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

Kalon K.L. Ho, MD, MSc; George B. Moody, Chung-Kang Peng, PhD; Joseph E. Mietus, BS; Martin G. Larson, ScD; Daniel Levy, MD; Amy L. Goldberger, MD

**Background**  Despite much recent interest in quantification of heart rate variability (HRV), the prognostic value of conventional measures of HRV and of newer indices based on nonlinear dynamics is not universally accepted.

**Methods and Results**  We have designed algorithms for analyzing ambulatory ECG recordings and measuring HRV without human intervention, using robust methods for obtaining time-domain measures (mean and SD of heart rate), frequency-domain measures (power in the bands of 0.003 to 0.04 Hz [VLF], 0.04 to 0.15 Hz [LF], and 0.15 to 0.5 Hz [HF]) and spectral power (TP) over all three of these bands), and measures based on nonlinear dynamics (approximate entropy [ApEn], a measure of complexity, and detrended fluctuation analysis [DFA], a measure of long-term correlations). The study population consisted of chronic congestive heart failure (CHF) case patients and sex- and age-matched control subjects in the Framingham Heart Study. After exclusion of technically inadequate studies and those with atrial fibrillation, we used these algorithms to study HRV in 2-hour ambulatory ECG recordings of 69 participants (mean age, 71.7±8.1 years). By use of separate Cox proportional-hazards models, the conventional measures SD (P<.01), LF (P<.01), VLF (P<.05), and TP (P<.01) and the nonlinear measure DFA (P<.05) were predictors of survival over a mean follow-up period of 1.9 years; other measures, including ApEn (P>.3), were not. In multivariable models, DFA was of borderline predictive significance (P=.06) after adjustment for the diagnosis of CHF and SD.

**Conclusions**  These results demonstrate that HRV analysis of ambulatory ECG recordings based on fully automated methods can have prognostic value in a population-based study and that nonlinear HRV indices may contribute prognostic value to complement traditional HRV measures. (Circulation. 1997;96:842-848.)

**Key Words**  • dynamics • Fourier analysis • heart failure • heart rate • survival

There has been much recent interest in analysis of HRV in a variety of clinical settings. The utility of conventional time- and frequency-domain HRV measures in assessing prognosis remains unclear, however, and the role of newer nonlinear indices of HRV is controversial. Often, subjective decisions (eg, related to treatment of noise and ectopy in the heart rate time series or to patient-specific selection of frequency bands) complicate the task of objective assessment of HRV measures. These subjective decisions not only obscure the utility of HRV measures but also add significantly to the cost of obtaining them. In the present investigation, therefore, we sought to address the following two questions in a feasibility study.

1. Can a fully automated analysis of HRV, requiring no subjective decisions, predict outcomes in a prospective, population-based study? For this purpose, we analyzed data obtained at the FHS from chronic CHF case patients and from sex- and age-matched control subjects, and we sought to develop HRV measures that are insensitive to the presence of typical errors in automated analysis of such data.

2. Do nonlinear indices of HRV have predictive value, in addition to time- and frequency-domain indices? To address this question, we studied ApEn, a measure of complexity, and DFA, an index of the presence or absence of long-term fractal correlations.

**Methods**

**Subjects: CHF Case Patients and Control Subjects**

Subjects for this analysis were drawn from participants in the FHS and the FOS. In 1948, 5209 residents of Framingham, Mass, were enrolled in the FHS, a prospective epidemiological study. These individuals have subsequently been evaluated at 2-year intervals with a review of their medical history, physical examination, ECG, blood tests, and various noninvasive pulmonary function and imaging studies. In 1971, children of the original study participants and spouses of these children (5124 in total) were entered in the FOS.

The primary group of interest was individuals with CHF. From information obtained at FHS and FOS examinations and from hospital and physician records, a diagnosis of CHF was established by the simultaneous presence of at least two major criteria or one major and two minor criteria for heart failure (Table I). Minor criteria for heart failure were acceptable
Selected Abbreviations and Acronyms
ApEn = approximate entropy
CHF = congestive heart failure
DFA = detrended fluctuation analysis
FHS = Framingham Heart Study
FOS = Framingham Offspring Study
HF = high frequency; 0.15 to 0.5 Hz
HRV = heart rate variability
LF = low frequency; 0.01 to 0.15 Hz
TP = total spectral power
VLF = very low frequency; 0.001 to 0.01 Hz

only if they could not be attributed to another medical condition. Information related to deaths was reviewed by three physicians who assigned a cause of death according to established FHS protocols.

Ambulatory ECG Recording
Ambulatory ECG monitors were placed on participants for 2 to 4 hours during their evaluations in the 18th and 19th examination cycles of the FHS (April 1983 to November 1985 and April 1985 to June 1988, respectively) and during the 3rd examination cycle of the FOS (December 1983 to September 1987). Modified ECG leads V1 and V6 were recorded simultaneously with standard ambulatory ECG recorders (Clinical Data, Inc.). During the recording period, all subjects underwent a standard battery of tests, including phlebotomy; measurements of height, weight, and blood pressure; pulmonary function tests; 12-lead ECGs; echocardiograms and other vascular ultrasound imaging; interviews and self-administered questionnaires; and elucidation of medical history and a physical examination performed by a physician. Participants were seated for most of their visit but walked from each test station to the next. Although most subjects received the complete battery of tests, the order of test administration was randomly assigned. Heart Study visits were scheduled in the morning for those in the FOS and in the early afternoon for the FHS cohort.

Case Patients and Control Subjects
All subjects with an antecedent diagnosis of CHF who attended one or more of the three specified examinations were identified. Ambulatory ECG recordings were available from 52 individuals. From the remainder of the FHS and FOS participants, we selected an equal number of control subjects matched for sex and age (±6 months). Those selected as control subjects had no documented history of coronary heart disease; CHF; hypertension; atrial arrhythmias; implanted cardiac pacemakers; or prior use of cardiac glycosides or antiarrhythmic, anti-ischemic, or antihypertensive medications.

We analyzed the first available recording for each of the 104 selected subjects (a few of those enrolled in the FHS had been recorded twice). The initial 2 hours of each of these 104 ECG recordings was processed in a fully automated manner to derive heart rate time series, time- and frequency-domain HRV measures, and nonlinear HRV measures. We excluded recordings that documented atrial fibrillation (from 12 subjects with CHF and 2 control subjects) and those that were technically inadequate (from 12 subjects with CHF and 9 control subjects) because of signal loss or insufficient duration.

Derivation of Heart Rate Time Series
Using a commercially available version of our Aristotle arrhythmia analysis software, 13 we obtained a beat annotation file (a list of the type and time of occurrence of each beat) for each ECG recording. Beats were automatically detected, and beat types were automatically determined. Using criteria based on timing and QRS morphology, Aristotle labels each detected beat as normal, ventricular ectopic, supraventricular ectopic, or unknown and identifies the location of the R-wave peak (with a resolution of 8 ms in the commercially available version used in this study). Although the software may be expected to make a small number of errors with respect to beat timing and type within a large and diverse data set, we did not attempt to identify or correct these errors; rather, we attempted to develop HRV measures that are robust with respect to the presence of such errors.

From the beat annotation file, we selected only those beats that Aristotle had labeled as normal and determined the intervals, NN(i), between those beats. When ectopic beats were present or when noise or loss of signal prevented detection of beats, the NN time series included intervals between nonconsecutive normal beats. These intervals were identified as outliers and rejected. We used the following strategy to identify additional outliers.

1. For each set of five contiguous NN intervals, NN(i−2), NN(i−1), NN(i), NN(i+1), NN(i+2), a local mean, NNmean(i), can be computed excluding the central interval: NNmean(i)=[NN(i−2)+NN(i−1)+NN(i)+NN(i+1)+NN(i+2)]/4.

2. The central interval, NN(i), is considered to be an outlier unless it lies within 20% of NNmean(i), i.e., unless 0.8×NNmean(i)<NN(i)<1.2×NNmean(i). This test was applied to each interval in the NN time series. All outliers identified in this way were rejected. From the remaining intervals, we constructed an instantaneous heart rate time series, HR(i): HR(i)=1/NN(i).

Segmentation of the Heart Rate Time Series
For each recording, we partitioned the heart rate time series into segments, each containing data corresponding to a 15-minute portion of the recording. The segments were constructed so that each segment after the first overlapped the final half of the previous segment. The number of heart rate values in each segment varied, because heart rate is not constant. We measured the number and the total duration of the interbeat intervals in each segment that were rejected as outliers. If >20% of the intervals or of the duration of any segment had been rejected, we excluded the segment from further analysis. For each remaining segment, we calculated three raw time-domain statistics [the mean, SD, and coefficient of variation of HR(i)], five raw frequency-domain statistics (described below), and one raw nonlinear statistic (ApEn, also described below). As noted below, we also used DFA to characterize each recording; unlike the other statistics described here, the DFA calculations were applied to the entire recording and not separately to each segment.
Fig. 1. a, DFA of interbeat interval time series from two representative subjects, one healthy and one with CHF. For both subjects, DFA curves are approximately linear over two regions, with a slope $\alpha_1$ for small values of $n$ (small time scales) and $\alpha_2$ for large values of $n$ (large time scales). Scaling exponent $\alpha_1$ measures strength of short-time correlations and $\alpha_2$ measures long-time correlations. Based on previous study, DFA index (a number between 0 and 1) is empirically derived by values of $\alpha_1$ and $\alpha_2$. Data points are for same subjects as shown in a. Note that DFA index for this CHF patient actually falls on flat region of graph corresponding to 0, the most abnormal possible value; healthy subject's index is close to 1.

Derivation of Time- and Frequency-Domain Measures

Both NN(i) and HR(i) may be regarded as samples of continuous functions obtained at irregular intervals (i.e., at the times of occurrence of the normal beats that terminate each NN interval). To use standard techniques for power spectral analysis, we resampled the heart rate time series at uniform 500-ms intervals, using linear interpolation between successive observations to fill in gaps as needed. Using standard fast Fourier transform methods, we derived power spectral density estimates of each 15-minute segment of these resampled time series (after subtracting the mean and removing any linear trend in each segment). From each such spectrum, we obtained the five raw frequency-domain statistics: the estimated power in each of three bands (VLF, LF, and HF), the TP over all three bands, and the scaling exponent, $\beta$ (the slope of a line fit by a least-squares criterion to a log-log plot of power versus frequency for frequencies between 0.001 and 0.1 Hz; this parameter characterizes the power-law scaling of the spectrum if it is of the form $P \sim f^{-\beta}$).

We characterized each recording in the time domain by the mean values of each of the three raw time-domain statistics. Because the mean values of the five raw frequency-domain statistics were not normally distributed across the study population, logarithmic transformations of these statistics were used to characterize each recording in the frequency domain.

Derivation of Nonlinear Measures

Detrended Fluctuation Analysis

DFA quantifies the presence or absence of long-range (fractal) correlations. This technique is a modification of root-mean-square analysis of random walks applied to nonstationary ("real-world") data. The details of this algorithm have been reported elsewhere. Unlike the other statistics used in this study, which are based on the segmented heart rate time series HR(i), the DFA index is based on the unsegmented NN interval series NN(i). Briefly, the DFA computation involves the following steps:

First, the NN time series (of total length $N$) is integrated, to yield

$$y(k) = \sum_{i=1}^{k} [NN(i) - NN_{ave}]$$

where $y(k)$ is the $k$th value of the integrated series, $NN(i)$ is the $i$th interbeat interval, and $NN_{ave}$ is the average interbeat interval over the entire series. Next, the integrated series is divided into windows of equal length, $n$. In each window of length $n$, a least-squares line is fitted to the data (representing the trend in that window) (Fig 1). The $y$ coordinates of the straight line segments are denoted by $y_s(k)$. Next we detrend the integrated time series, $y(k)$, by subtracting the local trend, $y_s(k)$, in each window. The root-mean-square fluctuation of this integrated and detrended series is calculated as

$$F(n) = \sqrt{\frac{1}{N} \sum_{k=1}^{N} [y(k) - y_s(k)]^2}$$

This computation is repeated over all time scales (window sizes) to obtain the relationship between $F(n)$ and the window size $n$ (i.e., the number of beats in a window that is the size of the window of observation). Typically, $F(n)$ will increase with window size. A linear relationship on a log-log plot indicates the presence of scaling. Under such conditions, the fluctuations can be characterized by a scaling exponent, $\alpha$, representing the slope of the line relating $log F(n)$ to $log n$.

A continuous DFA index was then used to summarize these analyses. In a previous study, the DFA scaling analysis was performed on one group of CHF patients as well as one normal control group. It was found that there is a region of scaling behavior over which normal (healthy) cardiac control operates. We used data sets from this prior study as "training sets" and estimated the probability that each individual heart-beat time series in the present study was operating in that normal region. This probability was called the DFA index, ranging between 0 and 1, with 1 indicating perfectly normal scaling behavior (Fig 1).

Approximate Entropy

ApEn is a "regularity statistic" that quantifies the unpredictability of fluctuations in a time series such as HR(i). Intuitively, one may reason that the presence of repetitive patterns of fluctuation in a time series renders it more predictable than a time series in which such patterns are absent. ApEn reflects the likelihood that "similar" patterns of observations will not be followed by additional "similar" observations. A time series containing many repetitive patterns has a relatively small ApEn; a less predictable (i.e., more complex) process has a higher ApEn.
The algorithm for computing ApEn has been published elsewhere\(^5\); here, we provide a brief summary of the calculations, as applied to the HR(t) time series. Given a sequence \( S_m \), consisting of \( N \) instantaneous heart rate measurements \( HR(1), HR(2), \ldots, HR(N) \), we must choose values for two input parameters, \( m \) and \( r \), to compute the ApEn of the sequence, \( \text{ApEn}(m,r,N) \). The first of these parameters, \( m \), specifies the pattern length, and the second, \( r \), defines the criterion of similarity. We denote a subsequence (or pattern) of \( m \) heart rate measurements, beginning at measurement \( i \) within \( S_m \), by the vector \( p_i(\cdot) \). Two patterns, \( p_i(\cdot) \) and \( p_j(\cdot) \), are similar if the difference between any pair of corresponding measurements in the patterns is less than \( r \), i.e., if \( |HR(i+k)−HR(j+k)|<r \), for \( 0≤k<m \). Now consider the set \( P_m \) of all patterns of length \( m \) [i.e., \( p_1(\cdot), p_2(\cdot), \ldots, p_m(\cdot) \)], within \( S_m \). We may now define \( \text{ApEn}(m,r,N) = \text{n}_{m,r}(N–m+1) \), where \( \text{n}_{m,r}(N) \) is the number of patterns in \( P_m \) that are similar to \( p_i(\cdot) \) (given the similarity criterion \( r \)). The quantity \( \text{ApEn}(m,r,N) \) is the fraction of patterns of length \( m \) that resemble the pattern of the same length that begins at interval \( i \). We can calculate \( \text{ApEn}(m,r,N) \) for each pattern \( p_i(\cdot) \) and define \( \text{ApEn}(r) \) as the mean of these \( \text{ApEn}(m,r,N) \) values. The quantity \( \text{ApEn}(r) \) expresses the prevalence of repetitive patterns of length \( m \) in \( S_m \). Finally, we define the ApEn of \( S_m \), for patterns of length \( m \) and similarity criterion \( r \), as \( \text{ApEn}(m,r,N) = \text{ApEn}(r) \), i.e., as the natural logarithm of the relative prevalence of repetitive patterns of length \( m \) compared with those of length \( m=1 \).

Thus, if we find similar patterns in a heart rate time series, ApEn estimates the logarithmic likelihood that the next intervals after each of the patterns will differ (i.e., that the similarity of the patterns is mere coincidence and lacks predictive value). Smaller values of ApEn imply a greater likelihood that similar patterns of measurements will be followed by additional similar measurements. If the time series is highly irregular, the occurrence of similar patterns will not be predictive for the following measurements, and ApEn will be relatively large.

In this study, we measured ApEn for each of the 15-minute segments of the heart rate time series. In segments with missing data (because of ectopy or noise), the measurements before and after each interval of missing data were treated as if they had been adjacent. As described above, segments were excluded if >20% of their intervals or duration were rejected; in addition, for calculation of ApEn only, segments that failed a stationarity test (<75%) were also excluded. (To assess the stationarity of heart rate in each segment, we computed the mean heart rate for each nonoverlapping subsequence of 100 measurements and calculated \( \sigma_{100} \), the SD of these 100-beat mean heart rates. The segment passed the stationarity test if \( \sigma_{100} < 90% \) of the SD of the heart rate measurements for the entire segment.) In previous work\(^{3,7,11} \), we explored various values for \( m \), the pattern length, and \( r \), the similarity criterion, and (on the basis of separability of similar populations of subjects not included in the present study) we chose \( m=2 \) and \( r=4.75 \) bpm (corresponding to 40% of the mean SD for all nonexcluded segments of all recordings in the study). Because the number of beats was not the same in each segment, \( n \) varied. The ApEn values obtained for each segment were averaged to obtain a mean value of ApEn to characterize each recording.

### Statistical Methods

Using a multivariate regression model adjusting for sex and age (in 5-year age groups), we compared the distributions of the HRV statistics (two mean time-domain statistics, the logarithms of the three mean frequency-domain statistics, the mean ApEn statistic, and the DFA index) between the subjects with CHF and the control subjects.\(^{12} \) Because five of the statistics (SD of heart rate, coefficient of variation of heart rate, TP, VLF power, and LF power) were highly correlated with one another (\( r=.72 \) to .95), the coefficient of variation of heart rate, TP, and LF power were excluded from the regression model, which examined all of the remaining statistics together as a function of sex, age, and CHF. Using Cox proportional-hazards regression models, we examined the relations of each of the HRV statistics with overall survival after the recording of the ambulatory ECG.\(^{18} \) Survival estimates were computed by Kaplan-Meier methods. Continuous measures are summarized as mean±SD. A value of \( P<.05 \) was required for statistical significance. All multivariable modeling was performed with the SAS System (SAS Institute).

### Results

The baseline characteristics of the study sample are shown in Table 2. The mean age of the study subjects was 71.7±8.1 years (range, 52 to 87 years); 43% of those with heart failure and 44% of the control subjects were women. The average time since diagnosis of heart failure was 6.9±6.2 years, with a median of 4.7 years and a range of 3 months to 26 years.

Table 3 presents the average values, by group, of the HRV measures we derived. After adjustment for age

### Table 2. Study Sample

<table>
<thead>
<tr>
<th></th>
<th>CHF Case Patients (n=28)</th>
<th>Control Subjects (n=41)</th>
<th>( P^* )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men. n</td>
<td>16</td>
<td>23</td>
<td></td>
</tr>
<tr>
<td>Woman. n</td>
<td>12</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>Age, y (mean±SD)</td>
<td>70.8±8.1</td>
<td>72.2±8.1</td>
<td></td>
</tr>
</tbody>
</table>

Values are unadjusted mean±SD.

* \( P \) values comparing the distribution between CHF case and control subjects are adjusted for age and sex. Coefficients of variation of heart rate, ln(TP), and ln(LF power) were excluded from the multivariate regression model because of high correlation with the other variables.

### Table 3. HRV Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>CHF Patients</th>
<th>Control Subjects</th>
<th>( P^* )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean heart rate</td>
<td>74.7±13.4</td>
<td>75.7±13.7</td>
<td>.756</td>
</tr>
<tr>
<td>SD of heart rate</td>
<td>4.15±1.17</td>
<td>5.45±1.74</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Coefficient of variation of heart rate</td>
<td>0.06±0.02</td>
<td>0.07±0.02</td>
<td></td>
</tr>
<tr>
<td>ln(TP)</td>
<td>2.54±0.64</td>
<td>3.13±0.64</td>
<td></td>
</tr>
<tr>
<td>ln(VLF power)</td>
<td>1.94±0.72</td>
<td>2.55±0.66</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>ln(LF power)</td>
<td>1.46±0.72</td>
<td>2.05±0.81</td>
<td></td>
</tr>
<tr>
<td>ln(HF power)</td>
<td>-0.12±0.71</td>
<td>0.08±0.75</td>
<td>.302</td>
</tr>
<tr>
<td>ln(ratio LF/HF power)</td>
<td>1.63±0.72</td>
<td>1.90±0.77</td>
<td>.036</td>
</tr>
<tr>
<td>Power-law scaling exponent, ( \beta )</td>
<td>-0.65±0.13</td>
<td>-0.61±0.14</td>
<td>.198</td>
</tr>
<tr>
<td>ApEn</td>
<td>1.18±0.16</td>
<td>1.24±0.21</td>
<td>.201</td>
</tr>
<tr>
<td>DFA index</td>
<td>0.30±0.40</td>
<td>0.55±0.40</td>
<td>.007</td>
</tr>
</tbody>
</table>
TABLE 4. Predictors of Survival

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Hazard Ratio†</th>
<th>95% CI</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Univariate predictors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean heart rate</td>
<td>1.22</td>
<td>0.71, 2.12</td>
<td>0.462</td>
</tr>
<tr>
<td>SD of heart rate</td>
<td>0.47</td>
<td>0.22, 1.03</td>
<td>0.58</td>
</tr>
<tr>
<td>Coefficient of variation of heart rate</td>
<td>0.37</td>
<td>0.16, 0.85</td>
<td>0.20</td>
</tr>
<tr>
<td>In(TP)</td>
<td>0.66</td>
<td>0.38, 1.15</td>
<td>0.143</td>
</tr>
<tr>
<td>In(VLF power)</td>
<td>0.74</td>
<td>0.44, 1.25</td>
<td>0.264</td>
</tr>
<tr>
<td>In(LF power)</td>
<td>0.61</td>
<td>0.35, 1.08</td>
<td>0.090</td>
</tr>
<tr>
<td>In(HF power)</td>
<td>0.87</td>
<td>0.46, 1.63</td>
<td>0.667</td>
</tr>
<tr>
<td>In(ratio LF/HF power)</td>
<td>0.74</td>
<td>0.46, 1.20</td>
<td>0.219</td>
</tr>
<tr>
<td>Power-law scaling exponent, β</td>
<td>0.80</td>
<td>0.48, 1.36</td>
<td>0.415</td>
</tr>
<tr>
<td>ApEn</td>
<td>0.91</td>
<td>0.51, 1.64</td>
<td>0.761</td>
</tr>
<tr>
<td>DFA index</td>
<td>0.39</td>
<td>0.15, 0.93</td>
<td>0.048</td>
</tr>
<tr>
<td>Multivariable predictors of survival*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coefficient of variation of heart rate</td>
<td>0.47</td>
<td>0.21, 1.02</td>
<td>0.56</td>
</tr>
<tr>
<td>DFA index</td>
<td>0.43</td>
<td>0.16, 1.17</td>
<td>0.097</td>
</tr>
<tr>
<td>Model 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD of heart rate</td>
<td>0.50</td>
<td>0.23, 1.10</td>
<td>0.85</td>
</tr>
<tr>
<td>DFA index</td>
<td>0.40</td>
<td>0.15, 1.05</td>
<td>0.062</td>
</tr>
</tbody>
</table>

*All univariate and multivariable models are stratified by the diagnosis of CHF.
†Hazard ratios are per SD of the respective variable.

and sex, there were statistically significant differences between the CHF case patients and control subjects. The SD of the heart rate, VLF power, LF power, and the ratio of LF to HF power were lower in the CHF case patients than in the control subjects. The DFA index, derived from the scaling exponent α, was also lower in the CHF case patients, indicating a lower amount of long-range correlations compared with the control subjects.

Over a mean follow-up period of 1.9 years, 12 deaths occurred (9 among the heart failure case patients and 3 among the control subjects). Causes of death among the CHF case patients included sudden death in 1, coronary heart disease in 1, cerebrovascular accident in 1, other cerebrovascular disease in another, cancer in 1, and other causes in 4; there were 2 cancer deaths among the control subjects and 1 death from other causes. Using several separate Cox proportional-hazards regression models, the diagnosis of CHF, the SD of the heart rate, the coefficient of variation of heart rate, the heart rate TP, VLF power, LF power, and the DFA index were predictors of survival. After stratification by the diagnosis of CHF, only coefficient of variation and the DFA index were significant predictors, whereas SD and LF power were of borderline significance (Table 4). When coefficient of variation or SD was combined with the DFA index in simple two-variable models, these variables were marginally significant predictors of survival after adjustment for the diagnosis of CHF (Table 4). This is illustrated in Fig 2, which depicts Kaplan-Meier estimates of survival. We determined the median values of SD and of the DFA index over the entire study population and used these values to establish "high" (median or greater) or "low" (submedian) ranges for these variables. As apparent from the graph, subjects with a low SD and a low DFA index (ie, mostly CHF case patients) did very poorly. Individuals who were at or above the median for both variables (ie, control subjects) did very well. The others had an intermediate prognosis.

Discussion

The major findings of the study are that (1) HRV analysis of ambulatory ECG recordings, based on fully automated methods, is capable of predicting outcomes in a small sample of CHF case patients and control subjects; (2) this investigation confirms the importance of time-domain measures (the coefficient of variation and SD of the heart rate) as diagnostic variables; (3) patients with CHF show a breakdown of the physiological long-range (or fractal) correlations of heart rate, as quantified by DFA; and (4) the DFA index is a significant diagnostic indicator that may complement traditional HRV measures in assessing risk.

Assessment of HRV

There has recently been an explosion of interest in quantification of HRV.19-24 Measurements of HRV can be classified into two general classes: "moment" statistics and dynamic statistics.3-7,24

The first class of measures, time-domain or moment statistics (eg, mean, SD, and coefficient of variation of heart rate), is based on linear analysis. These measures appear frequently in the literature. They are relatively straightforward to obtain (although careful attention to erroneous or missing data is necessary, because they tend to be highly sensitive to isolated outliers). Such measures do not depend on the order of the observations, however, and may therefore obscure significant information about heart rate dynamics. For example, two heart rate time series may have nearly identical mean rates and SDs but very different dynamics.4-10 Of note, in the present study, the mean heart rate for case patients and control subjects was not significantly different. Previous work has documented the very poor prognosis of CHF in participants in the FHS, with 3-year mortality rates after diagnosis of 43% in men and 36% in women.25 Only those subjects who survived to the next examination cycle (≥2 years later) were available for ambulatory ECG recordings. Thus, the CHF case patients included in these analyses (who had survived an average of 6.9 years since onset of CHF) were probably healthier than hospitalized CHF patients used in other
HRV studies, as evidenced by the absence of the anticipated relative sinus tachycardia in CHF case patients (mean heart rate, 75 bpm).

Dynamic statistics, a class that includes both frequency-domain measures (eg, power spectral density estimates derived from Fourier "all-zeroes" or autoregressive "all-poles" analyses) and new measures derived from nonlinear dynamics (chaos theory and fractal mathematics), do preserve information about the order of observations. The Fourier transform provides a useful representation of the component frequencies of heart rate time series, including HF heart rate oscillations associated with respiration and lower-frequency oscillations associated in part with baroreflex control.2,19,24 Nonstationarity in typical heart rate time series severely limits the range of frequencies that can be studied by conventional frequency-domain analytic methods. Furthermore, frequency-domain analysis, while retaining information relating to ordering of observations, is still based on linear models and may conceal the details of interactions between mechanisms. For example, a respiration-mediated change in heart rate may stimulate a response from another mechanism. In the time series, this phenomenon may be readily observable as a repetitive pattern of fluctuation, but in the frequency domain, it may be indistinguishable from uncorrelated fluctuations at different frequencies. Neither moment statistics nor frequency-domain measures reveal the presence or absence of such features in the time series.)

Nonlinear and Fractal Dynamics of the Heartbeat

Nonlinear and fractal dynamic analysis offers the prospect of revealing these details by providing direct measures of complexity8,7 and long-range correlations.8,20 We studied two indices of HRV derived from nonlinear and fractal dynamics: ApEn (a measure of complexity) and DFA (a measure of the presence of long-term fractal correlations).

Reduced sinus rhythm heart rate complexity has been reported with a variety of pathologic processes,7 with bed rest and deconditioning,16 and with aging.8 The present study did not indicate that ApEn was an independent predictor of survival, however. This result may be a consequence of the nonstationarity of the data sets and of intrinsic limitations of the measure in the context of very low overall variability, as in some case patients with CHF.

Healthy heart rate fluctuations show a complex type of variability that we and others have shown to be fractal, that is to say, having "self-similar" fluctuations on time scales ranging from seconds to hours.27 Furthermore, this fractal complexity generates long-range power-law correlations. The presence of such long-range fractal order indicates that fluctuations in heart rate are affected not only by the most recent value but also by much more remote events, in other words, a "memory" effect.27 In a previous study, we showed that a breakdown of these fractal scaling properties is noted with CHF and other disease conditions.8 The present study confirms the utility of DFA measurement and suggests that it may add prognostic information to that obtained from conventional SD and spectral measures of HRV.

Automated Analysis of HRV

As typically applied, HRV analysis requires interactive data "massaging" (with the risk of introducing bias) to clean up the heart rate time series sufficiently to obtain standard measures. Often the interpretation of results requires additional subjective decisions (eg, to determine which portion of the heart rate spectrum represents the contribution of respiration to HRV). These manipulations represent virtually the entire incremental cost of HRV analysis (beyond that of conventional long-term ECG analysis) while provoking skepticism with respect to the objectivity of the procedure.

Our method requires neither nonstandard instrumentation or recording protocols nor tedious, expensive, and potentially bias-inducing interactive data preparation. Its success results from careful design of robust methods for obtaining measures of HRV, including nonlinear and fractal dynamic measures that complement standard moment and frequency-domain statistics. The computational demands of our analysis, though significantly greater than those of standard analyses restricted to derivation of moment statistics, remain quite modest relative to those of conventional long-term ECG analysis. These encouraging results from a small sample with limited follow-up should be confirmed in a larger study.

Reliable measures of HRV should not be inordinately affected by isolated errors in the RR interval sequence resulting from noise, ectopy, or missing data. A comprehensive analysis of HRV must also look considerably beyond moment statistics to take account of the order of observations, of the presence or absence of repetitive patterns of heart rate fluctuation as markers of the dynamic processes that control heart rate, and of their interactions. Significantly, we have shown that a totally automated method for reliable and comprehensive HRV analysis can predict outcomes in a group of prospectively identified CHF case patients and control subjects.

Acknowledgments

This study was supported by grants from the G. Harold and Leila Y. Mathers Charitable Foundation and the National Aeronautics and Space Administration and by National Institutes of Health/National Heart, Lung, and Blood Institute contract N01-HC-38038. Dr Peng is a recipient of a National Institute of Mental Health FIRST Award (MH-54081). Dr Ho was supported in part by a Fellowship from the Pfizer Clinical Investigator Training Program.

References


