

In Sections 3.3–3.5, we established that six measures succeeded in completely separating the normal subjects from the CHF patients in our database for data lengths $L \geq 2^{15} = 32,768$ RR intervals: VLF, LF, $A(10)$, $\sigma_{\text{wav}}(32)$, $S_r(1/32)$, and DFA(32). We

TABLE 2 Computation Times (to the nearest 10 msec) for the 16 HRV Measures for Data Sets Comprising 75,821 RR Intervals^a

Measure	Execution time (ms)
VLF, LF, HF, and LF/HF	330
pNN50	40
SDANN	160
σ_{int}	190
$A(10)$	160
$\sigma_{\text{wav}}(32)$	20
$S_r(1/32)$	60
DFA(32)	650,090
α_D	650,110
α_W	220
α_S	920
α_A	610
α_R	570

^aThe long execution times for the two DFA measures result from the fact that it is an N^2 process whereas the 14 other methods are either N or $N \log(N)$.

now demonstrate from a fundamental point of view that these measures can all be viewed as transformations of the interval-based power spectral density and that all are therefore closely related.

Consider first the relationships of VLF and LF to $S_\tau(1/32)$. At a frequency $f = 1/32 \approx 0.031$ cycles/interval, $S_\tau(f)$ separates CHF patients from normal subjects; however, a range of values of f provide varying degrees of separability. For the data sets we analyzed, separation extends from $f = 0.02$ to $f = 0.07$ cycles/interval [49]. Recall that VLF and LF are simply integrals of $S_\tau(f)$ over particular ranges of f : from $f = 0.003$ to $f = 0.04$ cycles/interval for VLF and from $f = 0.04$ to $f = 0.15$ cycles/interval for LF. Because these ranges overlap with those for which the power spectral density itself provides separability, it is not surprising that these spectral integrals also exhibit this property. This illustrates the close relationship among VLF, LF, and $S_\tau(1/32)$.

We next turn to the relation between $\sigma_{\text{wav}}(32)$ and $S_\tau(1/32)$. In Section 2.5.3 the relationship between these two measures was established as

$$\sigma_{\text{wav}}^2(m) = m \int_{-\infty}^{\infty} S_\tau(f)H(mf) df \quad (23)$$

revealing that $\sigma_{\text{wav}}^2(m)$ is simply a bandpass-filtered version of the interval-based power spectral density for stationary signals. For the Haar wavelet, the bandpass filter $H(mf)$ has a center frequency near $1/m$ and a bandwidth about that center frequency also near $1/m$ [49]. Accordingly, $\sigma_{\text{wav}}^2(32)$ is equivalent to a weighted integral of $S_\tau(f)$ centered about $f = 1/32 \approx 0.031$, with an approximate range of integration of 0.031 ± 0.016 (0.016 to 0.047). Thus, the separability of $\sigma_{\text{wav}}^2(32)$ is expected to resemble that of $S_\tau(1/32)$ and therefore that of VLF and LF as well.

Now consider the count-based Allan factor evaluated at a counting time of 10 seconds, $A(10)$. Since the Allan factor is directly related to the variance of the Haar wavelet transform of the *counts* [36], $A(10)$ is proportional to the wavelet transform of the counting process evaluated over a wavelet duration of $T = 20$ seconds. The count-based and interval-based measures are approximately related via Eq. 7 so that $A(10)$ should be compared with $\sigma_{\text{wav}}(20/E[\tau]) \approx \sigma_{\text{wav}}(25)$, whose value, in turn, is determined by $S_\tau(1/25 = 0.04)$ and spectral components at nearby frequencies. Since this range of $S_\tau(f)$ is known to exhibit separability, so too should $A(10)$.

The relationship between detrended fluctuation analysis and the other measures proves more complex than those already delineated. Nevertheless, DFA(32) can also be broadly interpreted in terms of an underlying interval-based power spectral density. To forge this link, we revisit the calculation of $F(m)$ (see Section 2.5.4). The mean is first subtracted from the sequence of RR intervals $\{\tau_i\}$, resulting in a new sequence that has zero mean. This new random process has a power spectral density everywhere equal to $S_\tau(f)$, except at $f = 0$, where it assumes a value of zero. The summation of this zero-mean sequence generates the series $y(k)$, as shown in Eq. 21. Since summation is represented by a transfer function $1/j2\pi f$ for small frequencies f , with $j = \sqrt{-1}$, the power spectral density of $y(k)$ is approximately given by

$$S_y(f) \approx \begin{cases} 0 & f = 0 \\ \frac{S_\tau(f)}{4\pi^2 f^2} & f \neq 0 \end{cases} \quad (24)$$

Next, the sequence $y(k)$ is divided into segments of length m . For the first segment, $k = 1$ to m , the best fitting linear trend $y_m(k) = \beta + \gamma k$ is obtained; the residuals $y(k) - y_m(k)$ therefore have the minimum variance over all choices of β and γ . Over the duration of the entire sequence $y(k)$, the local trend function $y_m(k)$ will consist of a series of such linear segments and is therefore piecewise linear. Since $y_m(k)$ changes behavior over a time scale of m intervals, its power will be concentrated at frequencies below $1/m$. Because the residuals $y(k) - y_m(k)$ have minimum variance, the spectrum of $y_m(k)$ will resemble that of $y(k)$ as much as possible and in particular at frequencies below $1/m$, where most of its power is concentrated. Thus

$$S_{y_m}(f) \approx \begin{cases} S_y(f) & |f| < 1/m \\ 0 & |f| > 1/m \end{cases} \quad (25)$$

Equation 22 shows the mean-square fluctuation $F^2(m)$ to be an estimate of the variance of the residuals $y(k) - y_m(k)$. by Parseval's theorem, this variance equals the integral of the spectrum of the residuals over frequency. Therefore

$$F^2(m) \approx 2 \int_{0^+}^{\infty} [S_y(f) - S_{y_m}(f)] df = (2\pi^2)^{-1} \int_{1/m}^{\infty} S_{\tau}(f) f^{-2} df \quad (26)$$

so that the mean-square fluctuation $F^2(m)$ can be represented as a weighted integral of the RR-interval spectrum over the appropriate frequency range.

We conclude that all six measures that best distinguish CHF patients from normal subjects are related in terms of transformations of the interval-based power spectral density.

3.6.4. Selecting the Best Measures

In the previous section we showed that the six most effective measures are inter-related from a fundamental point of view. This conclusion is confirmed by the overall similarity of the ROC-area curves for these measures, as can be seen in Figure 8.

We now proceed to compare these ROC areas more carefully in Figure 9. This figure presents $1 - \text{ROC area}$ plotted against the data length analyzed L on a doubly logarithmic plot. This is in contrast to the usual presentation (such as Figure 8) in which ROC area is plotted against L on a semilogarithmic plot. The unconventional form of the plot in Figure 9 is designed to allow small deviations from unit area to be highlighted. Clearly, small values of the quantity $1 - \text{ROC area}$ are desirable. Of course, as the data length L diminishes, ROC area decreases (and therefore $1 - \text{ROC area}$ increases) for all six measures.

The trends exhibited by all six measures displayed in Figure 9 [VLF, $A(10)$, DFA(32), $\sigma_{\text{wav}}(32)$, $S_{\tau}(1/32)$, and LF] are indeed broadly similar, as we observed in Figure 8. For data sets that are sufficiently long, the six measures provide identical diagnostic capabilities (unity ROC area). At shorter data lengths, however, VLF, $A(10)$, and DFA(32) exhibit more degradation than do the other three measures. In accordance with the analysis provided in Section 3.6.3, VLF deviates the most at small values of L because it incorporates information from values of $S_{\tau}(f)$ well below the frequency range of separability. Similarly, DFA(32) favors low-frequency contributions by virtue

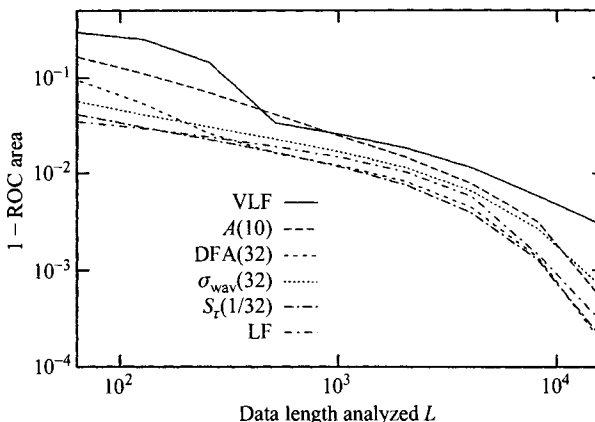


Figure 9 Doubly logarithmic plot of $1 - \text{ROC area}$ as a function of data length L for the six most effective measures: VLF, $A(10)$, DFA(32), $\sigma_{wav}(32)$, $S_{\tau}(1/32)$, and LF. The unconventional form of this ROC plot allows small deviations from unity area to be highlighted. The trends exhibited by all six measures are broadly similar, but VLF, $A(10)$, and DFA(32) exhibit more degradation at shorter data lengths than do the other three measures. Thus three measures emerge as superior: LF, $\sigma_{wav}(32)$, and $S_{\tau}(1/32)$. When computation time is taken into account in accordance with the results provided in Table 2, $\sigma_{wav}(32)$ and $S_{\tau}(1/32)$ are the two preferred measures.

of the $1/f^2$ weighting function in Eq. 26. Finally, the performance of $A(10)$ is suppressed because of the confounding effect of values of $S_{\tau}(f)$ for $f > 0.07$, as well as from possible errors deriving from our assumptions regarding the equivalence of the count-based and interval-based wavelet-transform variances.

The remaining three measures therefore emerge as superior: LF, $\sigma_{wav}(32)$, and $S_{\tau}(1/32)$. Judged on the basis of both performance (Figure 9) and computation time (Table 2), we conclude that two of these are best: $\sigma_{wav}(32)$ and $S_{\tau}(1/32)$.

It is desirable to confirm these conclusions for other records, comprising out-of-sample data sets from CHF patients and normal subjects. It will also be of interest to examine the correlation of ROC area with severity of cardiac dysfunction as judged, for example, by the New York Heart Association (NYHA) clinical scale. It is also important to conduct comparisons with other CHF studies that make use of HRV measures [3].

A sufficiently large database of normal and CHF records would make it possible to examine the detailed distinctions between the two classes of patients over a range of scales surrounding $m = 32$. A superior measure that draws on these distinctions in an optimal way could then be designed; for example, the weighting function in an integral over $S_{\tau}(f)$ could be customized. Other measures that simultaneously incorporate properties inherent in collections of the individual measures considered here could also be developed.

4. MARKERS FOR OTHER CARDIAC PATHOLOGIES

Returning to the individual-value plots in Figure 5, we briefly examine the behavior of the 16 measures for several other cardiovascular disorders for which these measures

may hold promise. Among these are three CHF patients who also suffer from atrial fibrillation (CHF with AF, third column, \triangle), one heart-transplant patient (TRANSPLANT, fourth column, $*$), one ventricular-tachycardia patient (VT, fifth column, \diamond), one sudden-cardiac-death patient (SCD, sixth column, \square), and two sleep-apnea patients (APNEA, last column, $+$). A summary of the more salient statistics of the original RR recordings for these patients is presented in Table A1. Prior to analysis, the RR records were truncated at $L_{\max} = 75,821$ RR intervals to match the lengths of the records of CHF patients and normal subjects. The two sleep-apnea data records were drawn from the MIT-BIH Polysomnographic Database (Harvard-MIT Division of Health Sciences and Technology) and comprised 16,874 and 15,751 RR intervals, respectively. Because of the limited duration of the sleep period, records containing larger numbers of RR intervals are not available for this disorder.

Too few patient recordings are available to us to reasonably evaluate the potential of these HRV measures for the diagnosis of these forms of cardiovascular disease, but the results presented here might suggest which measures would prove useful given large numbers of records.

We first note that for most of the 16 measures, the values for the 3 CHF patients with atrial fibrillation (\triangle) tend to fall within the range set by the other 12 CHF patients without atrial fibrillation. In particular, the same six measures that completely separate normal subjects from CHF patients without AF continue to do so when CHF patients with AF are included. The inclusion of the AF patients reduces the separability only for 3 of the 16 measures: pNN50, HF, and α_W .

Results for the heart-transplant patient ($*$) appear toward the pathological end of the CHF range for many measures, and for some measures the transplant value extends beyond the range of all of the CHF patients. Conversely, the sleep apnea values ($+$) typically fall at the end of, or beyond, the range spanned by the normal subjects. Indeed, three thresholds can be chosen for $S_\tau(1/32)$ (i.e., 0.000008, 0.0006, and 0.00436) that completely separate four classes of patients: sleep apnea [largest values of $S_\tau(1/32)$], normal subjects (next largest), CHF with or without atrial fibrillation (next largest), and heart transplant (smallest). Although such separation will no doubt disappear as large numbers of additional patient records are included, we nevertheless expect the largest variability (measured at 32 intervals as well as at other scales) to be associated with sleep-apnea patients because their fluctuating breathing patterns induce greater fluctuations in the heart rate. Conversely, heart-transplant patients (especially in the absence of reinnervation) are expected to exhibit the lowest values of variability because the autonomic nervous system, which contributes to the modulation of heart rate at these time scales, is subfunctional.

Results for the ventricular-tachycardia patient (\diamond) lie within the normal range for most measures, although the value for LF/HF is well beyond the pathological end of the range for CHF patients. Interestingly, four measures [VLF, HF, SDANN, and $A(10)$] yield the largest values for the ventricular-tachycardia patient of the 32 data sets examined. Perhaps $A(10)$, which exhibits the largest separation between the value for this patient and those for the other 31 patients, will prove useful in detecting ventricular tachycardia.

Finally, we consider results for the sudden-cardiac-death patient (\square), which lie between the ranges of normal subjects and CHF patients for all but one (LF) of the six measures that separate these two patient classes. The pNN50 and HF results, on the other hand, reveal values beyond the normal set, thereby indicating the presence of

greater variability on a short time scale than for other patients. The value of SDANN for the SCD patient is greater than that for any of the other 31 patients, so it is possible that SDANN could prove useful in predicting sudden cardiac death.

5. DOES DETERMINISTIC CHAOS PLAY A ROLE IN HEART RATE VARIABILITY?

The question of whether normal HRV arises from a low-dimensional attractor associated with a deterministic dynamical system or has stochastic underpinnings has continued to entice researchers [61]. Previous studies devoted to this question have come to conflicting conclusions. The notion that the normal human-heartbeat sequence is chaotic appears to have been first set forth by Goldberger and West [67]. Papers in support of this notion [68], as well as in opposition to it [69, 70], have appeared recently. An important factor that pertains to this issue is the recognition that correlation, rather than nonlinear dynamics, can lie behind these seemingly chaotic recordings [8].

We address this question in this section and conclude, for both normal subjects and patients with several forms of cardiac dysfunction, that the sequence of heartbeat intervals does not exhibit chaos.

5.1. Methods

5.1.1. Phase-Space Reconstruction

One way of approaching the question of chaos in heartbeat recordings is in terms of a phase-space reconstruction and the use of an algorithm to estimate one or more of its fractal dimensions D_q [6, 59, 71]. The box-counting algorithm [72] provides a method for estimating the capacity (or box-counting) dimension D_0 of the attractor [6, 8, 71, 73].

The phase-space reconstruction approach operates on the basis that the topological properties of the attractor for a dynamical system can be determined from the time series of a single observable [74–76]. A p -dimensional vector $\vec{T} \equiv \{\tau_i, \tau_{i+1}, \dots, \tau_{i+(p-1)l}\}$ is formed from the sequence of RR intervals $\{\tau_i\}$. The parameter p is the embedding dimension, and l is the lag, usually taken to be the location of the first zero crossing of the autocorrelation function of the time series under study. As the time index i progresses, the vector \vec{T} traces out a trajectory in the p -dimensional embedding space.

The box-counting algorithm estimates the capacity dimension of this trajectory. Specifically, the negative slope of the logarithm of the number of full boxes versus the logarithm of the width of the boxes, over a range of $4/[\max(\tau) - \min(\tau)]$ to $32/[\max(\tau) - \min(\tau)]$ boxes, in powers of two, provides an estimate of D_0 . For convenience in the generation of phase-randomized surrogates, data sets were limited to powers of two: $2^{16} = 65,536$ intervals for all but the two sleep-apnea data sets, for which $2^{14} = 16,384$ and $2^{13} = 8192$ intervals were used respectively (the sleep-apnea data sets are much shorter than the others as indicated in Table A1). For uncorrelated noise, the capacity dimension D_0 continues to increase as the embedding dimension p increases. For an attractor in a deterministic system, in contrast, D_0 saturates as p becomes larger than $2D_0 + 1$ [73]. Such saturation, however, is not a definitive signature of deterministic dynamics; temporal correlation in a stochastic process can also be responsible for what appears to be underlying deterministic dynamics [8, 77–79]. The distribution of the data can also play a role.

Surrogate data analysis, discussed in Section 5.1.3, is useful in establishing the underlying cause, by proving or disproving various null hypotheses that the measured behavior arises from other, nonchaotic causes.

5.1.2. Removing Correlations in the Data

Adjacent intervals in heartbeat sequences prove to exhibit large degrees of correlation. The serial correlation coefficient for RR intervals

$$\rho(\tau) \equiv \frac{E[\tau_{i+1}\tau_i] - E^2[\tau_i]}{E[\tau_i^2] - E^2[\tau_i]}$$

exceeds 0.9 for two thirds of the data sets examined, with an average value of 0.83; the theoretical maximum is unity, with zero representing a lack of serial correlation. Since such large correlation is known to interfere with the detection of deterministic dynamics, we instead employ the first difference of the RR intervals: $v_i \equiv \tau_i - \tau_{i+1}$. Such a simple transformation does not change the topological properties of dynamical system behavior, and a related but more involved procedure is known to reduce error in estimating other fractal dimensions [80].

The serial correlation coefficient for the differenced sequence $\{v_i\}$ turns out to have a mean value of -0.084 , and none exceeds 0.6 in magnitude. This sequence will therefore be used to generate the embedding vector $\vec{T} \equiv \{v_i, v_{i+1}, \dots, v_{i+p-1}\}$, with the lag l set to unity since $\rho(v) \approx 0$.

5.1.3. Surrogate Data Analysis

Information about the nature of an interval or segment series may be obtained by applying various statistical measures to surrogate data sets. These are processes constructed from the original data in ways designed to preserve certain characteristics of the original data while eliminating (or modifying) others. Surrogate data analysis provides a way of determining whether a given result arises from a particular property of the data set.

We make use of two kinds of surrogate data sets: shuffled intervals and randomized phases. In particular, we compare statistical measures calculated from both the original data and its surrogates to distinguish those properties of the data set that arise from correlations (such as from long-term rate fluctuations) from those properties inherent in the form of the original data.

Shuffled Intervals. The first class of surrogate data is formed by shuffling (randomly reordering) the sequence of RR intervals of the original data set. Such random reordering destroys dependencies among the intervals, and therefore the correlation properties of the data, while exactly preserving the interval histogram. As a refinement, it is possible to divide the data into blocks and shuffle intervals within each block while keeping each block separate. In this case dependencies remain over durations much larger than the block size, while being destroyed for durations smaller than it.

Randomized Phases. The other class of surrogate data we consider is obtained by Fourier transforming the original interval data and then randomizing the phases while

leaving the spectral magnitudes intact. The modified function is then inverse transformed to return to a time-domain representation of the intervals. This technique exactly preserves the second-order correlation properties of the interval sequence while removing other temporal structure, for example that needed for phase-space reconstruction. The interval histogram typically becomes Gaussian for this surrogate as a result of the central limit theorem.

5.2. Absence of Chaos

In Figure 10 we present plots of the capacity dimension D_0 against the embedding dimension p for the 12 normal subjects. Similar plots are presented in Figure 11 for CHF patients without atrial fibrillation and in Figure 12 for patients with other cardiac pathologies. For most of the panels in Figure 10, the D_0 estimate for the original data (solid curve) is indeed seen to rise with increasing embedding dimension p and shows evidence of approaching an asymptotic value of about 2. The phase-randomized data (dotted curve) yield somewhat higher results, consistent with the presence of a deterministic attractor.

However, estimates of D_0 for the shuffled data (dashed curve) nearly coincide with those of the original data. This surrogate precisely maintains the distribution of the relative sizes of the differenced RR intervals but destroys any correlation or other

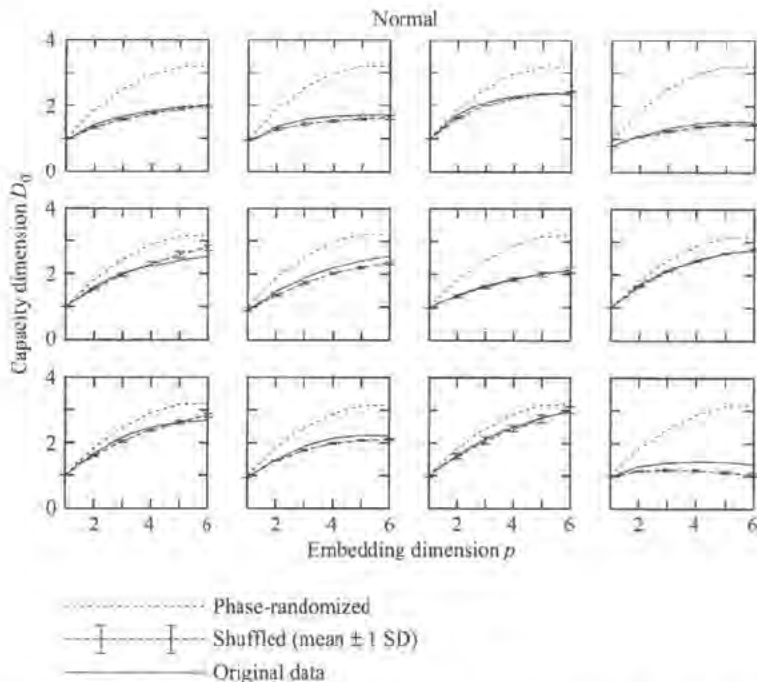


Figure 10 Capacity dimension D_0 as a function of embedding dimension p for the 12 normal subjects. The three curves shown for each subject correspond to the original differenced intervals (solid curves), the shuffled surrogates (dashed curves), and the phase-randomized surrogates (dotted curves). The dashed curves are quite similar to the solid curves in each panel, while the 12 dotted curves closely resemble each other.

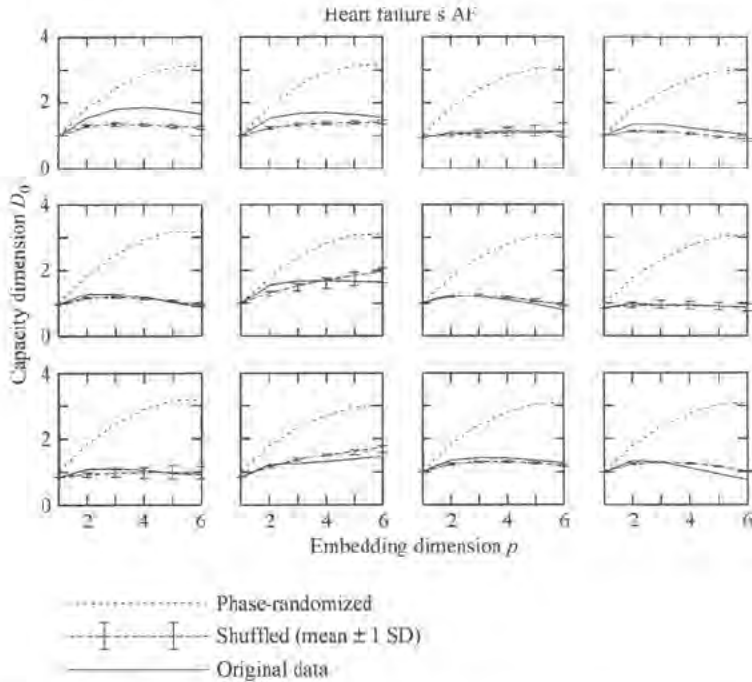


Figure 11 Capacity dimension D_0 as a function of embedding dimension p for the 12 CHF patients without atrial fibrillation. The three curves shown for each patient correspond to the original differenced intervals (solid curves), the shuffled surrogates (dashed curves), and the phase-randomized surrogates (dotted curves). The dashed curves are reasonably similar to the solid curves in most panels, while the 12 dotted curves closely resemble each other.

dependencies in the data. Specifically, no deterministic structure remains in the shuffled surrogate. Therefore, the apparent asymptote in D_0 must derive from the particular form of the RR-interval distribution and not from a nonlinear dynamical attractor. To quantify the closeness between results for original and shuffled data, the shuffling process was performed 10 times with 10 different random seeds. The average of these 10 independent shufflings forms the dashed curve, with error bars representing the largest and smallest estimates obtained for D_0 at each corresponding embedding dimension. Results for the original data indeed lie within the error bars. Therefore, the original data do not differ from the shuffled data at the $[(10 - 1)/(10 + 1)]100\% = 82\%$ significance level.

Results for a good number of the other 31 data sets, normal and pathological alike, are nearly the same; the original data lie within the error bars of the shuffled surrogates and consequently do not exhibit significant evidence of deterministic dynamics. However, for some records the original data exceed the limits set by the 10 shufflings. But attributing this behavior to low-dimensional chaos proves difficult for almost all the data records for three reasons. First, the surrogates often yield *smaller* results for D_0 than the original data, whereas shuffling data from a true chaotic system will theoretically yield *larger* D_0 estimates. Second, of the data sets for which the original results

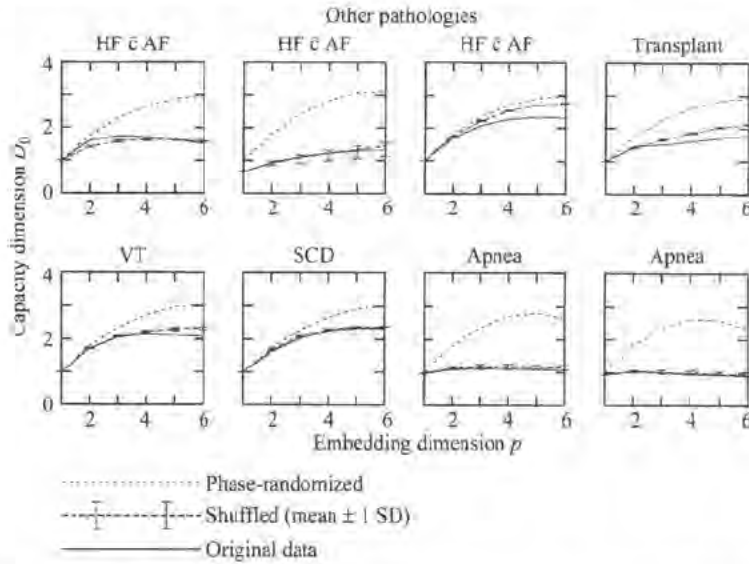


Figure 12 Capacity dimension D_0 as a function of embedding dimension p for the eight patients with other cardiac pathologies. The three curves shown for each patient correspond to the original differenced intervals (solid curves), the shuffled surrogates (dashed curves), and the phase-randomized surrogates (dotted curves). The dashed curves are reasonably similar to the solid curves in most panels while the eight dotted curves closely resemble each other.

fall below those of their shuffled surrogates, few achieve an asymptotic value for D_0 , but rather continue to climb steadily with increasing embedding dimension. Finally, for the remaining data sets, the original and shuffled data yield curves that do not differ nearly as much as do those from simulated systems known to be chaotic, such as the Hénon attractor (not shown) [81]. Taken together, these observations suggest that finite data length, residual correlations, and/or other unknown effects in these RR-interval recordings give rise to behavior in the phase-space reconstruction that can masquerade as chaos.

Results for the phase-randomized surrogates (dotted curves in Figures 10–12), while exhibiting a putative capacity dimension in excess of the original data and of the shuffled surrogates for every patient, exhibit a remarkable similarity to each other. This occurs because the differenced-interval sequence $\{v_i\}$ displays little correlation and the phase-randomization process preserves this lack of correlation while imposing a Gaussian differenced-interval distribution. The result is nearly white Gaussian noise regardless of the original data set, and indeed results for independently generated white Gaussian noise closely follow these curves (not shown).

Time series measured from the heart under distinctly nonnormal operating conditions such as ventricular fibrillation [82], or those obtained under nonphysiological conditions such as from excised hearts [83], may exhibit chaotic behavior under some circumstances. This is precisely the situation for cellular vibrations in the mammalian cochlea [84–86]. When the exciting sound pressure level (SPL) is in the normal physiological regime, the cellular-vibration velocities are nonlinear but not chaotic. When the

SPL is increased beyond the normal range, however, routes to chaos emerge. It has been suggested that since chaos-control methods prove effective for removing such behavior, the heartbeat sequence must be chaotic. This is an unwarranted conclusion because such methods often work equally well for purely stochastic systems, which cannot be chaotic [87].

6. MATHEMATICAL MODELS FOR HEART RATE VARIABILITY

The emphasis in this chapter has been on HRV analysis. The diagnostic capabilities of various measures have been presented and compared. However, the results that emerge from these studies also serve to expose the mathematical–statistical character of the RR time series. Such information can be of benefit in the development of mathematical models that can be used to simulate RR sequences. Such models may prove useful for generating realistic heartbeat sequences in pacemakers, for example, and for developing improved physiological models of the cardiovascular system. In this section we evaluate the suitability of several integrate-and-fire constructs for modeling heart rate variability.

6.1. Integrate-and-Fire Model

Integrate-and-fire models are widely used in the neurosciences [88] as well as in cardiology [8, 89–92]. These models are attractive in part because they capture known physiology in a simple way. The integration of a rate process can represent the cumulative effect of neurotransmitter on the postsynaptic membrane of a neuron or the currents responsible for the pacemaker potential in the sino-atrial node of the heart. The crossing of a preset threshold by the integrated rate then gives rise to an action potential or a heart contraction.

The sequence of discrete events constituting the human heartbeat can be viewed as a point process deriving from a continuous rate function. The integrate-and-fire method is perhaps the simplest means for generating a point process from a rate process [8, 17]. In this model, illustrated schematically in Figure 13, the rate function $\lambda(t)$ is integrated until it reaches a fixed threshold θ , whereupon a point event is generated and the integrator is reset to zero. Thus the occurrence time of the $(i + 1)$ st event is implicitly obtained from the first occurrence of

$$\int_{t_i}^{t_{i+1}} \lambda(t) dt = \theta \quad (27)$$

The mean heart rate λ (beats/s) is given by $\lambda = E[\lambda(t)] = 1/E[\tau]$. Because the conversion from the rate process to the point process is so direct, most statistics of the point

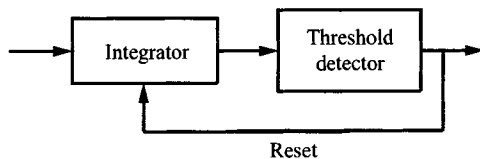


Figure 13 Simple integrate-and-fire model often used in cardiology and neurophysiology. A rate function $\lambda(t)$ is integrated until it reaches a preset threshold θ , whereupon a point event is generated and the integrator is reset to zero.

process closely mimic those of the underlying rate process $\lambda(t)$. In particular, for frequencies substantially lower than the mean heart rate, the theoretical power spectral densities of the point process and the underlying rate coincide [17].

6.2. Kernel of the Integrate-and-Fire Model

The detailed statistical behavior of the point process generated by the integrate-and-fire construct depends on the statistical character of the rate function in the integrand of Eq. 27 [8]. The fractal nature of the heartbeat sequence requires that the rate function $\lambda(t)$ itself be fractal. Assuming that the underlying rate is stochastic noise, there are three plausible candidates for $\lambda(t)$ [17]: fractal Gaussian noise, fractal lognormal noise, and fractal binomial noise. We briefly discuss these three forms of noise in turn.

6.2.1. Fractal Gaussian Noise

Fractal Gaussian noise served admirably in our initial efforts to develop a suitable integrate-and-fire model for the heartbeat sequence [8]. A Gaussian process has values that, at any set of times, form a jointly Gaussian random vector. This property, along with its mean and spectrum, completely defines the process. We consider a stationary rate process with a mean equal to the expected heart rate and a $1/f$ -type *rate*-based spectrum that takes the form [17]

$$S_\lambda(f_{\text{time}}) = \lambda^2 \delta(f_{\text{time}}) + \lambda [1 + (f_{\text{time}}/f_0)^{-\alpha}] \quad (28)$$

Here, λ (beats/s) and α are the previously defined heart rate and fractal exponent, respectively, $\delta(\cdot)$ is again the Dirac delta function, and f_0 is the cutoff frequency (cycles/s). Mathematically, cutoffs at low or high frequencies (or sometimes both) are required to ensure that τ has a finite variance. In practice, the finite duration of the simulation guarantees the former cutoff, and the finite time resolution of the simulated samples of τ guarantees the latter. It is to be noted that the designation fractal Gaussian noise properly applies only for $\alpha < 1$; the range $1 < \alpha < 3$ is generated by fractal Brownian motion [19]. In the interest of simplicity, however, we employ the term fractal Gaussian noise for all values of α .

There are a number of recipes in the literature for generating fractal Gaussian noise [93–96]. All typically result in a sampled version of fractal Gaussian noise, with equal spacing between samples. A continuous version of the signal is obtained by using a zeroth-order interpolation between the samples.

The use of a fractal Gaussian noise kernel in the integrate-and-fire construct results in the so-called fractal-Gaussian-noise driven integrate-and-fire model [8, 17]. The mean of the rate process is generally required to be much larger than its standard deviation, in part so that the times when the rate is negative (during which no events may be generated) remain small.

The *interval*-based spectrum is obtained by applying Eq. 7 to Eq. 28 (see Section 2.2 and Ref. 9). The following approximate result, in terms of interval-based frequency f , is obtained:

$$S_\tau(f) \approx \lambda^{-2} [\delta(f) + 1 + (\lambda f/f_0)^{-\alpha}] \quad (29)$$

We proceed to discuss two other physiologically based noise inputs [17]; both reduce to fractal Gaussian noise in appropriate limits.

6.2.2. Fractal Lognormal Noise

A related process results from passing fractal Gaussian noise through a memory-less exponential transform. This process plays a role in neurotransmitter exocytosis [97]. Since the exponential of a Gaussian is lognormal, we refer to this process as fractal lognormal noise [17, 97]. If $X(t)$ denotes a fractal Gaussian noise process with mean $E[X]$, variance $\text{Var}[X]$, and autocovariance function $K_X(\tau)$, then $\lambda(t) \equiv \exp[X(t)]$ is a fractal lognormal noise process with moments $E[\lambda^n] = \exp(nE[X] + n^2\text{Var}[X]/2)$ and autocovariance function $K_\lambda(\tau) = E^2[\lambda](\exp[K_X(\tau)] - 1)$ [97].

By virtue of the exponential transform, the autocorrelation functions of the Gaussian process $X(t)$ and the lognormal process $\lambda(t)$ differ; thus their spectra differ. In particular, the power spectral density of the resulting point process does not follow the exact form of Eq. 28, although for small values of $\text{Var}[X]$ the experimental periodogram closely resembles this ideal form [97]. In the limit of very small values of $\text{Var}[X]$, the exponential operation approaches a linear transform, and the rate process reduces to fractal Gaussian noise.

6.2.3. Fractal Binomial Noise

A third possible kernel for the integrate-and-fire model is fractal binomial noise. This process results from the addition of a number of independent identical alternating fractal renewal processes. The sum is a binomial process with the same fractal exponent as each of the individual alternating fractal processes [98, 99]. This binomial process can serve as a rate function for an integrate-and-fire process; the result is the fractal-binomial-noise driven integrate-and-fire model [17]. This construct was initially designed to model the superposition of alternating currents from a number of ion channels with fractal behavior [98–101]. As the number of constituent processes increases, fractal binomial noise converges to fractal Gaussian noise with the same fractal exponent therefore leading to the fractal-Gaussian-noise driven integrate-and-fire point process [17].

6.3. Jittered Integrate-and-Fire Model

Conventional integrate-and-fire constructs have only a single source of randomness, the rate process $\lambda(t)$. It is sometimes useful to incorporate a second source of randomness into such models. One way to carry this out is to impose random jitter on the interevent times generated in the integrate-and-fire process [17].

The procedure used for imposing this jitter is as follows. After generating the time of the $(i + 1)$ st event t_{i+1} in accordance with Eq. 27, the $(i + 1)$ st interevent time $\tau_{i+1} = t_{i+1} - t_i$ is multiplied by a dimensionless Gaussian-distributed random variable with unit mean and variance σ^2 . This results in the replacement of t_{i+1} by

$$t_{i+1} + \sigma(t_{i+1} - t_i)\mathcal{N}_i \quad (30)$$

before the simulation of subsequent events $(t_{i+2}, t_{i+3}, \dots)$. The quantity $\{\mathcal{N}_i\}$ represents a sequence of zero-mean, unity-variance independent Gaussian random variables and

the standard deviation σ is a free parameter that controls the strength of the proportional jitter. In accordance with the construction of the model, events at subsequent times experience jitter imposed by all previous events (see Figure 14).

The overall result is the jittered integrate-and-fire model. The jitter serves to introduce a source of scale-independent noise into the point process. Depending on the character of the input rate function, this model can give rise to the fractal-Gaussian-noise, fractal-lognormal-noise, or fractal-binomial-noise driven jittered integrate-and-fire processes, among others. Two limiting behaviors readily emerge from this model. When $\sigma \rightarrow 0$, the jitter disappears and the jittered integrate-and-fire model reduces to the ordinary integrate-and-fire model. At the other extreme, when $\sigma \rightarrow \infty$, the jitter dominates and the result is a homogeneous Poisson process. The fractal behavior present in the rate function $\lambda(t)$ then disappears and $\lambda(t)$ behaves as if it were constant. Between these two limits, as σ increases the fractal onset time $1/f_0$ of the resulting point process increases and the fractal characteristics of the point process are progressively lost, first at higher frequencies (shorter times) and subsequently at lower frequencies (longer times) [17].

6.3.1. Simulating the Jittered Integrate-and-Fire Point Process

We use a fractal Gaussian noise kernel in the jittered integrate-and-fire construct to simulate the human heartbeat. Fractal Gaussian noise with the appropriate mean, standard deviation, and fractal exponent α_S is generated using Fourier-transform methods. This noise serves as the kernel in an integrate-and-fire element (Eq. 27), the output of which is jittered in accordance with Eq. 30. For each of the 12 normal subjects and 15 CHF (without and with AF) patients, the model parameters were adjusted iteratively to obtain a simulated data set that matched the patient recording as closely as possible. In particular, each of the 27 simulations contained exactly $L_{\max} = 75,821$ RR intervals, displayed mean heart rates that fell within 1% of the corresponding data recordings, and exhibited wavelet-transform standard deviations that were nearly identical to those of the data. The simulation results for CHF patients with and without atrial fibrillation were very similar.

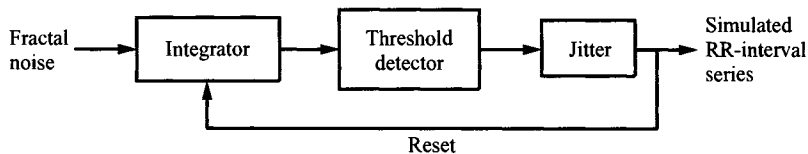


Figure 14 Schematic representation of the jittered integrate-and-fire model for generating a simulated RR-interval series. An underlying rate process $\lambda(t)$, assumed to be band-limited fractal Gaussian noise, is integrated until it reaches a fixed threshold θ , whereupon a point event is generated. The occurrence time of the point event is jittered in accordance with a Gaussian distribution and the integrator is then reset. The continuous rate process is thereby converted into a point process representing the sequence of R phases in the human heartbeat.

6.3.2. Statistics of the Simulated Point Process for Normal Subjects and CHF Patients

Having developed the jittered integrate-and-fire model and simulated sequences of RR intervals, we proceed to examine some of the statistical features of the resulting point processes.

Figure 15 displays representative results in which the data (solid curves) and simulations (dotted curves) are compared for a single normal subject (left column) and a single CHF (without AF) patient (right column). Figure 15a illustrates that the RR-interval sequences obtained from the model and the data are qualitatively similar for both the normal and heart-failure cases. Furthermore, the agreement between the simulated and actual wavelet-transform standard-deviation curves is excellent in both cases, as is evident in Figure 15b. Even the gentle increase of σ_{wav} for the heart-failure patient at small scale values is reproduced, by virtue of the jitter introduced into the model. It is apparent in Figure 15c, however, that the simulated RR-interval histogram for the heart-failure patient does not fit the actual histogram well; it is too broad, which indicates that σ_{int} of the simulation is too large. Finally, Figure 15d shows that the simulated spectra of the RR intervals match the experimental spectra quite nicely, including the whitening of the heart-failure spectrum at high frequencies. This latter effect is the counterpart of the flattening of σ_{wav} at small scales and results from the white noise introduced by the jitter at high frequencies.

6.3.3. Simulated Individual Value Plots and ROC-Area Curves

To examine how well the jittered integrate-and-fire model mimics the collection of data from a global perspective, we constructed individual-value plots and ROC-area curves using the 27 simulated data sets. The results are presented in Figures 16 and 17, respectively. Direct comparison should be made with Figures 5 and 8, respectively, which provide the same plots for the actual heart-failure patient and normal-subject data.

Comparing Figure 16 with Figure 5 and Figure 17 with Figure 8, we see that in most cases the simulated and actual results are remarkably similar. Based on their overall ability to distinguish between CHF and normal simulations for the full set of 75,821 RR intervals, the 16 measures portrayed in Figure 16 roughly fall into three classes. A similar division was observed for the actual data, as discussed in Section 3.6.

As is evident in Figure 16, five measures (indicated by boldface font) succeed in fully separating the two kinds of simulations and, by definition, fall into the first class: VLF, LF, $\sigma_{\text{wav}}(32)$, $S_{\tau}(1/32)$, and DFA(32). These measures share a dependence on a single scale, or small range of scales near 32 heartbeat intervals. However, of these five measures only two (LF and VLF) successfully separate the simulated results for the next smaller number of RR intervals ($L = 65, 536$) and none succeeds in doing so for yet smaller values of L .

The same five measures also fully distinguish the normal subjects from the CHF patients (see Figure 5); however, a sixth measure, $A(10)$, also achieves this. For the actual data all six of these measures successfully separate the actual heart-failure patients from the normal subjects for data lengths down to $L = 32,768$ RR intervals.

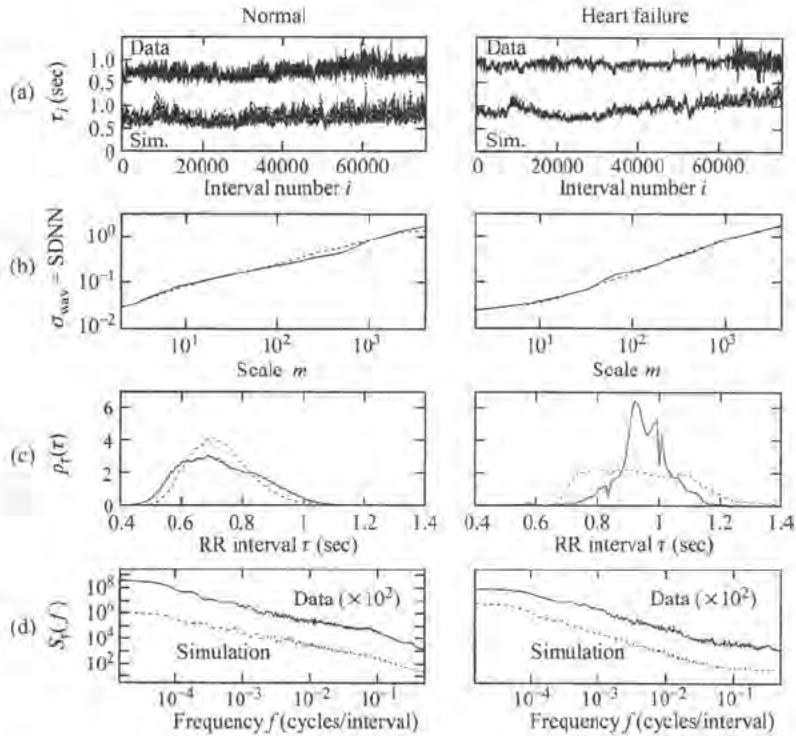


Figure 15 Comparison of data (solid curves) and simulations (dotted curves) for a single normal subject (left column, data set n16265) and a single CHF & AF patient (right column, data set a6796). The parameters $E[\tau_i]$ (representing the mean interbeat interval) and α_w (representing the fractal exponent) used in the simulations were derived from the data for the two individuals. The normal-subject simulation parameters were $1/\lambda = 0.74$ s, $\alpha_S = 1.0$, $f_0 = 0.0012$ cycles/s, and $\sigma = 0.022$; the CHF-patient simulation parameters were $1/\lambda = 0.99$ s, $\alpha_S = 1.5$, $f_0 = 0.00055$ cycles/s, and $\sigma = 0.025$. The lengths of the simulated data sets were identical to those of the experimental records analyzed. (a) RR-interval sequence over the entire data set. Qualitative agreement is apparent in both the normal and heart-failure panels. (b) Wavelet-transform standard deviation versus scale. The simulations mimic the scaling properties of the data in both cases, as well as the gentle flattening of σ_{wav} for the heart-failure patient at small scale values. (c) Interbeat-interval histogram. The model is satisfactory for the normal subject but fails to capture the narrowing of the histogram (reduction of σ_{int}) for the heart-failure patient. (d) Spectrum of the sequence of RR intervals (the data are displaced upward by a factor of 10^2 for clarity). The simulations capture the subtleties in the spectra quite well, including the whitening of the heart-failure spectrum at high frequencies.

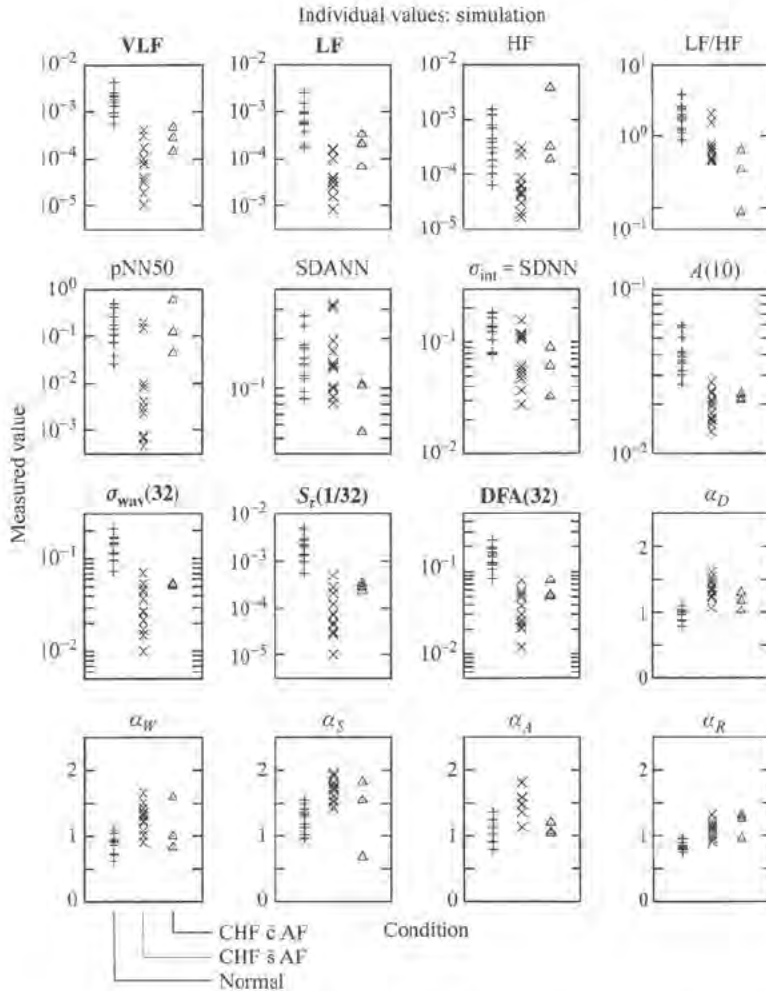


Figure 16 Individual value plots (jittered integrate-and-fire-simulations) for the 16 measures. Each panel corresponds to a different HRV measure. The three columns in each panel, from left to right, comprise data for (1) 12 simulations using normal-subject parameters (+), (2) 12 simulations using CHF-patient (\bar{s} AF) parameters (\times), and (3) 3 simulations using CHF-patient (\bar{c} AF) parameters (\triangle). Each simulation comprises 75,821 RR intervals and is carried out using parameters drawn from a single actual data set. The five measures highlighted in boldface font succeed in completely separating normal-subject and CHF-patient (\bar{s} and \bar{c} atrial fibrillation) simulations: **VLF**, **LF**, **$\sigma_{wav}(32)$** , **$S_r(1/32)$** , and **DFA(32)**. Each panel should be compared with the corresponding panel for the actual normal and CHF data in Figure 5 (leftmost three columns).

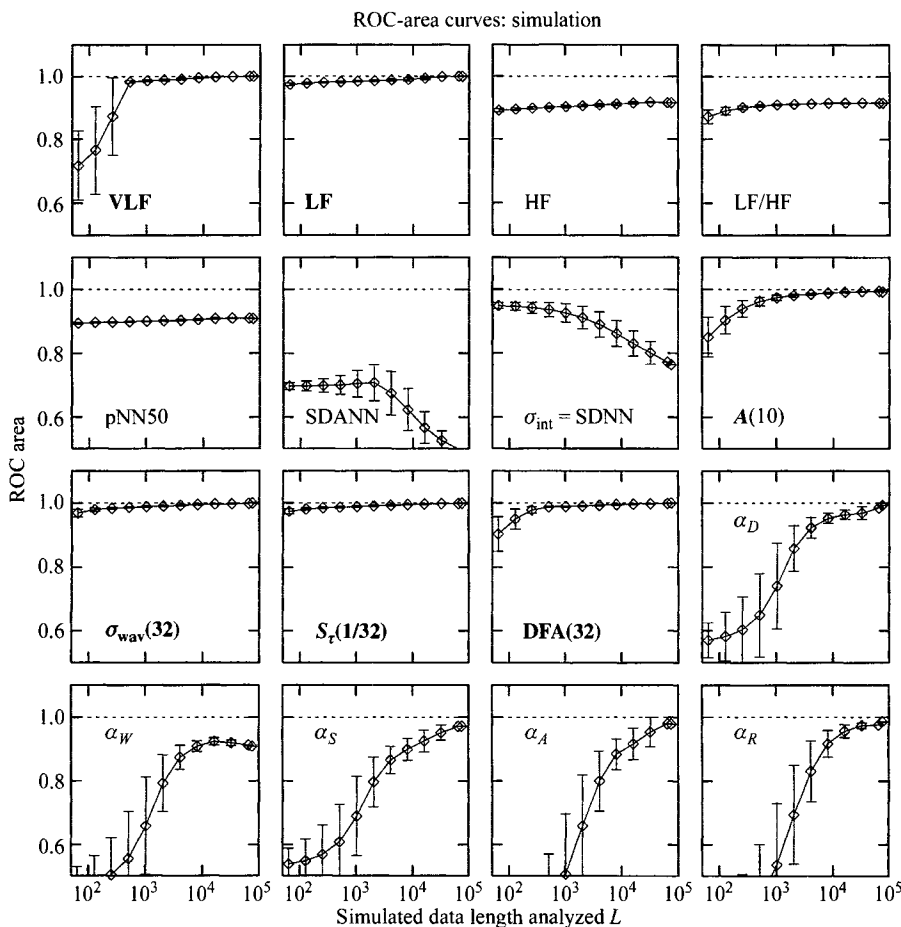


Figure 17 Area under the ROC curve (jittered integrate-and-fire simulations) as a function of number of RR intervals analyzed (L) for the 16 HRV measures (mean \pm 1 SD). An area of unity corresponds to perfect separability of the two classes of simulations. The five measures highlighted in boldface font [VLF, LF, $\sigma_{wav}(32)$, $S_r(1/32)$, and DFA(32)] provide such perfect separability but only for the largest number of RR intervals analyzed. Each panel should be compared with the corresponding panel for the actual normal and CHF data in Figure 8. The ROC-area simulations differ from those for the data in two principal respects: the performance of the two interval measures SDANN and σ_{int} severely degrades as the number of RR intervals increases, and the performance of the five fractal-exponent measures is substantially enhanced as the number of RR intervals increases.

The simulated ROC-area curves presented in Figure 17 resemble the curves for the normal and CHF data shown in Figure 8, but there are a few notable distinctions. The simulated and actual ROC areas (Figures 17 and 8, respectively) for the interval measures SDANN and σ_{int} are just about the same for $L = 64$ RR intervals. However, both of these measures exhibit *decreasing* ROC areas as the number of simulated RR intervals (L) increases (Figure 17). In contrast, the ROC areas based on the actual data

increase with increasing data length L in normal fashion (Figure 8). The decrease in ROC area observed in Figure 17 means that performance is degraded as the simulated data length increases. This paradoxical result probably arises from a deficiency of the model; it proved impossible to fit simultaneously the wavelet-transform standard deviation at all scales, the mean heart rate, and the RR-interval standard deviation. Choosing to fit the two former quantities closely to the data rendered inaccurate the simulated interval standard deviation. Since shorter files have insufficient length to exhibit the longer term fluctuations that dominate fractal activity, the interval standard deviation for small values of L is not influenced by these fluctuations. Apparently, the manner in which the long-term effects are expressed differs in the original data and in the simulation.

The most surprising distinction between the actual and simulated results, perhaps, is the dramatic putative improvement in the ability of the five scale-independent fractal-exponent measures to separate simulated heart-failure and normal data as the number of RR intervals increases. The Allan factor exponent α_A appears to offer the greatest "improvement," as seen by comparing Figures 17 and 8. All of the exponents except α_W attain an ROC area exceeding 0.97 for $L_{\max} = 75,821$ RR intervals. Nevertheless, the improved separation yields performance that remains inferior to that of the fixed-scale measures. The apparent improved separation in the simulated data may therefore arise from a nonlinear interaction in the model between the fractal behavior and the signal component near a scale of 32 heartbeat intervals. Or it may be that some aspect of the data, not reliably detected by fractal-exponent measures applied to the original data, emerges more clearly following the modeling and simulation. This could mean that another measure of the fractal exponent of the actual data, perhaps incorporating in some way the processing provided by the modeling procedure, might yield better results than the more straightforward fractal-exponent measures that we have used to this point.

6.3.4. Limitations of the Jittered Integrate-and-Fire Model

Figure 18 presents a direct comparison between four measures derived from the original data and from the corresponding simulations. Each panel contains 12 normal results (+), 12 CHF without AF results (\times), and 3 CHF with AF results (Δ). The diagonal lines correspond to perfect agreement.

The results for $S_\tau(1/32)$ and $\sigma_{\text{wav}}(32)$ are excellent, illustrating the remarkable ability of the model to mimic the data at a time scale of 32 interbeat intervals. The correlation coefficients for these two measures lie very close to unity, indicating almost perfect correspondence between the original and simulated data.

The upper left panel presents values for σ_{int} , the standard deviation of the RR intervals. Agreement between the simulation and original data is only fair, with simulated values of σ_{int} generally somewhat higher than desired. The correlation coefficient $\rho = 0.71$ for all 27 simulations indicates significant correlation ($p < 0.0003$ that 27 random pairs of numbers with the same distribution would achieve a magnitude of $\rho \geq 0.71$) but not extremely close agreement between simulated and original RR-interval sequences.

This disagreement highlights a problem with the simple jittered integrate-and-fire model. The fractal-Gaussian-noise integrate-and-fire model *without* jitter yields sub-

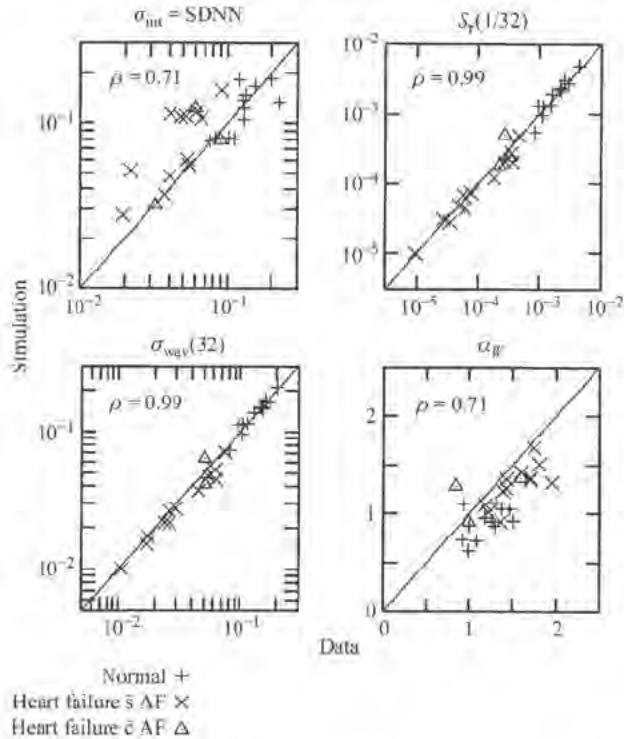


Figure 18 Simulation accuracy for four measures with their correlation coefficients ρ . The jittered integrate-and-fire model performs remarkably well for the measures $S_r(1/32)$ and $\sigma_{wav}(32)$; however, it does not perform nearly as well for the measures σ_{int} and α_W . These disagreements highlight problems with the simple jittered integrate-and-fire model.

stantially closer agreement between the data and simulation for this statistic [8]. Although the introduction of jitter leads to improved agreement for σ_{wav} , it brings the model results farther away from the data as measured by σ_{int} . Apparently, a method other than jitter must be devised to forge increased agreement with σ_{wav} while not degrading the agreement with σ_{int} . One possibility is the reordering of the individual RR intervals over short time scales.

The simulated exponent α_W also fails to show close agreement with the data, suggesting that there are other features of the model that are less than ideal. However, since this measure is of little use in separating CHF patients from normal subjects, this disagreement takes on reduced importance.

6.4. Toward an Improved Model of Heart Rate Variability

Agreement between the simulations and data would most likely be improved by adding other inputs to the model aside from fractal noise, as illustrated schematically in Figure 19. Input signals related to respiration and blood pressure are two likely candidates because it is well known that variations in these physiological functions influence

HRV. Experience also teaches us that there is a measure of white Gaussian noise present, possibly associated with the measurement process.

As indicated in Section 6.3.4, the introduction of jitter in the integrate-and-fire construct improves the agreement of the model with certain features of the data but degrades it for others. An adaptive reset will provide a more flexible alternative to the jitter. Moreover, the threshold θ could be converted into a stochastic process as a way of incorporating variability in other elements of the system [17].

All things considered, the jittered integrate-and-fire construct does a rather good job of mimicking the actual data, although the introduction of a more physiologically based model would be a welcome addition.

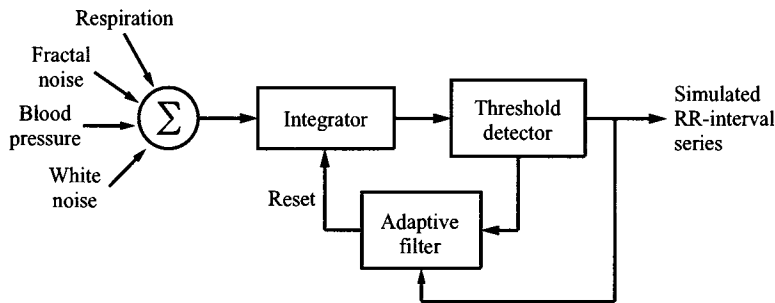


Figure 19 Potential modifications to the simple integrate-and-fire model. Multiple physiologically based inputs and adaptive feedback could serve to produce more realistic RR-interval simulations.

7. CONCLUSION

For the purposes of HRV analysis, the occurrence times of the sequence of R phases in the electrocardiogram can be represented as a fractal-rate stochastic point process. Using a collection of standard and novel HRV measures, we have examined the sequence of times between the R phases, namely the RR time series, of ECGs recorded from normal subjects and patients with congestive heart failure, as well as from several patients with other cardiovascular disorders. CHF patients who also suffered from atrial fibrillation were treated as a separate class. We examined 16 HRV measures comprising frequency-domain, time-domain, scaling-exponent, Allan-factor, detrended-fluctuation-analysis, and wavelet-transform methods.

ROC analysis was used to compare and contrast the performance of these 16 measures with respect to their abilities to distinguish properly heart-failure patients from normal subjects over a broad range of record lengths (minutes to hours). Scale-dependent measures rendered performance that was substantially superior to that of scale-independent ones. Judged on the basis of performance and computation time, 2 measures of the 16 emerged at the top of the list: the wavelet-transform standard deviation $\sigma_{\text{wav}}(32)$ and the RR interval-based power spectral density $S_{\tau}(1/32)$. They share in common the dependence on a single scale, or small range of scales, near 32

heartbeat intervals. The behavior of the ECG at this particular scale, corresponding to about 25 seconds, turns out to be a significant marker for the presence of cardiac dysfunction. The scale value itself is far more important than the details of the measures used to examine it.

Application of these techniques to a large database of normal and CHF records will make it possible to uncover just how the ECGs of pathological patients differ from those of normal subjects in the vicinity of this special scale. This would facilitate the development of an optimal diagnostic measure, which might take the form of a specialized weighting function over the interval-based power spectral density or over some combination of elementary measures.

We also addressed an issue of fundamental importance in cardiac physiology: the determination of whether the RR time series arises from a chaotic attractor or has an underlying stochastic origin. Using nonlinear-dynamics theory, together with surrogate-data analysis, we established that the RR sequences from both normal subjects and pathological patients have stochastic, rather than deterministic, origins. The use of a special embedding of the *differences* between adjacent RR intervals enabled us to reach this conclusion. This technique substantially reduced the natural correlations inherent in the interbeat-interval time series, which can masquerade as chaos and confound simpler analyses.

Finally, we developed a jittered integrate-and-fire model built around a fractal-Gaussian-noise kernel. Simulations based on this model provided realistic, although not perfect, replicas of real heartbeat sequences. The model was least successful in predicting interbeat-interval statistics and fractal-exponent values. Nevertheless, it could find use in some applications such as pacemaker excitation.

To confirm the observations we have reported, it will be desirable to carry out continuation studies using large numbers of out-of-sample records. It will also be useful to examine how the performance of the various measures correlates with the severity of cardiac dysfunction as judged, for example, by the NYHA clinical scale. Comparisons with existing CHF studies that make use of HRV measures should also be conducted.

REFERENCES

- [1] J. R. Levick, *An Introduction to Cardiovascular Physiology*. London: Butterworths, 1991.
- [2] R. Kitney, D. Linkens, A. Selman, and A. McDonald, The interaction between heart rate and respiration: Part II—Nonlinear analysis based on computer modeling. *Automedica* 4:141–153, 1982.
- [3] M. Malik, J. T. Bigger, A. J. Camm, R. E. Kleiger, A. Malliani, A. J. Moss, P. J. Schwartz, and the Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology, Heart rate variability—Standards of measurement, physiological interpretation, and clinical use. *Eur. Heart J.* 17:354–381, 1996. Also published in *Circulation* 93:1043–1065, 1996.
- [4] E. H. Hon and S. T. Lee, Electronic evaluations of the fetal heart rate patterns preceding fetal death, further observations. *Am. J. Obstet. Gynec.* 87:814–826, 1965.
- [5] R. E. Kleiger, J. P. Miller, J. T. Bigger, A. J. Moss, and the Multicenter Post-Infarction Research Group, Decreased heart rate variability and its association with increased mortality after acute myocardial infarction. *Am. J. Cardiol.* 59:256–262, 1987.

- [6] J. B. Bassingthwaite and G. M. Raymond, Evaluating rescaled range analysis for time series. *Ann. Biomed. Eng.* 22:432-444, 1994.
- [7] R. G. Turcott and M. C. Teich, Long-duration correlation and attractor topology of the heartbeat rate differ for healthy patients and those with heart failure. *Proc. SPIE* 2036:22-39, 1993.
- [8] R. G. Turcott and M. C. Teich, Fractal character of the electrocardiogram: Distinguishing heart-failure and normal patients. *Ann. Biomed. Eng.* 24:269-293, 1996.
- [9] R. W. DeBoer, J. M. Karemaker, and J. Strackee, Comparing spectra of a series of point events particularly for heart rate variability data. *IEEE Trans. Biomed. Eng.* BME-31:384-387, 1984.
- [10] D. R. Cox and P. A. W. Lewis, *The Statistical Analysis of Series of Events*. London: Methuen, 1966.
- [11] P. A. W. Lewis, ed., *Stochastic Point Processes: Statistical Analysis, Theory, and Applications*. New York: Wiley-Interscience, 1972.
- [12] B. E. A. Saleh, *Photoelectron Statistics*. New York: Springer-Verlag, 1978.
- [13] M. C. Teich and S. M. Khanna, Pulse-number distribution for the neural spike train in the cat's auditory nerve. *J. Acoust. Soc. Am.* 77:1110-1128, 1985.
- [14] P. R. Prucnal and M. C. Teich, Refractory effects in neural counting processes with exponentially decaying rates. *IEEE Trans. Syst. Man Cybern.* SMC-13:1028-1033, 1983.
- [15] M. C. Teich, C. Heneghan, S. B. Lowen, T. Ozaki, and E. Kaplan, Fractal character of the neural spike train in the visual system of the cat, *J. Opt. Soc. Am. A* 14:529-546, 1997.
- [16] S. B. Lowen and M. C. Teich, Estimation and simulation of fractal stochastic point processes. *Fractals* 3:183-210, 1995.
- [17] S. Thurner, S. B. Lowen, M. C. Feurstein, C. Heneghan, H. G. Feichtinger, and M. C. Teich, Analysis, synthesis, and estimation of fractal-rate stochastic point processes. *Fractals* 5:565-595, 1997.
- [18] W. Rudin, *Principles of Mathematical Analysis*, 3rd ed. New York: McGraw-Hill, 1976.
- [19] B. B. Mandelbrot, *The Fractal Geometry of Nature*, 2nd ed. New York: W. H. Freeman, 1983.
- [20] J. Feder, *Fractals*. New York: Plenum, 1988.
- [21] L. S. Liebovitch, *Fractals and Chaos Simplified for the Life Sciences*. New York: Oxford, 1998.
- [22] M. F. Shlesinger and B. J. West, Complex fractal dimension of the bronchial tree. *Phys. Rev. Lett.* 67:2106-2108, 1991.
- [23] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed. New York: McGraw-Hill, 1991.
- [24] A. V. Oppenheim and R. W. Schaffer, *Discrete-Time Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1989.
- [25] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C*, 2nd ed. New York: Cambridge University Press, 1992.
- [26] S. Akselrod, D. Gordon, F. A. Ubel, D. C. Shannon, A. C. Barger, and R. J. Cohen, Power spectrum analysis of heart rate fluctuation: a quantitative probe of beat to beat cardiovascular control. *Science* 213:220-222, 1981.
- [27] B. Pomeranz, R. J. Macaulay, M. A. Caudill, I. Kutz, D. Adam, D. Gordon, K. M. Kilborn, A. C. Barger, D. C. Shannon, and R. J. Cohen, Assessment of autonomic function in humans by heart rate spectral analysis. *Am. J. Physiol.* 248:H151-H153, 1985.
- [28] A. Malliani, M. Pagani, F. Lombardi, and S. Cerutti, Cardiovascular neural regulation explored in the frequency domain. *Circulation* 84:482-492, 1991.
- [29] M. M. Wolf, G. A. Varigos, D. Hunt, and J. G. Sloman, Sinus arrhythmia in acute myocardial infarction. *Med. J. Australia.* 2:52-53, 1978.

- [30] J. P. Saul, P. Albrecht, R. D. Berger, and R. J. Cohen, Analysis of long term heart rate variability: Methods, $1/f$ scaling and implications. In *Computers in Cardiology 1987*, pp. 419–422. Washington: IEEE Computer Society Press, 1988.
- [31] D. R. Cox and V. Isham, *Point Processes*. London: Chapman & Hall, 1980.
- [32] S. B. Lowen and M. C. Teich, The periodogram and Allan variance reveal fractal exponents greater than unity in auditory-nerve spike trains. *J. Acoust. Soc. Am.* 99:3585–3591, 1996.
- [33] D. W. Allan, Statistics of atomic frequency standards. *Proc. IEEE* 54:221–230, 1966.
- [34] J. A. Barnes and D. W. Allan, A statistical model of flicker noise. *Proc. IEEE* 54:176–178, 1966.
- [35] M. C. Teich, Fractal behavior of the electrocardiogram: Distinguishing heart-failure and normal patients using wavelet analysis. *Proc. Int. Conf. IEEE Eng. Med. Biol. Soc.* 18:1128–1129, 1996.
- [36] M. C. Teich, C. Heneghan, S. B. Lowen, and R. G. Turcott, Estimating the fractal exponent of point processes in biological systems using wavelet- and Fourier-transform methods. In *Wavelets in Medicine and Biology*, A. Aldroubi and M. Unser, eds., pp. 383–412. Boca Raton, FL: CRC Press, 1996.
- [37] P. Abry and P. Flandrin, Point processes, long-range dependence and wavelets. In *Wavelets in Medicine and Biology*, A. Aldroubi and M. Unser, eds., pp. 413–437. Boca Raton, FL: CRC Press, 1996.
- [38] C. Heneghan, S. B. Lowen, and M. C. Teich, Wavelet analysis for estimating the fractal properties of neural firing patterns. In *Computational Neuroscience: Trends in Research 1995*, J. M. Bower, ed., pp. 441–446. San Diego: Academic Press, 1996.
- [39] S. Mallat, A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Trans. Pattern Anal. Mach. Intel.* 11:674–693, 1989.
- [40] Y. Meyer, *Ondelettes et opérateurs*. Paris: Hermann, 1990.
- [41] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia: Society for Industrial and Applied Mathematics, 1992.
- [42] A. Aldroubi and M. Unser, eds., *Wavelets in Medicine and Biology*. Boca Raton, FL: CRC Press, 1996.
- [43] M. Akay, ed., *Time Frequency and Wavelets in Biomedical Signal Processing*. New York: IEEE Press, 1997.
- [44] M. Akay, Wavelet applications in medicine. *IEEE Spectrum* 34(5):50–56, 1997.
- [45] A. Arneodo, G. Grasseau, and M. Holschneider, Wavelet transform of multifractals. *Phys. Rev. Lett.* 61:2281–2284, 1988.
- [46] S. Thurner, M. C. Feurstein, and M. C. Teich, Multiresolution wavelet analysis of heart-beat intervals discriminates healthy patients from those with cardiac pathology. *Phys. Rev. Lett.* 80:1544–1547, 1988.
- [47] S. Thurner, M. C. Feurstein, S. B. Lowen, and M. C. Teich, Receiver-operating-characteristic analysis reveals superiority of scale-dependent wavelet and spectral measures for assessing cardiac dysfunction. *Phys. Rev. Lett.* 81:5688–5691, 1998.
- [48] M. C. Teich, Multiresolution wavelet analysis of heart rate variability for heart-failure and heart-transplant patients. *Proc. Int. Conf. IEEE Eng. Med. Biol. Soc.* 20:1136–1141, 1998.
- [49] C. Heneghan, S. B. Lowen, and M. C. Teich, Analysis of spectral and wavelet-based measures used to assess cardiac pathology. *Proceedings 1999 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 1999, paper no. SPTM-8.2.
- [50] Y. Ashkenazy, M. Lewkowicz, J. Levitan, H. Moelgaard, P. E. Bloch Thomsen, and K. Saermark, Discrimination of the healthy and sick cardiac autonomic nervous system by a new wavelet analysis of heartbeat intervals. *Fractals* 6:197–203, 1998.

- [51] C.-K. Peng, S. Havlin, H. E. Stanley, and A. L. Goldberger, Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. *Chaos* 5:82–87, 1995.
- [52] C.-K. Peng, S. Havlin, J. M. Hausdorff, J. Mietus, H. E. Stanley, and A. L. Goldberger, Fractal mechanisms and heart rate dynamics: Long-range correlations and their breakdown with disease. *J. Electrocardiol.* 28:59–65, 1996.
- [53] K. K. L. Ho, G. B. Moody, C.-K. Peng, J. E. Mietus, M. G. Larson, D. Levy, and A. L. Goldberger, Predicting survival in heart failure case and control subjects by use of fully automated methods for deriving nonlinear and conventional indices of heart rate dynamics. *Circulation* 96:842–848, 1997.
- [54] M. Kobayashi and T. Musha, $1/f$ fluctuations of heartbeat period. *IEEE Trans. Biomed. Eng.* BME-29:456–457, 1982.
- [55] C.-K. Peng, J. Mietus, J. M. Hausdorff, S. Havlin, H. E. Stanley, and A. L. Goldberger, Long-range anticorrelations and non-Gaussian behavior of the heartbeat. *Phys. Rev. Lett.* 70:1343–1346, 1993.
- [56] L. A. Nunes Amaral, A. L. Goldberger, P. Ch. Ivanov, and H. E. Stanley, Scale-independent measures and pathologic cardiac dynamics. *Phys. Rev. Lett.* 81:2388–2391, 1998.
- [57] H. E. Hurst, Long-term storage capacity of reservoirs. *Trans. Am. Soc. Civ. Eng.* 116:770–808, 1951.
- [58] W. Feller, The asymptotic distribution of the range of sums of independent random variables. *Ann. Math. Stat.* 22:427–432, 1951.
- [59] H. E. Schepers, J. H. G. M. van Beek, and J. B. Bassingthwaighte, Comparison of four methods to estimate the fractal dimension of self-affine signals. *IEEE Eng. Med. Biol. Mag.* 11:57–64, 1992.
- [60] J. Beran, *Statistics for Long-Memory Processes*. New York: Chapman & Hall, 1994.
- [61] J. B. Bassingthwaighte, L. S. Liebovitch, and B. J. West, *Fractal Physiology*. New York: Oxford, 1994.
- [62] H. L. Van Trees, *Detection Estimation, and Modulation Theory*, Part I. New York: Wiley, 1968.
- [63] A. Hald, *Statistical Theory with Engineering Applications*. New York: Wiley, 1952.
- [64] D. M. Green and J. A. Swets, *Signal Detection Theory and Psychophysics*. Los Altos, CA: Peninsula Press, 1988.
- [65] Y. Ashkenazy, M. Lewkowicz, J. Levitan, S. Havlin, K. Saermark, H. Moelgaard, P. E. Bloch Thomsen, M. Moller, U. Hintze, and H. Huikuri, Scale specific and scale independent measures of heart rate variability as risk indicators. e-print <http://xxx.lanl.gov/abs/physics/9909029>.
- [66] S. V. Buldyrev, A. L. Goldberger, S. Havlin, R. N. Mantegna, M. E. Matsa, C.-K. Peng, M. Simons, and H. E. Stanley, Long-range correlation properties of coding and noncoding DNA sequences: GenBank analysis. *Phys. Rev. E* 51:5084–5091, 1995.
- [67] A. L. Goldberger and B. J. West, Applications of nonlinear dynamics to clinical cardiology. *Ann. N. Y. Acad. Sci.* 504:195–213, 1987.
- [68] C.-S. Poon and C. K. Merrill, Decrease of cardiac chaos in congestive heart failure. *Nature* 389:492–495, 1997.
- [69] J. K. Kanters, N.-H. Holstein-Rathlou, and E. Agner, Lack of evidence for low-dimensional chaos in heart rate variability. *J. Cardiovasc. Electrophysiol.* 5:591–601, 1994.
- [70] D. E. Roach and R. S. Sheldon, Information scaling properties of heart rate variability. *Am. J. Physiol.* 274:H1970–H1978, 1998.
- [71] J. Theiler, Estimating fractal dimension. *J. Opt. Soc. Am. A* 7:1055–1073, 1990.
- [72] L. S. Liebovitch and T. Tóth, A fast algorithm to determine fractal dimensions by box counting. *Phys. Lett. A* 141:386–390, 1989.
- [73] M. Ding, C. Grebogi, E. Ott, T. Sauer, and J. A. Yorke, Plateau onset for correlation dimension: When does it occur? *Phys. Rev. Lett.* 70:3872–3875, 1993.

- [74] P. Grassberger and I. Procaccia, Measuring the strangeness of strange attractors. *Physica D* 9:189–208, 1983.
- [75] L. S. Liebovitch, Introduction to the properties and analysis of fractal objects, processes, and data. In *Advanced Methods of Physiological System Modeling*, V. Z. Marmarelis, ed., Vol. 2, pp. 225–239. New York: Plenum, 1989.
- [76] A. Wolf, J. B. Swift, H. L. Swinney, and J. A. Vastano, Determining Lyapunov exponents from a time series. *Physica D* 16:285–317, 1985.
- [77] A. R. Osborne and A. Provenzale, Finite correlation dimension for stochastic systems with power-law spectra. *Physica D* 35:357–381, 1989.
- [78] S. J. Schiff and T. Chang, Differentiation of linearly correlated noise from chaos in a biologic system using surrogate data. *Biol. Cybern.* 67:387–393, 1992.
- [79] J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, and J. D. Farmer, Testing for nonlinearity in time series: The method of surrogate data. *Physica D* 58:77–94, 1992.
- [80] A. M. Albano, J. Muench, C. Schwartz, A. I. Mees, and P. E. Rapp, Singular-value decomposition and the Grassberger–Procaccia algorithm. *Phys. Rev. A* 38:3017–3026, 1988.
- [81] C. Grebogi, E. Ott, and J. A. Yorke, Chaos, strange attractors, and fractal basin boundaries in nonlinear dynamics. *Science* 238:632–638, 1987.
- [82] F. X. Witkowski, K. M. Kavanagh, P. A. Penkoske, R. Plonsey, M. L. Spano, W. L. Ditto, and D. T. Kaplan, Evidence for determinism in ventricular fibrillation. *Phys. Rev. Lett.* 75:1230–1233, 1995.
- [83] A. Garfinkel, J. N. Weiss, W. L. Ditto, and M. L. Spano, Chaos control of cardiac arrhythmias. *Trends Cardiovasc. Med.* 5:76–80, 1995.
- [84] C. Heneghan, S. M. Khanna, Å. Flock, M. Ulfendahl, L. Brundin, and M. C. Teich, Investigating the nonlinear dynamics of cellular motion in the inner ear using the short-time Fourier and continuous wavelet transforms. *IEEE Trans. Signal Process.* 42:3335–3352, 1994.
- [85] M. C. Teich, C. Heneghan, S. M. Khanna, Å. Flock, M. Ulfendahl, and L. Brundin, Investigating routes to chaos in the guinea-pig cochlea using the continuous wavelet transform and the short-time Fourier transform. *Ann. Biomed. Eng.* 23:583–607, 1995.
- [86] M. C. Teich, C. Heneghan, and S. M. Khanna, Analysis of cellular vibrations in the living cochlea using the continuous wavelet transform and the short-time Fourier transform. In *Time Frequency and Wavelets in Biomedical Engineering*, M. Akay, ed., pp. 243–269. New York: IEEE Press, 1997.
- [87] D. J. Christini and J. J. Collins, Controlling nonchaotic neuronal noise using chaos control techniques. *Phys. Rev. Lett.* 75: 2782–2785, 1995.
- [88] H. C. Tuckwell, *Stochastic Processes in the Neurosciences*. Philadelphia: Society for Industrial and Applied Mathematics, 1989.
- [89] R. D. Berger, S. Akselrod, D. Gordon, and R. J. Cohen, An efficient algorithm for spectral analysis of heart rate variability. *IEEE Trans. Biomed. Eng.* BME-33:900–904, 1986.
- [90] B. W. Hyndman and R. K. Mohn, A pulse modulator model for pacemaker activity. *Digest 10th Int. Conf. Med. Biol. Eng.*, Dresden, Germany, p. 223, 1973.
- [91] B. W. Hyndman and R. K. Mohn, Model of the cardiac pacemaker and its use in decoding the information content of cardiac intervals. *Automedica* 1:239–252, 1975.
- [92] O. Rompelman, J. B. I. M. Snijders, and C. J. van Spronsen, The measurement of heart rate variability spectra with the help of a personal computer. *IEEE Trans. Biomed. Eng.* BME-29:503–510, 1982.
- [93] P. Flandrin, Wavelet analysis and synthesis of fractional Brownian motion. *IEEE Trans. Inform. Theory* 38:910–917, 1992.
- [94] B. B. Mandelbrot, A fast fractional Gaussian noise generator. *Water Resour. Res.* 7:543–553, 1971.

- [95] T. Lundahl, W. J. Ohley, S. M. Kay, and R. Siffert, Fractional Brownian motion: A maximum likelihood estimator and its application to image texture. *IEEE Trans. Med. Imaging* 5:152–161, 1986.
- [96] M. A. Stoksik, R. G. Lane, and D. T. Nguyen, Accurate synthesis of fractional Brownian motion using wavelets. *Electron. Lett.* 30:383–384, 1994.
- [97] S. B. Lowen, S. S. Cash, M.-M. Poo, and M. C. Teich, Quantal neurotransmitter secretion rate exhibits fractal behavior. *J. Neurosci.* 17:5666–5677, 1997.
- [98] S. B. Lowen and M. C. Teich, Fractal renewal processes generate $1/f$ noise. *Phys. Rev. E* 47:992–1001, 1993.
- [99] S. B. Lowen and M. C. Teich, Fractal renewal processes, *IEEE Trans. Inform. Theory* 39: 1669–1671, 1993.
- [100] S. B. Lowen and M. C. Teich, Fractal auditory-nerve firing patterns may derive from fractal switching in sensory hair-cell ion channels. In *Noise in Physical Systems and 1/f Fluctuations*, P. H. Handel and A. L. Chung, eds., pp. 781–784. New York: American Institute of Physics, 1993. AIP Conference Proceedings 285.
- [101] L. S. Liebovitch and T. I. Tóth, Using fractals to understand the opening and closing of ion channels. *Ann. Biomed. Eng.* 18:177–194, 1990.

APPENDIX A

Table A1 provides a summary of some significant statistics of the RR records analyzed in this chapter, before truncation to 75,821 intervals.

TABLE A1 Elementary Statistics of the Original RR-Interval Recordings before Truncation to 75,821 RR Intervals

File name	Condition	Number of RR intervals	Recording duration (s)	RR-interval mean (s)	Mean rate (s) ⁻¹	RR interval SD (s)
n16265	Normal	100,460	80,061	0.797	1.255	0.171
n16272	Normal	93,177	84,396	0.906	1.104	0.142
n16273	Normal	89,846	74,348	0.828	1.208	0.146
n16420	Normal	102,081	77,761	0.762	1.313	0.101
n16483	Normal	104,338	76,099	0.729	1.371	0.089
n16539	Normal	108,331	84,669	0.782	1.279	0.150
n16773	Normal	82,160	78,141	0.951	1.051	0.245
n16786	Normal	101,630	84,052	0.827	1.209	0.116
n16795	Normal	87,061	74,735	0.858	1.165	0.212
n17052	Normal	87,548	76,400	0.873	1.146	0.159
n17453	Normal	100,674	74,482	0.740	1.352	0.103
nc4	Normal	88,140	71,396	0.810	1.235	0.132
a6796	HF \bar{s} AF	75,821	71,934	0.949	1.054	0.091
a8519	HF \bar{s} AF	80,878	71,948	0.890	1.124	0.062
a8679	HF \bar{s} AF	119,094	71,180	0.598	1.673	0.051
a9049	HF \bar{s} AF	92,497	71,959	0.778	1.285	0.058
a9377	HF \bar{s} AF	90,644	71,952	0.794	1.260	0.060
a9435	HF \bar{s} AF	114,959	71,158	0.619	1.616	0.034
a9643	HF \bar{s} AF	148,111	71,958	0.486	2.058	0.024
a9674	HF \bar{s} AF	115,542	71,968	0.623	1.605	0.084
a9706	HF \bar{s} AF	115,064	71,339	0.620	1.613	0.100
a9723	HF \bar{s} AF	115,597	71,956	0.622	1.607	0.027
a9778	HF \bar{s} AF	93,607	71,955	0.769	1.301	0.070
a9837	HF \bar{s} AF	115,205	71,944	0.624	1.601	0.066
a7257	HF \bar{c} AF	118,376	71,140	0.601	1.664	0.039
a8552	HF \bar{c} AF	111,826	71,833	0.642	1.557	0.062
a8988	HF \bar{c} AF	118,058	71,131	0.603	1.660	0.091
tp987	Transplant	106,394	67,217	0.632	1.583	0.083
vt973	VT	86,992	70,044	0.805	1.242	0.202
sd984	SCD	77,511	62,786	0.810	1.234	0.089
slp59	Apnea	16,874	14,399	0.853	1.172	0.103
slp66	Apnea	15,751	13,199	0.838	1.193	0.103