Annals of Biomedical Engineering, Vol. 29, pp. 764–774, 2001
Printed in the USA. All rights reserved.
0090-6964/2001/29(9)/764/11/$15.00

Comparison and Clinical Application of Frequency Domain Methods in Analysis of Neonatal Heart Rate Time Series

KEVIN L. CHANG,[1,4] KENNETH J. MONAHAN,[1,3] M. PAMELA GRIFFIN,[2] DOUGLAS LAKE,[4] and J. RANDALL MOORMAN[1,5]

[1]Department of Internal Medicine (Cardiovascular Division), [2]Department of Pediatrics (Neonatology Division), [3]Department of Biomedical Engineering, [4]Department of Mathematics, and the [5]Cardiovascular Research Center, University of Virginia Health Sciences Center, University of Virginia, Charlottesville, VA

(Received 2 November 2000; accepted 13 June 2001)

Abstract—The frequency content of the heart rate (HR) series contains information regarding the state of the autonomic nervous system. Of particular importance is respiratory sinus arrhythmia (RSA), the high-frequency fluctuation in HR attributable to respiration. The unevenly sampled nature of heart rate data, however, presents a problem for the discrete Fourier transform. Interpolation of the HR series allows even sampling, but filters high-frequency content. The Lomb periodogram (LP) is a regression-based method that addresses these issues. To evaluate the efficacy of the LP and Fourier techniques in detecting RSA, we compared the spectrum of intervals, the spectrum of HR samples, and the LP of simulated and clinical neonatal time series. We found the LP was superior to the spectrum of intervals and the spectrum of HR samples in analysis near the critical frequency of one half the average sampling rate. Applying the LP to clinical data, we found (1) evidence of stochastic resonance, an enhancement of periodicity with the addition of small amounts of noise, and (2) reduced power at all frequencies prior to clinical diagnosis of neonatal sepsis.

[DOI: 10.1114/1.1397791]

Keywords—Lomb periodogram, Spectrum of intervals, Spectrum of heart rate samples, Low-pass filter, Stochastic resonance, Neonatal sepsis.

INTRODUCTION

The activity of the autonomic nervous system (ANS) contains important information regarding organ and system function. Parasympathetic tone can be estimated noninvasively by evaluation of respiratory sinus arrhythmia (RSA), the variation of heart rate (HR) with respiration.[16,21,29] In newborn infants, loss of RSA is a characteristic of asphyxia[9] and the degree of RSA predicts long-term outcomes.[11] The variations in intrathoracic pressure, venous return, left-ventricular end-diastolic volume, and arterial blood pressure brought on by each respiratory cycle produce characteristic changes in the HR. The frequency at which these changes occur corresponds to the respiratory rate. Therefore, measurement of RSA is tantamount to sampling the respiratory rate with the HR, measured using a series of the times between consecutive heartbeats, or RR intervals. The presence of RSA is typically identified by frequency domain analysis, where it appears as a peak in the power spectrum of the corresponding HR time series.

In addition to detection of RSA in studies of autonomic function, the early diagnosis of sepsis is potentially another important application for frequency domain analysis of neonatal HR time series. We have found that heart rate characteristics (HRC) change early in the course of this common and potentially catastrophic illness[14] for which premature infants are at particularly high risk. The abnormalities of the RR interval time series were reduced variability and transient decelerations, similar to the findings of fetal distress. Our initial analysis utilized the time domain, and we found that the third moment, or skewness, as well as quantiles or percentiles of normalized data sets could be used to distinguish normal from abnormal time series. Interestingly, the standard deviation, which reports only on the degree of variability, did not discriminate between normal and abnormal data. We might therefore expect that the total power of the signal is a poor marker of abnormality, as the total power is identical to the variance. This observation does not, however, rule out the possibility that frequency band-specific information might discriminate between normal and abnormal records.

HR data possess several unique properties not usually encountered by traditional frequency domain techniques. Chief among these properties is the fact that the RR intervals are not evenly sampled in time. The variation in the time between heartbeats is exactly the phenomenon that heart rate variability (HRV) analysis attempts to characterize. For even sampling to occur, the time be-

Address correspondence to J. Randall Moorman, Box 6012, MR4 Building, UVAHSC, Charlottesville, VA 22908. Electronic mail: moorman@virginia.edu. Both Kevin L. Chang and Kenneth J. Monahan contributed equally to this manuscript and should be considered the first author.

tween beats would have to be constant (as in the case of an artificial pacemaker) and there would be no variability to analyze. While unevenly sampled data may harbor periodic components and thus can be analyzed by Fourier theory, the familiar discrete Fourier transform (DFT) is not an accurate probe of the periodicity when the sampling is not even. To further complicate matters, the inaccuracy increases as the variability increases.[5] Berger and co-workers have described an algorithm that uses an interpolation scheme to create an evenly sampled series of heart rates from RR interval data.[5] They showed that this method captures accurately the frequency content of simulated RR data generated by an integrated pulse frequency modulation (IPFM) model. Although this approach has been widely used[1,17,23] and recommended by an international task force,[36] several possible disadvantages exist. First, all data are interpolated, and no directly measured values are used in the analysis. Second, Moody suggested that the interpolation process acts as a low-pass filter and attenuates the higher-frequency content of the original RR series.[27] As a result of this effect, subsequent analysis of the interpolated series may not entirely reflect the characteristics of the actual data.

The issues raised by interpolation of series of RR intervals apply to all HR data. The study of neonatal HR series in particular presents a more fundamental challenge to traditional spectral methods. Specifically, neonatal RSA may occur in bandwidths that are inaccessible via Fourier analysis. The relationship between the typical neonatal HR (120–160 beats/min) and respiratory rate (40–80 breaths/min) is such that the Nyquist criterion is often unsatisfied, resulting in aliasing of high-frequency information to lower frequencies and subsequent distortion of the frequency spectrum.[32,37] In addition, the magnitude of RSA in newborns and adults decreases with increasing respiratory rate.[8,19] Therefore, neonatal RSA is expected to be more difficult to detect than adult RSA, regardless of the method used.

The Lomb periodogram (LP), a method of spectral analysis that utilizes least-squares fitting of data to sinusoids of varying frequencies,[20,26,30,33] has several properties that make it an attractive alternative to the DFT for neonatal HR analysis. The LP can effectively analyze unevenly sampled data, as sinusoids can be fit as closely to unevenly sampled data as to evenly sampled points. The method requires no interpolation—the RR intervals are analyzed directly. In addition, the LP method allows the possibility of reporting on content above the average Nyquist frequency.[31] This possibility only exists for special circumstances in which the sampling is increased at the same time that the signal frequency is increased, an occurrence favored by chance and long records. In the case of RSA analysis, these conditions are not guaranteed as an increase in respiratory rate (signal of interest) does not necessarily correspond to an increase in HR (sampling rate).

Given the challenges posed by neonatal HR series to traditional Fourier methods and the potential for the LP to meet these challenges, we applied two Fourier methods and the LP to simulated and clinical neonatal HR data for the purpose of comparing how well each method detected RSA. We found the LP to be a superior method, and used this method to test the hypotheses that neonatal HR (1) displays stochastic resonance and (2) displays changes in the frequency domain prior to the clinical diagnosis of sepsis.

THEORY

Methods of Spectral Analysis

We call the periodogram of the DFT of a RR interval time series the *spectrum of intervals*. The sampling frequency is 1 sample/heartbeat and is multiplied by the mean HR (heartbeats/s) to yield units of equivalent Hz. This approach was improved by Berger and co-workers, who suggested an interpolation procedure yielding evenly spaced HR samples.[5] We call the periodogram of the DFT of the interpolated HR signal the *spectrum of HR samples* and note that the frequency unit is cycles/s (Hz).

A completely different approach to the problem of estimating frequency content of an unevenly sampled time series was developed by Lomb[26] and expanded upon by Scargle and others.[33] The calculation of the Lomb periodogram is equivalent to the least-squares fitting of the data to the sinusoid $A\cos[2\pi f(t-\tau)] + B\sin[2\pi f(t-\tau)]$, for which the power is proportional to A^2+B^2.[33] This procedure minimizes the mean-square error between the sinusoid and each data point.

Bandwidth Limitations of the Methods

For any evenly sampled signal, the range for which the frequency content is available is limited by the sampling rate. In theory, for a periodogram to report meaningful information at a given frequency, the corresponding signal must be sampled at least at twice that frequency. This critical sampling rate is known as the Nyquist frequency. In cases where the sampled signal contains frequency components greater than the Nyquist frequency, the power at those frequencies is spuriously added to the periodogram ordinates at frequencies below the Nyquist frequency, a process known as aliasing.

In healthy adults, the HR varies typically from 60 to 100 beats/min (1 to 1.67 Hz) and the respiratory rate ranges from 6 to 20 breaths/min (0.1 to 0.33 Hz). Therefore, in attempts to sample the respiratory rate with the heart rate, the Nyquist criterion will be satisfied in the majority of cases. In contrast, neonates generally have a

HR between 120 and 160 beats/min (2 and 2.66 Hz) and breathe at 40 and 80 breaths/min (0.67 and 1.33 Hz). In this case, the HR approaches the theoretical limit for accurate detection of the respiratory signal and may often times fail to satisfy the Nyquist criterion. In practice, the sampling frequency must be several-fold higher than the Nyquist frequency to prevent distortion of the resultant spectrum. For each of the three methods under investigation, there are limitations to the frequency content that can be reported on without aliasing.

For the spectrum of intervals, the frequencies are reported in units of equivalent Hz.[36] This measure is obtained from the product of the sampling rate (constant at one sample/beat) and the mean HR (beats/s) of the corresponding series of RR intervals. In this scheme, the mean HR becomes the average sampling frequency. The spectrum of intervals reports accordingly on frequencies up to $\frac{1}{2}$ of the mean heart rate, the average Nyquist frequency.

The spectrum of HR samples features a user-determined Nyquist frequency that is derived from the resampling rate used to create the interpolated HR series. The resampling rate is usually taken to be 4 Hz, which corresponds to a Nyquist frequency of 2 Hz.[5] The resampling frequency can be set arbitrarily, but disadvantages exist at the extremes of selection. Setting the frequency too low risks missing a substantial portion of the signal's meaningful content. A high resampling frequency increases the degree of interpolation and thus increases the amount of nonphysiological information present in the resultant signal.

Laguna and co-workers demonstrated that the maximum frequency that the LP can report on without aliasing is one half the mean Nyquist frequency, which they define as the mean HR.[25] This limitation is analogous to the Nyquist criterion of evenly sampled series. Of note, that study did not analyze separately epochs when the HR was greater than the average in order to assess the possibility of characterizing frequency content greater than the average Nyquist frequency.[20,30,31]

Significance of Peaks in the Periodograms

To assess the statistical significance of the peaks produced by the three methods, we modified a standard test used for this purpose, the Fisher test.[7] This test begins with the assumption that the Fuller test statistics ξ_m of the M periodogram ordinates follow an exponential distribution. The ξ_m are the individual periodogram ordinates normalized to the average of all ordinates in a given periodogram. Moreover, the largest ξ_m taken from a population of periodograms of white noise follow a separate distribution described by Fuller.[7] To calculate the significance of the peaks, we determine the probability that the largest observed Fuller test statistic from a given periodogram belongs to that distribution. Our modification was necessary because we have averaged periodogram ordinates, and we employed the fact that sums of exponential distributions are described by an incomplete gamma distribution.

More formally, we assume the null hypothesis that our M Fuller statistics $\{\xi_m : 1 \leqslant m \leqslant M\}$ (excluding the ordinate corresponding to the frequency zero) are derived from the DFT of Gaussian white noise. Since the periodogram ordinates are exponentially distributed, the probability that ξ_m is less than a threshold ξ_{thres} beyond which the probability of ξ_m occurring is less than a specified level of significance is

$$\text{prob}(\xi_m < \xi_{\text{thres}}) = 1 - e^{-\xi_{\text{thres}}}.$$

We define $\xi_{\max}$ to be the largest Fuller statistic of the periodogram. It follows that the probability that $\xi_{\max}$ is less than ξ_{thres} (i.e., that all ξ_m are less than ξ_{thres}) is

$$\text{prob}(\xi_{\max} < \xi_{\text{thres}}) = (1 - e^{-\xi_{\text{thres}}})^M.$$

We interpret this as the probability that no periodic component exists at the frequency of that which corresponds to $\xi_{\max}$. Chiu provided an improved test[7] defining the same probability as

$$\text{prob}(\xi_{\max} < \xi_{\text{thres}}) \approx 1 - \exp(-M e^{-\xi'}),$$

where

$$\xi' = \xi_{\text{thres}} \frac{M - 1 - \log M}{M - \xi_{\text{thres}}}.$$

We need to modify this approach because we average the periodogram ordinates in order to reduce the error of the spectral density estimate.[4] If we have M periodogram ordinates $\{P_n : 0 \leqslant n \leqslant M-1\}$, and average groups of α of these into one periodogram ordinate, we are left with M/α periodogram ordinates $\{P'_n : 0 \leqslant n \leqslant (M-1)/\alpha\}$. We calculate the Fuller statistic $\xi_{\max}$ of the largest averaged periodogram ordinate. Recall that the fundamental assumption underlying the Fisher test is that the periodogram ordinates of Gaussian white noise are exponentially distributed. For a statistical test of the averaged periodogram ordinates, we utilize the fact that sums of α exponentially distributed numbers follow a gamma distribution. Thus,

$$\text{prob}(\xi_{\max} \leqslant \xi_{\text{thres}}) = 1 - P(\alpha, \xi_{\text{thres}})^{M/\alpha},$$

where

$$P(\alpha,x)=\int_0^x \frac{1}{\Gamma(\alpha)} u^{\alpha-1}e^{-u}du$$

and

$$\Gamma(\alpha)=(\alpha-1)! \quad \text{when} \quad \alpha\in \mathbf{Z}.$$

As defined above, these statistical tests apply only to the largest periodogram ordinate. We used the same test on other ordinates as conservative approximations of their statistical significance. Indeed, in the case of the null hypothesis that the time series is Gaussian white noise, the probability that the second (or third, fourth, etc.) highest peak exceeds ξ_{thres} must be less than the probability that the highest peak exceeds ξ_{thres}.

Scargle showed analytically that power in the LP of Gaussian white noise also follows an approximate exponential distribution,[33] justifying our application of the Fisher test to LP ordinates in addition to ordinates of the other spectra to determine an approximate significance level.

METHODS AND DATA SETS

Frequency Domain Methods

For a time series with N points, the standard deviation of each periodogram ordinate may be reduced by a factor of $1/\sqrt{N}$ by averaging groups of N ordinates into one ordinate. Alternatively, the standard deviation may be reduced by $1/\sqrt{N}$ by dividing a data set into N segments of equal length, calculating periodograms for each segment, then averaging all N periodograms.[4,31]

To generate the spectrum of intervals, series of 2^m consecutive RR intervals were divided into seven segments of length 2^{m-2}, with each segment overlapping the previous one by 50%. These segments were analyzed with a fast Fourier transform (FFT) algorithm[31] and the resulting spectra were averaged to produce a representative periodogram. Following the treatment above, this maneuver reduces the error of the spectral estimate by a factor of $1/\sqrt{7}$.[4] As part of the FFT analysis, the series mean was removed and a 10% split cosine bell window was applied to the data.

The strategy used for the spectrum of HR samples is similar to that used for the spectrum of intervals. A series of consecutive RR intervals of length 2^m were resampled at 4 Hz using the algorithm described by Berger and co-workers.[5] The resulting series of HR samples was analyzed using a similar procedure as for the spectrum of intervals. The resampling rate of 4 Hz was greater than the mean HR of any of the simulated or clinical data sets, resulting in a resampled series with a greater number of data points than the original series for any given sampling period. The nearest multiple of 2^m resampled points was used to generate the spectrum of HR samples using a FFT technique.

To generate the LP, a series of consecutive RR intervals were analyzed using the algorithm set forth by Lomb[26] and modified by Press and co-workers.[30,31] As in the other spectral methods, an averaging procedure was used to reduce the standard deviation of the periodogram ordinates.

Simulated Data

We followed the established practice of using the IPFM model to simulate HR data with known frequency content. The IPFM model is derived from the formula[5]

$$T=\int_{t_k}^{t_{k+1}}[1+m(t)]dt.$$

Given the time t_k of the kth point of the series, the $k+1$st point is defined at time t_{k+1}, when the integral reaches the value T, and the ordinate of that point is taken as the interval $t_{k+1}-t_k$. We evaluated the integral at time steps of 0.01 s, and set T to be 0.4 s to simulate neonatal RR intervals. Since our chosen $m(t)$ had an expected value of zero, the mean sampling rate of the IPFM model as such was T. The Nyquist frequency was thus $1/(2T)$, or 1.25 Hz.

Our purpose was to evaluate power over a range of frequencies, so we chose $m(t)$ as follows:

$$m(t)=\sum_{n=3}^{12} 0.02\cos[2\pi(0.1n+0.01)t].$$

Theoretically, the periodograms generated from the above model would have peaks of equal height at, and only at, 0.31, 0.41, 0.51, 0.61, 0.71, 0.81, 0.91, 1.01, 1.11, and 1.21 Hz. These ten frequencies were chosen so that power at the harmonics of one frequency did not contribute to power at any of the other frequencies. Our initial implementation of this model yielded periodograms with a multiplicity of peaks at other frequencies. We attribute this finding to the fact that the IPFM model has been shown analytically to have power at many of the frequencies of the form $nf+k/2T$, for integer $n\{-\infty<n<\infty\}$ and $k\{1\leqslant k<\infty\}$.[3] A similar result exists when $m(t)$ is a sum of sinusoids.[28] In order to remove these extra peaks, we altered the threshold T of the model by integrating to $T+t_r$, where t_r is a normally distributed pseudorandom number with mean 0 s and standard deviation 0.033 s. These parameters correspond to a coefficient of variation of 8.3%, a reasonable approximation of neonatal HRV.[15]

Clinical Data

RR interval time series data were collected in the Neonatal Intensive Care Unit (NICU) of the University of Virginia Hospital. The University of Virginia Human Investigations Committee approved the protocol. The continuous analog output of a bedside EKG monitor was input to an 80486-based microcomputer equipped with a digital signal processor board (National Instruments AT-2200DSP) and sampled at 4 kHz. QRS complexes were identified using amplitude and duration criteria.

We studied two clinical data sets. The first was a set of 2551 consecutive RR intervals from a NICU patient paralyzed with the nondepolarizing neuromuscular blocking agent pancuronium and mechanically ventilated at 40 breaths/min (0.67 Hz). Normally distributed noise of varying standard deviations, in multiples of 0.5 ms, was added to assess the effect of added noise on RSA peak height. The local respiratory peak height was defined to be the largest periodogram ordinate between 0.66 and 0.68 Hz. The local signal-to-noise ratio was calculated as the quotient of the respiratory peak height and the average power in the bands 0.57–0.66 and 0.68–0.77 Hz. The second clinical data set consisted of records from 83 consecutive admissions to the NICU. Twenty of those patients had a total of 24 episodes of neonatal sepsis and sepsis-like illness. We defined this illness to be present when a physician suspected sepsis and obtained a blood culture and initiated antibiotic therapy.[14] This outcome measure is limited because different physicians might have different criteria for suspecting the illness. It has, however, the advantage of being a clearly defined time point. Despite its limitation, we have found that heart rate characteristics change significantly prior to this clinical endpoint. From the EKG signal of the bedside monitors, we recorded sets of 4096 RR intervals (approximately 20–30 min of data) as previously described.[14] There were 3863 6 h control epochs and 94 epochs from times prior to 24 episodes of sepsis and sepsis-like illness.

Statistical Analysis

For all spectra shown, we calculated Fuller statistics corresponding to significance levels of $p<0.05$ and $p<10^{-10}$, based on the relevant values of M and α. Any periodogram ordinate that exceeded the first threshold was considered a significant peak. Ordinates that exceeded the second threshold were considered highly significant. In addition, we consider that using the Fuller statistics of periodogram ordinates in lieu of the periodogram ordinates themselves normalizes the periodograms, allowing direct comparisons.

For analysis of the large clinical data set, the statistical approach was based on multivariable logistic regression analysis of infants near the time of episodes of sepsis, as diagnosed by the clinician.[14] We first determined the *a priori* risk of sepsis based of clinical features of birth weight and gestational age, and sought to establish whether knowledge of frequency domain characteristics of RR interval time series added a significant amount of additional information. We used the same approach to compare the LP and the spectra of HR samples. Each regression model yielded estimates of the probability of sepsis in the next 24 h for each 6 h epoch for each infant. We fit the data to a multivariable logistic regression model and adjusted the variance–covariance matrix of the maximum likelihood fit to adjust for the replication of measures from the same subject. We used receiver-operating characteristic (ROC) curve analysis to compare the effectiveness of models to estimate the risk of impending sepsis. This standard biostatistical metric is the area of the plot relating the true positive rate (or sensitivity) to the false positive rate (or 1 minus specificity) for many possible test threshold values, and is equal to 1.0 for a test that discriminates perfectly and to 0.5 that discriminates not at all. Testing the hypothesis that a measure of RR interval time series adds significant additional information to that already available to the clinical requires several steps. We (1) determined the *a priori* risk using a regression model that relates birth weight and gestational age to the outcome of sepsis, (2) we determined a new estimate of risk when frequency domain measures were added as input variables for the model, and (3) we used a Wald chi-square test of whether the improvement in ROC area was significant.

RESULTS

Periodograms of Noisy IPFM Model Series

We tested the three spectral methods on simulated data from a ten frequency IPFM model. We judged the methods' relative merits based on two criteria: (1) the periodograms should display a peak at the input frequencies of the model and (2) the heights of these peaks should be equal.

Figure 1 shows a segment of the time series (A), the spectrum of intervals (B), the spectrum of HR samples (C), and the Lomb periodogram. The peaks of the spectrum of intervals at 0.31, 0.41, 0.51, and 0.81 Hz exceeded the $p<0.05$ threshold, while only the peaks at 0.31 and 0.41 Hz exceeded the $p<10^{-10}$ threshold. No other periodogram ordinates exceeded either of the thresholds. In the spectrum of HR samples, the peaks at 0.31, 0.41, 0.51, 0.61, 0.71, 0.81, 0.91, and 1.01 Hz exceeded the $p<10^{-10}$ threshold while the peak at 1.11 Hz only exceeded the $p<0.05$ level. The Fuller statistic of the periodogram ordinate at 1.21 Hz did not reach either threshold. Sixteen spurious peaks exceeded the $p<0.05$ threshold; all but two (at 0.359 and 0.486 Hz) of these occurred at frequencies less than 0.3 Hz. In gen-

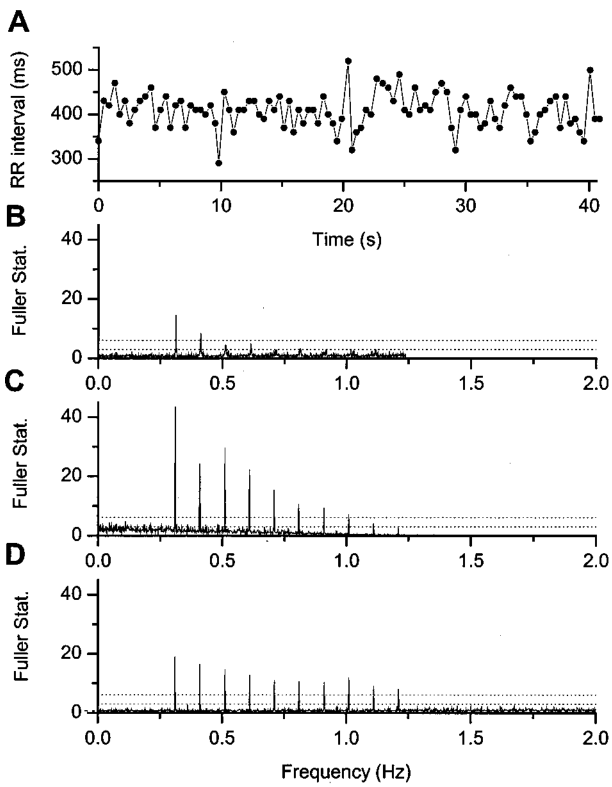

FIGURE 1. Noisy IPFM model with ten periodic components. (A) The first 100 points of the series. (B) The spectrum of intervals of 16,384(2^{14}) points in the series. The spectrum originally had 8192(2^{13}) periodogram ordinates that were condensed into 1024(2^{10}) ordinates up to the average Nyquist frequency of 1.25 equivalent Hz. (C) The spectrum of HR samples of the first 16,384 points of the resampled times series, which corresponds to the first 10,103 points of the original series. The spectrum originally had 8192 ordinates that were condensed into 1024 ordinates with frequencies up to the Nyquist frequency of 2 Hz. (D) The LP of the first 10,103 points of the original time series. The number of ordinates and the frequency range are the same as for the spectrum of HR samples. The periodogram ordinates are displayed as Fuller statistics. Peaks that exceed these lines are interpreted as statistically significant. The dotted lines in panels (B), (C), and (D) correspond to $p<0.05$ and $p<10^{-10}$ for $M=1024$ and $\alpha=8$.

eral, peak height decreased with increasing frequency. The peak at 0.41 Hz does not follow this trend as it is less than the peak at 0.51 Hz. This behavior is due to the effective splitting of the power at 0.41 Hz into two nearby frequencies. The Fuller statistic is 24.41 at 0.4082 Hz and 20.34 at 0.4116 Hz. If the power at these two frequencies is lumped together at 0.41 Hz, the trend of decreasing power with increasing frequency holds throughout the tested frequency range. The decreasing power with increasing frequency suggests a low-pass filtering effect of the interpolation process. Peaks at all ten frequencies in the Lomb periodogram exceeded the $p<10^{-10}$ level. There were no spurious peaks present in the Lomb periodogram. Note that near the average Nyquist frequency of 1.25 Hz, the LF detects correctly a highly significant peak at 1.21 Hz, whereas the spectrum of intervals fails to detect a peak and the spectrum of HR samples detects a statistically insignificant peak.

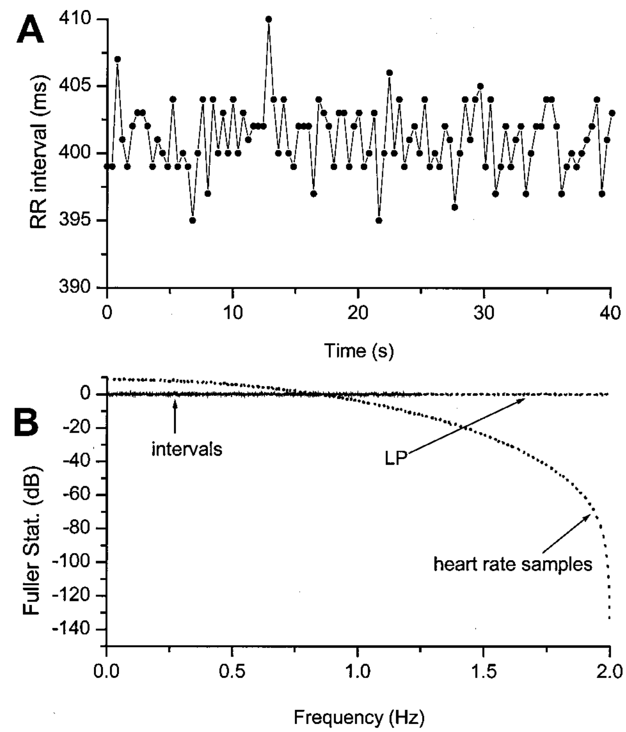

FIGURE 2. IPFM white-noise model. (A) The first 100 points of the series. (B) The spectrum of intervals, the spectrum of HR samples, and the LP. Each spectrum shown is the average of 1000 spectra. Note that the spectrum of intervals ends at 1.25 equivalent Hz.

To further demonstrate the low-pass filter effect of the interpolation process, we analyzed a noisy IPFM model that contained equal power at all frequencies. Figure 2 shows analysis of the IPFM model series with $m(t)=1+[0.125*\text{rand}(t)]$, where rand($t$) returned normally distributed pseudorandom numbers with mean 0 and variance 1. As expected, all Fuller statistics of the spectrum of intervals and the Lomb periodogram, regardless of frequency, were close to 1. The spectrum of HR samples, however, shows higher than expected power at low frequencies and lower than expected power at higher frequencies. The interpolation process distorts power over the entire range of frequencies, acting as a low-pass filter with a -3 dB cutoff of approximately 1 Hz.

Detection of Neonatal RSA

We next compared the three spectral methods' ability to detect RSA in neonatal HR data. Figure 3 shows the spectral analysis of a RR interval series from a preterm infant treated with the paralytic agent pancuronium who was breathing at a rate of 40 breaths/min (0.67 Hz)

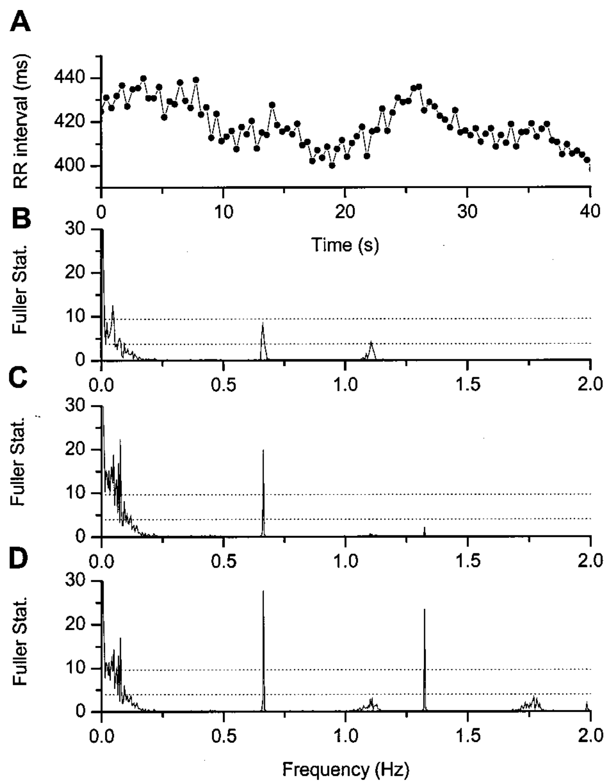

FIGURE 3. Mechanically ventilated neonate. (A) The first 100 points of the series. (B) The spectrum intervals of the first 2048(2^{11}) points of the time series. The spectrum originally had 1024 ordinates which were condensed into 256 ordinates with frequencies up to the average Nyquist frequency of 1.25 equivalent Hz. (C) The spectrum of HR samples of the first 4096(2^{12}) points of the resampled series, which corresponds to the first 2480 points of the original time series. The spectrum originally had 2048 ordinates, which were condensed into 512 ordinates with frequencies up to the Nyquist frequency of 2 Hz. (D) The LP of the first 2480 points of the original time series, with the same number of ordinates and the same frequency range as the spectrum of HR samples. The periodogram ordinates are displayed as Fuller statistics. Peaks that exceed these lines are interpreted as statistically significant. The dotted lines in each spectrum correspond to $p<0.05$ and $p<10^{-10}$ for $M=256$ (B) or 512 [(C) and (D)] and $\alpha=4$ [(B), (C), and (D)].

delivered by a mechanical ventilator. The mean HR was 145 beats/min (2.42 Hz). Since all respiratory activity was due to the ventilator, we expect to find a RSA peak in the periodogram at 0.67 Hz.

The spectrum of intervals shows a peak at 0.67 Hz that exceeds the $p<0.05$ but not the $p<10^{-10}$ threshold. There is another significant peak at 1.11 Hz that represents aliasing of the first harmonic of the RSA peak. Since the mean HR of the time series was 2.42 Hz, the Nyquist frequency is 1.21 Hz, less than the 1.32 Hz at which the first harmonic occurs. The power at this harmonic is aliased to a frequency symmetric about the Nyquist frequency, resulting in the peak at 1.11 Hz. The spectrum of HR samples shows a peak at 0.67 Hz that exceeds both thresholds and an insignificant peak at the first harmonic of 1.32 Hz. The LP shows two peaks that exceeded both significance thresholds, one at 0.67 Hz and one at the first harmonic of 1.32 Hz.

The LP and the spectrum of HR samples detected the periodic component at the respiratory frequency with higher Fuller statistics than the spectrum of intervals. The LP had higher power at the RSA peak than the spectrum of HR samples. Conversely, the low-frequency peaks of the LP had less power than the corresponding low-frequency peaks of the spectrum of HR samples. Both the LP and the spectrum of HR samples had increased power at 1.32 Hz, twice the infant's respiratory rate. This peak was highly significant in the Lomb periodogram but failed to reach significance in the spectrum of HR samples. We attribute these results to the low-pass filtering effect of interpolation.

Stochastic Resonance of Neonatal HRV

Having detected RSA with all three spectral methods, we could use these data to determine how the heights of peaks change relative to one another as the strength of the periodic component decreases relative to the variance of the entire data series. To this end, we added normally distributed noise to the RR intervals derived from the infant on the ventilator. We created many data sets of this type, each with an incrementally higher magnitude of noise. We observed how the RSA signal-to-noise ratio in the periodograms changed with each increase in noise. Figure 4 shows this analysis of the RSA signal-to-noise ratio. As described earlier, the signal-to-noise ratio was calculated relative to the heights of periodogram ordinates in a small, local window rather than the global window of the entire periodogram. Note that the local ordinate ratio is actually higher with added noise of small standard deviation. This enhancement of the signal-to-noise ratio with the addition of noise to the system is called stochastic resonance and has been noted in other biological and physical systems.[10]

Frequency Domain Analysis of Neonatal HRV Prior to Sepsis

We investigated the usefulness of frequency domain analysis in an important clinical setting. We have found that RR interval time series display abnormal characteristics prior to the clinical diagnosis of neonatal sepsis and sepsis-like illness,[14] a potentially catastrophic infectious illness that is the leading cause of mortality in very low birth weight infants in the NICU.[35] The heart rate characteristics abnormalities were reduced variability and transient decelerations, similar to those seen in fetal distress. This is different from the characteristic abnormality of HRV in adults with poor prognosis after myocardial infarction, or with congestive heart failure, where there is

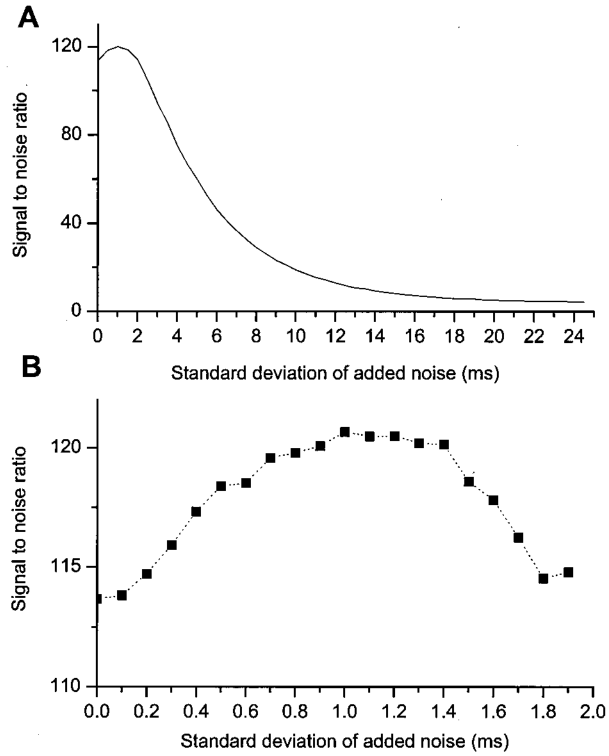

FIGURE 4. Effect of the addition of noise to data from the mechanically ventilated neonate. (A) Average of 1000 plots of the signal-to-noise ratio of the RSA peak as a function of the standard deviation of the added noise. (B) Expansion of the low standard deviation section of (A). Respiratory peak height was defined as the largest periodogram ordinate between 0.66 and 0.68 Hz. The signal-to-noise ratio was calculated as the quotient of respiratory peak height and the average power (excluding the respiratory peak) in the band 0.57–0.77 Hz.

reduced variability but apparently no transient decelerations.[22] In those settings, frequency domain analysis has consistently shown reduced total variance and reduced power at all frequencies.[6,29] In the neonatal setting, however, we found that the standard deviation had little predictive information compared with measures optimized to detect the presence of decelerations from a base line of otherwise reduced variability. Thus, it was not known if frequency domain measures are useful in making the early diagnosis of neonatal sepsis and sepsis-like illness.

We tested the hypothesis that frequency domain measures indeed change prior to this important clinical diagnosis. We continuously monitored 83 consecutive admissions to the NICU using previously described methods.[14] The 20 infants who had one or more episodes of sepsis and sepsis-like illness had significantly lower birth weight (BW; median 1211 g compared to 2835 g, $p < 0.001$, rank sum test) and gestational age (GA; 29.5

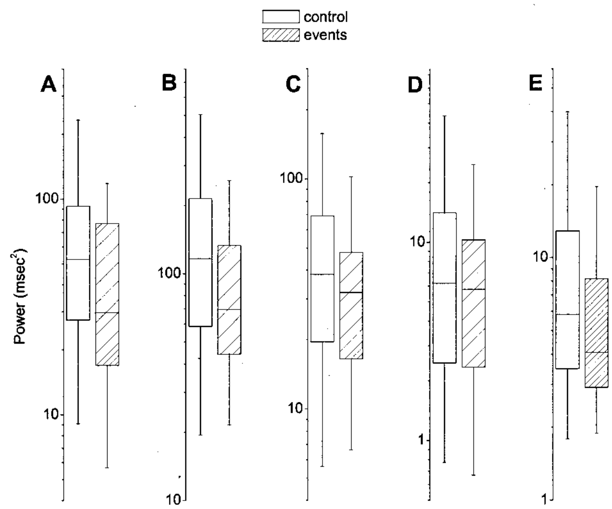

FIGURE 5. Comparison of LP prior to and remote from diagnosis of sepsis. The RR series were collected at times remote from clinical events (control, clear boxes) or in the 24 h prior to a diagnosis of sepsis (events, hatched boxes). LPs were calculated and subsequently integrated over five bandwidths: (A) 0–0.004 Hz, (B) 0.004–0.04 Hz, (C) 0.04–0.15 Hz, (D) 0.15–0.4 Hz, and (E) 0.4–3.0 Hz. In all five bandwidths, differences between control and event values were statistically significant ($p<0.05$, rank-sum test). In the box and whisker plot, the horizontal line marks the median value, the box encloses 50%, and the whiskers enclose 90% of the data.

weeks compared with 36 weeks, $p<0.001$), as expected.[12]

We calculated the LP on each 4096 beat record, and integrated the variance in five bandwidths: 0–0.004, 0.004–0.04, 0.04–0.15, 0.15–0.4, and 0.4–3.0 Hz. We summarized 6 h epochs beginning each midnight by the median values. We performed the analysis on 94 records of 4096 beats summarizing 6 h epochs occurring within 24 h of an episode of sepsis and sepsis-like illness, and on 3863 records that were remote from overt illness.

Figure 5 shows box plots of the summary data from the LP analysis. In each of the frequency bands, the box plots show a small but significant reduction in power for the records prior to sepsis. One explanation is that these measures indeed change prior to sepsis. Another explanation is that they are surrogate measures for the well-known clinical risk factors of prematurity such as low BW and low GA. To test the hypothesis that frequency domain measures add information to more readily available clinical data in the early detection of neonatal sepsis and sepsis-like illness, we used multivariable logistic regression. The steps were to determine the diagnostic usefulness of BW, GA, and the number of days in the hospital, then to determine if significant new information was gained when frequency domain measures were added as input variables. The results are shown in Table 1.

TABLE 1. Regression analysis of frequency domain measures in the diagnosis of neonatal sepsis and sepsis-like illness. BW=birth weight; days=days of postnatal age; d.f.=degrees of freedom; GA=gestational age; ROC=receiver–operator characteristic, overall *p* is for the entire model, *p* for added information relates to the significance of the information added by the frequency domain measures once the BW. GA, and days are known.

BW, GA, log (days)	ROC area	Overall *p*	*p* for added information	χ^2	d.f.
	0.65	0.034		8.6	3
0–0.004 Hz	0.69	0.00010	0.0037	23	4
0.004–0.04 Hz	0.69	0.0010	0.027	18	4
0.04–0.15 Hz	0.66	0.040	0.53	10	4
0.15–0.4 Hz	0.66	0.067	0.96	8.8	4
0.4–3.0 Hz	0.65	0.066	0.73	8.8	4

The most important findings are that only low-frequency components added significantly to clinical parameters in detecting early phases of neonatal sepsis and sepsis-like illness. This result is surprising, as it suggests that the degree of RSA is not of importance in this particular clinical setting, while it is of obvious importance in many others. The contributions of the previously described measures of moments and percentiles were larger. The third and fourth moments, as well as the 10th, 25th, 75th, and 90th percentiles of normalized RR intervals all contributed significantly to the clinical information in predicting sepsis in this data set (not shown). This result confirms our earlier findings comparing selected low- and high-risk groups in an unselected population of consecutive admissions.[14]

DISCUSSION

We analyzed simulated and clinical neonatal HR series in the frequency domain using traditional Fourier techniques (spectrum of RR intervals and spectrum of interpolated HR samples) and the Lomb periodogram, a method based on fitting data to sinusoids. Our most important findings are (1) the LP is better suited to detection of neonatal RSA than Fourier-based methods; (2) in a clinical data set in which respiration was at a single frequency, neonatal HR series can display stochastic resonance; and (3) spectral power is reduced at all frequencies in the 24 h prior to the clinical diagnosis of neonatal sepsis.

Comparison of the Lomb Periodogram with Fourier Methods

The spectrum of intervals detected RSA, but displayed significant aliasing of power beyond the Nyquist frequency. In the case of neonatal HR analysis, where the relationship between the HR and the respiratory rate may not satisfy the Nyquist criterion, the restriction to frequencies below the Nyquist frequency could be a liability. This limitation, along with the recognized difficulty of uneven sampling, suggests that this method may not be ideally suited for detection of neonatal RSA.

The spectrum of HR samples also detected RSA. However, our analysis confirms that the interpolation process used to create the HR samples acts as a low-pass filter and removes some of the high-frequency power from the original signal.[27] This property is evident upon analysis of white noise and simulated as well as clinical data. Thus, for infants breathing at very high respiratory rates, the interpolation process may remove the bandwidths necessary for detection of RSA.

The LP detected accurately RSA and displayed neither aliasing nor any filtering of time series data. The LP also returned the appropriate spectrum of white noise.

RSA has been used previously to report noninvasively on the function of the neonatal ANS. A consistent finding is that the sleep state has a profound effect, but the nature of that effect varies in the literature. Hathorn used the cross correlation of HR and respiratory rate signals to show enhanced entrainment of HR by respiration during quiet sleep.[18] Using spectra of intervals, Baldzer and co-workers found an inverse correlation of RSA with respiratory rate during quiet sleep.[2] This finding is consistent with the results of Dykes and co-workers, who showed that relatively little power was contained in the respiratory frequency range of sleeping neonates.[12] Our study differs in that our patient was premature, mechanically ventilated, and we used alternative methods of spectral analysis. The former difference may explain partially our results in light of the findings of Schechtman and co-workers, who showed that RSA decreases over the first month of life, then increases over the next five months.[34] With regards to mode of ventilation, Koh and co-workers found that RSA decreased with paralysis and mechanical ventilation compared to spontaneous breathing, although those results were obtained from healthy young adults undergoing elective surgery.[24] The sleep state of our patient during the recording interval is unknown, although our patient may provide a model of RSA in REM sleep, during which all skeletal muscles with the exception of the diaphragm are paralyzed.

Stochastic Resonance in Neonatal HR Series

One set of clinical data presented here is from a paralyzed infant with no respiratory effort beyond that delivered by mechanical ventilation. Therefore, all of the variance due to RSA lies in one very narrow frequency band. However, the HR signal is subject to noise from other, nonrespiratory sources such as sympathetic nervous system activity and environmental temperature. Despite the presence of these influences, RSA was easily detected. We hypothesized that the inherent biological noise present in the HR signal may have enhanced the detection of RSA, and we found that the addition of small amounts of artificial noise to this HR signal further enhanced the RSA peak. This finding is in agreement with the hypothesis that detection of RSA is facilitated by the presence of a noisy HR signal.

A similar stochastic resonance effect has been demonstrated in crayfish mechanoreceptors.[10] Douglass and co-workers postulate that sensory systems are ideal candidates for stochastic resonance given their need to detect weak signals in noisy environments. Applying that hypothesis to our system, information regarding respiration may be more efficiently transmitted to regulators of cardiac parameters (baroreceptors, carotid bodies, sinus node) if other, noisy signals influencing those parameters are present as well.

Frequency Domain Analysis of HR Prior to Clinical Detection of Neonatal Sepsis

Power was reduced in all the frequency ranges prior to sepsis. The regression analysis showed that the reduction in power at high frequency did not add information to that present in the BW and age, but that reduction in power at low frequency did. Interpretation of these results within the framework of autonomic nervous system physiology requires, in our opinion, more settled guidelines and a consensus on the assignments of high- and low-frequency power to the activity of the parasympathetic and sympathetic nervous systems, and the balance between the two (see Ref. 13 and responses). It is possible, moreover, that control of neonatal HR differs from that of adults in whom most of the studies have been carried out. From a clinical standpoint, the results suggest that frequency domain analysis may be a useful tool in the clinical exercise to detect neonatal sepsis and sepsis-like illness prior to clinical symptoms. Further investigation of a larger data set, though, is required.

REFERENCES

[1] Aarimaa, T., R. Oja, K. Antila, and I. Valimaki. Interaction of heart rate and respiration in newborn babies. *Pediatr. Res.* 24:745–750, 1988.

[2] Baldzer, K., F. D. Dykes, S. A. Jones, M. Brogan, T. A. Carrigan, and D. P. Giddens. Heart rate variability analysis in full-term infants: Spectral indices of study of neonatal cardiorespiratory control. *Pediatr. Res.* 26:188–195, 1989.

[3] Bayly, E. Spectral analysis of pulse frequency modulation in the nervous systems. *IEEE Trans. Biomed. Eng.* 15:257–265, 1968.

[4] Bendat, J. S., and A. G. Piersol. Random Data. Analysis and Measurement Procedures. New York: Wiley, 1986.

[5] Berger, R. D., S. Akselrod, D. Gordon, and R. J. Cohen. An efficient algorithm for spectral analysis of heart rate variability. *IEEE Trans. Biomed. Eng.* 33:900–904, 1986.

[6] Bigger, Jr., J. T., J. L. Fleiss, R. C. Steinman, L. M. Rolnitzky, R. E. Kleiger, and J. N. Rottman. Frequency domain measures of heart period variability and mortality after myocardial infarction. *Circulation* 85:164–171, 1992.

[7] Bloomfield, P. Fourier Analysis of Time Series. New York: Wiley, 2000.

[8] Brown, T. E., L. A. Beightol, J. Koh, and D. L. Eckberg. Important influence of respiration on human *R–R* interval power spectra is largely ignored. *J. Appl. Physiol.* 75:2310–2317, 1993.

[9] Divon, M. Y., H. Winkler, S. Y. Yeh, L. D. Platt, O. Langer, and I. R. Merkatz. Diminished respiratory sinus arrhythmia in asphyxiated term infants. *Am. J. Obstet. Gynecol.* 155:1263–1266, 1986.

[10] Douglass, J. K., L. Wilkens, E. Pantazelou, and F. Moss. Noise enhancement of information transfer in crayfish mechanoreceptors by stochastic resonance. *Nature (London)* 365:337–340, 1993.

[11] Doussard-Roosevelt, J. A., S. W. Porges, J. W. Scanlon, B. Alemi, and K. B. Scanlon. Vagal regulation of heart rate in the prediction of developmental outcome for very low birth weight preterm infants. *Child Dev.* 68:173–186, 1997.

[12] Dykes, F. D., P. A. Ahmann, K. Baldzer, T. A. Carrigan, R. Kitney, and D. P. Giddens. Breath amplitude modulation of heart rate variability in normal full term neonates. *Pediatr. Res.* 20:301–308, 1986.

[13] Eckberg, D. L. Sympathovagal balance: A critical appraisal. *Circulation* 96:3224–3232, 1997.

[14] Griffin, M. P., and J. R. Moorman. Toward the early diagnosis of neonatal sepsis and sepsis-like illness using novel heart rate analysis. *Pediatrics* 107:97–104, 2001.

[15] Griffin, M. P., D. F. Scollan, and J. R. Moorman. The dynamic range of neonatal heart rate variability. *J. Cardiovasc. Electrophysiol.* 5:112–124, 1994.

[16] Grossman, P., G. Stemmler, and E. Meinhardt. Paced respiratory sinus arrhythmia as an index of cardiac parasympathetic tone during varying behavioral tasks. *Psychophysiology* 27:404–416, 1990.

[17] Grossman, P., B. J. van, and C. Wientjes. A comparison of three quantification methods for estimation of respiratory sinus arrhythmia. *Psychophysiology* 27:702–714, 1990.

[18] Hathom, M. K. Respiratory sinus arrhythmia in new-born infants. *J. Physiol. (London)* 385:1–12, 1987.

[19] Hirsch, J. A., and B. Bishop. Respiratory sinus arrhythmia in humans: How breathing pattern affects heart rate. *Am. J. Physiol.* 10:H620–H629, 1981.

[20] Horne, J. H., and S. L. Baliunas. A prescription for period analysis of unevenly sampled time series. *Astrophys. J.* 302:757–763, 1986.

[21] Katona, P. G., and F. Jih. Respiratory sinus arrhythmia: Noninvasive measure of parasympathetic cardiac control. *J. Appl. Physiol.* 39:801–805, 1975.

[22] Kleiger, R. E., J. P. Miller, J. T. Bigger, Jr., A. J. Moss, and

Multicenter Post Infarction Research Group. Decreased heart rate variability and its association with increased mortality after acute myocardial infarction. *Am. J. Cardiol.* 59:256–262, 1987.

[23] Kluge, K. A., R. M. Harper, V. L. Schechtman, A. J. Wilson, H. J. Hoffman, and D. P. Southall. Spectral analysis assessment of respiratory sinus arrhythmia in normal infants and infants who subsequently died of sudden infant death syndrome. *Pediatr. Res.* 24:677–682, 1988.

[24] Koh, J., T. E. Brown, L. A. Beightol, and D. L. Eckberg. Contributions of tidal lung inflation to human $R-R$ interval and arterial pressure fluctuations. *J. Autonomic Nervous System* 68:89–95, 1998.

[25] Laguna, P., G. B. Moody, and R. G. Mark. Power spectral density of unevenly sampled data by least-square analysis: Performance and application to heart rate signals. *IEEE Trans. Biomed. Eng.* 45:698–715, 1998.

[26] Lomb, N. R. Least-squares frequency analysis of unequally spaced data. *Astrophys. Space Sci.* 39:447–462, 1976.

[27] Moody, G. B. Spectral analysis of heart rate without resampling. Compututers in Cardiology. IEEE Computer Society Press, 1993, pp. 715–718.

[28] Nakao, M., M. Norimatsu, Y. Mizutani, and M. Yamamoto. Spectral distortion properties of the integral pulse frequency modulation model. *IEEE Trans. Biomed. Eng.* 44:419–426, 1997.

[29] Ori, Z., G. Monir, J. Weiss, X. Sayhouni, and D. H. Singer. Heart rate variability: Frequency domain analysis. *Cardiol. Clinics* 10:499–533, 1992.

[30] Press, W. H., and G. B. Rybicki. Fast algorithm for spectral analysis of unevenly sampled data. *Astrophys. J.* 338:277–280, 1989.

[31] Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. Numerical Recipes in FORTRAN, The Art of Scientific Computing. New York: Cambridge University Press, 1994.

[32] Rother, M., H. Witte, U. Zwiener, M. Eiselt. and P. Fischer. Cardiac aliasing—A possible cause for the misinterpretation of cardiorespirographic data in neonates. *Early Human Development* 20:1–12, 1989.

[33] Scargle, J. D. Studies in astronomical time series analysis. II. Statistical aspects of spectral analysis of unevenly spaced data. *Astrophys. J.* 263:835–853, 1982.

[34] Schechtman, V. L., R. M. Harper, and K. A. Kluge. Development of heart rate variation over the first six months of life in normal infants. *Pediatr. Res.* 26:343–346, 1989.

[35] Stoll, B. J., *et al.* Late-onset sepsis in very low birth weight neonates: A report from the National Institute of Child Health and Human Development Neonatal Research Network. *J. Pediatr. (St. Louis)* 129:63–71, 1996.

[36] Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology Heart Rate Variability: Standards of measurement, physiological interpretation, and clinical use. *Circulation* 93:1043–1065, 1996.

[37] Witte, H., U. Zwiener, M. Rother, and S. Glaser. Evidence of a previously undescribed form of respiratory sinus arrhythmia (RSA)—The physiological manifestation of ‘‘Cardiac aliasing.’’ *Pflugers Archiv. Eur. J. Physiol.* 412:442–444, 1988.