Published in final edited form as:

Comput Biol Med. 2013 September 1; 43(9): 1154–1166. doi:10.1016/j.compbimed.2013.05.018.

Path Length Entropy Analysis of Diastolic Heart Sounds

B. Griffel^{a,*}, M. K. Zia^c, V. Fridman^b, C. Saponieri^b, and J. L. Semmlow^a

^aDepartment of Surgery, University of Medicine and Dentistry of New Jersey Clinical Academic Building, 125 Patterson Street, New Brunswick, NJ 08901

^bDepartment of Cardiology, Long Island College Hospital, 339 Hicks Street Brooklyn, NY 11201

^cDepartment of Psychology and Behavioral Science, SUNY Stonybrook 101 Nicolls Road, Stony Brook, NY 11794

Abstract

Early detection of coronary artery disease (CAD) using the acoustic approach, a noninvasive and cost-effective method, would greatly improve the outcome of CAD patients. To detect CAD, we analyze diastolic sounds for possible CAD murmurs. We observed diastolic sounds to exhibit $1/f$ structure and developed a new method, path length entropy (PLE) and a scaled version (SPLE), to characterize this structure to improve CAD detection. We compare SPLE results to Hurst exponent, Sample entropy and Multi-scale entropy for distinguishing between normal and CAD patients. SPLE achieved a sensitivity-specificity of 80%–81%, the best of the tested methods. However, PLE and SPLE are not sufficient to prove nonlinearity, and evaluation using surrogate data suggests that our cardiovascular sound recordings do not contain significant nonlinear properties.

Keywords

Entropy; Coronary Artery Disease; Nonlinear Signal Processing; Bioacoustics; Diastolic Murmurs; Noninvasive Measurements; $1/f$ Noise

1. Introduction

Coronary artery disease (CAD) is the leading cause of death in the United States among men and women, with over 800,000 deaths each year [1]. CAD development is a long-term process that involves the narrowing of the coronary arteries due to buildup of plaque. This may cause angina or myocardial infarction [2][3], but because of compensatory mechanisms within the vasculature, symptoms of CAD may not present until stenosis is severe (greater than 80% occlusion) [2]. Additionally, current diagnostic exams for CAD such as the catheter angiogram, the cardiac stress test, and the CT angiogram, are not suitable for early screening because of the associated risk and expense [2]. Compensatory mechanisms and screening limitations both limit early detection of CAD, when its progression can be

© 2013 Elsevier Ltd. All rights reserved.

*Corresponding author: benjamin.griffel@gmail.com (B. Griffel).

Conflicts of Interest

John L. Semmlow holds a financial interest in SonoMedica.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

controlled using drugs and diet. Therefore, there is a great need for an easy to administer, noninvasive diagnostic test that can evaluate patients before symptoms occur.

Ideally, a screening tool should be noninvasive, easy to administer, low cost, and risk free. Acoustic-based detection methods that simply involve placing microphones on the chest would fit this ideal, provided acoustic signatures of CAD could be detected with high accuracy. Turbulent blood flow in other arteries, such as the carotid, generate audible signals, termed bruits. Therefore, it should be possible to acoustically detect bruits caused by turbulence in the coronary arteries [4][5][6]. Dock [7] Sangster [8], and Duncan [9] have all reported detection of bruits associated with CAD through standard auscultation, but CAD bruits are typically inaudible and thus electronic recording and signal processing are needed to detect their signatures.

The acoustic approach was first proposed by Semmlow *et al.* [10] who used sound recordings from the chest to examine differences in the spectral domain of normal and CAD subjects. This study found greater power in the frequency range of 100 to 180 Hz for subjects with known coronary artery disease. Akay M. *et al.* and Akay Y.M *et al.* used more advanced spectral methods such as adaptive line enhancement [11], parametric modeling [12], autoregressive modeling [13], and wavelet analysis [14]. These analyses all examined spectral properties, which is appropriate for linear signals, but turbulence is a nonlinear process [15] and acoustic signals caused by turbulence may contain nonlinear structures. Nonlinear analysis methods could provide improved detection of acoustic signatures containing turbulence.

Padmanabhan [16] was the first to apply a nonlinear method to detect CAD. He successfully used the Grassberger Procaccia algorithm [17] to estimate the correlation dimension for differentiating normal and CAD subjects. Akay used approximate entropy (ApEn) [18], a different nonlinear measurement that detects repeatability, and found greater entropy in diastolic segments recorded from CAD subjects compared to normal subjects. Schmidt *et al.* [19] showed success using Sample entropy (SampEn), an improvement of the ApEn algorithm, and autoregressive (AR) modeling to diagnose CAD.

Of the research described previously, no group has demonstrated clinically relevant accuracy. If CAD signals contain nonlinearity, nonlinear methods may pick up additional differences in signal characteristics and push accuracy levels within reach of clinical relevance. However, it has yet to be proven that the acoustic signatures of CAD actually contain nonlinearity. Schmidt [20] concluded that they do not, but it is possible this is simply due to limitations in their cardiac microphone. A more sensitive sensor may pick up nonlinear signatures that were previously missed.

We have previously analyzed the data in this study in Griffel et al. [21]. In that study we used mutual information analysis to discriminate between normal and diseased subjects. Here we have developed our own analysis method and applied it to our data set. A spectral analysis of the diastolic heart sound recordings revealed a seemingly $1/f$ type power spectrum. This suggested the possibility of nonlinear properties within our measured signal. As Osborne and Provenzale [22] showed that a $1/f$ process may result in a fractional correlation dimension, Padmanabhan's correlation dimension estimates may have reflected a $1/f$ process and not dynamical chaos. Thus in this study we devised a signal measurement that probes for long-term correlations typical of $1/f$ noise signal. A surrogate data analysis was also performed to test for nonlinearity. Our analysis did not demonstrate the presence of nonlinear dynamics in our measured signals, but our developed measurement was effective for separating normal and diseased subjects.

The proposed measure is similar to the Hurst exponent [23], a classic measurement used to examine a $1/f$ process. The Hurst exponent characterizes the predictability of a signal, which gives a sense of the probability that a signal will increase or decrease at the next sample. This predictability is a property of long-term correlation within a signal, however the Hurst exponent does not measure the long-term correlation directly. Here we develop a nonlinear measurement, termed path length entropy (PLE) which measures the correlation in a more direct way through the analysis of signal predictability. PLE examines signal predictability through the measurement of the distribution and lengths of monotonic segments within a signal. By measuring predictability more directly, this method is more sensitive to subtle differences in the structure of $1/f$ signals [24], differences which may not be apparent through Hurst exponent analysis. We will show that this improves classification in normal and diseased subjects over our previous study [21]. This measurement can also be used to probe for scaling behavior of a signal by estimating PLE after a series of coarse graining steps, a measurement we term SPLE.

To demonstrate the sensitivity of SPLE to nonlinear signatures, we used it to evaluate the logistic map, a well known nonlinear system [25], as well as surrogate data created from logistic map series. Surrogate data has the same linear properties as the original time series, but none of the nonlinear properties. Therefore we can also use surrogate data to probe for nonlinear properties within our heart sound recordings.

To compare SPLE with other common nonlinear measurements, we also analyzed heart sounds using the Hurst exponent, SampEn and Multiscale entropy (MSE). We show that SPLE effectively discriminates between normal and diseased subjects and showed superior discriminatory ability compared to all other measures. We also show statistically significant differences between scaling behavior of sounds from normal and diseased subjects. However, both normal and diseased subjects show similar trends with respect to length scale, and neither is scale invariant with respect to entropy. We conclude that SPLE is an effective measurement of cardiac sounds and may be useful to examine properties of normal and diseased subjects. Surrogate data analysis indicates there may be some weak nonlinearity present in our signals, however this nonlinearity appears unable to aid in discrimination of normal and diseased subjects and therefore may not be related to CAD.

2. Methods

2.1. Development of Path Length Entropy

Path length entropy is a weighted and log scaled average of the probabilities that a sample in the trajectory of a signal is on a rising or falling segment of a given length. It is similar to the Hurst exponent in that it characterizes the tendency for a signal rise or fall, and thus is a measure of its predictability and long-term correlation. Results of the first three steps of the PLE calculation are shown in Fig. 1. To calculate PLE, we begin by differentiating the signal (solid line, Fig. 1) with respect to time. We take the SGN function of the differentiated signal to yield $PL(n)$ (dashed line, Fig. 1), as the sign of the slope indicates whether the segment is rising or falling.

$$PL(n) = \text{SGN}\left(\frac{dx(n)}{dn}\right). \quad (1)$$

Next we determine $C(i)$, which is the count of lengths of monotonic segments $PL(n)$ within our signal as a function of i , where i is the i th start of a monotonic interval (Fig. 1). $C(i)$ is thus a series of path lengths of monotonic segments within our original time series, and is defined in equation (2)

$$C(i) = \sum_i^{i+1} |PL(n)| \quad (2)$$

which is the sum of values of $PL(n)$ between i and $i + 1$. Summing the counts $C(i)$ of each length yields a histogram in which each bin represents a monotonic path length. Bin sizes are chosen such that there is a bin for any path length that occurs at least once. Multiplying the counts by their path length gives the total number of points on a monotonic segment of given length. For example, if there are 20 counts of monotonic segments of length 5, then there are 100 samples contained within those segments. After dividing the weighted histogram by the total number of samples in $x(n)$, we arrive at a probability distribution function $p(C)$. This represents the probabilities of a given sample being on a segment of points which belong to a rising or falling segment of a given length. An example of histograms $C(i)$ for four examples of artificial signals is shown in Fig. 2. Artificial signals were generated in MATLAB (Natick, MA).

Figure 3 shows the difference between the probability distribution function $p(C)$ and the histogram function $C(i)$. Since the probability distribution $p(C)$ is weighted by the path lengths themselves, $p(C)$ is not simply $C(i)$ scaled by the number of path length counts. This operation would give the probability function for the individual path lengths themselves. However, what we are interested in is not the probability of individual path lengths, but rather the tendency of values along the signal. By weighting the path length counts by their path lengths and then dividing by the number of samples, $p(C)$ gives the probability that a sample exists on a path length of a particular length. Hence, a larger number of samples may exist on a segment of very long but uncommon path length, than exist on segments of shorter but more common lengths. This is the case as shown as Fig. 3, which shows $p(C)$ and $C(i)$ for Gaussian noise and fractional Brownian motion (fBm). In Fig. 3(A), $C(i)$ for Gaussian noise, the longest path length is 1, but most samples exist on segments of 4 samples long as shown in $p(C)$ (Fig. 3(B)). Likewise $C(i)$ for the fBm signal shows the most common path lengths are 1 and 2 samples (Fig. 3(C)), but $p(C)$ shows most samples in the signal exist along segments 7 samples in length (Fig. 3(D)).

Using Shannon's entropy equation,

$$PLE(x(n)) = - \sum_C p(C) \log_2 p(C), \quad (3)$$

we can calculate an entropy associated with path length, PLE. PLE will have a maximum value for a signal in which the frequency of path lengths is inversely proportional to their length, somewhat approximated by Fig. 2(C). PLE takes into account not only the values of the average path lengths in a signal, but the distribution of the path lengths. A signal such as fractional noise will tend to have a distribution of path lengths with a wider range compared to Gaussian or identically individually distributed noise (Fig. 2). A minimal entropy calculation will occur for data that monotonically increases or decreases, a totally non-random signal.

PLE can also be used in conjunction with a scaling analysis to probe scaling behavior. We used the modified coarse graining procedure given by Valencia *et al.* [26]. Here a signal is coarse grained by resampling such that its length is reduced by a scale factor j . The resampling steps are preceded by filtering with a lowpass Butterworth IIR filter. This filter prevents aliasing of higher frequencies which would normally be caused by resampling.

Here we use an 4th-order filter. For every coarse graining step j , the cutoff frequency is reduced to $0.75 \cdot F_s/j$, where F_s/j is the new sampling frequency after the resampling step.

After every coarse graining step, PLE is recalculated. The curve of PLE Vs scale is then integrated, performed as a summation, equation 4, to give a second PLE parameter, which we term scaled PLE (SPLE). This is determined as,

$$SPLE = \sum_{s=1}^S PLE(s, x(n)), \quad (4)$$

where s is the scale factor and ranges from $(1 < s < S)$.

In this study, we used 30 scaling steps ($S = 30$). Thirty steps were chosen because our average data length is approximately 6000 samples and it was undesirable to decrease the resampled diastolic segments below 100 samples in length. In addition, statistically significant differences between normal and diseased subjects were not seen for coarse graining steps greater than 30 (Fig. 8(B)).

2.2. Traditional Nonlinear Methods

2.2.1. Hurst Exponent—Hurst exponent analysis was performed using a method given by Bruce [27]. This method is similar to a rescaled range analysis given by Mandelbrot and Taqqu [28] with two modifications. Before calculations, each segment was subtracted by its minimum value so that calculated mean values were always greater than zero, as negative values would yield imaginary numbers when taking logarithms. Secondly, instead of grouping segments of the signal and taking the average (a weak filtering and resampling), we use the coarse graining procedure as described by Valencia and used in SPLE. This procedure uses a filter and resample step, and therefore minimizes the signal aliasing that might occur with the grouping and average method of Mandelbrot and Taqqu. As part of determining the Hurst exponents, the signal was first differentiated with respect to n . This yielded Hurst exponents of a more typical value (between 0 to 1). The rescaled range $R(i)$ of samples is calculated from equation (5) where S is the standard deviation and M is the mean. $R(i)$ is calculated for scale values i for i coarse graining steps, where coarse graining is as defined for PLE. Scales factors of 11 to 41 were used as these were experimentally seen to give the best linear region in a log scaled plot (Fig. 8(C)).

$$R(i) = \frac{S(x_i(n))}{M(x_i(n))} \quad (5)$$

The Hurst exponent is given in equation (6) where H is 1 plus the slope of the log $R(i)$ plotted Vs log i .

$$H = \frac{d \log R(i)}{d \log i} + 1 \quad (6)$$

A comparison of the traditional Hurst exponent estimation (group-and-average method) and the estimation used in this study (filter-and-resample) is shown in figure 4. Here the Hurst exponent (H) of 20 fBm signals of 1,000 samples each, ranging from $H = 0.5$ to $H = .99$, were determined using both the group-and-average and filter-and-resample methods. The fractional noise was developed using a method given in Bruce [27]. The nominal Hurst exponent is shown as a dash-dot line, the measured Hurst exponent using the group-and-

average method is shown as a solid line, and the filter-and-resample method is shown as a dashed line. Generating fractional noise of a specific Hurst exponent is difficult, which is why neither method exactly matches the nominal line, but in general the traditional and resample-filter methods agree with each other.

2.2.2. Sample Entropy and Multiscale Entropy—SampEn was introduced by Richman and Moorman [29] as an improvement to ApEn. To compute SampEn, first we transform our one dimensional signal to a multidimensional signal. For a one dimensional time series, we may use delay embedding to create an m dimensional signal. Each vector is the original signal delayed by an integer multiple of delay τ . Sample entropy is based on a ratio of probability of matches within dimension m and $m+1$, where matches are defined as points within radius r of each other. Here an r of 0.1 times the standard deviation of each segment was used. SampEn is calculated as,

$$SampEn = -\log \frac{A}{B}, \quad (7)$$

$$B = \frac{\sum_{i=1}^{N-m} B_i}{N-m}, \quad (8)$$

$$A = \frac{\sum_{i=1}^{N-m-1} A_i}{N-m-1}, \quad (9)$$

where B_i is the number of points close to point x_i of the data in the embedded time series with one point subtracted to account for the self-match. A_i is calculated similarly at a dimension $m+1$. We used an m of 3 and τ of 20 samples. Both values were determined by calculating SampEn for increasing values of τ and m and determining the value at which SampEn reaches a horizontal asymptote.

Multiscale entropy [30] is the sum of the SampEn calculated for a range of scales. Coarse graining was performed as defined for PLE and thirty coarse graining steps were used.

2.3. Subject Population

The data used in this study were the same as for the study in Griffel *et al.* [21], however the subsequent analysis methods are different in this study, including the new method SPLE. Our subject pool was taken from patient volunteers from the private cardiology practice of Dr. Cesare Saponieri in Brooklyn, NY. All subjects were informed and consented to have data collected and be used for this study and the study was approved by the IRB of the Robert Wood Johnson Medical School (New Brunswick, New Jersey). Subjects were taken from a pool of patients at risk for CAD as determined via cardiac stress testing. The presence or absence of coronary artery blockages was verified via a catheter angiogram either shortly before or after data acquisition. Subjects with systolic or diastolic murmurs were excluded from the study. Our study had three classes, normal, marginal and diseased. Diseased was defined as having greater than 50% occlusion [6] (11 male and 4 female, age 63.7 ± 6.0 years). Subjects with some occlusion but less than 50% are classified as marginal (3 females and 1 male, age 59.7 ± 10.6 years) and subjects with no occlusions were classified as normal (11 females and 1 males, age 61.8 ± 9.3 years), for a total of 16 clinically normal subjects and 15 diseased.

2.4. Data acquisition

The SonoMedica (McLean, Virginia) cardiac microphone system was used for data acquisition. This system uses highly sensitive PZT crystal microphones specially designed for detecting heart sounds. The electrocardiogram (ECG) was recorded to determine the timing of the cardiac cycle. All data recordings took place in a quiet room and microphone placement was adjusted until strong S1 and S2 heart sounds (lubb and dubb) were observed in a real time recording display. Microphones were held in place with a strap to reduce breathing artifacts. Subjects were seated in a reclining position at approximately 45 degrees. Heart sounds were recorded over a 10 second period during breath-hold. The subject was instructed to exhale before the breathhold and not inhale until after the recording had completed. A 10 second recording window was chosen because some of the subjects may have found it difficult to hold their breath for a longer period. In addition, a stomach microphone was placed on the side of the abdomen to record stomach sounds and ambient noise. We have shown that a significant number of abnormal sounds in the diastolic window are produced by the stomach and therefore non-cardiovascular in nature [31].

During each 10 second recording, 4 channels of data were recorded. Two channels come from readings from microphones placed side by side on the chest. A third channel comes from a microphone placed on the abdomen. The 4th channel comes from an electrocardiogram signal. A pair of 10 second recordings were made using the side by side microphone pair, after which the microphone pair was moved to the next position and another pair of recordings were made. There were three positions used, corresponding to base (right of the sternum between 3rd and 4th rib), middle (at the sternum between 4th and 5th rib), and apex (left of the sternum between 5th and 6th rib) positions of the heart. These positions are shown in Fig. 5. Data was recorded at 22099 Hz sampling frequency and 16 bits of resolution. A lowpass filter with a cutoff of 2000 Hz was used to eliminate aliasing.

2.5. Preprocessing

Figs. 6 and 7 display examples of typical recordings from a normal (Figs. 6(A–D), 7(A–D)) and diseased (Figs. 6(E–H), 7(E–H)) subject. At this scale diseased subjects are visually indistinguishable from normal subjects so their data appear similar. Fig. 7 shows the signals shown in Fig. 6 after filtering with a 4th-order Butterworth bandpass filter with a passband of 200 to 500 Hz (the same filter used for preprocessing signals before evaluation). While there are large scale differences between the normal and diseased subjects, these low frequency differences, in particular in S1 and S2 sounds, are not of interest in this study, as only the diastolic window was considered. Within the diastolic window, the differences between the normal and diseased subjects is not apparent visually either before or after filtering.

Figs. 6 and 7 show signals from each of the 4 recording channels. Signals from the first two channels are shown in Fig. 6(A,B) and 7(E,F). Signals from the stomach channel (channel 3) are shown in Fig. 6(C,G) and 7(C,G). Figs. 6(D,H) and 7(D,H) show a recording of ECG signals with the PQRST segments labeled. The time period shown in Fig. 6 and 7 includes the first two heart sounds, S1 and S2. These sounds correspond with valve closures and are not of interest to us except as indicators of proper microphone positioning. Our region of interest is marked as the diastolic window in Fig. 6 and 7. This is the quiet period of the heart cycle, which corresponds to the heart filling with blood and is the period of greatest coronary artery blood flow.

For each subject, individual beats were isolated using a method given by Pan [32]. In each beat, the diastolic segments were further segmented by hand. Beats that showed excessive noise in the stomach channel were eliminated. For each subject there are two recordings

made over the base, middle and apex positions for a total of 6 recordings per subject. Normal and diseased subjects had a minimum of 30 beats each after editing. Before processing, all beats were digitally bandpass filtered from 200 to 500 Hz using a 4th-order Butterworth IIR filter. PLE, SPLE, Hurst, SampEn and MSE measurements were calculated on every beat over both cardiac channels and the mean and standard deviation (STD) of the total set of these values were determined, providing two parameters for each subject and measurement type.

2.6. Classification

Steve Gunn's [33] support vector machine package for MATLAB was used to generate our classifier. A first order fit was used. For PLE, SPLE, Hurst, SampEn and MSE, the classification features were the mean and STD of each measure taken over all diastolic segments per subject.

2.7. Surrogate Data Analysis

Surrogate data analysis was performed to establish the presence (or lack thereof) of nonlinearity in heart sound samples. Surrogates were created using a method given by Kantz and Schreiber [34]. Here the Fourier transform of the signal is determined and separated into magnitude M and phase φ components. Phase components above half the sampling frequency are eliminated and the remaining phase components are randomized (φ_r). A new phase is constructed using φ_{tr} , a time reversed version of φ_r . The surrogate phase is $\varphi' = [\varphi_r, -\varphi_{tr}]$, φ_r and $-\varphi_{tr}$ concatenated. A surrogate time series can then be constructed by computing the inverse FFT of $Me^{i\varphi'}$. The resulting surrogate is the same as a Gaussian signal filtered through an AR model of the measured data and has the same spectral properties as the measured data segments but does not contain any nonlinearity.

Surrogate data was evaluated using the Hurst exponent and SPLE with an analysis identical to that of the measured data. Analyses were carried out on 100 unique sets of surrogate series. Hurst exponent and SPLE were chosen because they were the only tested measurements in our study that showed the ability to discriminate normal and diseased subjects well above chance.

To determine significance of surrogate data results, a rank order analysis was performed on the classification accuracy obtained from our surrogate series and measured data. We also compared the average measurements of our surrogate sets from normal and diseased subjects with respect to the Hurst exponent and SPLE. By comparing the accuracy of discrimination between surrogates of normal and diseased subjects, we can test whether or not the Hurst exponent and SPLE are discriminating based on linear or nonlinear features, while examining the values of the metrics for surrogate data and comparing them to the metrics from measured data gives insight into the presence of any nonlinearity.

2.8. SPLE Sensitivity to Nonlinearity

In order to determine the sensitivity of SPLE to nonlinear signatures, we evaluated SPLE using the logistic map [25], a well known nonlinear mapping function. This function is given as $x[n+1] = Ax[n](1 - x[n])$. As A varies from 3 to 4, the behavior of the logistic map undergoes period doubling and becomes chaotic at approximately $A > 3.56995$. We evaluated the SPLE of the logistic map for $3 < A < 4$ and compared this to the SPLE of surrogates for the logistic map time series.

2.9. Cross-validation

Cross-validation was performed on our Hurst exponent analysis and our SPLE analysis using a 5 x 2-fold method. More than 5 runs is not advisable for a small data set because it

may underestimate the variance estimate [35]. For both folds, a random sample of 16 subjects were used as a training set for the remaining 15. Class balance was maintained for the training and test sets. For each run, the test and training sets were also swapped, this gives 10 total classifications. The mean and variance of the accuracy, sensitivity, and specificity, were taken on the set of the 10 classifications.

3. Results

Mean power spectra of normal and diseased diastolic segments are shown in Fig. 8(A). The spectra are not significantly different from each other as determined via two-tailed t-test ($p < 0.05$). The $1/f$ nature of the power spectrum is evident from this figure.

A plot of PLE Vs scale is shown in Fig. 8(B). Here statistically significant points along the curve are shown as blue triangles, where significance is defined as ($p < 0.05$) on a two tailed T-test. PLE at the base scale is not significant but becomes significant as the scale increases. At all points along the curves, there is greater PLE measured in diseased subjects than normal subjects.

Scaling behavior for the Hurst exponent and MSE for normal and diseased subjects are shown in Figs. 8(C) and 8(D). While PLE Vs scale factor shows a large region of scale values with statistically significant differences between normal and diseased subject mean values, the Hurst exponent shows significance for only a small set of scales. SampEn does not show significance at any scale value.

A separation of normal and diseased subjects is shown in Figs. 9(A) and 9(B). Fig. 9(A) shows our classifier applied to mean and STD values for PLE, while Fig. 9(B) shows the classification for segments using SPLE. The sensitivity-specificity were 80%–81% for SPLE and 60%–75% for PLE. In both plots, normal subjects are shown in black circles and diseased subjects are shown in red. Marginal subjects are shown as yellow so that they can be compared with subjects that are free of stenosis. Marginal subjects are successfully classified as clinically normal.

Hurst Exponent analysis gives a classification sensitivity-specificity of 80%–75% (Fig. 9(C)). SampEn and MSE result in classifications of, 53%–65% and 40%–69% respectively (Fig. 10).

The receiver-operator characteristic (ROC) curves for SPLE and the Hurst exponent are shown in Fig. 11. The curve for PLE is not shown but is similar to that of chance. For SPLE and Hurst exponent, the respective areas under the curve (AUC) were approximately 82% and 77% respectively. As with unscaled PLE, SampEn and MSE had ROC AUCs similar to chance.

Hurst exponent analysis and SPLE were seen to give the highest overall accuracies of our tested measurements. To investigate whether Hurst exponent analysis yielded different information than SPLE, we examined the false positives and false negatives for both measures. These are given in Table 1. Two of the four unique false negatives and two of the five unique false positives are shared between SPLE and the Hurst exponent analysis.

The best combination of mean and STD values of SPLE and Hurst exponent yielded sensitivity and specificity values of 80% and 81%. This is the same as for scaled PLE alone but slightly better than the Hurst Exponent.

3.1. Surrogate Data Analysis

The surrogate data results for both discrimination of normal and diseased subjects and normal/diseased subject values are summarized in Table 2. Here the range of the metrics for a set of 100 surrogate values is given, as well as the metric for the measured data, and the rank of the measured data. Numbers in parentheses indicate the number of surrogate sets that shared this rank. For 96/100 surrogate data sets, Hurst exponent analysis discriminated between normal and diseased subjects with higher accuracy than for measured data. For SPLE however, only two surrogate sets were able to be discriminated with better accuracy than the measured data. In both cases the range of accuracy between the surrogate sets and the measured data was only 6%. The measured data SPLE value had ranks of one and 101 for the average normal and diseased subjects respectively. For the Hurst exponent value of measured data, these were one and twelve respectively.

3.2. Cross-Validation

The results of the cross-validation analysis for both the Hurst exponent and SPLE analysis are shown in Table 3. The mean and STD of the accuracy of each of our ten classifications using data subsets for SPLE was $78 \pm 8.4\%$, with STD being the value after $\pm$. The mean and STD of the sensitivity were $83 \pm 15\%$. Mean and STD of the specificity were $74 \pm 7.1\%$. For the Hurst exponent the mean and STD of accuracy were $77 \pm 10\%$. The mean and STD of the sensitivity and specificity for Hurst exponent were $73 \pm 17\%$ and $82 \pm 23\%$.

3.3. SPLE Detection of Nonlinearity

In order to analyze the ability of SPLE to pick up nonlinear characteristics, we evaluated the logistic equation using SPLE for values of A between 3 and 4. This was compared to an analysis of surrogates of the logistic equation for the same values. The results of this analysis are given in Fig. 12. The first trace (Fig. 12(A)), shows a bifurcation plot for the logistic map. This map shows the period doubling behavior of the map, which occurs until approximately $A = 3.56995$, after which the logistic map becomes chaotic. The SPLE as a function of A is shown in Fig. 12(B). Here we can clearly see a change in the SPLE at the point in which chaos ensues, with a clear increase in SPLE. The SPLE also decreases in the short period three windows that appear at approximately $A = 3.63$, $A = 3.72$ and $A = 3.84$. The SPLE value is also very stable for $A < 3.56995$. This is different from the behavior we see in Fig. 12(C), in which the SPLE for surrogate data appears to be approximately the same until $A = 3.9$, well into the chaotic region.

4. Discussion

PLE is an entropy method that gives a quantified measurement of the randomness in a signal and has been shown to be substantially faster than MSE and SampEn [24]. This method looks at the tendency within a signal to develop monotonic segments, a tendency that reflects the presence of long-term correlations within a signal. Signals with high frequency content will tend to have shorter path lengths than signals with lower frequencies. However, PLE is not simply a measure of the number of monotonic segments/sample but also a measure of how the monotonic segments are distributed in both length and frequency. This distribution will be wider in data that contains many scales, which would result in high PLE values. PLE was therefore combined with a coarse graining algorithm to probe for scaling behavior.

PLE was motivated by the observation that diastolic segments appear to have a $1/f$ power spectrum, which is shown in Fig 8(A). We were additionally motivated by recent literature which employed entropy methods successfully for characterizing diastolic murmurs [18] [19]. Somewhat surprisingly given the previous work of Akay *et al.* and Semmlow *et al.*,

there is no significant difference observed between the power spectra of normal and diseased subjects. The FFT alone is likely not sufficient to elucidate the differences in spectral characteristics between diseased and normal subjects.

To validate the ability of SPLE to detect nonlinearities, we tested the algorithm on series from the logistic map, as well as surrogates derived from the logistic map (Fig. 12). This analysis showed that there is a sharp distinction in SPLE when the transition to chaos occurs. For surrogate data, there is no transition to chaos, and we see no significant change in the measured SPLE as a function of A . This appears to indicate that SPLE is sensitive to nonlinearity.

Examining Fig. 8(B), the curve of PLE values as a function of scale, we see no major differences between normal or diseased subjects for the general shape of the PLE Vs scale curve. However, for scale factors between 2 and 24, there is a statistically significant difference in the values of either curve. This explains why there is a stronger degree of separation between normal and diseased subjects when comparing PLE with SPLE (Figs. 9(A) and 9(B)). For all measurements shown, PLE is better able to differentiate between normal and diseased subjects when scaling is taken into account.

For Hurst exponent analysis, overall separation was nearly as good as with SPLE (Fig. 9(C)), but had fewer statistically significant differences between normal and diseased subjects when scaling was examined. This may account for the decrease in accuracy for the Hurst exponent. For SampEn and MSE, we found poor sensitivity-specificity measurements as well as AUC from the receiver operator characteristics. Unlike PLE, the addition of a scaling component did not improve results, as both SampEn and MSE give separations and ROC curves similar to chance. These results contradict Schmidt and Akay, who both found good separation of normal and diseased subjects using entropy methods. A major difference between data acquisition techniques used here versus those of Schmidt *et al.* [20] and Akay *et al.* [18] is the type of microphone used. It is likely that differences in microphone response may account for discrepancies in our results and between those of others.

We assessed our data using surrogate data to probe for evidence of nonlinearity. We used surrogate data in two ways. The first way was to compare the measured metrics (in this case Hurst exponent and SPLE) using the measured and surrogate data. The second method was to attempt to discriminate between surrogate data signals from both normal and diseased subjects. Each of these tests gives a different aspect of the nature of surrogate data. The first test is evidence for whether or not our tested metric can detect any differences between the surrogate sets and our measured test. The second test shows whether or not any potential nonlinearity in our measured signals is useful for discrimination between normal and diseased subjects.

The results of the two surrogate data tests are given in Table 2. Surrogate data and our measured data gave distinct metric values for the Hurst exponent and SPLE. The rank of the mean SPLE value for measured normal subjects was 1, while the rank for measured diseased subjects was 101. This implies that SPLE may be measuring a genuine nonlinear aspect of our signal. The mean Hurst exponent for normal subjects had a rank of 1, while for diseased subjects, the mean Hurst exponent had a rank of 12. The Hurst exponent measurement ranks also suggests that the Hurst exponent may be sensitive to a nonlinearity present in our signals, albeit somewhat less so than SPLE.

The respective rank order of Hurst and SPLE values for normal and diseased subjects is suggestive that SPLE and Hurst exponent analysis may detect some nonlinearity in our measured data; however, they also indicate that both normal and diseased subjects may contain nonlinearities. If both normal and diseased subject data contains nonlinearity, it is

unlikely that the nonlinearity is related to disease state. When we attempt to discriminate normal and diseased surrogate sets, the obtained results also fail to support that nonlinearity is related to disease state, as both Hurst exponent and SPLE were able to successfully discriminate our surrogate data sets. The Hurst exponent in fact was able to discriminate 96/100 surrogate data sets with better accuracy than for the measured data. SPLE was not able to discriminate any surrogate data set with better accuracy than for the measured data, but 3/100 sets were able to be discriminated with equal accuracy.

Our surrogate data analysis suggests that while our data may contain some nonlinear characteristics, these characteristics do not aid in discrimination of normal and diseased subjects. Therefore the detected nonlinearity may not be related to CAD. This may be because the nonlinear signatures come from a cardioacoustic source unrelated to coronary artery flow. We also can't rule out an external artifact as the source of the nonlinearity.

Since we found SPLE and the Hurst exponent to be similarly effective, we investigated if they were measuring similar signal characteristics. If Hurst exponent and SPLE share the same false positives and false negatives, this would be indicative that they are measuring the same properties within our signals. However, Hurst exponent and SPLE only shared 2/4 unique false negatives and 2/5 unique false positives, Table 1. This indicates the Hurst exponent and SPLE are measuring different characteristics within our signals. No combination of Hurst exponent, PLE, SampEn and MSE improved accuracies above that for scaled PLE. This may be because some subjects are simply not amenable to the acoustic method of CAD detection, due to high body mass index or other physiological variations.

PLE measures the randomness in a signal and how it is distributed across scales. Higher PLE in diastolic segments from CAD positive subjects indicates there is a greater degree of randomness in these signals. Under the assumption that diastolic bruits are caused by turbulent blood flow, this aligns with expectations. Marginal subjects, those with some stenosis but less than the 50% threshold, were classified along with normal subjects who were free from occlusions. However, no clear trend exists among the 4 marginal subjects. This may indicate that below a certain stenosis threshold, CAD bruits may not exist or be high enough in amplitude to be detected with this method.

Due to the small sample size in this study, we performed a cross-validation procedure to try to get a sense of the variance of our results. By sub-sampling the data into random subsets to be used as training and test sets, we can get a better sense of how well our data can be generalized and the variance of our measurements. The method we used has some known issues [35], in particular, since there will be some repeats of data points in our subsamples for every run, the variance will tend to be underestimated. One way to help remedy this is issue is to use fewer runs and to make sure that the subsets contain similar numbers of normal and diseased subjects. We have done so in our tests. The results of cross-validation gave a mean and STD for discrimination accuracy with SPLE of 78% and 8.4%. This is close to the maximum achieved accuracy of 81%. The STD is approximately an order of magnitude less than the mean, therefore it is reasonable to assume that our discrimination results are not merely due to small subject size or lack of variance in our subject pool.

Since the Hurst exponent analysis gave us a classification accuracy close to that of SPLE, at 77%, we also performed a cross-validation analysis for the Hurst exponent results, which are shown in table 3. The cross-validation results for the Hurst exponent were comparable to that of SPLE, implying that the Hurst exponent and SPLE would give similar results if there were a greater number of subjects. A two-tailed t-test of the accuracy measurements from the cross-validation of Hurst analysis and SPLE analysis gave a p value of 0.6, well above the typical level of significance. However, the mean accuracy for SPLE was still slightly

higher than for the Hurst exponent even after cross-validation. It may be that there is a small but genuine advantage to SPLE compared to the Hurst exponent that could be revealed if we had access to a greater number of subjects.

In this study we evaluated 4 methods of data analysis and found two of them lacking. It is possible that SPLE performed well solely due to chance and that MSE and SampEn failed to discriminate between normal and diseased subjects because there really aren't any true differences between our data classes. This would be a classic type one error. However, there are several reasons to believe that our discrimination is due to genuine differences between the measured data of our classes. We found statistically significant ($p < 0.05$) differences in the scaling behavior between normal and diseased subjects using a two-tailed T-test for a short range of scales using the Hurst exponent, and nearly every tested scale for SPLE. It is likely that this is the reason that Hurst exponent and SPLE were our best performing methods. Hurst exponent and SPLE are similar and elucidate similar signal characteristics, so it is reasonable that they should both be able to discriminate normal and diseased subjects. We know from past studies that heart sound recordings show differences between normal and diseased subjects so there is an a-priori reason to believe that differentiating normal and diseased subjects should be possible. A previous study using the same data as this study [21] showed the ability to discriminate subjects with 77% accuracy using mutual information. Lastly, we also showed the ability to discriminate subjects using surrogate data. If SPLE only discriminated subjects because of chance, it is unlikely that discrimination of surrogate data using SPLE would be possible. As shown in table 2, all hundred surrogate data sets were able to be discriminated using SPLE and Hurst exponent analysis with at least 74% accuracy. It is for these reasons we feel it is unlikely that the success of SPLE is due to a type-one error.

PLE here is implemented as a block function to the entire diastolic window, under the assumption that turbulent signals will be present throughout this window. This is not likely be the case, as it is known that blood flow is maximum early in diastole. A short-term PLE analysis in which PLE is developed as a function of time over the diastolic window will be performed in the future. Another future modification would be to develop a multidimensional PLE. PLE is currently implemented in one dimension, unlike SampEn and other methods that can operate in multiple dimensions. It is possible to implement this algorithm multi-dimensionally, using a multidimensional definition of a monotonic segment as either up or down. A simple definition could be Euclidean distance from the centroid of the multidimensional trajectory of the signal. This method may yield improved results as it would take into consideration information across several dimensions instead of just the time dimension.

5. Conclusion

We have demonstrated a new method, PLE and its scaled version SPLE, for nonlinear analysis. We used these methods to differentiate between normal and diseased subjects with respect to CAD. While SPLE was effective for discrimination between normal and diseased subjects, neither PLE or SPLE are sufficient to prove the presence of nonlinearity. For SPLE, Hurst exponent and MSE, both normal and diseased subjects show similar trends with respect to scaling behavior. However, small but measurable differences in these scaling behaviors allow us to improve differentiation between normal and diseased subjects using SPLE. SPLE was compared to SampEn, MSE and Hurst exponent, three popular methods for nonlinear analysis. These methods were not as accurate on our data set as SPLE. SPLE showed improved differentiation over PLE without scaling, indicating that there are differences between the time scaling behavior of diastolic signals from normal and diseased subjects. Hurst exponent is another measure of $1/f$ behavior and also gave high accuracy

results, further indicating the presence of $1/f$ structural differences between normal and diseased subjects. Discrimination of normal and diseased subjects was able to be performed on surrogate data approximately as well as measured data, but surrogates showed somewhat different values of our metrics when compared with measured data. This appears to indicate that there is some small degree of weak nonlinearity present in our data set, but that this nonlinearity is not related to CAD.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the NIH-HIHLB Grant HL079672-02.

References

1. Kochanek KD, Xu J, Murphy S, Minio AM, Kung H-C. Deaths: Preliminary data for 2009. National Vital Statistics Reports. 59(4)
2. Wilson PWF, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. *Circulation*. 1998; 97(18):1837–1847. [PubMed: 9603539]
3. Semmlow J, Rahalkar K. Acoustic detection of coronary artery disease. *Annual Review of Biomedical Engineering*. 2007; 9(1):449–469.
4. Wang JZ, Tie B, Welkowitz W, Semmlow JL, Kostis JB. Modeling sound generation in stenosed coronary arteries. *IEEE transactions on bio-medical engineering*. 1990; 37(11):1087–1094. [PubMed: 2276756]
5. Wang JZ, Tie B, Welkowitz W, Kostis J, Semmlow J. Incremental network analogue model of the coronary artery. *Medical and Biological Engineering and Computing*. 1989; 27(4):416–422. URL <http://dx.doi.org/10.1007/BF02441434>. [PubMed: 2601469]
6. Ku DN. Blood flow in arteries. *Annual Review of Fluid Mechanics*. 1997; 29(1):399–434.
7. Dock W, Zoneraich S. A diastolic murmur arising in a stenosed coronary artery. *The American Journal of Medicine*. 1967; 42(4):617–619. [PubMed: 6023006]
8. Sangster JF. Diastolic murmur of coronary artery stenosis. *Heart*. 1973; 35(8):840–844.
9. Duncan G, Gruber J, Dewey C, Myers G, Lees R. Evaluation of carotid stenosis by phonoangiography. *The New England journal of medicine*. 1975; 293(22):1124–1128. [PubMed: 127121]
10. Semmlow J, Welkowitz W, Kostis J, Mackenzie JW. Coronary artery disease—correlates between diastolic auditory characteristics and coronary artery stenoses. *IEEE transactions on bio-medical engineering*. 1983; 30(2):136–139. [PubMed: 6832793]
11. Akay M, Akay YM, Welkowitz W, Semmlow JL, Kostis JB. Application of adaptive filters to noninvasive acoustical detection of coronary occlusions before and after angioplasty. *IEEE transactions on biomedical engineering*. 1992; 39(2):176–184. [PubMed: 1612621]
12. Semmlow JL, Akay M, Welkowitz W. Noninvasive detection of coronary artery disease using parametric spectral analysis methods. *IEEE engineering in medicine and biology magazine*. 1990; 9(1):33–36. [PubMed: 18238314]
13. Akay M, Welkowitz W, Semmlow JL, Kostis J. Application of the arma method to acoustic detection of coronary artery disease. *Medical and biological engineering & computing*. 1991; 29(4):365–372. [PubMed: 1787751]
14. Akay M. Wavelet applications in medicine. *IEEE Spectrum*. 1997; 34(5):50.
15. Ruelle, D. World Scientific Series on Nonlinear Science, Series a. Vol. 16. World Scientific Pub Co Inc; 1996. Turbulence, Strange Attractors, and Chaos.

16. Padmanabhan V, Semmlow J. Dynamical analysis of diastolic heart sounds associated with coronary artery disease. *Annals of Biomedical Engineering*. 1994; 22(3):264–271. URL <http://dx.doi.org/10.1007/BF02368233>. [PubMed: 7978547]
17. Grassberger P, Procaccia I. Characterization of strange attractors. *Physical Review Letters*. 1983; 50(5):346–349.
18. Akay M, Akay YM, Gauthier D, Paden RG, Pavlicek W, Fortuin FD, Sweeney JP, Lee RW. Dynamics of diastolic sounds caused by partially occluded coronary arteries. *IEEE transactions on bio-medical engineering*. 2009; 56(2):513–517. [PubMed: 19272945]
19. Schmidt S, Hansen J, Hansen C, Toft E, Struijk J. Comparison of sample entropy and ar-models for heart sound-based detection of coronary artery disease. *Computers in Cardiology*. 2010; 37:385–388.
20. Schmidt S, Graebe M, Toft E, Struijk J. No evidence of nonlinear or chaotic behavior of cardiovascular murmurs. *Biomedical Signal Processing and Control*. 2011; 6(2):157–163. special Issue: The Advance of Signal Processing for Bioelectronics - The Advance of Signal Processing for Bioelectronics. 10.1016/j.bspc.2010.07.003
21. Griffel, B.; Zia, MK.; Fridman, V.; Saponieri, C.; Semmlow, JL. Detection of coronary artery disease using automutual information; *Cardiovascular Engineering and Technology*. 2012. p. 1-12. URL <http://dx.doi.org/10.1007/s13239-012-0094-6>
22. Osborne AR, Provenzale A. Finite correlation dimension for stochastic systems with power-law spectra. *Physica D: Nonlinear Phenomena*. 1989; 35(3):357–381.
23. Gisiger T. Scale invariance in biology: coincidence or footprint of a universal mechanism? *Biological reviews of the Cambridge Philosophical Society*. 2001; 76(2):161–209. [PubMed: 11396846]
24. Griffel, B.; Zia, MK.; Fridman, V.; Saponieri, C.; Semmlow, JL. A novel nonlinear predictability based entropy method for evaluation of noisy signals. *Proceedings of the 2011 IASTED International Symposia Imaging and Signal Processing in Healthcare*; May 2011;
25. May RM. Simple mathematical models with very complicated dynamics. *Nature*. 1976; 261:459–467. 10.1038/261459a0 [PubMed: 934280]
26. Valencia J, Porta A, Vallverdu M, Claria F, Baranowski R, Orlowska-Baranowska E, Caminal P. Refined multiscale entropy: Application to 24-h holter recordings of heart period variability in healthy and aortic stenosis subjects. *Biomedical Engineering, IEEE Transactions on*. 2009; 56(9): 2202–2213. 10.1109/TBME.2009.2021986
27. Bruce, EN. *Biomedical Signal Processing and Signal Modeling*. Wiley-Interscience; New York, NY: 2000.
28. Mandelbrot, BB.; Taqqu, MS. Robust R/S Analysis of Long Run Serial Correlation. *Proceedings of the 42nd Session of the International Statistical Institute*; 1979. p. 69-104.
29. Richman JS, Moorman JR. Physiological time-series analysis using approximate entropy and sample entropy. *American Journal of Physiology - Heart and Circulatory Physiology*. 2000; 278(6):H2039–H2049. [PubMed: 10843903]
30. Costa M, Goldberger AL, Peng CK. Multiscale entropy analysis of complex physiologic time series. *Physical Review Letters*. 2002; 89(6):068102–1–4. URL <http://link.aps.org/doi/10.1103/PhysRevLett.89.068102>. [PubMed: 12190613]
31. Zia MK, Griffel B, Fridman V, Saponieri C, Semmlow JL. Noise detection and elimination for improved acoustic detection of coronary artery disease. *Journal of Mechanics in Medicine and Biology*. 12(2) URL <http://dx.doi.org/10.1142/S0219519412400076>. 10.1142/S0219519412400076
32. Pan J, Tompkins WJ. A real-time qrs detection algorithm. *Biomedical Engineering, IEEE Transactions on*. 1985; BME-32(3):230–236.
33. Gunn, SR. Tech rep. Image Speech and Intelligent Systems Research Group, University of Southampton; 1997. Support vector machines for classification and regression.
34. Kantz, H.; Schreiber, T. *Nonlinear time series analysis*. Cambridge University Press; Cambridge, U.K.: 2004.

35. Dietterich TG. Approximate statistical tests for comparing supervised classification learning algorithms. *Neural Computation*. 1998; 10(7):1895–1923. URL <http://dx.doi.org/10.1162/089976698300017197>. 10.1162/089976698300017197 [PubMed: 9744903]

Figure 1.

Results of the first three steps of PLE calculation: The original signal (solid line), the time derivative (dashed line), and the SGN of the time derivative (stem plot) are shown. Consecutive positive or negative segments are counted and binned into a histogram, the first four segments are labeled $i=1 \dots 4$

Figure 2.

Histogram of counts of monotonic segments of increasing length of 4 artificial noise segments of 10,000 samples each. (A) Identically individually distributed noise (IID) showing path lengths from 1 to 6, (B) Gaussian distributed noise showing path lengths from one 1 to 5, (C) fractional Brownian motion showing path lengths from 1 to 20 and (D) fractional Gaussian noise showing path lengths from 1 to 7

Figure 3.

The histogram $C(i)$ and probability distribution function $p(C)$ for Gaussian noise (A,B) and fBm signal of $H=0.5$ (C,D). For Gaussian noise the longest path length is 1 (A), but most samples exist on segments of 4 samples long (B). Likewise for the fBm signal, the most common path lengths are 1 and 2 samples (C), but most samples in the signal exist along segments of 7 samples in length (D).

Figure 4.

A comparison of the group-and-average (solid line) and filter-and-resample (dashed line) methods for determining the Hurst exponent. The nominal Hurst exponent is shown as a dash-dot line. The two methods show agreement with each other. Deviations from the nominal line are likely due to the test data themselves and not deficiencies in either method.

Figure 5.

Chest placement for microphones showing the Base, Middle and Apex position. The stomach microphone is shown on the lower left. The microphones are held in place with straps. Subjects were seated and reclined at approximately 45 degrees.

Figure 6.

Four channels of unfiltered signals for a normal (A–D) and a diseased subjects (E–H), including both active chest microphones (AB and EF), the stomach microphone for ambient noise (C and G), and the ECG channel (D and H). The diastolic window is also shown. For normal subjects the diastolic window was 6210 ± 2130 samples and diseased subjects 5813 ± 1800 samples. This corresponds to 281 ± 104 msec for normal subjects and 263 ± 80 msec for diseased subjects

Figure 7. Four channels of signals for a normal (A–D) and a diseased subjects (E–H) after filtering with a 4th order Butterworth filter of passband 200–500 Hz, including both active chest microphones (AB and EF), the stomach microphone for ambient noise (C and G), and the ECG channel (D and H). The diastolic window is also shown.

Figure 8.

A) Averaged spectra shown as log power Vs log frequency for normal (black) and diseased (red) subjects. B) PLE C) Range $R(i)$ and D) Sample entropy versus scale for normal (black) and diseased (red) subjects. Blue triangles indicate scale values at which the differences are significant. Error bars show 95% confidence interval

Figure 9.

Classification using a linear SVM classifier applied to PLE and Hurst exponent data obtained from normal (black circles), marginal (yellow circles) and diseased (red Xs) subjects. A) PLE without scaling. B) PLE with scaling. D) Hurst exponent.

Figure 10.

Classification using A) MSE shows slightly better performance than B) SampEn at base level scale using a linear SVM classifier of normal (black circles) and diseased (red Xs) subjects. Marginal subjects shown in yellow

(A) ROC curve for Scaled PLE

(B) ROC curve for Hurst Exponent

Figure 11.

ROC curve for a) Scaled PLE and b) Hurst exponent shows comparable accuracy compared but SPLE is slightly higher. The red X marks the point on the curve with the highest overall accuracy

Figure 12.

A) The logistic equation has a bifurcation plot that shows period doubling until approximately $A = 3.56995$. B) SPLE analysis of the logistic map for increasing values of A show a clear change at the transition to chaos. C) SPLE analysis of surrogates for the logistic map do not show a particular trend until approximately 3.9 well into the chaotic region.

Table 1

False Negatives and Positives for Hurst exponent and SPLE

Measurement	False Negatives (subject)	False Positives (subject)
Hurst exponent	47, 64, 73	31, 50, 62, 70
SPLE	37, 47, 64	31, 32, 70

Table 2

Surrogate data analysis shows small differences between measured and surrogate data values, but not for discrimination of normal and diseased subjects. Numbers in parentheses indicate the number of values that shared the rank.

Measure	Range of Surrogate Values	Measured Data Value	Measured Data Rank
Hurst exponent Accuracy	77–83%	77%	1 (4)
SPLE Accuracy	74–81%	81%	97 (3)
Mean H Normal	0.976–0.982	0.976	1
Mean H Diseased	0.971–0.989	0.984	12 (2)
Mean SPLE Normal	81.7–82.6	81.7	1
Mean SPLE Diseased	85.9–86.6	86.6	101

Table 3

Results of a 5 x 2 fold cross-validation for the Hurst exponent and SPLE. The $\pm$ indicates the STD

Measurement	Accuracy (%)	Sensitivity (%)	Specificity (%)
SPLE	78 $\pm$ 8.4	83 $\pm$ 15	74 $\pm$ 7.1
Hurst	77 $\pm$ 10	73 $\pm$ 17	82 $\pm$ 23