HRV Spectral Analysis by the Averaged Periodogram: *Does the Total Power of the Spectrum Really Match with the Variance of the Tachogram?*

Fabio Badilini, Ph.D. and Pierre Maison Blanche, M.D.

From the Hôpital Lariboisière, Paris, France

Heart rate variability (HRV) by spectral analysis can provide information about autonomic nervous system activity. This is mainly achieved by focusing on the low and high frequency components of the tachogram power spectrum. Initially, this technique was applied to short recording periods in very well- controlled conditions. Recently, the use of HRV spectral techniques have been extended to the ambulatory environment, and currently many Holter systems permit analysis of long recording periods. Use of 24-hour Holter recordings soon elicited concerns essentially related to the need for stationary data with conventional spectral techniques. In addition, the recent finding of strong correlations between high frequency spectral parameters and some time-domain indices seems to discourage the long-term approach.

According to Fourier theory (Parseval Theorem), the spectral total power should correspond to the time-domain variance. However, this can rarely be confirmed with the use of commercially available systems, especially for Fast Fourier Transform (FFT) approaches based on averaged Periodograms. For example, a recent study demonstrated that 24-hour total power calculated with three commercial Holter systems (all FFT based) on the same set of analog tapes was completely discordant.[1] In addition, the 24-hour total power of each of the three systems was not comparable to the corresponding time-domain variances. Most probably, the three Holter systems are valid (the low frequency/high frequency ratios were very similar) and the discrepancies found could be related to different preprocessing FFT features or to an improper application of the method.

We are strongly persuaded that the same comparison performed on short-term, fully validated, and stationary data may have led to much more consistent results. Unfortunately, available documentation rarely provides detailed description of the steps taken for spectral analysis. In this regard, the blame for inappropriate use of the technique and consequent increased skepticism for using Holter spectral HRV can be shared between the manufacturers and the users.

The aim of this article is to review the technical reasons that may lead to a discrepancy between total power of averaged Periodograms and time-domain variance. We will then discuss the importance of these discrepancies with respect to Holter spectral analysis.

THE POWER SPECTRUM OF THE TACHOGRAM

Spectral analysis is based on a mathematical transformation that leads from the time domain (second units) to the frequency domain (Hz units).[2] For finite length, discrete time signals such transformation is called Discrete Fourier Transform (DFT).[3] When the number of time-domain samples is an integer power of 2, the efficient algorithm of Fast Fourier Transform (FFT) can be applied to calculate the DFT.

The DFT makes sense only when applied to deterministic signals, i.e., for signals determined by a mathematical expression or by a precise rule of some type. Like all biological signals, the time series of RR intervals, or tachogram, is not a deterministic but rather a stochastic signal, i.e., it is char-

Address for reprints: Fabio Badilini, Ph.D., Hôpital Lariboisière, Service de Cardiologie, 2 rue Ambroise-Pare, 75475 Paris Cedex 10, France. Fax: (33) 149-956981.

acterized by phenomena that are intrinsically aleatory. For stochastic signals that are further stationary, it is more common to evaluate the power spectral density (PSD), often referred to as power spectrum.[4] The most important requirement for a stationary time series is that both its mean and variance should be invariant of the observation period, a condition that the tachogram can fulfill only during specific short periods of time.

One of the most common approaches is to estimate the PSD directly from the DFT.[3-5] This estimate is called Periodogram and basically consists of taking the square magnitude of the DFT. Generally, before the calculation of the Periodogram, the tachogram is multiplied by a smoothing window. This operation is aimed to reduce the effect of spectral leakage, which consists of a smearing of all the frequency peaks caused by abrupt signal changes at the boundary of the time series. The leakage effect can be diminished by tapering the window smoothly to zero at each end, at the price of a resolution loss. Thus, the choice of the right window will depend upon which of these two effects is more crucial.[6] For example, the rectangular window is the best in term of resolution but the worst in terms of leakage; the Blackman window has a very strong tapering (minimal leakage) but also a very reduced resolution. The Hanning window represents a good trade-off between the two effects, and it is generally the one used for HRV spectral analysis.

It has been demonstrated that the Periodogram is an unbiased estimator, but that it is not consistent (which means that its variance does not approach zero with increasing window length but rather is approximately the same size of the power spectrum). A well-accepted solution to this problem is to segment the finite length time series in a number P of sequences (eventually overlapping) and to define as an estimate of the PSD the average of the Periodograms of each single sequence.[7] The averaged Periodogram is a consistent and unbiased estimator, as the variance of the PSD tends to zero with increasing P. The total number of subsegments will clearly depend on the total tachogram length, the subsequent length, and the percent of overlapping (which typically is set to 50%).

TOTAL POWER AND VARIANCE

Differences Due to Lack of Stationarity

Apart from some short time intervals, the tachogram is not a stationary signal. The main consequence is that the calculated spectrum is a poor representation of the frequency components characterizing the time series. In the case of averaged Periodogram, a further consequence is that the total power of the PSD does not match with the originating variance. This is due to two main effects: the effect of windowing; and the effect of averaging.

Effect of Windowing

As a consequence of the Parseval Theorem, the integral (area) of the Periodogram corresponds to the variance of the windowed time series. Clearly, this variance does not necessarily match that of the original signal. In order to compensate for the change of power determined by the windowing effect, the window should incorporate an appropriate scale factor, which can be obtained as follows[2]:

$$s = \frac{1}{\sqrt{\frac{1}{N}\sum_{n=0}^{N-1} (w[n])^2}} \quad (1)$$

where w[n] is the normalized window sequence (w[n/2] = 1) and N is the window length. The more the tapering effect of the window, the larger will be the factor. For the rectangular window the scale factor is equal to 1; for the Hanning window it is $\sqrt{8/3}$. The scale factor can be directly incorporated within the window coefficients, or alternatively applied in the frequency-domain. In the latter case, the factor has to be squared (a constant in the time-domain is square transformed in the PSD). Unfortunately, the full compensation of the scale factor takes place only in perfectly stationary conditions, when the oscillatory components uniformly characterizes the time series. To understand this, we can imagine a component only present at the beginning of the time series, which will be practically eliminated with the application of a (nonrectangular) window. In this case, even with the proper compensating factor, the variance of the windowed data will not correspond to the variance of the original tachogram; it will be larger or smaller depending on the magnitude of the components that had been canceled out (at the boundaries of the segment) with respect to the components that had been magnified (at the center of the segment). Sometimes the scale factor is completely omitted with the result that (independently of stationarity) the PSD will be constantly biased by the missing factor.

Effect of Averaging

As previously stated, the purpose of the averaged Periodogram is to obtain a consistent estimation. Because of stationarity, each subspectrum should have the same total power (the single Periodogram is unbiased). If stationarity is not satisfied, the total power of the averaged Periodogram will not necessarily correspond to the variance of the original tachogram, but rather (disregarding the effect of windowing), to the average of the variances of all segments. This effect is apparent for a small number of segments and tends to disappear when a large number of segments are used, as the errors due to each Periodogram compensate for each other.

In addition, there are three differences not related to stationarity, i.e., effect of detrending, resampling, and of integration mode.

Effect of Detrending

The most common alternative method to estimate the PSD is based on a modelization of the tachogram with an autoregressive (AR) model. Independently of the series of advantages that this method offers (which are not the aim of this article), but rather focusing once again on the issue of total power, we can say that the typical methdology followed by AR models require neither windowing nor segmentation of the data. Thus, the total power of the PSD obtained with AR modeling is generally closer to the variance of the original signal. However, one preprocessing feature that does affect the variance is the so-called detrending. This feature consists of removing very slow components in order to decrease the relative importance of the power close to 0 frequency (DC). Detrending is generally achieved by removing the regression line fitted to the data, but higher order possibilities have also been proposed. Detrending is not at all an exclusivity of AR approaches; nevertheless, it has almost always been proposed in parallel with this method. Independently of other factors, the effect of detrending will intuitively be a reduction of total power.

Effect of Resampling

Since the clock of the tachogram is the sinus node (one observation per beat), its natural frequency units are cycles/beat rather than cycles/s (Hz). In other words, if we directly apply the average Periodogram to a tachogram, the frequency units obtained will be in cycles/beat and the Nyquist frequency of the system will be of 0.5 cycles/beat (i.e., half of the sampling period).[2] In order to obtain Hz units we could either divide the cycles/beat by the

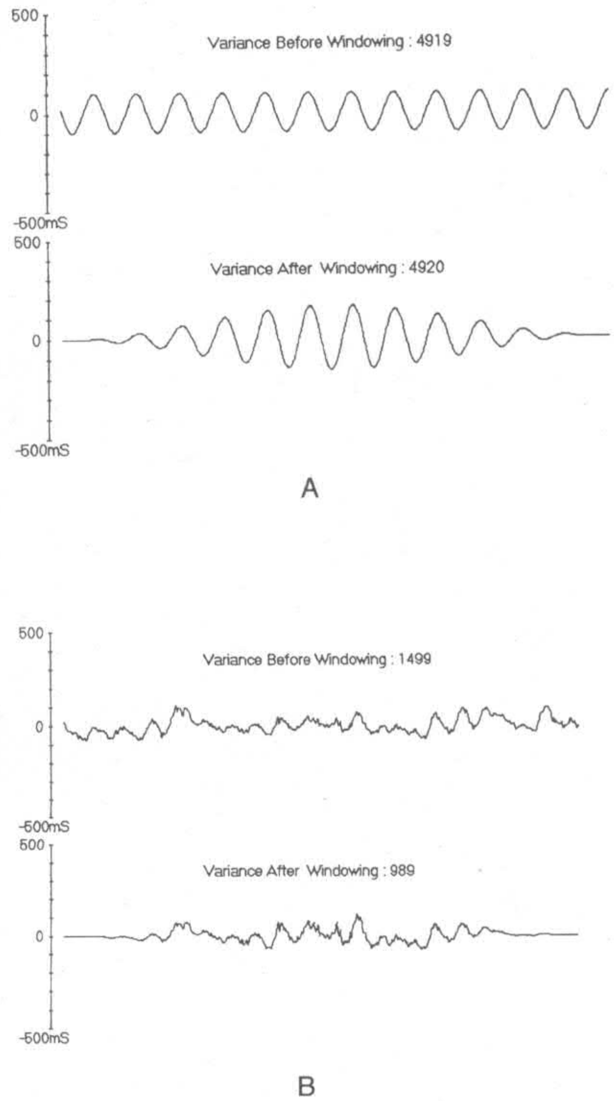

Figure 1. Effect of Hanning window on stationary and nonstationary data. Panel A depicts a sinusoid (0.05 Hz) before and after application of Hanning window (with scale factor incorporated); the variance of the sequence is conserved. Panel B shows the effect of the same window applied to 256 seconds of a nonstationary tachogram (resampled at 4 Hz). Despite the correcting scale factor, the variance of the windowed segment is quite different (reduced) due to the elimination of the components at the boundaries of the segment, which have a larger amplitude with respect to the central ones.

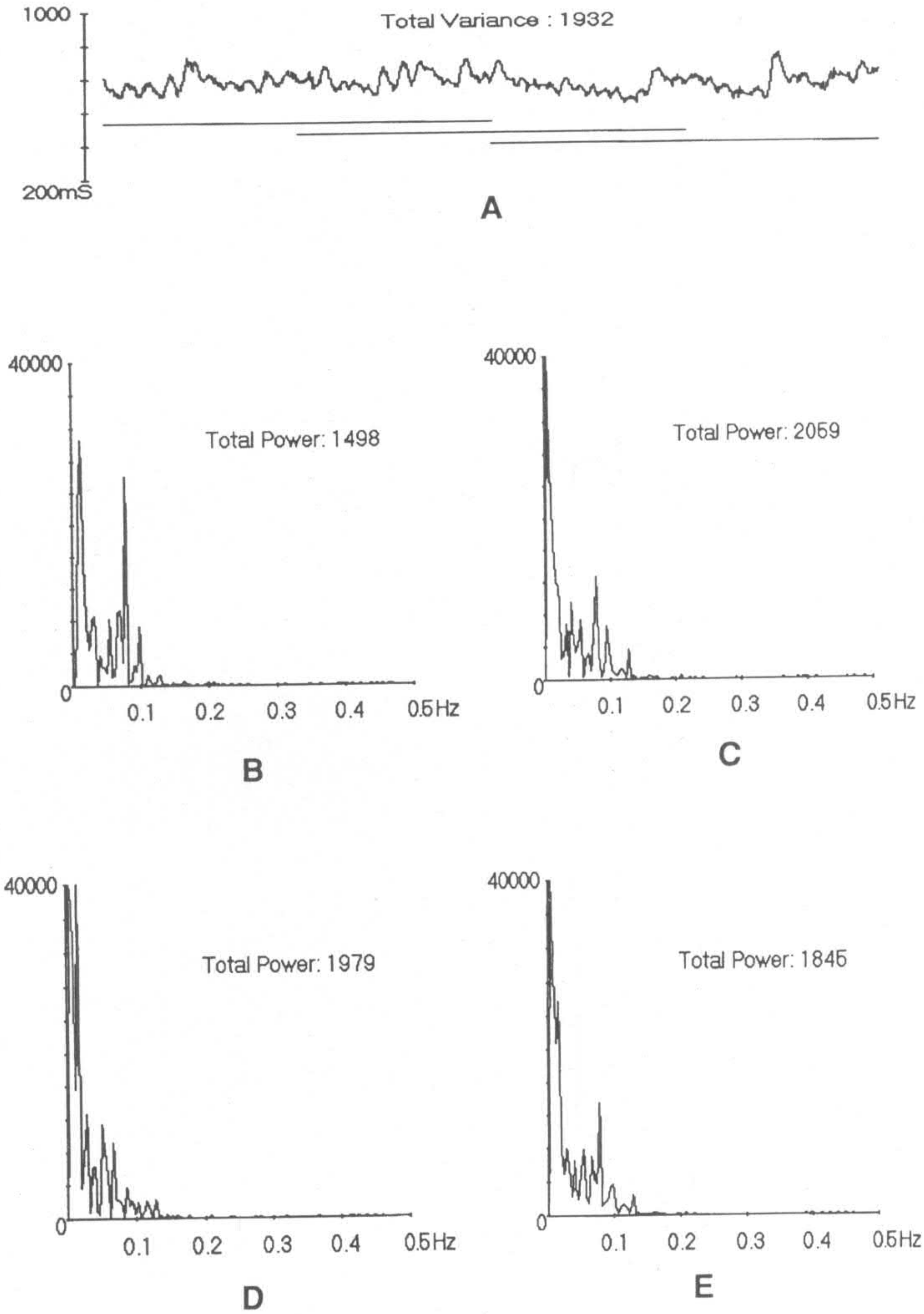

Figure 2. Effect of averaging Periodogram on nonstationary data. Panel A shows a 512-second tachogram (resampled at 4 Hz) with a total variance of 1932 ms^2. By applying averaged Periodogram with 50% overlap three segments are obtained (indicated with solid lines under the tachogram). Rectangular window is applied so that the window effect of Figure 1 is annulled. Panels B, C, and D depict the PSDs obtained for each of the segments. Panel E is the averaged

average RR interval at the end of the procedure (obtaining the so-called equivalent Hz), or to interpolate the tachogram beforehand to obtain a uniformly samples time series (resampled tachogram). The former option is typical of AR approach, whereas the latter is more common in FFT approaches. From a theoretical point of view, the interpolated tachogram should be obtained by uniform sampling of the continuous time tachogram reconstructed with the application of an ideal low pass filter.[8] However, this procedure is tedious and more practical algorithms, such as linear interpolation or the more accurate boxcar filter,[9] had been proposed. Depending on the resampling method applied, the variance of the resampled tachogram could be sensitively modified. Then, it should be clear that the total power of the calculated PSD will have to be matched to the variance of the interpolated tachogram.

Effect of Integration Mode

The total power can also be affected by the way the spectrum is integrated. Due to the symmetric properties of the PSD, the integration is generally performed only on the positive frequency axes and subsequently doubled. Since the spectrum itself is a discrete function, the area is an approximated measure generally performed with either a zero hold method (total area is the sum of all rectangular bins) or with a first order function (total area is the sum of all trapezoids). Clearly, the calculated total power will be different in the two cases. If the spectrum is characterized by a very high DC that decreases quickly (as it is often the case), the difference between the two methods can be significant.

Of course, the precision of integration is also directly influenced by the frequency resolution, which depends on the window length. Nevertheless, the trapezoid rule yields a smaller error, as it better fits the shape of the PSD in the frequencies very close to DC.

All the effects described are somewhat independent from each other. As a result PSD total power may differ significantly from the initial tachogram variance. In Figures 1 and 2 we show two examples that clearly reflect the first two described effects, which are the ones that largely influence Holter data. Figure 1 shows the example of windowing nonstationary data with Hanning window. Figure 2 gives an example of a three segment averaged Periodogram. In order to isolate only the averaging effect, the window applied in this example is the rectangular one, which by definition does not affect the variance.

DISCUSSION AND RECOMMENDATIONS

The main purpose of spectral analysis is to discern the frequency content of signals in order to perform opportune processing within specific subbands. In this regard, it does not really matter if the spectrum is missing a scaling factor, especially if this is the consequence of appropriate mathematical operations applied to improve the discrimination of the various components. After all, the variance is a parameter very easily calculated in the time-domain, and the spectral analysis was never conceived to recalculate this index.

Despite the existence of a mismatch between variance and total power, a very strong correlation between the two has been observed in both normal and postinfarction populations.[10,11] However these results were obtained with a single 24-hour spectrum performed with an alternative methodology (in-toto method)[12] which, by its definition, does not implement either time-domain windowing or averaging of subspectra. Good correlations were also found by comparing powers obtained again with 24-hour in-toto method, and in averages of 5-minute Periodograms over the same period.[13] Authors were not exhaustive in terms of windows used. Nevertheless, the correlations obtained confirmed that the average of many Periodograms (288 5-min segments) somewhat compensates the mismatches of each single spectrum.

Even if, with an opportune description of the method, the mismatch does not represent an obstacle, an elegant solution to the dilemma is to use

Periodogram. Total powers are calculated with the trapezoid rule. All spectra are plotted with the same scale. As it can be observed, the three single spectra do not have the same components, as it should be for stationary data. In particular the strong peak around 0.08 Hz of the first segment is completely absent in the third one. Total power of the averaged Periodogram (1845 ms^2) is about 5% different with respect to the total variance of the original tachogram.

normalized units, i.e., to normalize the power calculated within a band with respect to the total power, or with respect to the power above a certain cutoff.[14] In this case, the concept of total power is intentionally left out. Normalized units are particularly useful when analyzing short-term tachograms, where the variance can be largely determined by very low trends outside the range of interest. Indeed, the behavior of the relative importance of low frequency and high frequency bands during 24-hour Holter recordings has clinical relevance. The main limitation associated with the a priori use of normalized units is in that they do not allow access to the absolute amplitude of the oscillations within such bands.

In conclusion, the most convenient solution, as recommended by the ESC task force document recently published in this Journal,[15] should be to provide both raw power and normalized power. In addition, Holter commercial manuals should include detailed description of all the technical features implemented. The perfect match between a PSD total power and variance of the signal may still be recommended for simulated stationary data during the validation procedure of each specific methodology.

Independently of total power, it may be worthwhile to identify Holter "best portions" to perform spectral analysis. The choice of best portions should be based on stationarity by imposing, for instance, the conservation of mean and variance (within a certain tolerance) in subperiods of these best portions. Clearly, in a Holter environment, this would mean to isolate the analysis on a very selected amount of time, and to raise the concerns of those who consider ambulatory monitoring a tool for analyzing the largest amount of information. Nevertheless, the selection of best periods, by optimizing the environment, may highlight more important frequency-domain features, and by so doing, to reclaim the importance of Holter spectral analysis.

APPENDIX

The Fourier Transform (FT) of a signal is a mathematical transformation that gives information on the frequency content of the signal. This transformation leads from the time-domain (second units) to the frequency-domain (Hz units). For discrete time signals, the FT is a periodic continuous function with periods equal to the sampling frequency.[2] All the information can then be obtained by retaining one period, which in general is the one centered around 0 Hz. In order to make computer analysis possible, the FT must also be discretized (spectral sampling). If the discrete time signal x[n] has finite length, the Discrete Fourier Transform (DFT) is the transformation that provides spectral sampling.[3] For the finite length sequence x[n], the DFT is defined as follows:

$$V[k] = \sum_{n=0}^{L-1} x[n]e^{-j(2\pi/L)kn} \qquad k = 0, \ldots, L\text{-}1 \quad (A.1)$$

where n is the time-domain index and k is the frequency-domain index. In Equation (A.1) both x[n] and V[k] are assumed to have length L. If the length of x[n] is $M < L$, the sequence x[n] will be enlarged of $(L - M)$ samples of zero amplitude (zero padding). The situation with $M > L$ is unpractical because in that case some of the discrete time samples would be ignored by the transformation of Equation (A.1).

Both the FT and the DFT make sense only when applied to deterministic signals. For nondeterministic signals (such as the tachogram) that can be further considered as realizations of the stationary random processes, it is more common to evaluate the power spectral density (PSD).[4] One of the most common approaches is to estimate the PSD directly from the DFT of a finite length segment of a stationary time series. The concept of finite length implicitly assumes that the infinite length time series is multiplied ("windowed") by a finite length sequence. The general expression of a Periodogram is the following:

$$I[k] = \frac{T}{L}\,|V[k]|^2 = \frac{T}{L}\left|\sum_{n=0}^{L-1} w[n]x[n]e^{-j(2\pi/L)kn}\right|^2 \qquad k = 0, \ldots, L\text{-}1 \quad (A.2)$$

where x[n] is the original signal, w[n] is the applied window, T is the sampling period, and L the window length. Basically, the Periodogram is the square magnitude of the DFT of the windowed time series normalized by the scale factor T/L, which assures proper units for the power spectrum (ms^2 / Hz if x[n] is expressed in ms). The PSD is a real function (i.e., it has no phase information) and for each frequency it yields the relative density of power. In addition, the PSD is an even function, which means that it is symmetric with respect to the 0 Hz axis. As a consequence of the Parseval Theorem, the area of the PSD (total power) corresponds to the variance of the time-domain sequence.

REFERENCES

1. Jung J, Heisel A, Tscholl D, et al. Heart rate variability: A comparison of commercially available systems. (abstract) PACE 1995; 18:805.
2. Oppenheim AV, Willsky AS. Signals and Systems. Englewood Cliffs, New Jersey, Prentice Hall, 1983.
3. Oppenheim AV, Schafer RW. Discrete-time signal processing. Englewood Cliffs, New Jersey, Prentice Hall, 1975.
4. Papoulis A. Probability, Random Variables, and Stochastic Processes. New York, New York, McGraw-Hill, 1965.
5. Kay SM, Marple SL. Spectrum analysis: A modern perspective. Proc IEEE 1981; 69:1380-1419.
6. Harris FJ. On the use of windows for harmonic analysis with the Discrete Fourier Transform. Proc IEEE 1978; 66: 51-83.
7. Welch PD. The use of Fast Fourier Transform for the estimation of power spectra: A method based on time averaging over short, modified Periodograms. IEEE Trans Audio and Electroac 1967; 15:70-74.
8. DeBoer R, Karemaker JM, Strackee J. Comparing spectra of a series of point events particularly for heart rate variability data. IEEE Trans Biomed Eng 1984; 31:384-387.
9. Berger RD, Akselrod S, Gordon D, et al. An efficient algorithm for spectral analysis of heart rate variability. IEEE Trans Biomed Eng 1986; 33:900-904.
10. Bigger JT, Albrecht P, Steinman RC, et al. Comparison of time- and frequency-domain based measures of cardiac parasympathetic activity in Holter recordings after myocardial infarction. Am J Cardiol 1989; 64:536-538.
11. Kleiger RE, Bigger TJ, Bosner MS, et al. Stability over time of variables measuring heart rate variability in normal subjects. Am J Cardiol 1991; 68:626-630.
12. Albrecht P, Cohen RI. Estimation of heart rate power spectrum bands from real-world data: Dealing with ectopic and noisy data. Comput Cardiol 1988; 15:311-314.
13. Rottman JN, Steinman RC, Albrecht P, et al. Efficient estimation of the heart period power spectrum suitable for physiologic or pharmacologic studies. Am J Cardiol 1990; 66:1522-1524.
14. Malliani A, Pagani M, Lombardi F, et al. Cardiovascular neural regulation explored in the frequency domain. Circulation 1991; 84:1482-1492.
15. Heart Rate Variability: Standards ofmeasurement, physiological interpretation, and clinical use. Report from the ESC/NASPI Task Force. A.N.E. 1996; 1(2 Pt. 1):151-181.