ABSTRACT

CHARACTERIZATION OF HYPERTENSION THROUGH MULTIVARIATE ANALYSIS UTILIZING LINEAR AND NONLINEAR METHODS

by
Diane L. Donnelly

Analysis of blood pressure by nonlinear methods is vastly underutilized in current research. As such, 24-hour ambulatory blood pressure data from a small cohort of borderline hypertensive and normotensive subjects were analyzed using linear and nonlinear methods. Data were collected and provided by researchers from Columbia University. The cohort size was twelve subjects, consisting of two groups of six each. Although disease state was known, group membership for individual subjects was not. Therefore, one aspect of this research was to separate the cohort, a-priori, into two distinct, evenly sized groups, based solely on analysis results. Separation was accomplished by the long-term scaling exponent α_2 from detrended fluctuation analysis and parameter SD2 from Poincaré analysis. Linear results did not aid in subject separation. Linear analyses consisted of heart rate and blood pressure variability and baroreflex response. The nonlinear analysis methods included approximate entropy (ApEn), detrended fluctuation analysis (DFA) and Poincaré mapping. Analysis was performed hourly, averaged by group and presented as the mean +/- the standard deviation.

Results from linear analysis agree with previous published reports of elevated blood pressure variability, decreased heart rate variability and decreased baroreflex response in hypertension. Results from nonlinear analyses of systolic blood pressure data

revealed elevated approximate entropy values in borderline hypertension indicating an increased randomness. The long-term scaling exponent, $\alpha 2$, from the detrended fluctuation was less in borderline hypertension indicating a break down of the scaling properties of systolic blood pressure in the very early stages of hypertension. Poincaré plots alone revealed little difference between subject groups however, the quantitative parameter SD2, which is an indication of the long-term variability, was on average greater in borderline hypertension. Approximate entropy, the long-term scaling exponent $\alpha 2$ from detrended fluctuation analysis and SD2 from Poincaré analysis were statistically significant to $p < 0.05$ between groups. Statistical significance was determined by paired t-tests over the 24-hour recording period.

The broader impact of this work was the finding that nonlinear analysis methods alone facilitated a-priori subject separation. Characterization of hypertension in a close physiologically cohort was achieved through application of nonlinear analysis methods. Linear analysis methods did not aid in determining group membership during any phase of this research. This work and the results that follow are unique due to the use of nonlinear methods in the analysis of systolic blood pressure, specifically in a cohort of borderline hypertensive and normotensive subjects. In addition to the novel use of nonlinear methods in blood pressure analysis, the presentation of results is also unique. Activity plots of approximate entropy, detrended fluctuation analysis and Poincaré parameters have not been seen in literature researched to date. Generally, presentation of these parameters is one gross measure versus temporal changes over the course of 24-hour recordings.

CHARACTERIZATION OF HYPERTENSION THROUGH MULTIVARIATE ANALYSIS UTILIZING LINEAR AND NONLINEAR METHODS

by
Diane L. Donnelly

A Dissertation
Submitted to the Faculty of
New Jersey Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Biomedical Engineering
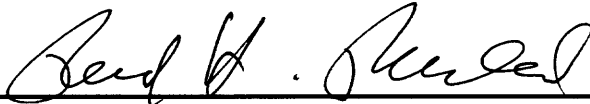
Department of Biomedical Engineering

August 2006

CHARACTERIZATION OF HYPERTENSION THROUGH MULTIVARIATE ANALYSIS UTILIZING LINEAR AND NONLINEAR METHODS

Diane L. Donnelly
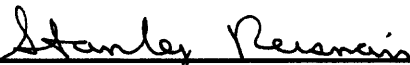
Dr. Ronald H. Rockland, Dissertation Advisor   Date
Associate Dean, Newark College of Engineering, NJIT
Associate Professor of Engineering Technology/Biomedical Engineering, NJIT
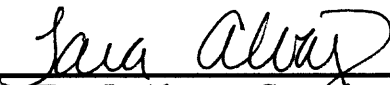
Dr. Stanley S. Reisman, Committee Member   Date
Professor of Biomedical Engineering, NJIT

Dr. Tara L. Alvarez, Committee Member   Date
Associate Professor of Biomedical Engineering, NJIT

Dr. Ronald E. De Meersman, Committee Member   Date
Professor, Columbia University, College of Physicians and Surgeons

Dr. Matthew N. Bartels, Committee Member   Date
John A. Downey Associate Professor of Clinical Rehabilitation Medicine,
Columbia University, College of Physicians and Surgeons

BIOGRAPHICAL SKETCH

Author: Diane L. Donnelly

Degree: Doctor of Philosophy

Date: August 2006

Undergraduate and Graduate Education:

- Doctor of Philosophy in Biomedical Engineering,
  New Jersey Institute of Technology, Newark, NJ, 2006

- Master of Science in Biomedical Engineering,
  New Jersey Institute of Technology, Newark, NJ, 2003

- Bachelor of Science in Electrical and Computer Engineering Technology,
  New Jersey Institute of Technology, Newark, NJ, 2002
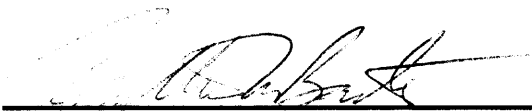
Major: Biomedical Engineering

Presentations and Publications:

Donnelly D. L., Rockland R. H., Reisman S. S., Quigley K. S.,
      "Baroreflex Sensitivity Index in Chronic Fatigue Syndrome,"
      Proceedings of the 30th Annual Northeast Bioengineering Conference, NEBC
      2004, Springfield, MA, USA, pp. 85-87, 2004.

Donnelly D. L., Rockland R. H., Reisman S. S., Quigley K. S.,
      "Continuous Measurement of BRSI in Chronic Fatigue Syndrome,"
      Proceedings of the 26th Annual International Conference of the IEEE Engineering
      in Medicine and Biology Society, EMBS 2004, San Francisco, CA, USA., vol. 1,
      pp. 906-908, 2004.

To those who believed in me and shared in my dream and for the support and
encouragement throughout these years

Thank you

This thesis is dedicated to my parents Arthur and Adelaide Donnelly and to the memory
of my brothers Daniel and Arthur

ACKNOWLEDGMENT

TABLE OF CONTENTS

TABLE OF CONTENTS
(Continued)

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES
(Continued)

CHAPTER 1

INTRODUCTION

Hypertension is a leading risk element in stroke, kidney failure and congestive heart failure (CHF), as well as an implicating factor in myocardial infarction [1]. Primary hypertension, also called essential or idiopathic hypertension, is characterized by a persistently elevated blood pressure that has no known organic cause. There are several categories of hypertension; however, approximately 95 % of all hypertensive subjects are characterized as primary. Due to the unknown or idiopathic nature of this disorder, additional and innovative research methodologies are necessary. Further, it has been reported that there exists increasing evidence that analysis of scaling and complexity properties of heart rate dynamics may provide valuable information for clinical use [2]. It is hypothesized that this could be true for blood pressure dynamics.

It has been reported that blood pressure variability in hypertension increases with increasing blood pressure and correlates closely with end-organ damage [3, 4]. As such, application of nonlinear measures can aid in the characterization of hypertension. Approximate entropy (ApEn) is a nonlinear measure that is sensitive to the degree of regularity in a time series. A high degree of regularity produces a small ApEn value, while reduced regularity produces a large ApEn value. Detrended fluctuation analysis (DFA) has the capability of detecting embedded correlations in a seemingly non-stationary time series [5]. Peng et al. have shown that in subjects with congestive heart failure there is a breakdown of this correlation behavior [5]. Quantifying parameters SD1 and SD2 from Poincaré maps represent short and long-term variability, respectively, of the data. Additionally, Poincaré maps provide a graphic portrayal of the behavior of a

1

time series. Standard linear measures of blood pressure and heart rate variability, along with baroreflex response determined in the frequency domain by the alpha index, provide expanded information in hypertension to compliment these nonlinear analysis methods.

1.1 Hypothesis

The National Heart Lung and Blood Institute (NHLBI) has for the past three decades administered the organization of a coalition of professional, public and voluntary organizations including seven federal agencies to improve awareness of, and issue guidelines for, the prevention and treatment of hypertension [6]. The "Seventh Report of the Joint National Committee on the Prevention, Detection and Treatment of Hypertension", or JNC VII provides the latest guidelines in which a pre-hypertensive state is defined.   Analysis that can elucidate autonomic nervous system behavior during this stage will greatly aid to further characterize and understand the pathophysiology of the disease.

Given the increased use of nonlinear methods in heart rate variability studies, along with references to their clinical value, it is hypothesized that these same analyses applied to blood pressure data, which has at best been performed on a limited basis, will provide valuable information with regard to blood pressure dynamics, thus expanding existing research.   Although current linear analysis techniques of blood pressure data have proven useful, there is a need to explore the utility of nonlinear analysis methods in an effort to broaden the field of research into blood pressure dynamics and the underlying pathophysiology in hypertension.

It is further hypothesized that the application of nonlinear analysis methods in conjunction with current linear methods holds the potential to reveal information present in blood pressure data that may not be apparent using linear methods alone. The combination of these analysis methods may lead to advancements in the characterization of hypertension and elucidate the associated autonomic dysfunction that has traditionally been described with linear methods.

The first area of research of this thesis is to apply nonlinear analyses to blood pressure data in a small cohort of unmedicated borderline hypertensive subjects and normal controls. The utility of these analyses in conjunction with linear methods is to aid in the characterization of hypertension. This is based on two areas of research that have been published in the literature. The first area states that hypertension is due to an underlying autonomic dysfunction [1, 7-11]. Therefore, techniques that can elucidate this dysfunction can aid in understanding the underlying pathophysiology of hypertension. The second area of research that exists in the literature is the recent work of applying nonlinear analysis methods to understand heart rate data, and its underlying autonomic mechanism [5, 12-22]. Therefore, the overall goals of this thesis are to:

- Utilize linear and nonlinear analysis methods with blood pressure and pulse interval data.

- Evaluate results of nonlinear analyses to determine their efficacy as an aid in the characterization of hypertension.

- Assess the significance and clinical utility of these methods.

In order to achieve the goals of this thesis three linear and three nonlinear analyses were used. These analysis methods are:

- Blood Pressure Variability   (linear)
- Heart Rate Variability   (linear)
- Alpha Method   (linear)
- Approximate Entropy   (nonlinear)
- Detrended Fluctuation Analysis   (nonlinear)
- Return Map/Poincaré plot   (nonlinear)

Blood pressure variability (BPV) and heart rate variability (HRV) are frequency analysis methods widely used in research to understand the activity of the parasympathetic and sympathetic nervous systems, the two subdivisions of the autonomic nervous system. Previously, it was stated that research has found that hypertension is due to an underlying autonomic dysfunction. Therefore, it is anticipated that these two analyses will assist in understanding the contribution of the nonlinear methods to characterize hypertension.

The alpha method, also called the spectral technique, is an approach to assess spontaneous baroreflex sensitivity based on estimations in the frequency domain to obtain high and low frequency alpha coefficients. This method is based on the fact that the systolic blood pressure (SBP) and pulse interval (PI) show a high degree of linear correlation at the respiratory frequency and at 0.1 Hz in normal subjects, and on the hypothesis that the correlation at these two frequencies is due to the baroreflex coupling [23-26]. This analysis will aid in understanding the baroreflex response in hypertension and provide an additional means for interpreting nonlinear results to understand how these techniques facilitate the characterization of hypertension.

Nonlinear analyses were chosen to provide an indication of the regularity, long and short term scaling characteristics and variability. These parameters were determined through the use of approximate entropy (ApEn), detrended fluctuation analysis (DFA) and Poincaré maps respectively. It is hypothesized that the combination of these analyses will enhance the current body of knowledge and aid in the characterization of hypertension.

1.2 Background

The use of nonlinear analysis is gaining acceptance in heart rate variability studies. A literature search uncovered numerous journal articles on the subject of nonlinear techniques applied to electrocardiogram (ECG) recordings. Approximate entropy (ApEn), developed by Steve Pincus [12], is a nonlinear measure that quantifies the regularity of a time series. Higher regularity produces a small ApEn value; conversely reduced regularity produces a large ApEn value. Given previous research findings that blood pressure variability increases with increasing blood pressure, [4,10] it is hypothesized that ApEn will be a valuable aid in the characterization of hypertension. It is anticipated that approximate entropy will be elevated in unmedicated borderline hypertension versus a normotensive subject. Pincus verified the differences in regularity of heart rate between a normal infant and a case of aborted sudden infant death syndrome (SIDS) during quiet sleep. This is illustrated in Figure 1.1.
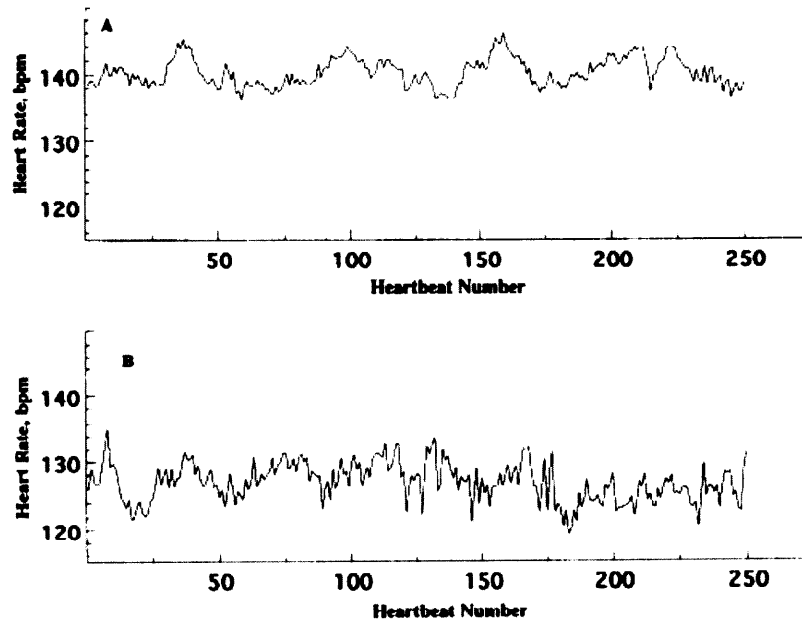
Figure 1.1 Top is aborted SIDS infant; Bottom normal infant. Tracing (a) appears more regular than (b), Pincus quantified differences between these tracings with ApEn [13].

Voss et al. in a 1996 paper applied methods of nonlinear dynamics in an attempt to improve risk stratification in sudden cardiac death. In this paper, Voss et al. utilized three dimensional Poincaré maps, or phase-space representations, to facilitate the visualization of beat to beat dynamics in heart rate between normal controls and heart failure subjects. This is illustrated in Figure 1.2. By plotting a three dimensional representation, with axes x, y and z represented as the points x (i), x (i+1), x (i+2) respectively of the time series, these plots were useful in that they were illustrative of the complexity of an electrocardiogram recording of a patient at high risk for sudden cardiac death [14].
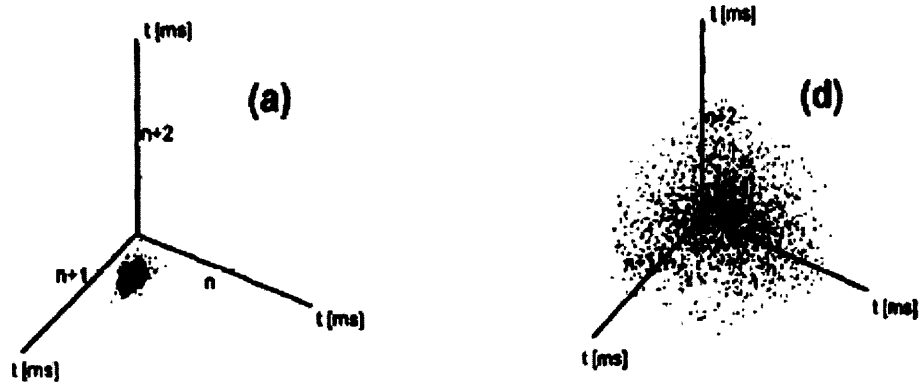
Figure 1.2 A 3-dimensional 30-min phase space plot of a healthy subject (a) versus that of a patient at high risk for sudden cardiac death with atrial fibrillation (d). Recordings are in milliseconds (ms) [14].

Based on their results, it was clear that the three-dimension phase-space plots (Poincaré maps) of a healthy person are quite different from those of patients with heart failure [14]. Voss et al. found the use of three dimensional Poincaré maps useful in a severe disease state, sudden cardiac death versus healthy controls. The relative health of the cohort in this study is more closely matched than in the case of Voss. Thus, this research has used two-dimensional Poincaré plots as an aid in the characterization of hypertension. It is hypothesized that application of two dimensional Poincaré plots can provide additional insight into the dynamics of blood pressure and aid in the discrimination of subjects. Characterization of hypertension based on Poincaré results were quantified through measures of dispersion SD1 and SD2 the short and long-term variability of the data respectively.

Detrended fluctuation analysis (DFA) was first introduced by Peng et al. in February 1994, where it was used to quantify long-range power-law correlations in DNA

sequences. Briefly, DFA is a modified root-mean-square analysis of a random walk. The data are first integrated; this is an important step in detrended fluctuation analysis as it maps the time series to a self-similar process. The integrated time series is then separated into equal size windows of size "n" as shown in Figure 1.3. The local trend within each window is calculated with a least squares-fit. The linear trend is then removed from the respective window. This process is repeated over various size windows.
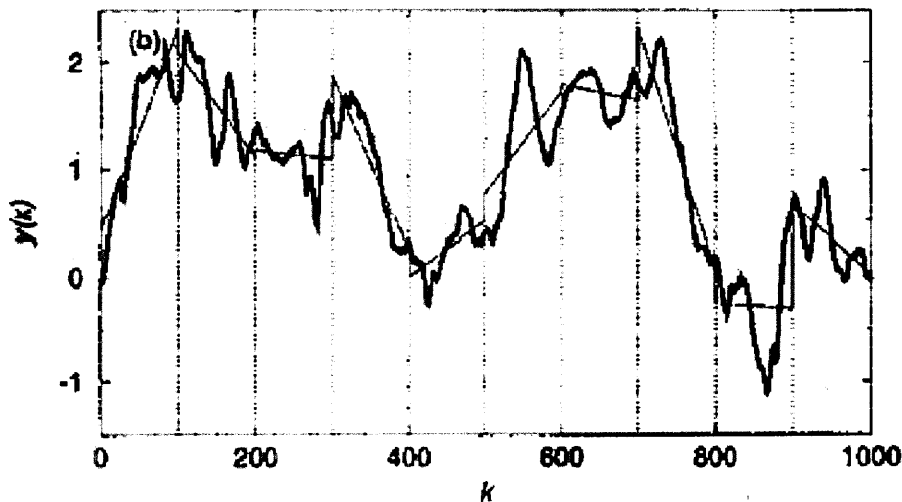
Figure 1.3 Integrated time series, vertical line indicates box size of n = 100, a least squares-fit represents the trend in each box and is shown with the straight line segments [5].

The characteristic fluctuation, termed F(n), for a given box size is calculated with a modified root-mean square equation. This calculation is repeated over all time scales providing a relationship between F(n), and window size "n". A linear relationship on a log-log plot indicates the presence of scaling which is shown in Figure 1.4 [5, 27]. The slope of the line termed alpha, relating log F(n) to log n is taken as a scaling exponent, the value of which is used to determine the status of the system under evaluation.
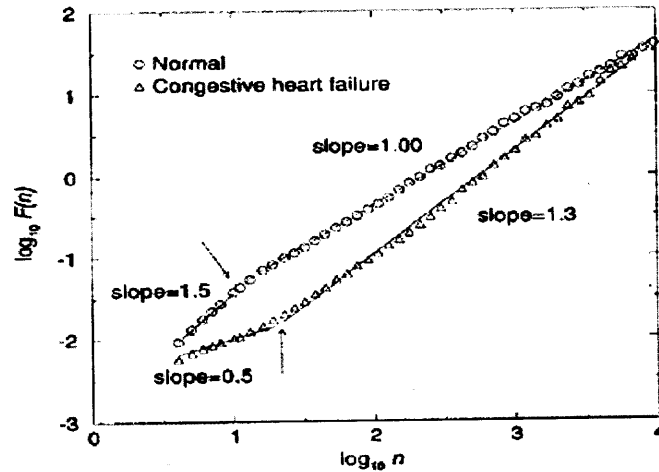
Figure 1.4 Log-log plot of normal vs. congestive heart failure subjects. The slope as noted above is the scaling exponent "alpha" of the time series. There are two slopes or alpha values indicating different scaling behavior for small box sizes versus larger box sizes. Peng et al. define this as crossover scaling that is evident in both normal and congestive heart failure subjects [5].

In an effort to determine a clinically reasonable data length, Peng et al. concluded that data sets of N = 8192 points to be a statistically reasonable choice. Using data of this length (N=8192), Peng calculated two scaling exponents alpha-1 and alpha-2. Alpha-1 is a short scaling exponent obtained from a least squares fit of log F(n) versus log n for box sizes 4<= n <=16; alpha-2 a long scaling exponent obtained from box sizes 16< n <= 64 [5]. Statistically significant results were obtained between normal and congestive heart failure (CHF) subjects for both short and long term alpha exponents [5]. Scatter plots were used to illustrate the spread of the two scaling exponents. The scatter plots indicated a tighter clustering of normal subjects versus their congestive hear failure counterparts. This is shown in Figure 1.5. It has been noted that DFA analysis has been useful in revealing the extent of long-range correlations in time series [28].
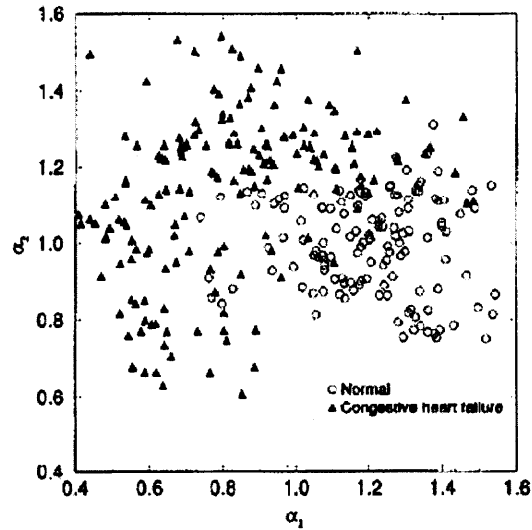
Figure 1.5 Scatter plot of alpha-1 versus alpha-2 for healthy subjects, circles, and congestive heart failure subjects, diamonds [5].

Vickman et al. combined the use of traditional time and frequency methods with detrended fluctuation (DFA) and approximate entropy (ApEn) analyses in a study of paroxysmal atrial fibrillation. He found that both DFA and ApEn detected abnormalities in heart rate dynamics that traditional methods were unable to measure [15]. Makikallio et al. in a 10-year follow-up study of elderly subjects to predict the risk for sudden cardiac death determined that the short alpha-1 scaling exponent from detrended fluctuation analysis was a powerful predictor of cardiac death and in particular sudden cardiac death [16]. Several other articles discussed or reviewed nonlinear dynamics as they apply to heart rate variability [17-22]. However none of them discussed these same analyses as applied to blood pressure analysis.

CHAPTER 2

AUTONOMIC DYSFUNCTION IN HYPERTENSION

2.1 Research Studies

In 1989, Oz et al. hypothesized that the development of hypertension is due to a defect in cardiovascular control [7]. They tested their hypothesis with Wistar-Kyoto rats (WKY) and WKY spontaneously pre-hypertensive rats (SHR) with and without alpha-sympathetic blockade to determine differences in spectral parameters between the two groups prior to the onset of hypertension. Alpha-sympathetic blockade was accomplished with prazosin, which is classified as an alpha-adrenergic blocker. Prazosin works by binding to alpha-1 receptors that control the constriction of blood vessel (veins and arteries), thus the name alpha-sympathetic blockade. Adrenergic refers to catecholamines, epinephrine and norepinephrine, which would normally bind to alpha-1 receptors to constrict veins and arteries. When Prazosin binds to the alpha-1 receptors, it prevents sympathetic nervous system catecholamines from binding and as a result sympathetic response is reduced.

Blood pressure perturbations were achieved by a reduction of blood volume through hemorrhage. The abrupt loss of volume and pressure elicited a sympathetic response, specifically the binding of catecholamines to alpha-1 receptors to constrict blood vessels. This is a measurable response between the two groups by spectral analysis. During hemorrhage, a larger increase in two low frequency ranges (0.004 – 0.07 Hz and 0.145 – 0.15 Hz) were seen in the SHR rats as compared to their normotensive counterparts, and during alpha-sympathetic blockade the response in these same ranges in the SHR rats were blunted [7]. The alpha-sympathetic branch of the

sympathetic nervous system along with the renin-angiotensin system are both reportedly involved in regulation of the low frequency range (0.004-0.07Hz) of arterial blood pressure fluctuations [8]. It was concluded by Oz et al. that the divergence between the two strains during hemorrhage with and without alpha-sympathetic blockade suggests that hypertension is linked not only to the alpha-sympathetic activity, but the remaining blood pressure control system as well [7].

Conclusive evidence of a disparity between borderline hypertensive and normotensive subjects was presented by Akselrod et al. while studying the response to a change in posture between the two groups [9]. In that study they reported the existence of an enhanced basal level of sympathetic activity in unmedicated borderline hypertensive subjects as compared to their normotensive counterparts. More recently, in an article describing the sympathetic nervous system's role in regulating blood pressure variability, it was noted that activity of the sympathetic nervous system provides one of the fundamental mechanisms in the control of arterial blood pressure [10].

In 2003, Davrath et al. researched autonomic malfunction in normotensive subjects with a genetic disposition to essential hypertension. Utilizing heart rate and blood pressure they investigated the possibility for early detection of essential hypertension. During an active change in posture, they found that low frequency fluctuations in heart rate were more prominent in young adult normotensive offspring of one hypertensive parent as opposed to normotensive offspring of two normotensive patents [11]. They also noted a reduced alpha-index which was attributed to a decreased baroreceptor activity.

The alpha-index is a frequency domain method for determining baroreflex sensitivity and relies on the assumption that a high degree of linear correlation exists at the respiratory frequency and 0.1 Hz in normal subjects and on the hypothesis that this correlation is due to the baroreflex coupling [23, 24, 26, 29, 30].

In 2001, Beevers et al. noted that although there have been research studies into hypertension there remains much uncertainty regarding the pathophysiology of hypertension [31]. Only a small percentage (2% - 5%) of all hypertensive subjects has an identifiable underlying cause such as renal or adrenal disease. These subjects are diagnosed as having secondary hypertension. Roughly 95% of all hypertensive subjects are diagnosed with essential or primary hypertension, which means there is no identifiable underlying cause; the disease is classified as idiopathic [31]. Based on the research to date, it is accepted that one of the primary pathologies associated with hypertension is a complex autonomic dysfunction with evidence of sympathetic hyperactivity and/or vagal withdrawal [1, 11]. Due to the reported complex nature of the autonomic dysfunction and the inherent nonlinearity of the physiological system, it is hypothesized that application of nonlinear methods may reveal additional information not available through linear analysis as was noted in the Vickman study.

2.2 Research Motivation

Although there have been nonlinear analyses performed on blood pressure, [1, 32-41], nonlinear analysis remains an under-utilized method. Of the eleven articles noted that utilize nonlinear analysis of blood pressure, only four have applied these methods to human subjects [32, 36, 40, 41]; the remaining studies were performed on rats and dogs.

Butler et al. [32] reviewed the fractal nature of short-term systolic blood pressure and heart rate variability during lower body negative pressure, by evaluating the slope beta of the $1/f^\beta$ power law scaling. In 1997, Mestivier et al. [36] looked at the relationship between diabetic autonomic dysfunction and heart rate variability using recurrence plots.

The Jartti et al. [40], and Kuusela et al. [41], studies appear to have used the same data with expanded nonlinear methods employed by Kuusela. One of the nonlinear measures utilized in both studies was approximate entropy. These are the only studies identified to date where this nonlinear measure was utilized to evaluate blood pressure data. The main difference between these two studies is that Kuusela used several nonlinear measures versus two utilized by Jartti.

In 1995, Wagner et al. utilized Lyapunov exponents and correlation dimension to investigate complexity in blood pressure following baroreceptor denervation in conscious dogs. A review article written in 1996, by Wagner et al. discussed chaos in blood pressure control. Almog et al. in 1996, utilized correlation dimension to characterize arterial blood pressure and surrogate data analysis to test for nonlinearity in rats. In 2001, Mestivier et al. used recurrence plots to study the effects of autonomic blockers on linear and nonlinear indexes of blood pressure and heart rate in spontaneously hypertensive rats. In 2000, Gonzalez et al. utilized recurrence quantification analysis to detect sources of nonlinearity in the variability of R-R intervals and blood pressure in rats. Dabire et al. in 1998, also utilized recurrence plots to quantify nonlinear indices in normotensive rats. In 2001, Eyal et al. investigated correlation dimension to quantify the complexity in blood pressure control in rats.

It is clear that linear analysis has improved our understanding of hypertension and its physiological correlates. It is also clear, however, that it is necessary to expand analysis techniques to advance research into blood pressure control mechanisms as they relate to hypertension. Compared to the volume of nonlinear research performed on heart rate data, there is a minimal amount of comparable work on blood pressure data and no studies identified that utilize approximate entropy (ApEn) or detrended fluctuation analysis (DFA) in a cohort of normotensive and unmedicated borderline hypertensive subjects.

As noted earlier, there have been only two studies identified which have used ApEn to analyze blood pressure data [40, 41]. These studies utilized a pharmacological agent (terbutaline) in order to identify changes in the variability of heart rate and blood pressure. There have been no studies identified which are completely noninvasive that utilize these methods and none in this particular cohort of subjects.

CHAPTER 3

METHODS

3.1 Approximate Entropy

Approximate entropy analysis has been widely used in various heart rate variability studies [2]; yet only two research studies have been identified that apply this technique to blood pressure. These two studies are virtually identical in their use of ApEn, reducing the number of independent studies identified to one. Additionally, analysis of blood pressure data utilizing approximate entropy has not been applied to data from a cohort of normotensive and unmedicated borderline hypertensive individuals.

Approximate entropy (ApEn) is a statistic that quantifies the regularity and complexity of a time series. When computing ApEn, two parameters "m" and "r" must be selected and fixed with "m" a positive integer and "r" a positive real number [12]. Values for "m" equal to 1 or 2, and "r" ranging between 0.1 to 0.25 times the standard deviation of the data have been found to produce good statistical validity [13]. Because the parameter "r" is expressed as a fraction of the standard deviation of the data, ApEn is a scale-invariant measure allowing measurements on data sets of different amplitudes to be compared. The parameter "m" defines a vector or pattern length, with "r" defining the acceptable scalar distance between vector elements.

In order to calculate ApEn, select parameters "m" and "r" as described above. Given N data points $x(1)$, $x(2)$,...$x(N)$, form vector sequences $u(1)$ through $u(N-m+1)$ defined as $x(i) = [u(i),..,u(i+m-1)]$. These vector sequences represent "m" consecutive u values beginning with the ith point. The distance between vectors $x(i)$ and $x(j)$ is the maximum distance in their respective scalar components. The vector sequences define

$C_i^m(r)$ values that measure within tolerance "r" the regularity or frequency of patterns similar to a given pattern of length "m". Mathematically $C_i^m(r)$ is defined as [12, 13, 43, 44]:

$$C_i^m(r) = x(j) \leq N - m + 1$$

$$d[x(i),x(j)] \leq r]/(N-m+1) \tag{3.0}$$

The natural log of the $C_i^m(r)$ values defines $\Phi^m(r)$ [13, 43]:

$$\Phi^m(r) = (N-m+1)^{-1} * \sum_{i=1}^{N-m+1} \ln C_i^m(r) \tag{3.1}$$

This process is repeated using vectors of length m+1 to verify the number of vectors that remain within tolerance "r" given the next incremental data point. This process allows the parameter ApEn to be defined in Eqn. 3.2 [12, 13, 42].

$$ApEn(m,r) = \lim_{N \to \infty} [\Phi^m(r) - \Phi^{m+1}(r)] \tag{3.2}$$

The statistic which approximates the parameter of Eqn. 3.2 is defined as [13, 42]:

$$-ApEn(m,r,N) = \Phi^{m+1}(r) - \Phi^m(r) \tag{3.3}$$

When there are recognizable patterns that repeat the conditional probability will be closer to one, producing logarithmic values closer to zero. With random behavior, conditional probabilities will be closer to zero producing higher logarithmic values. Therefore, when regularity is high ApEn is low; conversely when regularity is low ApEn is high. ApEn can be thought of as the negative natural logarithm of the conditional probability that sequences that are close for "m" points remain close for an additional point. An advantage of approximate entropy analysis is its demonstrated ability to discern changing complexity with a relatively small amount of data [2, 12, 13, 43].

3.2 Detrended Fluctuation Analysis

There have been no studies identified to date that utilize DFA in the analysis of blood pressure in a cohort of normotensive and unmedicated borderline hypertensive subjects. Detrended fluctuation analysis is a widely used technique for the detection of long-range correlations in noisy, nonstationary time series and has been successfully applied to diverse fields such as DNA sequencing, neuron spiking, human gait and heart rate dynamics [43]. When calculating DFA, the data series is first integrated over all N such that [5]:

$$Y(k) = \sum_{i=1}^{k}[B(i) - B(u)] \tag{3.4}$$

Where $B(i)$ equals the ith data interval and $B(u)$ is the average interval. This step is crucial to detrended fluctuation analysis because it maps the time series to a self-similar process allowing for the investigation of fractal properties of the accumulated time series versus the original signal.

Next, the data are divided into equal boxes of length "n". Within each box of length "n" the data are fit with a least squares line representing the local trend which is removed see Figure 1.3, (pg. 7), and the root-mean-square fluctuation of the integrated and detrended series is calculated. Mathematically this is shown in Equation 3.5 [5].

$$F(n) = \sqrt{\frac{1}{N}\sum_{k=1}^{N}[y(k) - y_n(k)]^2} \tag{3.5}$$

The value $F(n)$ represents the average fluctuation as a function of box size. This calculation is repeated over all box sizes to provide a relationship between $F(n)$ and the number of beats in a box, typically $F(n)$ increases with box size "n" [5]. A linear

relationship on a log-log graph (Figure 1.4 pg. 9), indicates the presence of scaling. Given this relationship, the fluctuations can be characterized by two scaling exponents α_1 and α_2, which are slopes of the line relating *log F(n)* to *log n* [5, 27, 45]. The short-term scaling exponent is α_1, and α_2 is the long-term scaling exponent. If data are completely uncorrelated, such as white noise, the value of α will be equal to 0.5. If there are only short-term correlations, the slope may be close to but not exactly 0.5, it will however approach a value of $\alpha = 0.5$ for large windows. A value of α greater than 0.5 and less than or equal to 1 is an indication of long range power-law correlations [5, 27, 45].

3.3   Poincaré Map

The Poincaré plot was developed by Henri Poincaré for analyzing complex systems and has been used in physics, astronomy, geophysics, mathematical biology and medical sciences [46]. In medical science, the predominant use is in the study of heart rate variability [47-50]. The Poincaré map, also called the return map, is a graphical tool as well as a quantitative analysis technique to describe the dynamics of a system. When constructing a Poincaré map, one state of the system is plotted versus its state in the next time step or simply x (n) versus x (n+1). In the lexicon of nonlinear dynamics, the plot represents a two-dimensional embedding of the time series. The utility of the return map is its ability to reveal the possibility of an underlying determinism. Meesmann et al. are one of several groups of researchers that have utilized return maps as one of a number of nonlinear dynamic techniques to study heart rate variability. Figure 3.1 was taken from the Meesmann 2000 study; it illustrates a tachogram of a 24-hour Holter record of a healthy individual, and the associated return map.
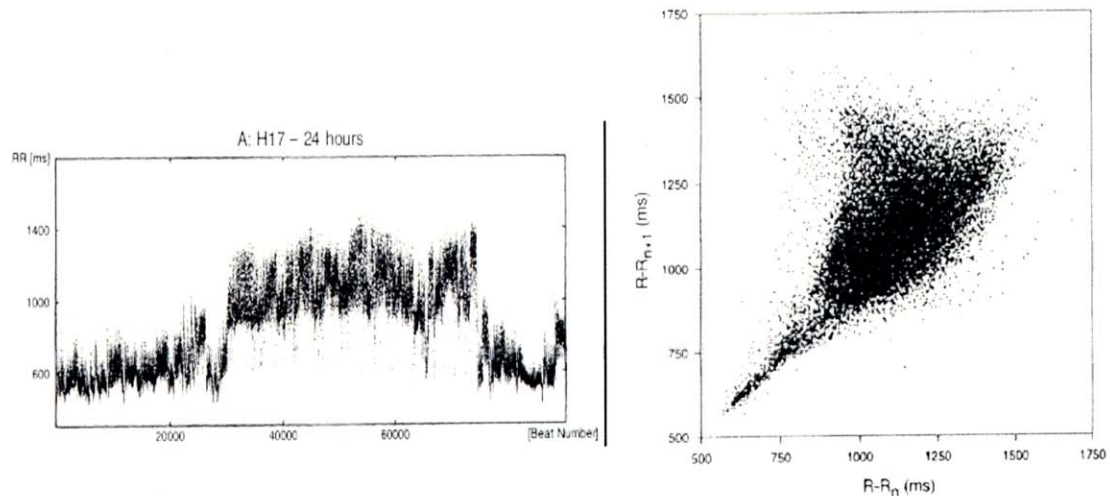
Figure 3.1 Left tachogram from a 24-hour Holter record of a healthy female. Right a Poincaré or return map of the data [47].

The long R-R intervals in the tachogram on the left in Figure 3.1 represent a prolonged period of sleep with longer R-R intervals. The large fluctuations appear to be interrupted by shorter beat-to-beat intervals indicating an increase in heart rate for short periods. The comet shaped return map represents this increased heart rate (reduced beat-to-beat intervals) in the thin portion or tail of the comet with the larger variability in the wider portion or head of the comet [47].

Statistically, the Poincaré map displays the correlation between consecutive intervals in a graphical manner. To characterize a Poincaré map mathematically, many researchers have adopted the practice of fitting a virtual ellipse to the plot with axes SD1 and SD2 as illustrated in Figure 3.2 [50]. SD1 and SD2 are common Poincaré descriptors defining the standard deviation along the line of identity (SD2) and the standard deviation of points along a line perpendicular to the line of identity (SD1) [46, 48, 50].
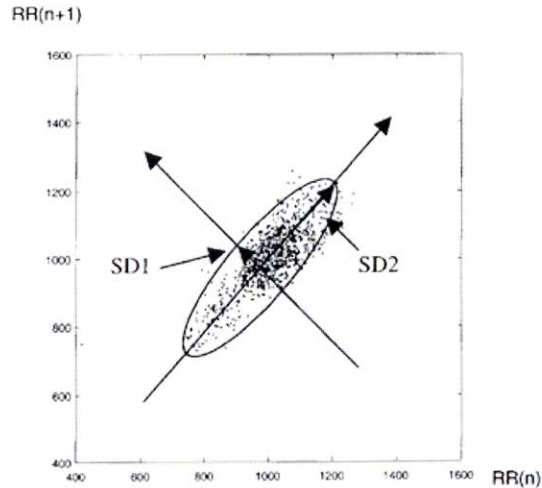
Figure 3.2 Illustration of Poincaré plot and quantification indicating SD1 and SD2 [46].

The ellipse in Figure 3.2 serves to help the reader visualize the dispersion of the data along the two axes. The parameters SD1 and SD2 are widely accepted as determinants of the short-term (SD1) and long-term (SD2) dispersion of data [46, 50]. These parameters have determined statistical differences between subject populations when used in heart rate data [47, 48]. These statistical descriptors and how to calculate them are described elsewhere [46, 50-53].

3.4   Alpha Index

In previous research, baroreflex sensitivity was determined with the sequence method on a cohort of chronic fatigue syndrome and closely matched normal subjects. The sequence method determines a baroreflex sensitivity index by first identifying systolic blood pressure ramps. That is systolic blood pressure (SBP) values that rise or fall for three or more consecutive beats. Once all SBP ramps have been identified the R-R intervals are

compared for concomitant rise or fall in value. If SBP and RRI have an associated increase or decrease for three or more consecutive beats the slope of a linear regression line between SBP and RRI is taken as an indication of baroreflex sensitivity. Because the sequence technique is based on ramp patterns, baroreflex sensitivity is assessed over a wide frequency range. The spectral technique focuses on systolic blood pressure and pulse interval oscillations that are limited to narrow frequency ranges. This means that the alpha coefficients derived from this approach reflect the baroreflex ability to modulate the sinus node at specific frequencies providing the possibility to separately assess sympathetic and parasympathetic contributions [10, 25]. Despite differences between the sequence and spectral techniques, the quantification of baroreflex sensitivity is reportedly quantitatively and qualitatively equivalent [25]. In this study, baroreflex sensitivity was analyzed in the frequency domain with the alpha method.

The alpha method is an approach to the assessment of spontaneous baroreflex sensitivity based on estimations in the frequency domain to obtain the alpha coefficients, α_{LF} and α_{HF}. This method is based on the fact that the systolic blood pressure (SBP) and pulse interval (PI) show a high degree of linear correlation at the high or respiratory frequency and at the low frequency of 0.1 Hz in normal subjects, and on the hypothesis that the correlation at these two frequencies is due to the baroreflex coupling [23-26].

Initially SBP and RRI data are subdivided into short segments ranging from 128 to 1024 beats. From these segments, interpolated interbeat intervals are created from which the fast Fourier transform (FFT) of each segment is then taken. This is followed by calculation of coherence. If the coherence in the high or low frequency range is >0.5, then the SBP and RRI spectra are integrated over these peaks [23-26]. Coherence is an

indication of the linear relationship between the two signals in the frequency domain. The threshold on coherence is traditionally imposed to increase the specificity of the baroreflex sensitivity estimation [23]. The α_{LF} and α_{HF} coefficients are calculated by taking the square root of the ratio between the RRI and the SBP spectral powers at the correlated frequencies as indicated in Eqn. 3.6.

$$\alpha_{LF} = \sqrt{\frac{RR_{LF}}{SBP_{LF}}} \quad \alpha_{HF} = \sqrt{\frac{RR_{HF}}{SBP_{HF}}} \tag{3.6}$$

Equation 3.6 is a calculation of the gain of the transfer function between systolic blood pressure and R-R interval changes. This relies on the assumption that the ratio between these data at the noted frequencies is a reflection of the baroreflex function [25]. Figure 3.3 illustrates RRI and SBP spectra and the calculated coherence.
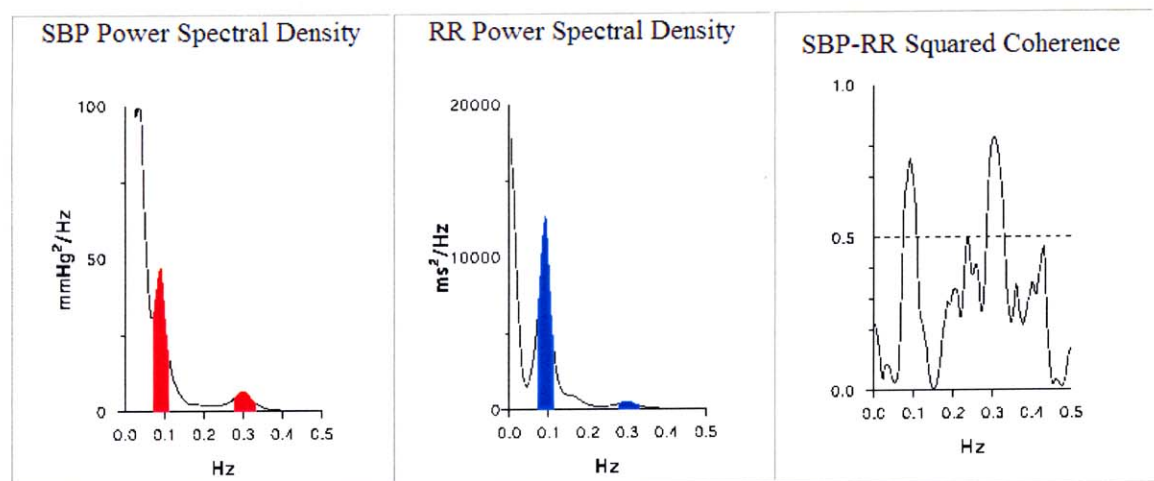
Figure 3.3 Spectra of SBP left, RRI middle, right is the SBP-RRI squared coherence, the dotted line in the coherence plot is the threshold value of 0.5 [54].

The respiratory frequency is considered the high frequency (HF) range providing an assessment of parasympathetic influence. The low frequency (LF) range is taken as 0.1 Hz and is considered to represent the sympathetic contribution [10, 24, 25]. Figure 3.4 is a graphic from a study by Parati et al. 2000 that illustrates the equivalence between the sequence and spectral techniques in assessment of baroreflex sensitivity.
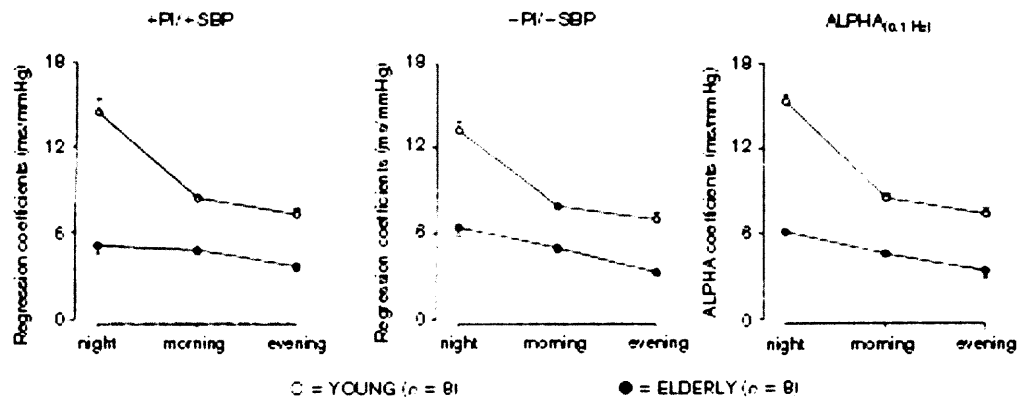
Figure 3.4 Time and frequency domain estimates of spontaneous baroreflex sensitivity. The data shown are the average (± SE) of 4 hour intervals selected in the morning, afternoon and night from two groups of eight young and elderly subjects. Left and middle panels are sequence technique of positive and negative sequences respectively. Far right panel are alpha coefficients computed around 0.1 Hz [25].

3.5   Heart Rate Variability

Heart rate variability (HRV) is a widely used marker of autonomic activity [54].   The parasympathetic and sympathetic nervous systems both divisions of the autonomic nervous system are largely responsible for controlling heart rate [55, 56].

In order to assess HRV, a signal is derived from an electrocardiogram, called an interbeat interval (IBI) which is a measure of time in seconds or milliseconds between consecutive R-waves. This is called the R-to-R interval and is illustrated in Figure 3.5.
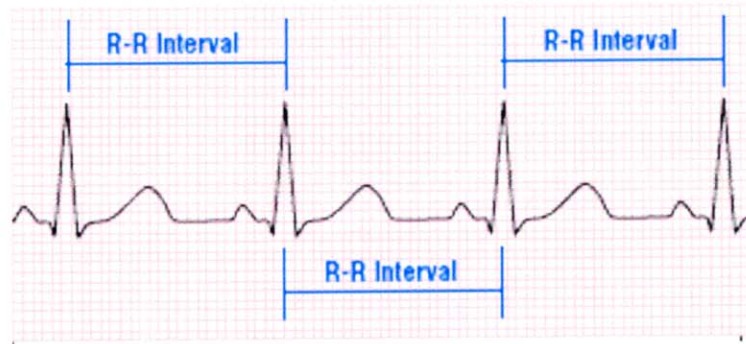
Figure 3.5 Illustration of R-to-R intervals from which an interbeat interval (IBI) is derived [57].

In this study, continuous blood pressure data are utilized to create the same signal. The use of pulse interval (PI) data has been previously confirmed as a statistically reliable substitute for electrocardiogram data [58]. In order to detect pulse interval data, the location of each systolic peak was determined utilizing custom software which was developed for this research. These intervals were stored in time intervals as milliseconds between successive peaks.   From the IBI, descriptive statistics such as mean heart rate, variance and standard deviation can be calculated.   Standard deviation provides a gross

estimate of overall HRV but it does not discriminate in the frequency domain [59]. In order to obtain this information, frequency analysis techniques are employed. Before frequency analysis is possible, the IBI signal is interpolated to create an IIBI or interpolated interbeat interval. The IIBI is a continuous wave with equally spaced samples created by backward step interpolation [60]. From the IIBI, power spectral data can be obtained which typically indicates three distinct frequency ranges, a very low frequency (0.003 Hz to 0.04 Hz) band, a low frequency (0.04 Hz to 0.15 Hz) band and a high frequency (0.15Hz to 0.4Hz) band [61, 59]. The very low frequency band has been associated with thermoregulatory systems [60, 61]. The low frequency band is associated with both sympathetic and parasympathetic control of heart rate via baroreflex feedback with the high frequency band linked to parasympathetic activity and respiration [59, 61, 62]. These frequency ranges are of particular significance because the parasympathetic and sympathetic nervous systems both divisions of the autonomic nervous system are largely responsible for controlling heart rate and blood pressure [3, 49, 55].

Sympathetic nervous system activity is not easily defined from the power spectra. As a result, a sympathovagal measure, which is the ratio of low frequency to high frequency, is used as an approximate measure of sympathetic activity.

3.6   Blood Pressure Variability

Blood pressure, measured in millimeters of mercury (mm Hg), is usually noted as systolic blood pressure over diastolic blood pressure. Systolic pressure represents the contraction and emptying of the ventricles when the heart is pumping blood to the body. Diastolic pressure is the relaxation and filling of the heart in preparation for the next cardiac cycle. A cardiac cycle consists of alternating systolic and diastolic periods.

Measurement of blood pressure reflects the total force per unit area that the blood exerts on the interior of a blood vessel. There is no gold standard identifying normal blood pressure values, rather there are acceptable ranges. In general, in the aorta of a resting, young healthy adult, blood pressure rises to roughly 120 mm Hg during systole, and drops to about 80 mm Hg during diastole [55, 56].

Blood pressure variability (BPV) is defined as the fluctuations around its mean value and/or the fluctuations in some predetermined frequency. In order to assess blood pressure variability, it is necessary to record an interbeat interval (IBI) signal as well as the corresponding systolic peaks. An interbeat interval represents the time in milliseconds (ms), between consecutive systolic peaks. The systolic peak is the maximum pressure attained during each beat or cardiac cycle. These measures can be obtained from continuous blood pressure recordings, using computer algorithms to detect the points of interest. The standard deviation of mean daytime and nighttime blood pressure is one of the more widely used approaches to the assessment of blood pressure during a 24-hour period [3, 4]. While standard deviation does provide an estimate of overall BPV, it does not discriminate frequency domain power distributions.

In order to obtain this information, frequency analysis techniques are performed. However, before frequency analysis is possible, an interpolated systolic blood pressure signal is derived. The derivation of this signal is identical to that described in Section 3.5. From the interpolated systolic blood pressure interval, power spectral data can be obtained. The same frequency ranges that are present in heart rate variability exist in blood pressure variability. However the mechanisms responsible for overall blood pressure variability remain unclear [10]. It has been noted that the low frequency range in blood pressure variability is due purely to sympathetic influences [9, 11]. This is in contrast to heart rate variability where it is noted that both parasympathetic and sympathetic activities play a role in low frequency power [59, 61, 62]. In 1997, Akselrod et al. noted differences in low frequency power between heart rate and blood pressure in response to postural change. They state that these findings concur with previous research and confirmed that the low frequency component of blood pressure variability is only slightly and indirectly influenced by parasympathetic activity [9]. High frequency is attributed to parasympathetic activity and the mechanical affects of respiration [11]. These frequency ranges are of particular significance because the parasympathetic and sympathetic nervous systems are largely responsible for controlling heart rate and blood pressure [3, 49, 55].

One of the mechanisms contributing to sympathetic and parasympathetic activity is the baroreceptor reflex response. Changes in mean arterial blood pressure are automatically mediated by the baroreceptor reflex. Baroreceptors are stretch receptors that influence the heart and blood vessels adjusting cardiac output and total peripheral resistance to restore blood pressure to normal. The most important baroreceptors

involved in the minute to minute regulation of blood pressure are located in the carotid sinus and the aortic arch. They are sensitive to both mean arterial blood and pulse pressures. These receptors are strategically located providing critical information regarding changes in arterial blood pressure in the vessels traveling directly to the brain before branching off to supply the rest of the body [55]. The baroreceptor reflex is the body's rapid response system that counteracts temporary disturbances in pressure. They do not however, regulate arterial pressure in the long term because they adapt to prolonged arterial pressure changes.

CHAPTER 4

PROGRAMMING

The data for this thesis are 24-hour blood pressure files collected by researchers at Columbia University. As such, it was necessary to develop a program that was capable of detecting, recording and storing consecutive systolic peaks along with the time interval between peaks for subsequent analysis. A custom LabVIEW® version 7.0, program that detected R-waves in electrocardiogram data had already been developed on the Signals Lab at NJIT. An illustration of R-waves and the R-to-R interval were previously shown in Figure 3.6. This custom LabVIEW® program was heavily modified to include the detection of blood pressure parameters and write to file capabilities. Regardless of the input data, electrocardiogram (ECG) or continuous blood pressure, points of interest are detected, can be dynamically edited to correct detection errors and written to file for subsequent analysis. Partial views of the front panel and the hierarchy of the program are shown in Appendix A.

Programming for blood pressure and heart rate variability were developed in LabVIEW® a graphical programming language. Programming is near identical for both of these programs. The difference lies in derivation of the interpolated interbeat interval. With heart rate variability, the interpolated interbeat interval is derived from the interbeat interval which is a record of the time, generally in milliseconds, between consecutive R-waves. For this work the interbeat interval is the time in milliseconds between consecutive systolic peaks, or pulse interval data. The use of pulse interval data, as noted in Section 3.6, has been shown to be an equivalent, and statistically sound substitute for R-to-R interval data obtained from electrocardiograph recordings [58]. The interbeat

interval data enters a loop which segments the data into blocks of 512 data points to satisfy stationarity. The program then derives an interpolated interbeat interval (IIBI) from each data segment and removes the mean. The fast Fourier transform is taken of each demeaned IIBI and the frequencies of interest are integrated and normalized per Malik et al. [62]. Data that has been normalized and the data from which the normalized values were derived are written to file for subsequent analysis. Programming for blood pressure variability is identical to that of heart rate variability with the exception of the derivation of the interpolated interbeat interval. Heart rate variability derives an IIBI from an interbeat interval. In order to create an IIBI for blood pressure data both the interbeat interval data and the respective systolic blood pressure values are needed. A partial view of the front panel of the blood pressure variability program is shown in Figure 4.1. The front panel for the heart rate variability program is similar.
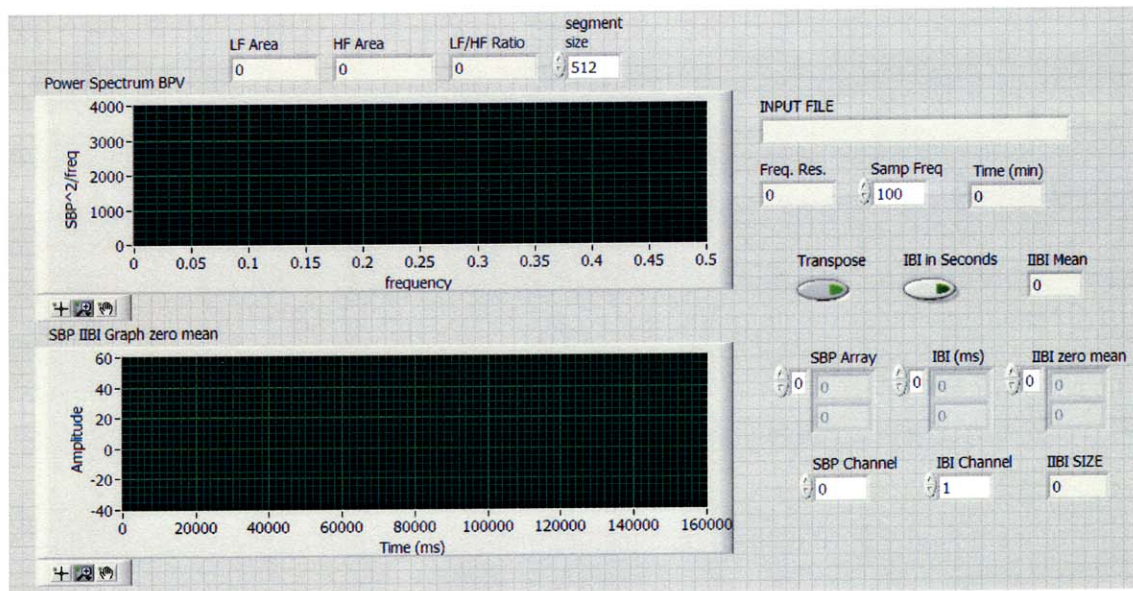
Figure 4.1 Front panel view of blood pressure variability program developed in LabVIEW®. The front panel of heart rate variability is similar and is not shown.

Custom programming for the alpha index was developed in MatLAB® version R2006a. The program takes in two columns of data, an array of systolic blood pressure in millimeters of mercury (mm Hg) and the associated interbeat interval in milliseconds. Interpolated interbeat intervals are created for the systolic blood pressure and the pulse interval data in blocks of 512 data points. The mean is removed from each IIBI and the spectra are determined by application of the Fourier transform. This is followed by calculation of coherence between the systolic blood pressure and pulse interval spectra. Coherence is then checked to determine if it is equal to or greater than 0.5 within the low and high frequency ranges of interest. When this occurs the spectra are integrated and the corresponding alpha index is calculated. The frequency ranges were identified following correspondence with Dr. Paolo Castiglioni, whose group in Italy has done extensive research on the baroreflex response and is widely published. The frequency ranges used by Castiglioni et al. with 24-hour data are, low frequency >= 0.07 Hz to < 0.14 Hz, and high frequency >= 0.14 Hz to 0.5 Hz. The low frequency range used is the same as suggested by Dr. Castiglioni, the high frequency range was set to >= .14 Hz to .4Hz for this work. Once the alpha indices were calculated they were written to file for further analysis.

Poincaré map, approximate entropy and detrended fluctuation analysis are all written in MatLAB®. The core MatLAB® programs for approximate entropy and detrended fluctuation analysis were downloaded from PhysioNet and customized to match the goals of this research study. The original code was written by Dr. Daniel Kaplan [28]. The core code for both approximate entropy (ApEn) and detrended fluctuation analysis (DFA) were modified such that they are now capable of processing

data from this thesis. The output for each program was tested. Auxiliary programs were developed for both DFA and ApEn programs to optimize runtime, facilitate analysis of hourly blocks of data and accumulate results. Once complete the results are then written to file for further analysis.

Poincaré map programming was developed from algorithmic descriptions in papers by Brennan et al. [50, 53] and Piskorski et al. [46]. The program creates data vectors, previously described in Section 3.3, calculates the quantitative measures SD1 and SD2 from the vectors and plots the results. An auxiliary program that runs the main Poincaré program was developed to optimize processing and segmenting large data sets.

Cluster analysis programming was written in MatLAB® using the kmeans algorithm. Three cluster analysis programs were developed specifically to analyze the three configurations of data available. These three programs and kmeans clustering are discussed in Chapter 6.

Principal component analysis (PCA) was written in MatLAB® with the princomp algorithm. As with cluster analysis, three separate programs were written to effectively analyze the three data configurations. The three programs and principal component analysis princomp are discussed in Chapter 6.

Numerous subroutines and auxiliary programs were developed in both MatLAB® and LabVIEW® to support the main programs discussed above and facilitate all analyses.

CHAPTER 5

DATA

5.1   Description

The study consists of 24-hour data from 12 subjects.   Columbia researchers have indicated that there exist six normotensive and six unmedicated borderline hypertensive subjects within this cohort.   The status of individual subjects is blind for this research in order to determine if it is possible to classify groupings a-priori based on combined analyses.   Once all analysis was complete, subjects were separated into groups. Separation of subjects was based on the individual analyses performed utilizing both cluster and principal component analysis, both of which are discussed later.   Once two groups were identified, statistical analysis between groups was assessed.

The data were recorded with a Portapres™ at a sampling frequency of 100 samples per second.   The Portapres™ is a noninvasive ambulatory finger arterial blood pressure monitor which is lightweight; battery operated and is worn by the subject in a waist belt [63]. The measurement is taken in two adjacent fingers alternately at selectable intervals, usually in the non-dominant hand and utilizes a height correction system to compensate for hydrostatic blood pressure changes in the finger [63].   The height adjustment allows the subject to have free hand movement during daily activities and sleep.   The Portapres™ is shown in Figure 5.1.
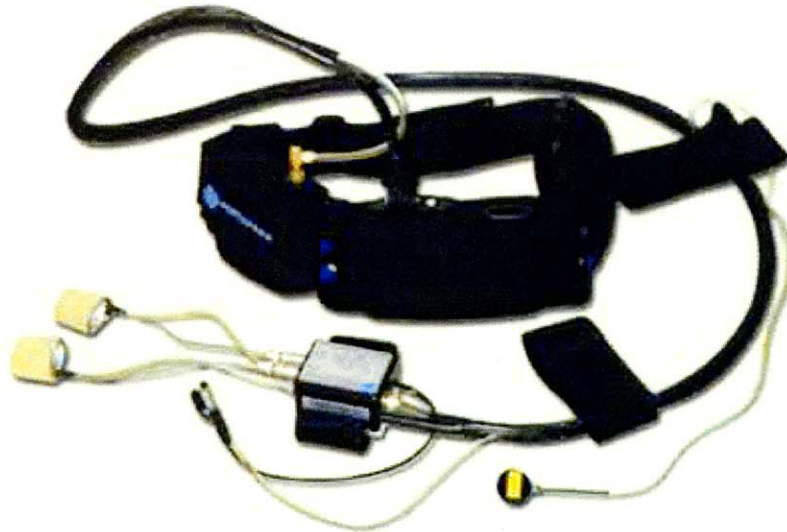
Figure 5.1 Portapres™ noninvasive finger arterial blood pressure monitor [63].

The Portapres™ is fully automatic. A measurement is automatically interrupted in the event of operational error such as disconnecting the air tube to the finger cuff; the measurement is automatically restarted after a waiting period [63]. The data are stored on a built in memory card whose contents can be transferred to a personal computer via a serial link. The transferred data consists of the pressure waveform, the height correction signal and status information. Status information also called an event marker consists of an eight bit binary coded message indicating artifact that may have been detected during the original recording. Beat analysis are not stored on Portapres™. These analyses are acquired with Beatscope™, proprietary software specific to the Portapres™ system [63]. Beatscope™ converts the data files to text files for analysis by other software systems. Additionally, Beatscope™ analysis provides the time of measurement, systolic, diastolic and mean blood pressures, heart rate, interbeat interval, stroke volume, cardiac output, ejection time, total peripheral resistance, status coding and height information. An illustration of Beatscope™ output is shown in Figure 5.2.

Figure 5.2 Output of Beatscope™ analyses [63].

Header information, which is not illustrated in Figure 5.2, indicates among other things, the time of day the measurement began. This output is called a beat file. There is one beat file for every measurement file.

As noted earlier, data were processed in one hour increments; each one hour increment was inspected for artifact. Artifact removal was limited to a maximum of 5% of the data being considered with the goal of less than 1% artifact filtering. If greater than 5% artifact exists in one measurement file it was not considered for analysis. Each subject has an individual log file. The log files are of a general nature and do not provide gross detail of subject activity. The log files were furnished by Columbia researchers. The log files provide a time reference, in military notation, and a general indication of activities. A typical log file is reproduced in Table 5.1.

Table 5.1 Typical Subject Log File Provided by Columbia Researchers Indicating General Activities Over Course of Blood Pressure Recording

Mil. Time 1-24 hrs	ACTIVITIES
22 – 7	Sleep
7 – 8	Wake up, bathroom, shower
8 – 9	Breakfast
9 – 14	Relaxed sitting; studying, reading paper, computer work, phone calls
14 – 15	Lunch
15 – 18	Relaxed; watching TV, reading, computer work, chatting, phone calls
18 - 19	Dinner
19 – 22	Watching TV, reading, computer work, chewing gum

The actual log files for each subject are shown in Appendix C. All log files are closely related as to activity. This is due to the direct supervision of the clinician during data recording to maintain similarity of schedule between subjects. The logs were utilized in conjunction with the various analysis methods to aid in the quantification of results. The hours in Table 5.1 are in military notation.

There are several areas where gaps exist in the data. These include but may not be limited to discontinuities due to transitions between the various measurement files, artifacts and Portapres™ calibrations. During data collection, embedded Portapres™ software performs a system calibration once every 60 minutes as well as prior to the beginning of all measurements and selectively during and after artifacts.

Portapres™ is a portable Finapres™ technology based system that utilizes the same arterial volume clamp measurement method of Finapres™. An infrared light-emitting diode transmits light through the tissue bed of the finger which is picked up by a photo sensor. Based on the amount of light transmitted, the system adjusts finger cuff pressure to maintain a constant blood volume in the finger. Initially, pressure is applied above systolic pressure to occlude blood flow. As cuff pressure is reduced, arterial blood flow causes oscillations in the cuff which are picked up by the photo sensor. Oscillations increase as the cuff pressure is reduced, eventually reaching a maximum before tapering off. The point of maximum oscillation is taken as true mean arterial pressure [64]. Once mean arterial pressure is identified this value is used as a set point for a servomotor which manipulates cuff pressure with the goal of maintaining the amount of light passing through the finger. The pressure changes necessary to maintain the set point mimic the arterial blood pressure waveform providing a continuous non-invasive blood pressure measurement [64].

Because of the hourly calibrations, data can be handled in one hour epochs for processing and analysis. Upon inspection of subject data files, the hourly calibration gaps appear to range between 10 to 30 seconds.

The embedded Portapres™ as well as the Beatscope™ software are proprietary packages, therefore it is impossible to determine pre or post processing of data, i.e. identification of artifacts such as premature ventricular contractions or how the system is equipped to deal with such events. Without prior knowledge of these software algorithms, specifically how Beatscope™ processes artifact the beat file data become largely immaterial other than providing time frames for data recording.

From the data files, systolic blood pressure (SBP) values in millimeters of mercury (mm Hg) and interbeat interval values in milliseconds (ms) were extracted using a custom LabVIEW® software program. This program was originally developed in the Signals Lab at NJIT to detect electrocardiogram R-waves. It has been heavily modified for this research to detect blood pressure values. These data, systolic blood pressure and interbeat intervals, served as the basis of all subsequent analyses. This program is discussed in Chapter 4 and Appendix A.

5.2   Data Preparation

Given the inherent gaps in this data and the fact that Peng et al. [5] recommends a minimum of 8000 data points to calculate detrended fluctuation analysis (DFA), it became necessary to determine the effect on analysis when using data that has been "stitched" together. Calculation of approximate entropy (ApEn) does not pose this problem as Pincus et al. [13] has stated that as few as 1000 data points are adequate to obtain statistically robust results.

Several research studies were identified utilizing detrended fluctuation as an analysis tool that have done so successfully with much smaller data sets [15, 65, 66]. The data sizes were 400 to 1000 points [15], 1000 points [65] and 500 points [66]. Attempts to contact the primary author of [65] led to correspondence with a colleague Dr. Peter P. Domitrovich who verified the use of 1000 data points in their research and noted that Peng has said himself that this was valid. Moreover, Dr. Domitrovich stated that their research uses only the normal-to-normal intervals skipping all time series intervals that do not fit this predefined criteria, thus the data are stitched together.

In an effort to further substantiate the use of data sets that have been "stitched" together, thus eliminating gaps, one of the larger continuous data sets (5603 points) from this cohort was selected for experimental analysis. These data were systematically analyzed with an increasing number of data segments removed. The data removed ranged from 9% to 47% of the total data set. Following data removal, the data were rejoined with no nearest neighbor averaging to smooth the discontinuity. Detrended fluctuation analysis was performed on the full data set as well as on all adjusted data sets. All results were compared to those from the full data set with Pearson correlation. In all cases, the correlation was 99% regardless of the amount of data removed. These results are consistent with previous research [67]. Figure 5.4 illustrates the effects of scaling behavior of detrended fluctuation analysis due to gradually increased data removal.

It is clear from Figure 5.4 that scaling behavior of detrended fluctuation analysis (DFA) is consistent regardless of the amount of data removed. Near identical results were identified in research by Chen et al. [67]. These results and those reported by Chen et al. are taken as clear evidence that combining a data set is a valid approach for the calculation of detrended fluctuation analysis. Detrended fluctuation analysis is a robust measure that identifies the fractal component, if it exists, in the data. Even when progressively larger gaps were carved out the data, DFA continued to identify a scaling exponent that was well correlated with the scaling exponent from the intact data.

Log-Log Fn vs. n
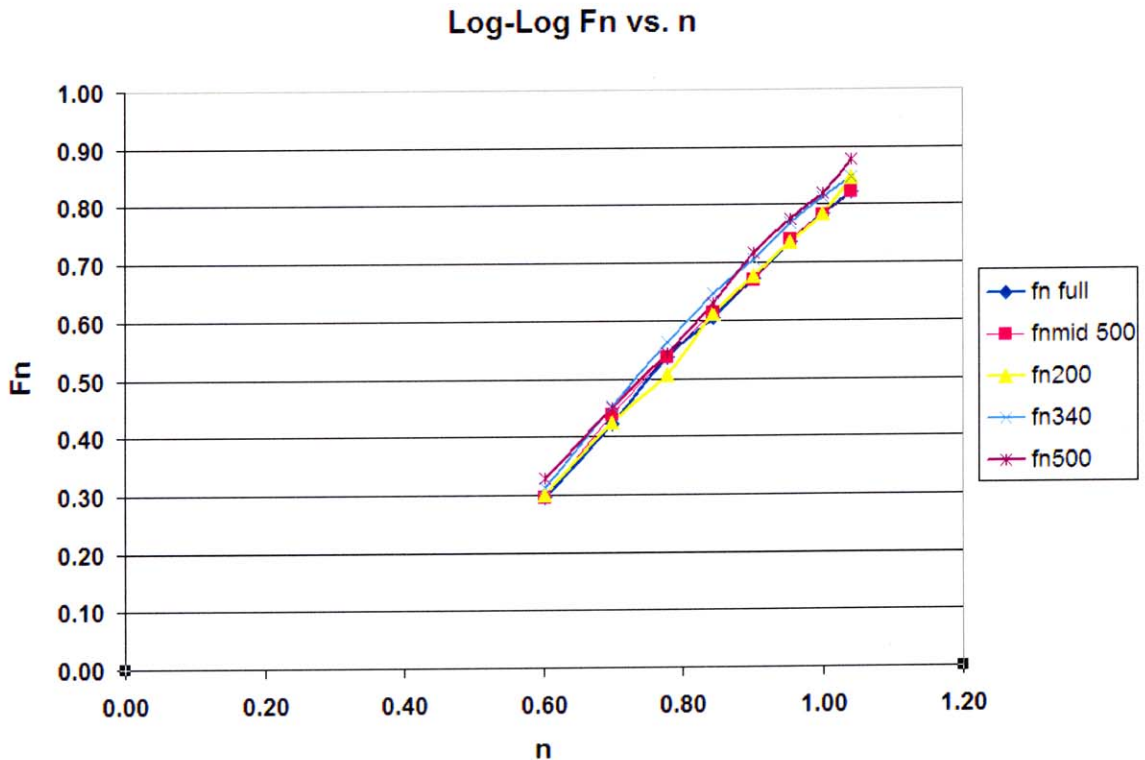
Figure 5.3 Log-log plot of Fn versus box size "n" illustrating the effect of the scaling behavior of detrended fluctuation analysis on the full data set (fn full) versus the same data with gradually larger gaps introduced.   The middle 500 data points were removed in fnmid 500, in fn200, fn340 and fn500 the number of data points removed (200, 340, 500) were removed from each block of 1000 data points.

The results for approximate entropy (ApEn) indicate changes occur when gaps in the data increase beyond 200 points removed per thousand data points. Figure 5.5 is a plot of approximate entropy values for the full data set versus the same data set with larger gaps introduced.   This is the same data that was used to test detrended fluctuation analysis. There has been no research identified that performed similar tests on approximate entropy. Based on these original results, it is clear that as gaps in the data increased beyond 200 points removed per thousand or > 18%, ApEn values diverge from

that of the full data set. These results were not surprising, as approximate entropy is a calculation of the regularity of a data set. In the context tested, ApEn is capable of identifying regularity similar to the full data results even when rather large sections of data are removed (18%) and the data are rejoined with no nearest neighbor averaging. Although ApEn is not as robust as DFA, gaps are well tolerated. These findings are evidence that splicing data is valid for the analysis of approximate entropy. However, care must be taken regarding the amount of missing data and appropriate judgment made when applying approximately entropy to data that has been stitched together.

ApEn Full Data vs. Data with Gaps
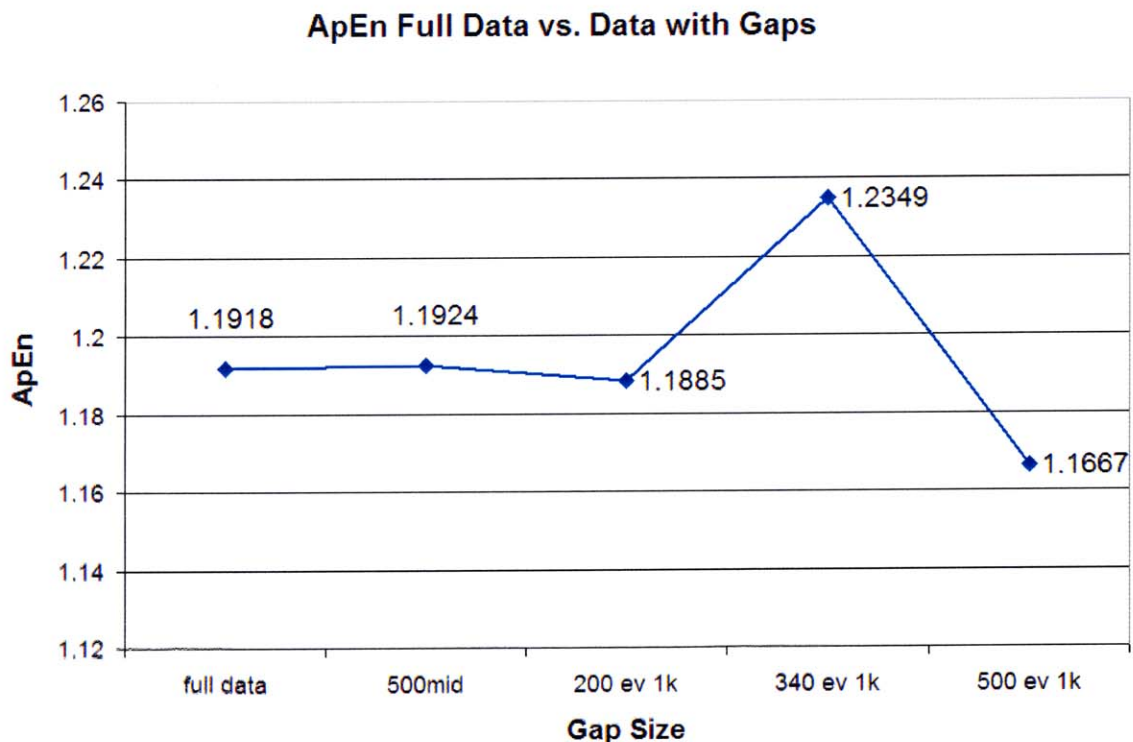
Figure 5.4 Approximate entropy (ApEn) values for full data set versus the same data set with larger gaps introduced. ApEn is relatively unaffected by gaps as large as 200 points removed from every 1000 data points.

Validating the use of data that has been combined due to gaps was necessary in the event data were combined for analysis and to verify the robustness of both measures. Detrended fluctuation analysis (DFA) and approximate entropy (ApEn) were calculated hourly. The results for ApEn were normalized to the largest number of data points for that hour. There was no normalization of DFA given the demonstrated robustness of this measure.

Poincaré analysis consisted of calculating SD1 and SD2 from hourly data. These results were normalized consistent with approximate entropy results. The parameters SD1 and SD2 are widely accepted as determinants of short-term and long-term data dispersion, respectively [46]. Research by Huikuri et al. in which they utilized Poincaré plots, reported that in experiments in which different numbers of R-to-R intervals were deleted at random, the quantitative analysis, SD1 and SD2, remained stable [48]. They note that the error was less than 5% providing there was less than 20% of the data deleted. These results are similar to those that were found when testing the affect of gaps in approximate entropy analysis. In addition to quantitative analysis Huikuri et al. analyzed the Poincaré plots by visual inspection classifying them by shape. In their research, they describe four distinct patterns or shapes which they used in conjunction with their quantitative analysis. Visual inspection of Poincaré plots for this research was noted. However, distinctly different patterns were not expected due to the close physiologic state of the cohort. One Poincaré map was generated for each subject for visual inspection and quantitative analysis. These maps are organized by group for visual comparison and representation. Quantitative measures were calculated hourly and normalized consistent with that of approximate entropy.

In order to assess baroreflex sensitivity, analysis was performed using the spectral technique. Current literature suggests analysis over a range of 128 to 1024 beats. For this research, alpha indices were calculated in blocks of 512 beats. Results were averaged to obtain hourly results and normalized to 60-minutes. Procedures outlined in Section 3.4 were followed.

The length of data analyzed for blood pressure and heart rate variability were blocks of 512 data points which allows the presumption of stationarity which is consistent with current and previous research. Blood pressure and heart rate variability were discussed in Sections 3.5 and 3.6. The criteria outlined in those sections were followed. The low frequency (LF) and high frequency (HF) components of blood pressure variability, and the reported mechanisms responsible for each were discussed in Section 3.6. Each frequency band was analyzed. In addition, analysis also included LF/HF ratio between subjects. All BPV and HRV parameters are shown in normalized units consistent with Malik et al. [62]. Should data handling for any of the analyses change, i.e. block size, changes will be addressed in discussion and conclusions.

Activity plots were generated for all variables for each group. Results were averaged within groups to obtain hourly results. All plots are shown as the mean +/- the standard deviation. Representation of the data in this manner is unique in that the use of activity plots for the nonlinear results have not been identified in previous literature.

5.3 Analysis Overview

Processing the raw data files through the custom LabVIEW® program was the first step to acquire systolic blood pressure and pulse interval data for further analysis.   A systematic approach was followed for each measurement file to acquire appropriate values in one hour epochs. All processed data from each measurement file are in individual folders for all subjects including start and end times.   In order to capture circadian effects, hourly results were utilized.

The size of one hour of data depends on the length of interbeat interval for each subject; however the average hourly result consists of 4000 data points.   The data were analyzed hourly with the results normalized.   For discrete results, approximate entropy and Poincaré parameters, results were normalized to the largest number of data points for that hour.   The continuous analyses, heart rate variability, blood pressure variability, and alpha index were normalized to 60-minutes. Blood pressure and heart rate variability are analyzed in one hour segments by averaging the results of blocks of 512 data points. Baroreflex sensitivity is analyzed in the frequency domain by the alpha index in blocks of 512 data points obtaining hourly results consistent with blood pressure and heart rate variability.

During the course of this research, it was known that within the twelve subjects there were two separate groups, one normotensive the other unmedicated borderline hypertensive. What was not known, was subject grouping, i.e. which subjects belonged to which group.   Thus, in addition to applying linear and nonlinear analysis methods to analyze blood pressure data, another facet of this research was to separate the twelve subjects, a-priori, into two distinct groups based on these results.   After grouping was

determined and properly documented, Columbia researchers revealed the correct subject grouping. Comparisons were then made between the grouping provided by Columbia versus the grouping determined by the linear and nonlinear analyses of this thesis. Because these two groupings did not completely agree, a-posteriori analysis was performed to determine if an optimal set of variables exist to correctly group this cohort.

Once two groups are established, t-tests were used to determine statistical significance. The null hypothesis (Ho) was that there is no difference between the mean of each group. The alternative hypothesis (H_1) was that there is a difference in the mean between groups.

CHAPTER 6

GROUPING METHODOLOGY

6.1   Overview

A large portion of this work was a-priori separation of the cohort based exclusively on the results of linear and nonlinear analyses.   As such, it was necessary to develop a methodology to achieve that goal and then implement it.   This required organization of the data, determination of techniques which were capable of separating a physiologically close cohort and custom programming.

The following sections describe how the data were organized and the configurations available for analysis, followed by discussion of how missing hours of data were handled in terms of analysis and storage.   The two methods chosen to separate the cohort were principal component and cluster analyses.   Each of these methods is discussed along with general implementation of each.

6.2   Data Matrices

Subject data were analyzed as outlined previously.   For every hour there are eight (8) linear and ten (10) nonlinear variables for a total of eighteen (18) variables per subject per hour.   Three 24-hour matrices were compiled for analysis.   The first contains all linear variables (12 x 8 x 24), the second contains all nonlinear variables (12 x 10 x 24) variables, the third combines these matrices forming one (12 x 18 x 24) nonlinear-linear combination matrix.   All matrices are formatted with subjects in rows and variables in columns.   Subject designations are shown in Table 6.1 below.

Table 6.1 Subject Number and Identification

Subject #	1	2	3	4	5	6	7	8	9	10	11	12
Subject ID	a3	a5	a6	a11	b12	m1	m2	m4	n10	o7	s8	z9

Numerous analyses were possible by selection of the different dimensions from within the above matrices. For instance, extraction of all rows over all dimensions for any one column removes a single variable for all subjects over the 24-hour period. Selection of all rows and columns for any one dimension removes one hour for all subjects and all variables. Choosing all columns and dimensions for one row provides an individual profile of one subject for all variables over the 24-hour period. Custom programming was developed to extract the data from any of the three matrices for analysis. Table 6.2 lists variable order in columns for linear and nonlinear matrices. The combined matrix is a concatenation of the nonlinear and linear matrices, respectively, from Table 6.2 resulting in eighteen (18) columns. The number of rows in all matrices is twelve, equal to the number of subjects as shown in Table 6.1. Columns contain variables, and the dimension of every matrix is 24 corresponding to 24-hr recording. Columns one through 10 of the combined matrix contains nonlinear variables as listed in Table 6.2, followed by the linear variables in columns 11 through 18.

Table 6.2 Matrix Configuration for Linear and Nonlinear Variables

Linear Matrix		Nonlinear Matrix	
Column	Variable	Column	Variable
1	Alpha index LF	1	Alpha1 SBP
2	Alpha index HF	2	Alpha2 SBP
3	HRV LF	3	Alpha1 PI
4	HRV HF	4	Alpha2 PI
5	HRV LF/HF	5	ApEn SBP
6	BPV LF	6	ApEn PI
7	BPV HF	7	SD1 SBP
8	BPV LF/HF	8	SD2 SBP
		9	SD1 PI
		10	SD2 PI

As often occurs with 24-hour data recording, there were gaps that led to missing hour(s) of data for some subjects. Subjects with missing hours were represented in the above matrices with zeros for all variables for that hour. As clustering and principal component analyses are sensitive to extended strings of zeros, which exist in the case of hourly analysis, it was necessary to remove any strings of zeros as well as to adjust subject identification and the data matrix prior to analysis. Custom programming was developed in MatLAB® to remove strings of zeros and ensure correct subject identification.

Reiterating from above, the dimensions available for analysis are, by variable over the course of 24-hours for all subjects, by hour for all variables and subjects, and by individual subject for all variables over the course of 24-hours. Each of these configurations required slightly different code to gather data, search for and remove strings of zeros, adjust the analysis matrix and store the results in a multi-dimensional matrix. As such, separate programming was developed to address the needs of each dimensional analysis. There are a total of six programs, three for cluster and three for principal component analysis. Each program performs specific data gathering and analysis, the results of which were saved in MatLAB® ".mat" files for further analysis. Saving results in this manner facilitated the task of subject separation and comparison of results. Cluster and principal component analysis methods are discussed in following sections.

6.3   Cluster Analysis

Cluster analysis is a method used to group objects into similar types or respective categories. It is an exploratory data analysis tool that aims to sort objects into groups such that there is a maximum degree of similarity between objects belonging to the same group and a minimal association to objects in other groups. What cluster analysis can do is aid in uncovering structure or structures in data. What it can not do is provide an explanation as to why the structure or structures exist. There are several clustering algorithms; however for this work MatLAB® kmeans clustering was utilized. The kmeans algorithm was accompanied by custom programming to obtain, format and store the data.

Part of this thesis work was to separate subjects into two groups a-priori based on the results of linear and nonlinear analyses. It was noted earlier that this cohort consists of normotensive and unmedicated borderline hypertensive subjects. However, subject grouping is unknown. It was hypothesized that through the use of linear and nonlinear analysis of data from this cohort, it would be possible to separate these subjects (a-priori) into two separate groups. When clustering by kmeans, the number of clusters must be identified, as it is known that two groups do exist, kmeans clustering is the logical clustering algorithm to utilize.

Data are formatted as noted in Section 6.2, prior to cluster analysis data were standardized by subtracting the mean and dividing by the standard deviation. As a result, all data have a mean of zero and a standard deviation of one. The program (kmeans) begins with k random clusters, in this case two, and iteratively moves data between clusters with the goal of minimizing variance within while maximizing variance between clusters. This iterative process ceases when continued shuffling of the data results in increased variance within clusters. Although there are several distance measures available, squared Euclidian distance, in which the centroid is the mean of the points in that cluster is used here [68].

The silhouette command is a feature of kmeans clustering which if used, can provide one or two outputs. One output returns only the silhouette values in a single column vector; the other returns both the silhouette value vector and then plots the results. The silhouette value vector is a measure of how similar each point in a cluster is to points in its own cluster as compared to points in another cluster. These values range from negative one (-1) to positive one (+1) [68]. A value close to positive one indicates

strong separation. As the value approaches zero, it is an indication that the separation strength is weakened. When a value is negative it is an indication that the algorithm, through all its iterations, is unsure of where to assign this particular point, thus it is possible that the object may be assigned to the wrong cluster. In this work, the silhouette value vector was used with the values stored for further analysis. Based on the above, it is apparent that cluster analysis is simply a guide to grouping variables. Strength was not always sufficient when utilizing cluster analysis alone therefore, subject separation required additional methods which is why principal component analysis was used.

It was difficult to determine an acceptable level of strength of separation for cluster analysis for this cohort. They are very closely related physiologically, and although it was hypothesized that separation is possible from results of linear and nonlinear analysis, the strength of separation was unknown. Therefore, the strongest clustering ultimately determined group separation. Once the two groups were determined, support for that grouping was assessed from all grouping results.

6.4   Principal Component Analysis

Principal component analysis (PCA) is a multivariate analysis technique which is generally used to simplify or reduce the dimension of a dataset. It is a linear transformation that selects a new coordinate system for the data set such that, the greatest variance by any projection of the data lies on the first axis or first principal component. The second largest variance lies on the projection of the second principal component and so on. This linear transformation permits for reduction of a data set while retaining the

important characteristics of the data that most contribute to its variance, i.e. information loss is minimized while reducing the dimension of the data [68, 69].

Principal component analysis describes the variation of multivariate data in terms of a set of uncorrelated or orthogonal projections each of which is a linear combination of the original variables [68, 69]. This work has used MatLAB® principal component analysis algorithm princomp which places the principal components (PC's) in decreasing order of importance. This algorithm is accompanied by custom programming to retrieve, format and store the data. Generally, the first few principal components are all that are necessary to summarize the data while minimizing loss of information; as a result the dimension of the data can be reduced, in some cases substantially reduced [69]. The strength of PCA is that it maintains the underlying structure of the data with fewer components. If cluster analysis alone does not provide strong separation of subjects, then in theory, application of principal component analysis prior to clustering reduces the data while preserving the majority of its structure. This data reduction allows cluster analysis to better discriminate between subjects and improve separation strength. This was found to be true in this work. As with cluster analysis, principal component analysis requires data standardization. MatLAB® suggests dividing the data by its standard deviation [68]. In this work, data are standardized by subtracting the mean and dividing by the standard deviation, in MatLAB® this is called zscore.

The outputs that were used from the princomp command were the coefficient matrix, the transformed data, and the variance. The coefficient matrix contains the principal component coefficients. These coefficients represent linear combinations of the original variables that are used to generate new variables. They are the eigenvectors of

the correlation matrix of the original data. This matrix is ordered such that the largest principal components are in descending order of importance by column. The largest values in any one column identify the prominent loadings or factors of that principal component. The transformed data, called Scores in MatLab®, represent the new variables of the reduced data, which are produced by the coefficient matrix. The values in the variance output are the eigenvalues of the coefficient matrix which reflect the variance of the respective principal component. In order to express each variance value as a percentage of the total variance, each variance value is multiplied by 100 and then divided by the sum of all the variances. This calculation aids in the decision as to the number of principal components to maintain for further analysis.

In this work, the number of principal components (PC's) used for clustering was the number of PC's that were necessary to improve strength of separation and discrimination between subjects. Therefore, the fewest principal components that improved cluster strength and adequately clustered the data into two distinct groups were used. The criterion used for group membership was the identification of the strongest, most discriminating cluster results that separated the cohort into two distinct and equally sized groups.

CHAPTER 7

SUBJECT SEPARATION

7.1 Overview

The development of a grouping methodology was followed by implementation and evaluation of the results. In the sections that follow, an evaluation of the results from each data configuration and the determination of groupings are discussed. The effect of zeros on principal component and cluster analysis is described followed by the approach to subject separation. Subject grouping and how it was accomplished is detailed. The configurations and the specific variables that eventually led to subject grouping are clarified.

A comparison between subject groupings determined by this analysis and the correct grouping revealed by Columbia researchers is then presented. As subject grouping determined by these analysis methods used was not 100% correct, a-posteriori analysis was explored to determine if there exists an optimal set of variables that would improve the accuracy of subject grouping. How variable groupings were determined is explained in detail as are the general results from a-posteriori analysis. The utility of each a-posteriori analysis was assessed in the same fashion as the original data configurations to determine its efficacy.

7.2   Grouping Methods

Results were formatted as discussed in Section 6.2 with subjects and variable listings

shown in Tables 6.1 and 6.2, respectively. Initial grouping for analysis is performed by

programmatically acquiring various data configurations as shown in Table 7.1.

Table 7.1 Analysis Configurations Available From Matrices

#	Configuration	Data
1	B = x(a,:,:)	By subject for all variables all hours
2	B = x(:,a,:)	All subjects for one variable all hours
3	B = x(:,:,a)	All subjects for all variables by hour

Cluster and principal component analyses were performed on all of the

configurations in Table 7.1.   As configurations one and three from Table 7.1 are subject

and hour specific respectively, each contain strings of zeros due to missing hours of data.

Initially analysis was performed with and without removing zeros. This was done more

as an exercise to definitively conclude what was intuitively known, strings of zeros

corrupt analysis results therefore, they require removal prior to both cluster and principal

component analyses.   If the missing hours, represented as strings of zeros, are not

removed too much weight is given to that subject/variable resulting in skewed clustering

and incorrect principal components. Of the twelve subjects, eight had gaps of at least one

full hour. The worst case of missing data was subject "a6" (see Table 6.1) who, for

various reasons, had five one hour gaps through the course of the 24-hour recording.

The approach to subject separation was to perform both cluster and principal

component analysis on all configurations from the three matrices described in Section 6.2

and listed in Table 7.1. This was done in order to determine the best cluster results to resolve group separation. Initially all configurations were clustered to determine if cluster analysis alone provided sufficient strength of separation and adequate discrimination between groups. In general, cluster strength varied widely within and between the various configurations of Table 7.1.

In order to improve cluster discrimination and strength, each data configuration was analyzed with principal component analysis. As described in Chapter 6, Section 6.4, principal component analysis reduces data dimensionality while maintaining its structure. Preprocessing the data with principal component analysis, followed by clustering improved cluster discrimination and strength however, groups were not evenly split for the vast majority of configurations. Following analysis, results were written to file for evaluation of each cluster configuration for group determinations. Columbia researchers were notified of the two groups determined by this analysis following final and full documentation of same. Once Columbia received this information, then and only then, did they reveal the correct subject grouping. The group results from this analysis and those revealed by Columbia are discussed in the following section.

7.3   Subject Grouping

Once all configurations from Table 7.1 were analyzed and all results written to file, each were evaluated separately. Analysis of configuration one in Table 7.1 took all variables one subject at a time and clustered the variables for that subject over the 24-hour period. Each matrix was analyzed. Principal component and cluster analysis were performed to determine how the variables clustered for each individual subject. It was believed that

there may be a similarity between subjects as to how the variables clustered, this proved false. Regardless of the number of clusters used or the blocks of time for which principal component and cluster analysis were applied, this configuration provided no information toward determination of groups.

Evaluation of the results from configuration two in Table 7.1 were in stark contrast to those of configuration one. This was an analysis for all subjects over the 24-hour period one variable at a time. Thus, the clustering was in terms of the subjects by variable over the full 24-hour recording period. There were eighteen clusters to evaluate from this analysis. Because of the structure of the configuration, it was only necessary to perform this analysis on the combined matrix to capture all variables. Performing the analysis on all three matrices would provide redundant information. Analysis of this configuration proved to be the most valuable of all other analyses because it alone, was responsible for determination of group membership.

The two variables, from configuration two, that lead to the determination of subject grouping were detrended fluctuation analysis (DFA) scaling exponent α_2 and SD2 derived from Poincaré mapping, α_2-SBP and SD2-SBP, respectively were analyses of 24-hour systolic blood pressure data over all subjects. The variable that provided the strongest separation was α_2-SBP; while this variable exhibited strength of separation the two groups were uneven, with seven in one group and five in the other. The variable SD2-SBP separated the subjects with near identical groupings and comparable strength with the exception that subject division for this variable was even with six subjects in each group. Cluster results for α_2-SBP and SD2-SBP are shown in Appendix D.

The only information revealed by Columbia researchers regarding this cohort was that it was comprised of two groups of six, one normotensive the other unmedicated borderline hypertensive. Therefore rather than select the grouping of α_2-SBP which consisted of a seven/five split, clustering from SD2-SBP, with almost identical clustering and similar strength, was selected as the final determination following exhaustive data analysis to determine the support for this grouping. Groups as determined by variable analysis are listed in Table 7.2.

Table 7.2 Group Identifications as Determined by Variable Analysis of All Subjects Over 24-hour Systolic Blood Pressure Recordings

GROUP 1		GROUP 2	
Subject #	Subject ID	Subject #	Subject ID
1	a3	3	a6
2	a5	6	m1
4	a11	7	m2
5	b12	8	m4
10	o7	9	n10
11	s8	12	z9

Evaluation of analysis results from the third configuration in Table 7.1, while not discriminate enough to definitely point to two separate groups, provided support for the grouping indicated in Table 7.2. The third configuration was analysis over all subjects and all variables by hour therefore; there were twenty-four clusters from this analysis to evaluate.

There were a total of eighteen variables analyzed for all subjects over the 24-hour period, two of these variables, α_2-SBP and SD2-SBP, determined subject grouping. Of the sixteen variables remaining, fourteen exhibited support of the groups shown in

Table 7.2. Support is quantified as agreement in clustering of four or more of the subjects from either group listed in Table 7.2, where satisfactory cluster discrimination between groups is established. Satisfactory cluster discrimination is defined as separation of subjects into two clusters such that, no one cluster contains four or more members of both groups in the same cluster, i.e. given two clusters sized eight and four where the cluster of eight contains four subjects from Group 1 and four subjects from Group 2 is unsatisfactory cluster discrimination and is not counted in support of established grouping.

Additional evidence supporting the groups listed in Table 7.2 was identified with hourly analysis of the data, configuration three in Table 7.1. As with analysis by variable it was found that analyzing hourly data by PCA prior to clustering provided improved strength of separation and better discrimination between subjects. Hourly analysis of the linear results indicated support in 20 of the 24 hours for the grouping. Hourly analysis of the nonlinear results indicated 15 of the 24 hours in support of the variable grouping. Finally, hourly analysis of the combination matrix containing both nonlinear and linear results indicated 12 of 24 hours in support of the identified groupings. Hours with inadequate cluster discrimination as defined above were discounted. The one analysis that provided no insight into subject grouping was subject profiling or configuration one in Table 7.1. Due to the lack of information provided by this configuration it was not considered when determining support for groupings listed in Table 7.2. The grouping revealed by Columbia researchers is shown in Table 7.3.

Table 7.3 Subject Grouping as Revealed by Columbia Researchers

COLUMBIA GROUP 1		COLUMBIA GROUP 2	
Subject #	Subject ID	Subject #	Subject ID
4	a11	1	a3
5	b12	2	a5
9	n10	3	a6
10	o7	6	m1
11	s8	7	m2
12	z9	8	m4

The group determinations shown in Table 7.2 were reported to Columbia researchers. Upon receipt of those group designations Columbia researchers revealed the correct group results which are shown in Table 7.3. Group 1 is reported as normotensive with Group 2 the unmedicated borderline hypertensive subjects. Linear and nonlinear analyses correctly discriminated the subject cohort with 67% accuracy, or four of the six subjects in each group were correct. The same definitions of group support and cluster discrimination as defined previously were used to identify linear and nonlinear analyses that supported the Columbia grouping indicated in Table 7.3. The two variables that were instrumental in the group designations of Table 7.2 obviously support the grouping in Table 7.3 therefore they were not be counted (as they were not previously) in support of the Columbia grouping. Of the sixteen remaining analyses by variable, 15 support the Columbia grouping, one more than support the grouping determination of Table 7.2. Hourly analysis of linear, nonlinear and combined matrices supports the Columbia grouping in 15, 17 and 13 hours, respectively out of the 24-hour recording. Group comparisons along with support by hour and variable analysis have been summarized in Table 7.4.

Table 7.4 Comparison Between Grouping Determined by Cluster Analysis and The Correct Grouping Provided by Columbia Researchers

GROUP COMPARISONS					
GROUP1					
COLUMBIA	4	5	9	10	11	12
ANALYSIS	1	2	4	5	10	11
GROUP2					
COLUMBIA	1	2	3	6	7	8
ANALYSIS	3	6	7	8	9	12
	COLUMBIA		ANALYSIS			
SUPPORT	HRS	VARS	HRS	VARS		
LINEAR	15	7	20	8		
NONLINEAR	17	8	15	6		
COMBINED	13	N/A	12	N/A		

The summary of results shown in Table 7.4 indicates greater support from linear variables both hourly and by variable for the grouping determined by this analysis than for the Columbia groups, where as slightly greater support is seen for the Columbia groups in both the nonlinear and combined matrices. Grouping was an exhaustive process for which there was only one clear choice in determining the two groups based on analysis results. Because the groupings from Columbia and those found here do not match exactly, further analysis was performed to determine if the accuracy of separation could be improved. This was accomplished by permutations of variables. Various

combinations of data were analyzed in an effort to improve upon the accuracy of grouping and to determine if an optimal set of variables exists to correctly separate this cohort. Blocks of data were also considered in an effort to improve upon grouping accuracy. The details of this post-grouping analysis are presented in Section 7.4.

7.4 A Posteriori Analysis

There are thousands of variable combinations given a total of eighteen variables from which to choose; as such, a more efficient method to identify variable groupings was necessary. The coefficient matrix from principal component analysis (PCA), discussed in Chapter 6 Section 6.3, was used to guide variable group selection as were the results from cluster analysis. When analyzing data with principal component analysis, the factor or component loadings in the coefficient matrix are in terms of the horizontal axis of the analyzed data. When analyzing the combined matrix by configurations two and three of Table 7.1, the coefficient matrix is in terms of hours and variable respectively for each analysis. The factor or component loadings of these coefficient matrices were used to identify variable groupings.

The combined matrix, sized 12 x 18 x 24, contains all variables nonlinear and linear as previously described in Chapter 6 with variable listings shown in Table 6.2. This matrix was analyzed by variable (configuration two Table 7.1) with principal component analysis. Configuration two is analysis of all subjects over the course of 24-hours one variable at a time. The dimension of each variable matrix is 12 x 24, which results in one 24 x 24 coefficient matrix for each variable representing the linear combinations of the original variables, in this case hours. The largest coefficients

(component or factor loadings) in each column correspond to the most important or most heavily weighted hours for that variable. Configuration three of Table 7.1 is analysis of all subjects over all variables one hour at a time. The dimension of each variable matrix for this analysis is 12 x 18 resulting in one 18 x18 coefficient matrix for each hour which represents the linear combinations of the original variables, in this case variables. Each of these analyses aided in identifying heavily weighted hours and variables from which to create new variable groupings. New variable groupings were also created based on the original analysis that resulted in determination of subject grouping.

Results of the PCA of the combined matrix by configuration two of Table 7.1 indicated the most heavily loaded factors (hours) were 2 PM, 8 PM and 3 AM occurring with 50, 45 and 41 percent frequency, respectively, across all variables. On individual analysis of those three hours only 3 AM indicated a measure of support (67%) of the two groupings shown in Table 7.4. Although this analysis proved too broad to narrow down variable groupings, three groups were formed which represented the variables whose factor loadings were highest for the hours noted above. When analyzing these three groups only those hours that clustered subjects evenly (6/6) were considered in determining the ability of any one group to resolve the correct grouping as described in Table 7.4. Upon hourly analysis of the three groups no one hour improved upon the 67% success in subject separation.

Principal component analysis of the combined matrix by configuration three of Table 7.1 provided improved information as to new variable groupings. This configuration is an hourly analysis of all subjects over all eighteen variables. The first principal component (PC) or first column of each coefficient matrix was used to identify

the more heavily loaded factors, in this case variables, to aid in determining new group combinations. Generally only the first few principal components are considered for further analysis. Given that the first principal component contains the largest variance of all the principal components, only it was used in identifying new variable groupings. In addition to the new groups identified from the combined matrix, groupings were formed from an identical analysis of the nonlinear matrix. This matrix is described in Chapter 6, with variable listings indicated in Table 6.2. Analysis of these new variable groupings consisted of principal component analysis, which was performed hourly, followed by cluster analysis of the transformed data. Only those clusters that separated subjects evenly were considered. Not one of these new groupings improved upon the 67% success rate in separating subjects.

The combined matrix was then analyzed in blocks of four; eight and twelve hours in an attempt to identify a time frame that may be more successful in subject separation. This analysis was performed by variable as the time frame with which to calculate principal components and determine clusters was now reduced. Analysis of the first five blocks of four hours did not improve beyond the 67% success rate to correctly separate subjects. In the last four hour block, between 8PM and midnight, variable number 10, the Poincaré measure SD2-PI separated the subjects with an 83% success rate, or five of six subjects correctly clustered. Analysis of the combined matrix in eight hour segments revealed two variables that separated the subjects with 83% success, one in the first eight hour block the other in the third eight hour block. Again only those clusters that separated subjects evenly were considered. The first variable to improve upon subject separation (83%) was detrended fluctuation analysis α_2-SBP, variable two in Table 6.2.

This is interesting given that this variable was extremely instrumental during a-priori subject grouping and was not considered (did not evenly separate subjects) during analysis of four hour blocks of the same data. The second variable that improved upon subject separation (83%) was approximate entropy of systolic blood pressure, ApEn-SBP, variable five in Table 6.2. This variable too was not considered in the four hour analyses of the same data. The two twelve hour block analyses did not improve upon the 67% success rate in subject separation.

While over 50 variable groupings were separately analyzed in an effort to identify an optimal set of variable groupings that would improve subject separation, this represents a small portion of the over 5,000 possible variable groupings. As noted earlier, with eighteen (18) variables the number of grouping variations is enormous, as such only the groupings identified by principal component and cluster analysis as the most probable optimal groups were investigated here. The majority of these analyses did not improve upon the 67% accuracy in subject separation as determined during a-priori analysis. When data were segmented into blocks of hours and analyzed, however there was intermittent improvement in grouping. Grouping improved from four of six (67%) to five of six (83%) among various nonlinear results during specific time intervals as noted earlier. Continued research of variable groupings may identify an optimal variable grouping that will consistently improve upon subject separation.

CHAPTER 8

DISCUSSION OF RESULTS

8.1   Overview

There are eighteen variables derived from six analysis methods for two groups of six subjects each, over the course of 24-hour blood pressure recordings. This research was performed in an effort to characterize hypertension through the utilization of three linear and three nonlinear methods of analysis. The three linear methods are heart rate variability (HRV), blood pressure variability (BPV) and the alpha index. The three nonlinear methods are detrended fluctuation analysis (DFA), approximate entropy (ApEn) and Poincaré plots. Data were collected by the researchers at Columbia University. Two of the three nonlinear methods, DFA and Poincaré plots, have not been identified as having being used in any prior research to study blood pressure data in humans. Approximate entropy is the only nonlinear method used here that has been identified in one previous research study of blood pressure in humans. No other research studies have been identified that have used these three nonlinear methods together as applied to blood pressure data in humans. Further no research studies been identified that have used the same nonlinear and linear methods in conjunction to analyze blood pressure data or characterize hypertension in humans.

In the graphs and discussions that follow, the two groups are referred to as the borderline hypertensive (BHT) group and the normotensive (NT) group.   All results, with the exception of overlay plots and sympathovagal ratio (LF/HF) for heart rate variability (HRV) and the LF/HF ratio for blood pressure variability (BPV) are shown hourly as the mean +/- the standard deviation for that group. The LF/HF ratios for both

67

HRV and BPV are calculated from the low frequency and high frequency hourly averages. The x-axis in all of the graphs is labeled one through 24 representing the 24-hour recording period. Sleep hours are marked on each plot.

As noted earlier, subject activity is near identical for all subjects. The common sleep time for all subjects begins at hour 22 and continues through hour seven. At hour eight all subjects wake-up and eat breakfast. Between hours nine and 21 activity is fundamentally the same for all subjects. Individual subject log files are shown in Appendix C for both groups. The first six log files in Appendix C are the borderline hypertensive group. The last six log files are the normotensive group.

8.2 Blood Pressure Variability

Blood pressure control is a complex process that is dependent upon many control systems. Mean arterial blood pressure is the main driving force supplying blood to the tissues in the body. The two determinants of mean arterial pressure are cardiac output and total peripheral resistance. Cardiac output depends on heart rate and stroke volume. Heart rate depends on the balance between sympathetic and parasympathetic activity. Sympathetic activity increases heart rate while parasympathetic activity decreases it. Stroke volume is controlled by sympathetic activity and end-diastolic volume and end-diastolic volume is controlled by venous return which is controlled by sympathetic activity [55].

The autonomic nervous system, a branch of the efferent division of the central nervous system, has two subdivisions. The sympathetic nervous system and the parasympathetic nervous system, both of which innervate smooth muscle, cardiac muscle

and glands [55]. One method of quantifying blood pressure variability is through spectral analysis of the different frequency components which can be derived from blood pressure recordings. The frequencies analyzed are the low frequency (LF) in the range of 0.04Hz to 0.15Hz, the high frequency (HF) ranging from 0.15Hz to 0.4Hz and the LF/HF ratio.

The low frequency range is attributed to activity of both the sympathetic and parasympathetic nervous systems. The high frequency is attributed to parasympathetic activity and the mechanical affects of respiration. Previous research has noted that in subjects with higher than normal blood pressure, the amplitude of their blood pressure variability is greater than normal and increases progressively with increasing blood pressure [4, 9-11].

8.2.1 Low Frequency BPV

Figure 8.1 illustrates the low frequency (LF) component of blood pressure variability (BPV) for both groups. The frequency values are represented in normalized units. Results are shown as the mean +/- the standard deviation over the 24-hour recording period. The low frequency power in the borderline hypertensive subjects (BHT) versus the normotensive subjects (NT) is higher in 17 of the 24-hours. Or 71% of the time the BHT subject group exhibited elevated sympathetic activity over their NT counterparts. Of the remaining seven hours five are lower in the BHT group than in the NT group. These hours are highlighted in Figure 8.2 which is an overlay plot of the low frequency component of blood pressure variability without the standard deviation indicators. Statistically the difference over 24-hours is significant to $p < 0.05$.
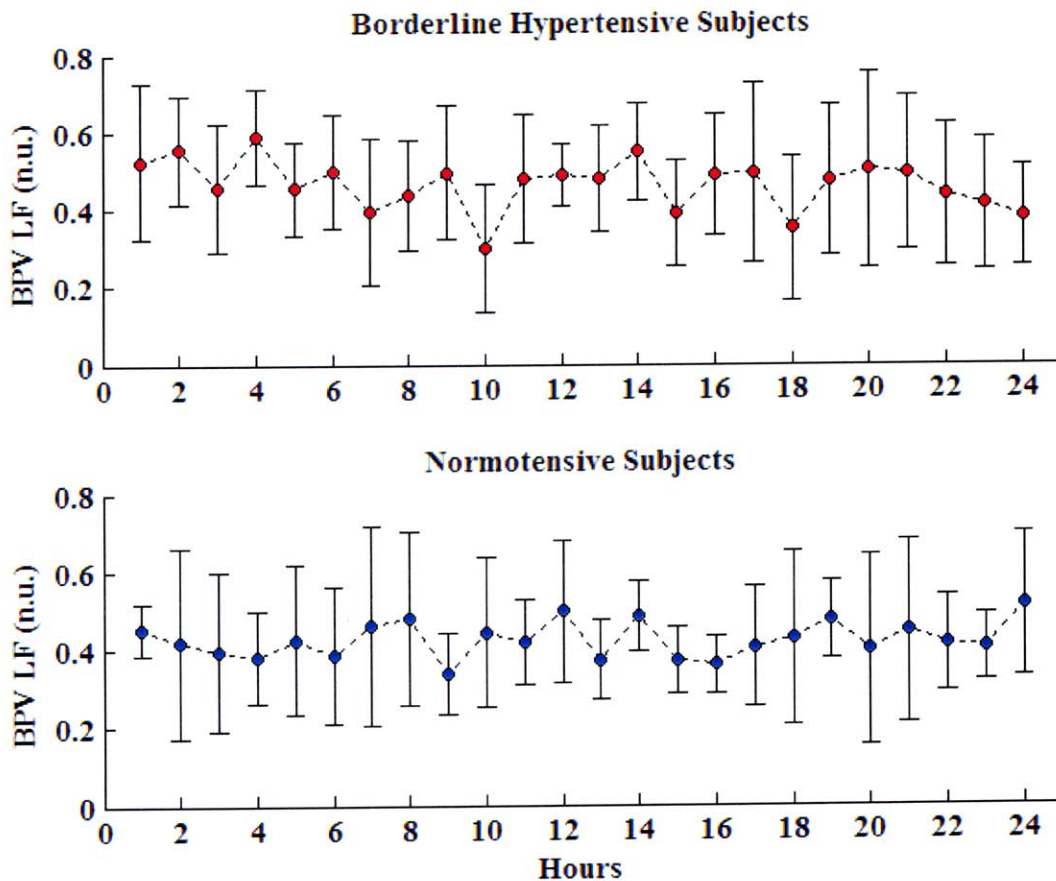
Figure 8.1 The low frequency component of blood pressure variability (BPV) for the two groups. Results are shown as the mean +/- standard deviation. LF is expressed in normalized units. Statistical significance between groups is p < 0.05 using paired t-test.
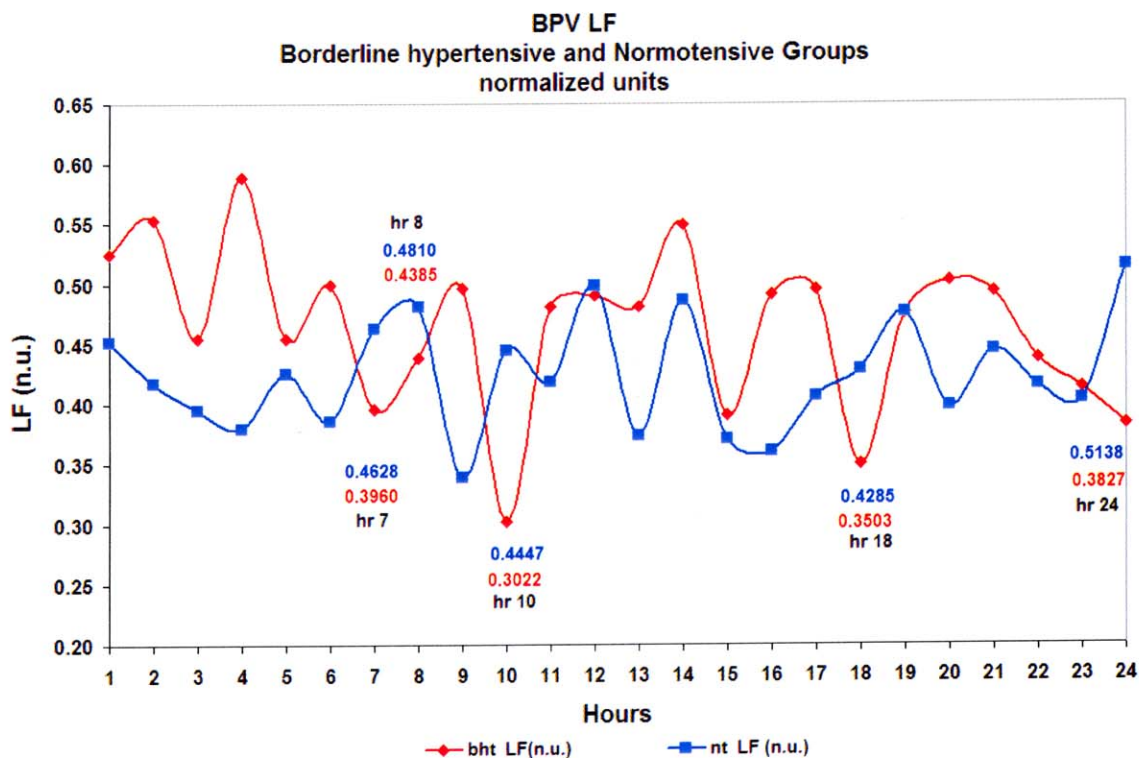
Figure 8.2 Illustration of hours when the low frequency power is lower in the BHT group than in the NT group. Low frequency power is represented in normalized units. This is the same data as Figure 8.1. Difference between groups using paired ttest was statistically significant to $p < 0.05$.

Figure 8.2 is the same information shown in Figure 8.1 without the standard deviation markers. This plot makes it easier to visualize mean group activity and the five hours where low frequency power of the borderline hypertensive group is lower than the power in the normotensive group. Note the change in scale of the y-axis. The oscillations in the low frequency power for the BHT group present higher peak values than in the normotensive group. While the normotensive group oscillates, the competition between sympathetic and parasympathetic activity may be better balanced than in the borderline hypertensive group. This may be become evident when evaluating high frequency behavior and the LF/HF ratio. The same plot as Figure 8.2 was generated

for the high frequency power in order to make valid comparisons between high and low frequencies.

Activity logs prepared during recording were provided by Columbia for each subject. Actual subject logs are shown in Appendix C. As noted earlier subject activity was near identical for all subjects over the course of the 24-hour recording period. As such, sympathetic activity can not be linked to unusual behavior of one subject or group of subjects during the recording period. Therefore, it is believed that the differences in low frequency power shown here are either due to elevated sympathetic activity, diminished parasympathetic outflow, or both in the borderline hypertensive group versus the normotensive group during near identical activity for the recording period. During initial analysis, the LF variable did not contribute to a-priori subject separation, nor did it aid in improving the accuracy of subject separation during permutation of variables. It is, however, not surprising that the differences are statistically significant given previous research that has shown higher sympathetic activity in hypertension [1, 7-11, 70, 71]. It is concluded that the results for low frequency blood pressure variability are consistent with previous research which reported elevated sympathetic activity and blood pressure variability in hypertensive subjects.

8.2.2 High Frequency BPV

Figure 8.3 illustrates the high frequency (HF) power distribution for both groups. In Figure 8.3, the pattern of behavior appears more closely matched between groups in the first twelve hours than in the last twelve when behavior between the two groups becomes more divergent. When the divergent behavior begins, it is the normotensive group that has higher maximum peak values than the borderline hypertensive group. This is

particularly true during hours 16 through 21, which is the only time there is a statistically significant difference between these two groups to $p < 0.05$ in high frequency power. Statistically there is no difference between the two groups for the full 24-hour recording period.
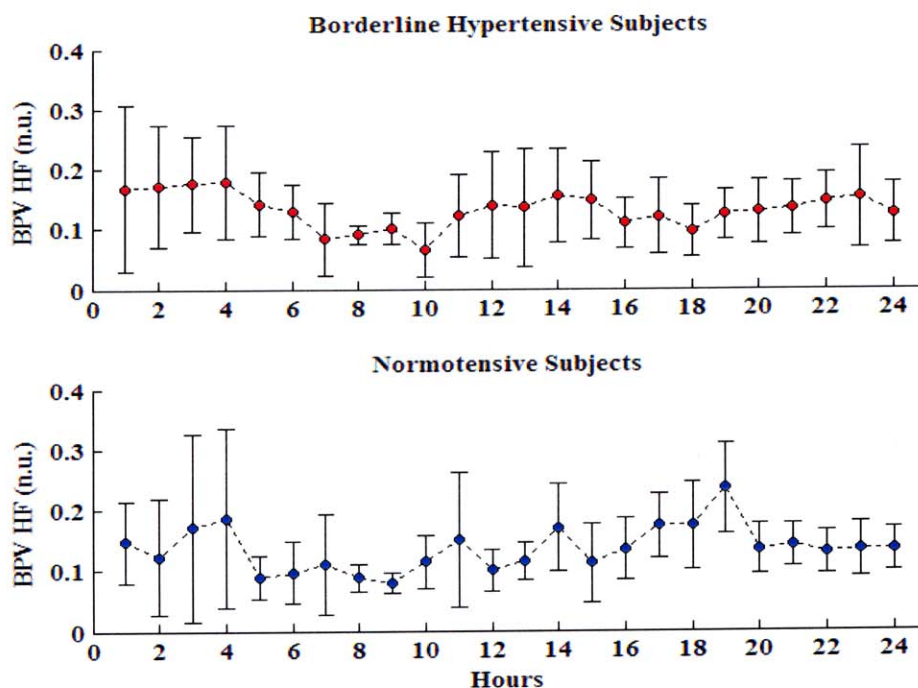
Figure 8.3 The high frequency component of blood pressure variability (BPV) for the two groups. Results are shown as the mean +/- standard deviation. HF is expressed in normalized units. Statistically there is no difference between groups for the full 24-hour period.

The mechanical effects of respiration on the high frequency power are impossible to define for either group as there was no paced breathing, no respiratory recordings and no references in the logs as to breathing patterns. This behavior, in high frequency power between the groups, is in contrast to the low frequency behavior. In the low frequency power, the borderline hypertensive group exhibited elevated sympathetic

outflow as compared to the normotensive group; yet, the parasympathetic response by the normotensive group to a lower magnitude sympathetic outflow is greater than that of the BHT group. The parasympathetic response in borderline hypertension may be too slow, too weak or both to counteract the strong sympathetic outflow.

Five hours were identified in the low frequency range, (Figure 8.2), where the borderline hypertensive group average was lower than the normotensive group average. In order to assess parasympathetic activity during these same hours, an identical plot was created for the high frequency power distribution and is shown in Figure 8.4. The same five hours are highlighted in Figure 8.4 with data labels to indicate the high frequency power at those times.
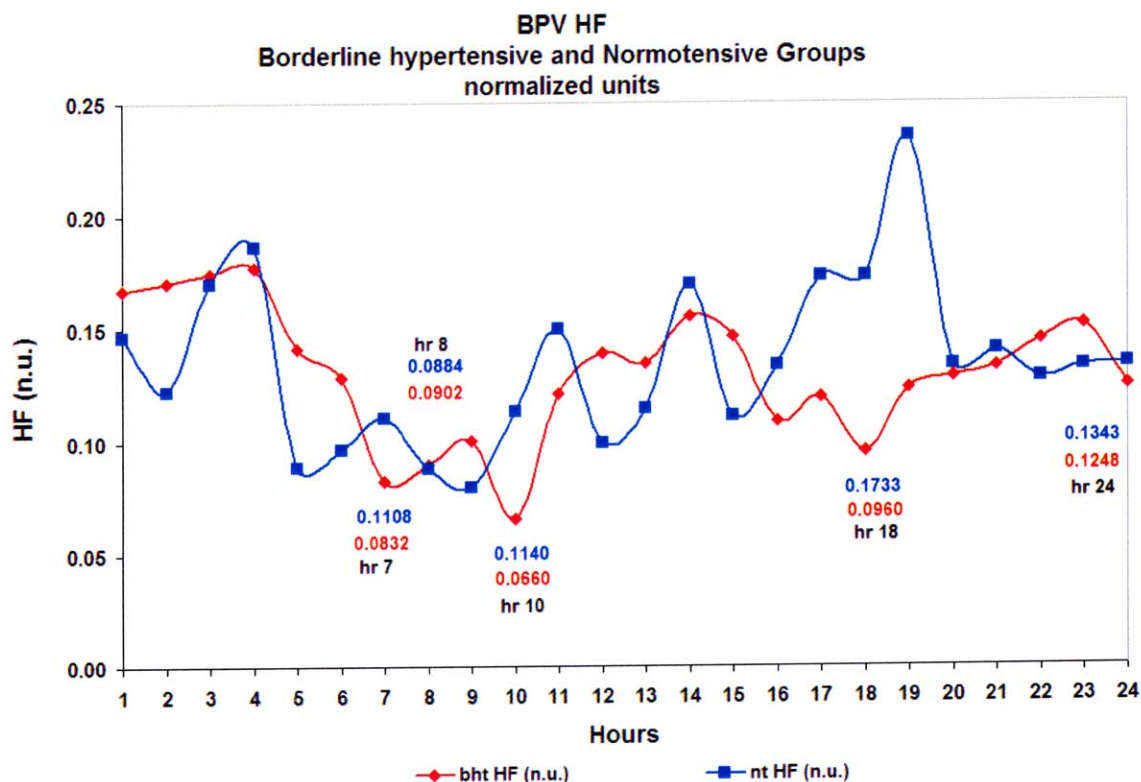
Figure 8.4 Plot of the high frequency power between groups. Data labels are consistent with those of Figure 8.2. High frequency power is shown in normalized units.

The five hours marked with data values are the same five hours marked in Figure 8.2 where the low frequency (LF) power in the borderline hypertensive group was lower than the normotensive group. Sympathetic activity of the borderline hypertensive group in four of the five hours, seven, 10, 18 and 24 decreased, in hour eight sympathetic activity increased. The same group's parasympathetic response paralleled the sympathetic response, decreasing and increasing respectively. This parallel course between LF and HF powers in the borderline hypertensive group is true 75% of the time. In other words there are only six hours where an antagonistic relationship between the two powers exists in the borderline hypertensive group. In the normotensive group parallel association between LF and HF powers occurs 58% of the time, or 14 of the 24 hours.

Figure 8.4 reinforces previous assessment of Figure 8.3 and provides improved visualization of the strong peaks in the normotensive group between the 16th and 21st hours. Figure 8.4 illustrates more clearly what can be interpreted to be a sluggish response by the borderline hypertensive group to elevated sympathetic outflow.

8.2.3 LF/HF Ratio BPV

The LF/HF ratio for blood pressure variability is shown in Figure 8.5. These results are the ratio of the hourly averages shown in Figures 8.1 and 8.3. There are statistically significant results from paired t-tests between groups during hours 13 through 24 to $p < 0.05$. Paired t-tests for the full 24-hour period did not return a statistically significant result.
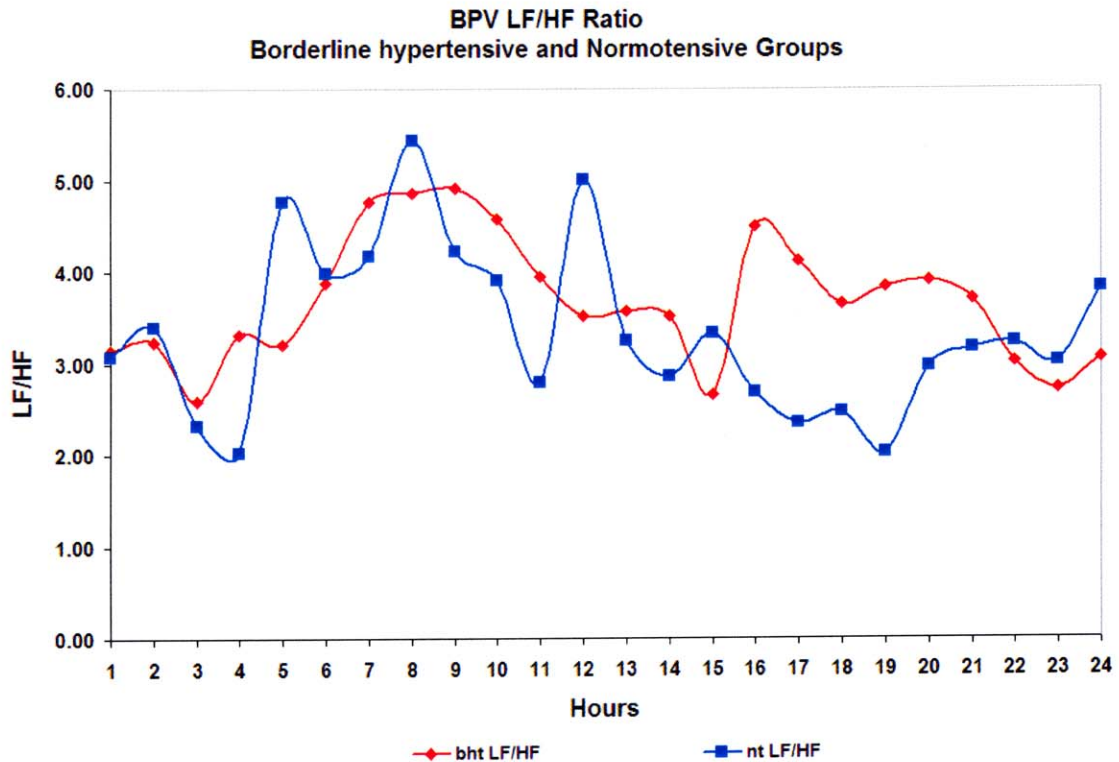
Figure 8.5 LF/HF ratio of BPV for both groups.

The LF/HF ratio provides a window into the reciprocal relationship between sympathetic and parasympathetic activity, the implication being when one is excited the other is inhibited to some degree. In the borderline hypertensive group, the LF/HF ratio appears to be dominated by sympathetic activity between the hours 16 to 21, where in the normotensive group the opposite is true. In the last four hours, the LF/HF ratio for the BHT group is dominated by parasympathetic activity while in the normotensive group the dominance appears to be sympathetic. On average, the borderline hypertensive group has a higher low frequency and lower high frequency component than the normotensive group in blood pressure variability.

8.3 Heart Rate Variability

Heart rate is determined primarily by sympathetic and parasympathetic influences on the sinoatrial (SA) node. The SA node is one of several clusters of autorhythmic cells that spontaneously depolarize exciting the contractile cells of the heart causing it to beat. Because the SA node has the fastest spontaneous rate of depolarization of all other autorhythmic cell clusters, normally it is the primary pacemaker. The heart is innervated by both the sympathetic and parasympathetic nervous systems both of which can modify heart rate. Specifically, the parasympathetic nervous system decreases heart rate while the sympathetic nervous system increases heart rate.

In 1981, Akselrod et al. introduced power spectral analysis of heart rate fluctuations in the adult conscious dog to quantitatively evaluate beat-to-beat cardiovascular control [72]. The clinical importance of heart rate variability (HRV) became apparent in the late 1980s when it was shown that decreased HRV was a strong predictor of mortality following an acute myocardial infarction [62]. Today, frequency analysis of heart rate variability is an extensively used noninvasive technique that is capable of providing information on autonomic function. The low frequency (LF) range, 0.04Hz to 0.15Hz, has been attributed to both sympathetic and parasympathetic influences. The high frequency range, 0.15Hz to 0.4Hz, is reportedly due to parasympathetic influence [62, 71].

One previous research study of heart rate variability in hypertension has reported similar HRV findings between normotensive and borderline hypertensive subjects with decreased variability in severe hypertension [71]. As this study cohort consists of borderline hypertensive and normotensive subjects, it will be interesting to determine if

the findings here coincide with that research study. Data for this study utilizes pulse interval data derived from 24-hour systolic blood pressure recordings. The use of pulse interval data has been shown to be a statistically robust substitute for R-to-R interval data [58].

8.3.1 Low Frequency HRV

Figure 8.6 illustrates the low frequency component of heart rate variability (HRV), in normalized units. Results are shown as the mean +/- the standard deviation.
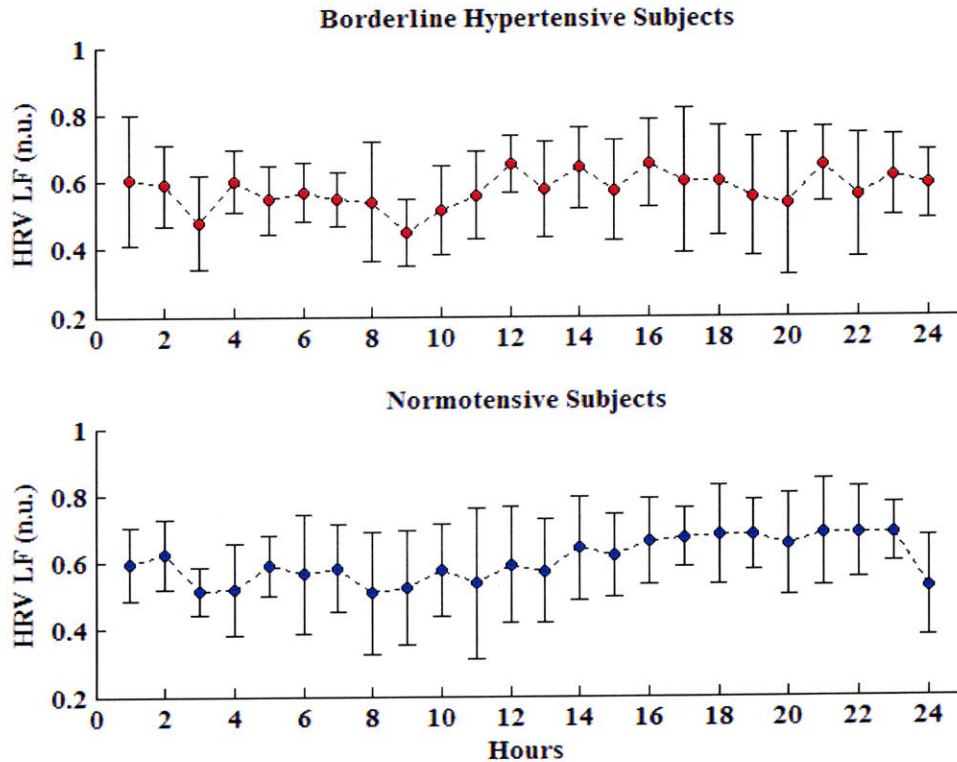
Figure 8.6 The low frequency component of heart rate variability (HRV) for the two groups. Results are shown as mean +/- standard deviation. LF is expressed in normalized units. Group differences are statistically significant to $p < 0.05$.

As is the case in blood pressure variability, the low frequency component for heart rate variability is statistically significant to $p < 0.05$ based on paired t-testing over the full 24-hour recording period. In Figure 8.6, the dispersion of power about the mean in the normotensive group appears more regular than the borderline hypertensive group. Figure 8.7 is an overlay plot of the data from Figure 8.6 without the standard deviation markers. Note that the vertical scale was adjusted to improve visualization of the hourly fluctuations. The borderline hypertensive group has more abrupt changes in power than the normotensive group. This is evident by the changes in the standard deviation markers of Figure 8.6. It is also clear from Figure 8.7, that the LF power in the normotensive group becomes more stable during the end of the day. Both groups retired during the 22nd hour, this explains the abrupt drop in power in the 24th hour.

HRV LF
Borderline hypertensive and Normotensive groups
normalized units
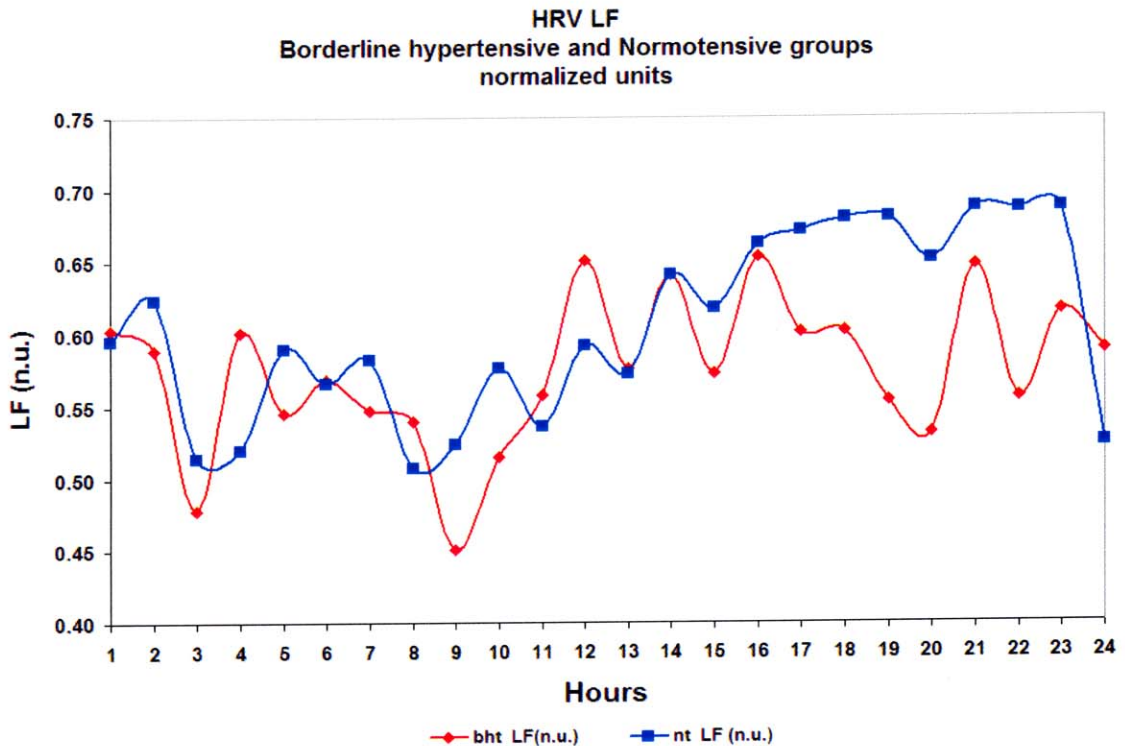
Figure 8.7 The low frequency component of heart rate variability (HRV) for the two groups. Note vertical scale begins at .40 versus zero.

8.3.2 High Frequency HRV

Figure 8.8 is the high frequency component of heart rate variability. Results are shown as the mean +/- the standard deviation and are in normalized units. Results are statistically significant between hours 1 through 18 to $p < 0.05$. This is in contrast to the high frequency power in blood pressure variability where statistically significant results were limited to hours 16 through 21.
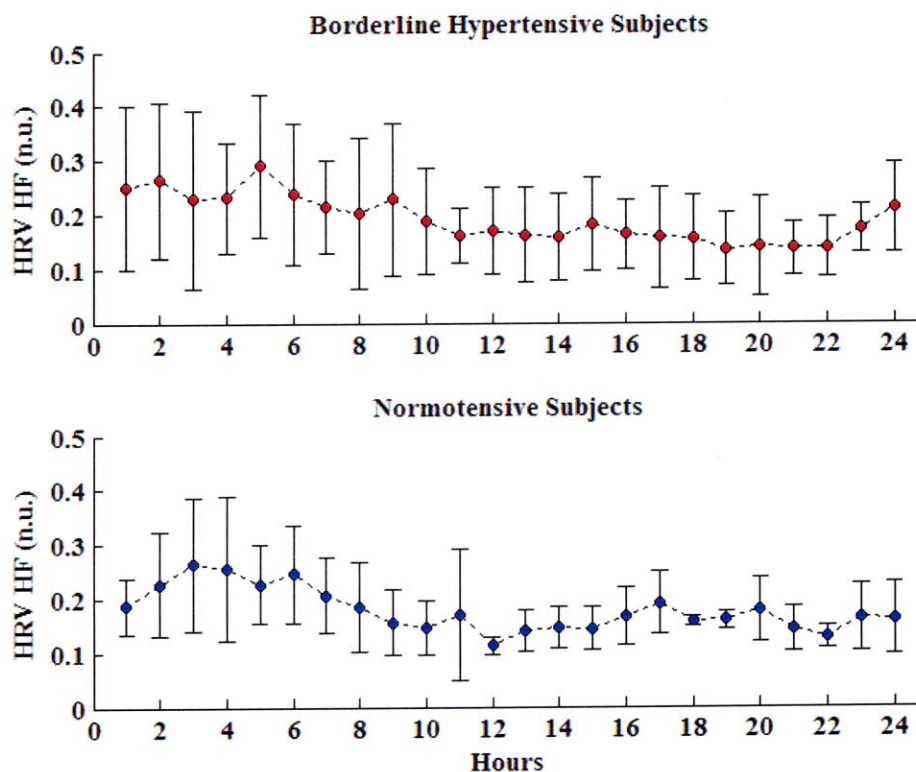
Figure 8.8 The high frequency component of heart rate variability (HRV) for the two groups. Results are shown as mean +/- standard deviation. HF is expressed in normalized units.

In Figure 8.8, it can be seen that the high frequency power in the borderline hypertensive group has a larger dispersion of data around the mean than the normotensive group through out the 24-hour period. This is indicated by the larger standard deviation bars for each hour. As in blood pressure variability it is impossible to assess the affect of respiration on the high frequency power. Low frequency power indicates decreased heart rate variability; the high frequency power in borderline hypertension is elevated over the normotensive group 75% of the time diminishing sympathetic activity. An overlay plot of the high frequency is shown in Figure 8.9.
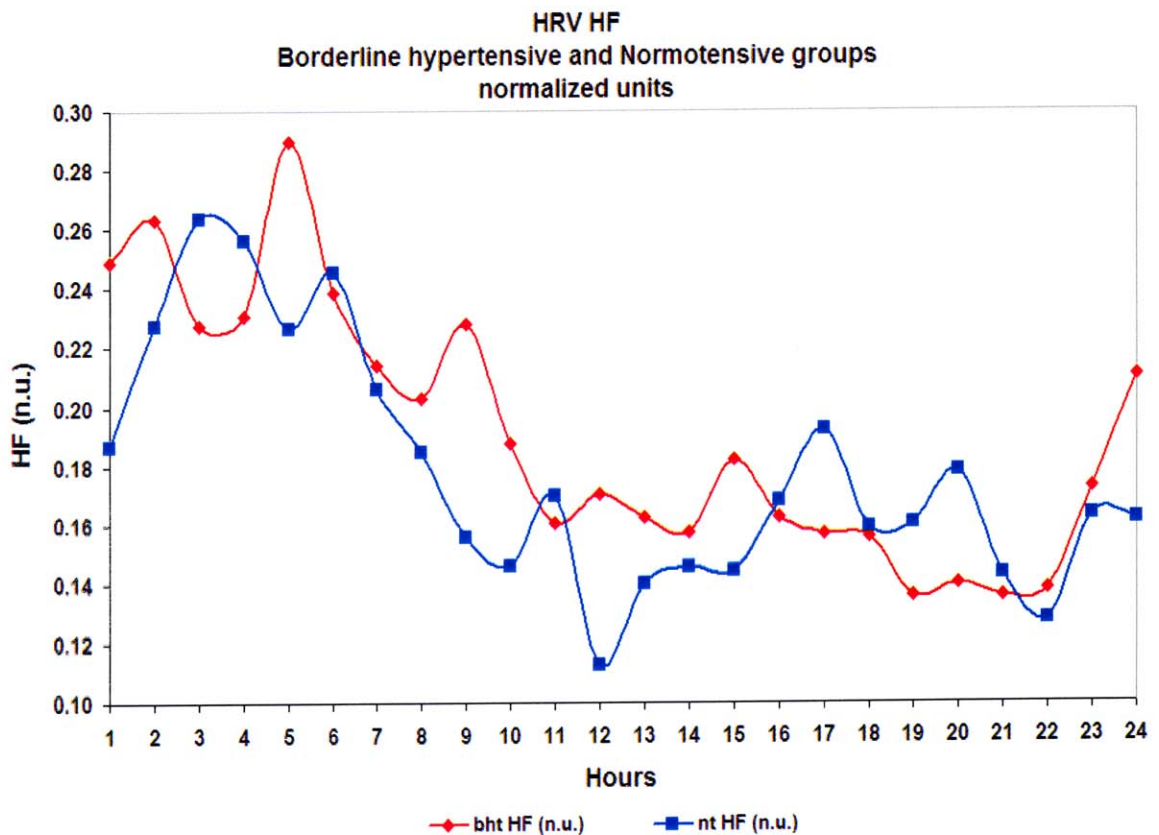
Figure 8.9 Heart rate variability (HRV) high frequency power for both groups. Note the change in vertical axis. Statistically the two groups differ between hours 1 to 18 to $p < 0.05$.

In Figure 8.9, the majority of the time, parasympathetic activity withdraws more rapidly in the normotensive group than in the borderline hypertensive subjects. The sharp drop in HF power between hours 6 and 10 aids in understanding the stable dispersion of power during those same hours in Figure 8.8. Although the overlay plot of Figure 8.9 does not provide additional information, it aids in explaining power dispersions shown in Figure 8.8.

8.3.3 Sympathovagal Ratio HRV

The sympathovagal ratio (LF/HF) for heart rate variability is shown in Figure 8.10. These results are the ratio LF/HF of the hourly averages shown in Figures 8.6 and 8.8. Results are shown as the mean +/- the standard deviation.

The sympathovagal ratio for heart rate variability is statistically significant to $p < 0.05$ between the two groups. This is in contrast to the LF/HF ratio in blood pressure variability where there statistical significance was limited to the hours 13 through 24. The peaks in Figure 8.10 for both borderline hypertensive and normotensive subjects indicate sympathetic dominance. This is based on comparison of low and high frequency power distributions in Figures 8.7 and 8.9. On average, the borderline hypertensive group has a lower low frequency and higher high frequency component than the normotensive group over the 24-hour period. This is in contrast to the LF/HF ratio in blood pressure variability where the exact opposite is true.
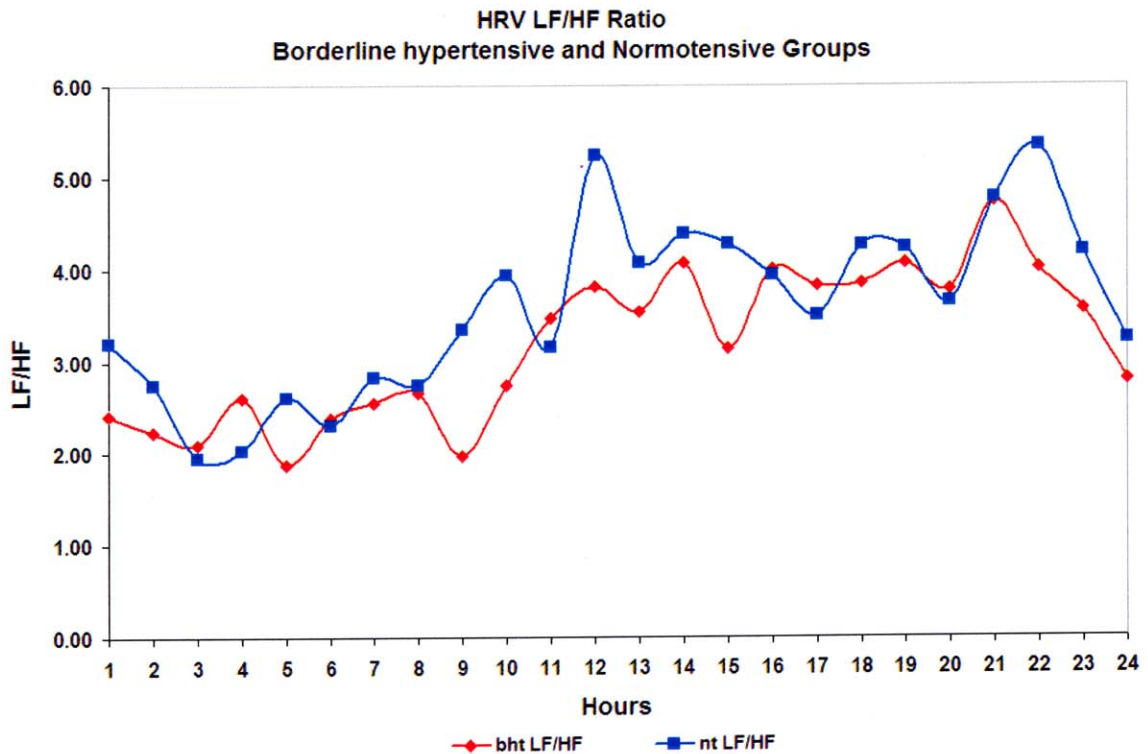
Figure 8.10 Sympathovagal ratio for heart rate variability (HRV) for both groups. Results are statistically significant to $p < 0.05$.

8.4   Alpha Index

Previous studies have shown that baroreflex sensitivity is negatively correlated with rising blood pressure and shows stepwise reductions in sensitivity in the transition from borderline to established hypertension [73-75].

Alpha index, also called spectral technique, is an estimate of the sensitivity of the baroreceptor reflex. It is based on the frequency domain analysis of the spontaneous variability of systolic blood pressure and the R-to-R interval. In this work, the R-to-R interval is replaced with the pulse interval (PI) obtained from systolic blood pressure recordings. This index is based on the fact that systolic blood pressure and heart rate

show a high degree of linear correlation in the respiratory and low frequency ranges in normal subjects, and on the hypothesis that the correlation at these two frequencies is due to the baroreflex coupling [23-26, 29, 32, 54]. Spectra of systolic blood pressure and pulse interval are calculated from which coherence is computed. The coherence function is used to evaluate the linear correlation between systolic blood pressure and the pulse interval. If coherence, at the frequencies of interest, is greater than or equal to 0.5, the spectra of systolic blood pressure and the pulse interval are integrated over those frequencies. Through integration, the low and high frequency powers of the systolic blood pressure (SBP-LF, SBP-HF) and pulse interval (PI-LF, PI-HF) are obtained. The square-root ratio of PI and SBP powers at these frequencies is then computed. The alpha index is discussed in greater detail in Section 3.4.

The frequency ranges used in this work for the alpha index are, low frequency (LF) between 0.07 and 0.14 Hz, and high frequency from 0.14 to .4 Hz. These ranges were selected following correspondence with Dr. Marco Di Rienzo and Dr. Paolo Castiglioni, renowned for their extensive research of the baroreceptor response. When using 24-hour data it is difficult to identify a clear respiratory peak in blood pressure spectra. During normal daily-life activities respiratory peak related power is more likely found in a high frequency broadband spectrum. Following their approach, the respiratory frequency may not be a single frequency, or a single spectral peak, rather it could be the set of all spectral components in the high frequency band where coherence is greater than 0.5. It is also possible, that occasionally, the breathing rate can fall completely within the low frequency band. In this case, it is very likely that the high frequency alpha estimate

will be zero due to low coherence. If this occurs the alpha index estimate will be considered as a missing estimate and not reported as 0 ms/mmHg.

8.4.1 Alpha Index Low Frequency

Low frequency alpha index (α_{LF}) response for both groups is shown in Figure 8.11. Results are shown as the mean +/- the standard deviation for each hour. The results for the low frequency alpha index are statistically significant to $p < 0.05$.
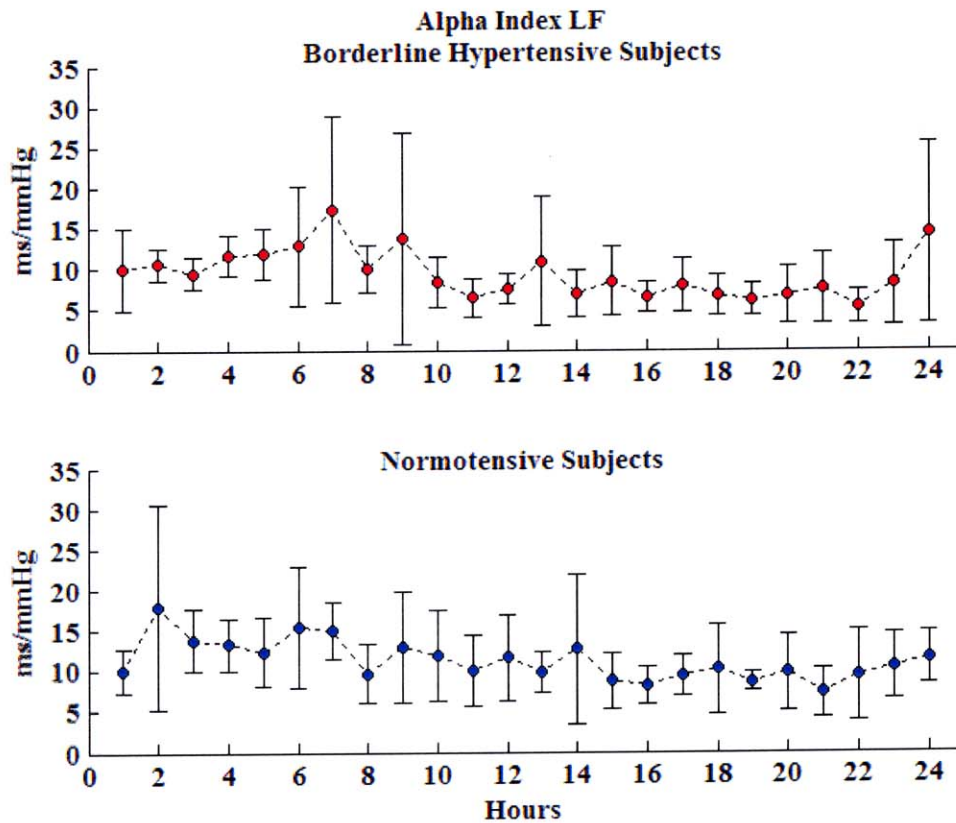
Figure 8.11 Low frequency alpha index (α_{LF}) for both groups. Results are shown as mean +/- the standard deviation. The results are statistically significant to $p < 0.05$.

Figure 8.11 indicates lower α_{LF} values in the borderline hypertensive group than in the normotensive group. With the exception of a few hours, the normotensive group has greater dispersion about the mean. Greater dispersion is interpreted as heightened responsiveness of the baroreceptor reflex in the normotensive group over the borderline hypertensive group. This is consistent with previous research that has indicated a blunted baroreflex response exists in hypertension [73-75]. An overlay plot of the low frequency alpha index is shown in Figure 8.12 which indicates a blunted baroreflex response in borderline hypertension.
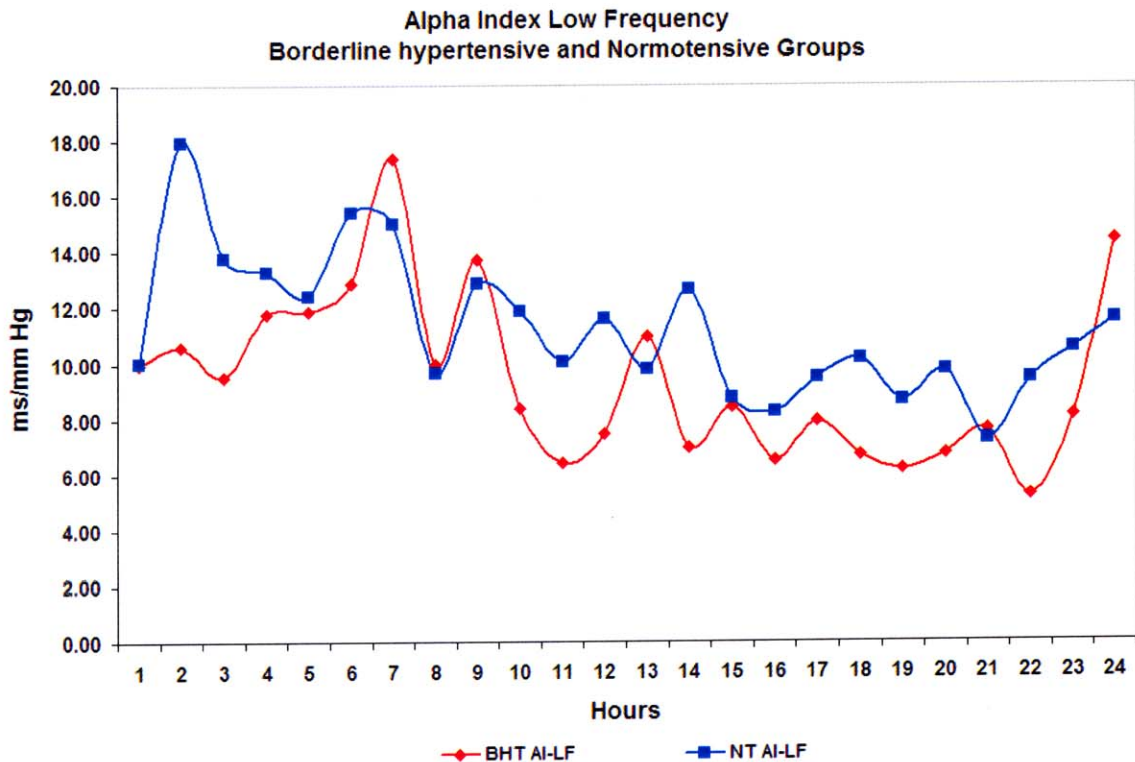
Figure 8.12 Overlay plot of alpha index low frequency. Results are statistically significant to $p < 0.05$.

8.4.2 Alpha Index High Frequency

High frequency alpha index (α_{HF}) response for both groups is shown in Figure 8.13. Results are shown as the mean +/- the standard deviation for each hour. The results are not statistically significant for this variable.
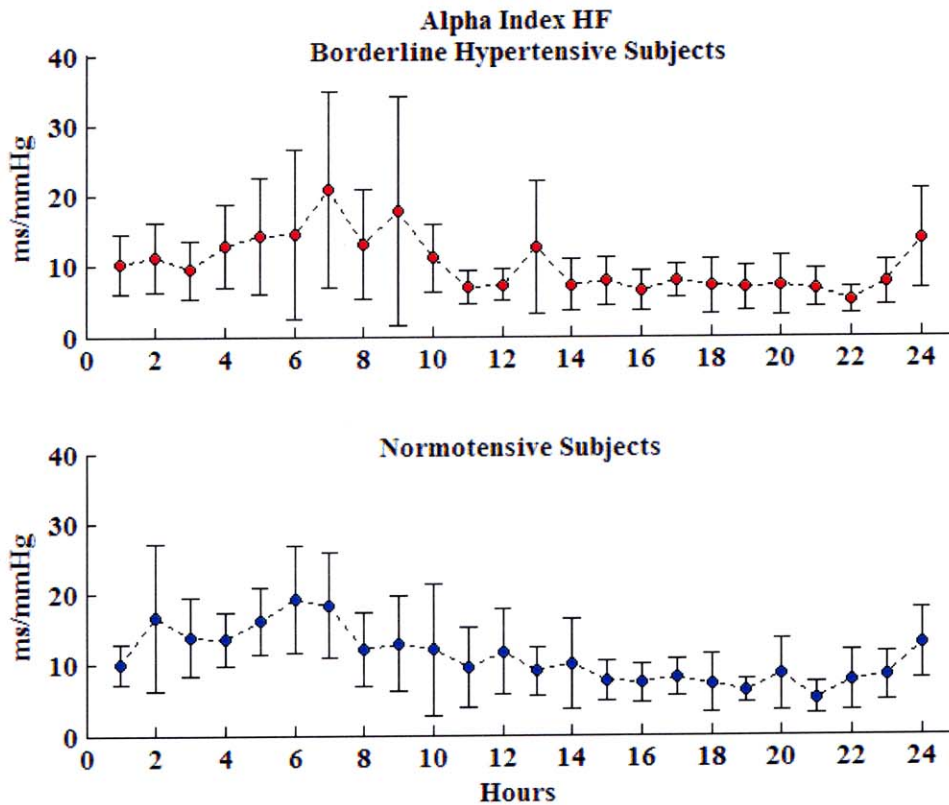
Figure 8.13 High frequency alpha index (α_{HF}) for both groups. Results are shown as mean +/- the standard deviation.

These results are consistent with the high frequency analysis of blood pressure and heart rate variability; there is no statistical difference over the 24-hour recording. However, both HRV and BPV did have blocks of hours that were significantly different statistically. That is not the case for the high frequency alpha index.

8.5 Approximate Entropy

Analysis of blood pressure data with approximate entropy has not been done in a cohort of normotensive and borderline hypertensive subjects. Approximate entropy (ApEn) is a statistic that quantifies the regularity and complexity of a time series. When there are recognizable patterns that repeat, approximate entropy returns a low value. Conversely, with random behavior approximate entropy results in a higher value. Therefore, when regularity is high then ApEn is low; conversely when regularity is low then ApEn is high. The parameters used for ApEn were $r = .15$ and $m = 2$, approximate entropy and these parameters are explained in greater detail in Section 3.1.

Approximate entropy analysis was performed on both systolic blood pressure and pulse interval data. Statistically significant results were obtained for the systolic blood pressure data to $p < .0.05$. Results for pulse interval data however, were not statistically significant. As ApEn is a regularity statistic, it was expected that approximate entropy values would be higher in borderline hypertension than in the normotensive group. This result was expected given that previous research has shown higher blood pressure variability in hypertension which would indicate increased randomness of the data. Results presented here concur with those previous findings. Figure 8.14 are results of approximate entropy analysis of systolic blood pressure data for both groups.

It can be seen in Figure 8.14 that the approximate entropy values are higher in the borderline hypertensive group than the normotensive group. Dispersion of data about the mean appears decreased in the normotensive group as compared to the borderline hypertensive group. This may indicate a more consistent pattern of variability in the normotensive group, versus the increased variability in the borderline hypertensive group.

Approximate entropy results of systolic blood pressure are presented in an overlay plot in Figure 8.15 which highlights the higher values of the borderline hypertensive group over their normotensive counterparts.
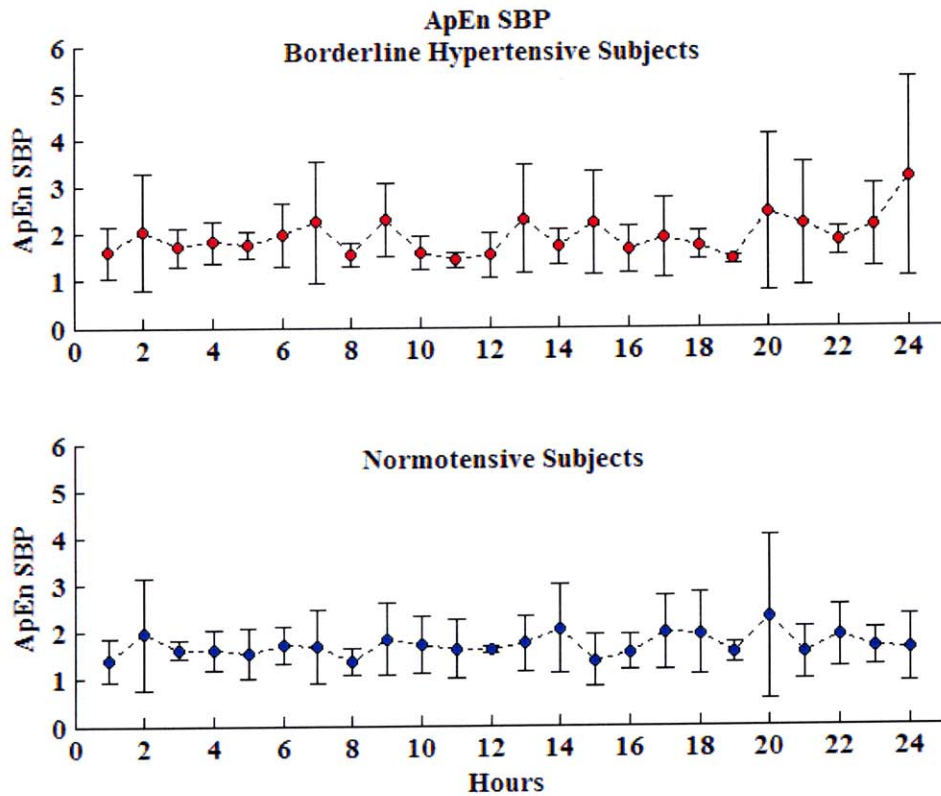
Figure 8.14 Results of approximate entropy (ApEn) analysis of systolic blood pressure data for borderline hypertensive and normotensive groups. Results are shown as the mean +/- the standard deviation for each hour. Results are significant to $p < 0.05$.
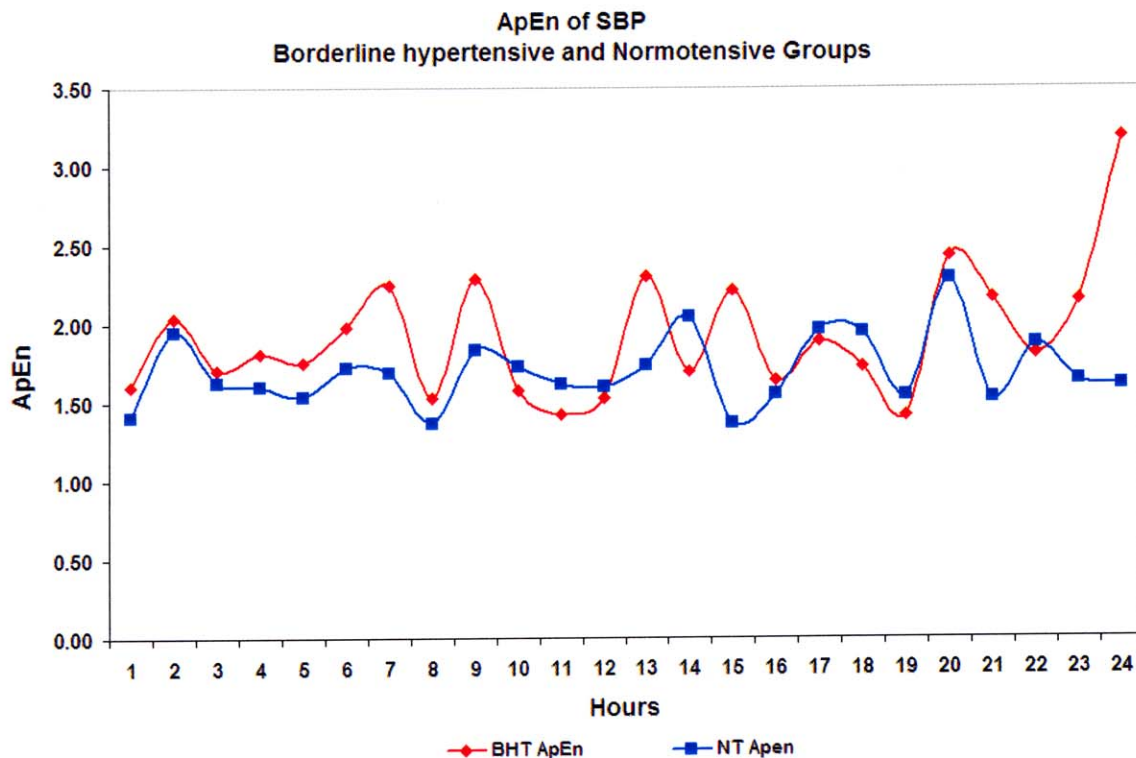
Figure 8.15 Overlay plot of approximate entropy results for systolic blood pressure data for both groups.

Figure 8.15 provides a clearer view of the elevated approximate entropy values in the borderline hypertensive group (BHT) over the normotensive group. The elevated ApEn values reinforce the results of elevated blood pressure variability in the BHT group indicating more random behavior over time. While the borderline hypertensive group ApEn values are not consistently above those of the normotensive group, they are greater for the majority of the 24-hour recording period. The close physiologic state of the cohort may be a contributory factor in this phenomenon. It is probable that in a more severe hypertensive state, approximate entropy values would be greater than presented here in the borderline case.

Figure 8.16 plots the approximate entropy values for the pulse interval data for both groups. These results are not statistically significant between the two groups. With the exception of several hours, the dispersion about the mean is very similar between groups. It is also apparent from Figure 8.16 that the ApEn values for the pulse interval data are more closely matched between groups than those of systolic blood pressure data in Figure 8.14. An overlay plot here serves no purpose and will not be shown.
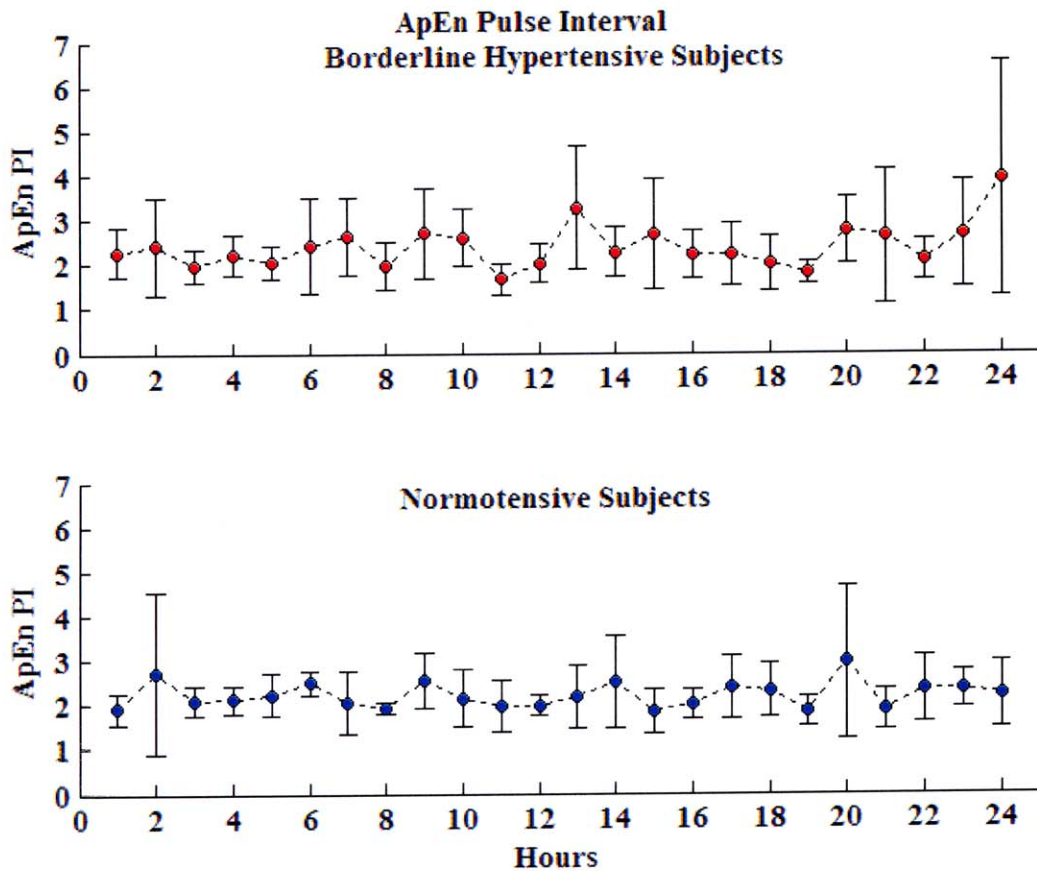
Figure 8.16 Approximate entropy results for pulse interval data for both groups. Results are shown as mean +/- the standard deviation. Statistically there is no difference between groups.

8.6   Detrended Fluctuation Analysis

Detrended fluctuation analysis is a widely used technique for the detection of long-range correlations in noisy, nonstationary time series and has been successfully applied to diverse fields such as DNA sequencing, neuron spiking, human gait and heart rate dynamics [43]. Although detrended fluctuation analysis has been widely utilized in very diverse fields as noted above, there have been no studies identified to date that have utilized DFA in the analysis of blood pressure or in a cohort of normotensive and borderline hypertensive subjects. Detrended fluctuation analysis is described in greater detail in Section 3.2.

Figure 8.17 illustrates the short term scaling exponent alpha-1 ($\alpha 1$) from detrended fluctuation analysis of systolic blood pressure. Statistically there is no difference between groups for this variable over the full 24-hour period. There is however, statistical significance $p < 0.05$ in the hours from 13 to 24. Results are shown as the mean +/- the standard deviation. Alpha-1 is the short-term scaling exponent and was calculated over box sizes in the range of $4 <= n < 11$. The scaling exponent represents the slope of the line when plotting log (Fn) versus log (n), these terms are explained in detail in Section 3.2. The slope quantifies the short-term scaling properties of the systolic blood pressure data.

As detrended fluctuation analysis has not been applied to blood pressure data, or utilized with a cohort of normotensive and borderline hypertensive subjects, results can not be compared with previous research. It is believed that due to an elevated variability, the scaling exponents for the borderline hypertensive group will be on average lower than those of the normotensive group. It is postulated that this is due to the break down of

the scaling behavior in hypertension as the system tends towards more random dynamics. The elevated low frequency power in blood pressure variability and the elevated approximate entropy values for the borderline hypertensive group support this premise.
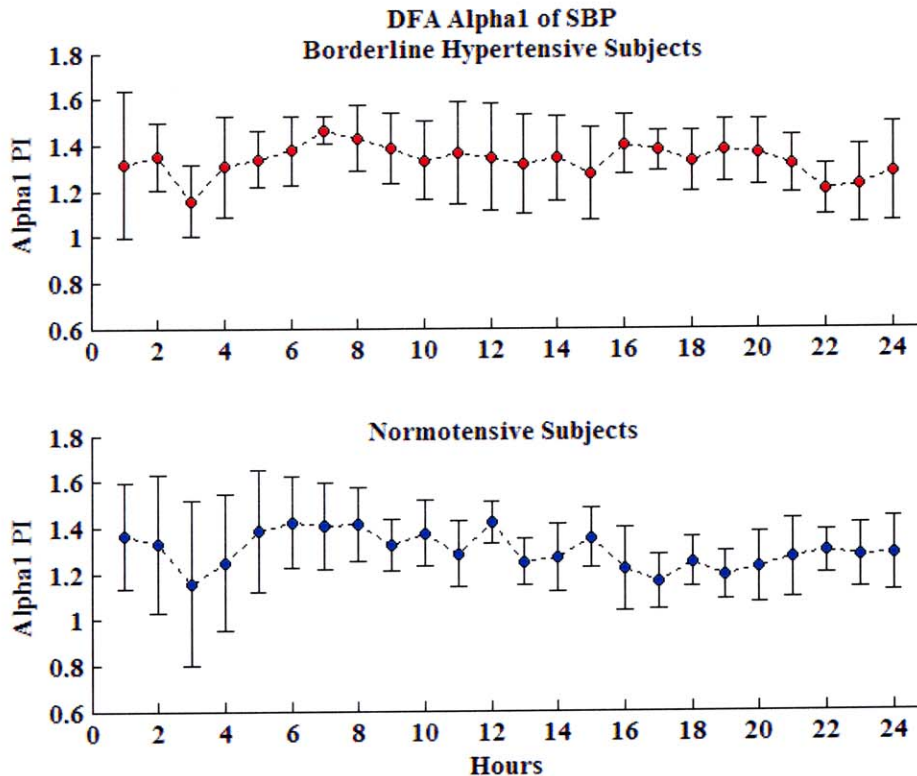
Figure 8.17 Short-term scaling exponent α_1 from systolic blood pressure data for both groups. Results are shown as the mean +/- the standard deviation. Statistically there is no difference between groups for the full 24-hour period. The hours between 13 and 24 are statistically significant to $p < 0.05$. Statistical relevance was tested by paired t-testing.

During the first twelve hours in Figure 8.17, the dispersion about the mean and fluctuations in the short-term scaling exponent values appear similar. During the final twelve hours oscillations in the α_1 values are opposite one another between the two groups. These differences in the last twelve hours are significant to $p < 0.05$. The

differences are best seen in the overlay plot of Figure 8.18, which illustrates the similarities in the first twelve hours and the divergent behavior in the final twelve hours. In the final twelve hours the BHT group values are elevated. On average, α_1 for the BHT group is greater than the NT group, which was unexpected. This may be due to the time-scale over which α_1 was calculated. It may be necessary to increase this scale to capture short-term scaling.
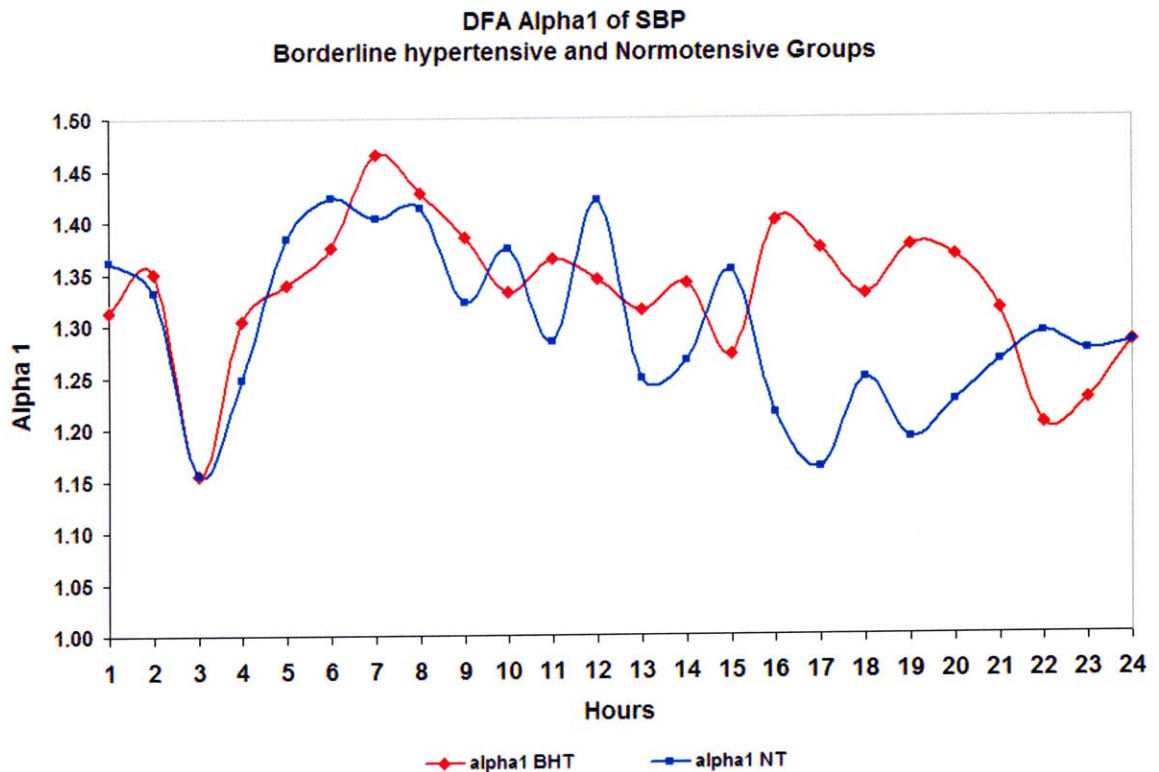
Figure 8.18 Overlay plot of short-term scaling exponent α_1 from systolic blood pressure data for both groups. Statistically there is no difference between groups for the full 24-hour period. The hours between 13 and 24 are however, statistically significant to $p < 0.05$. Statistical relevance was tested by paired t-testing.

Figure 8.19 illustrates the long-term scaling exponent, α_2, of detrended fluctuation analysis from systolic blood pressure which was calculated over box sized $11 <= n <= 64$. Results for this variable are statistically significant to $p < 0.05$ over the full 24-hour recording period. This variable is one of the two that were highly instrumental in a-priori subject grouping. Figure 8.19 indicates greater dispersion about the mean in the normotensive group versus the borderline hypertensive group for the majority of the 24-hours and on average, higher long-term scaling exponent values.
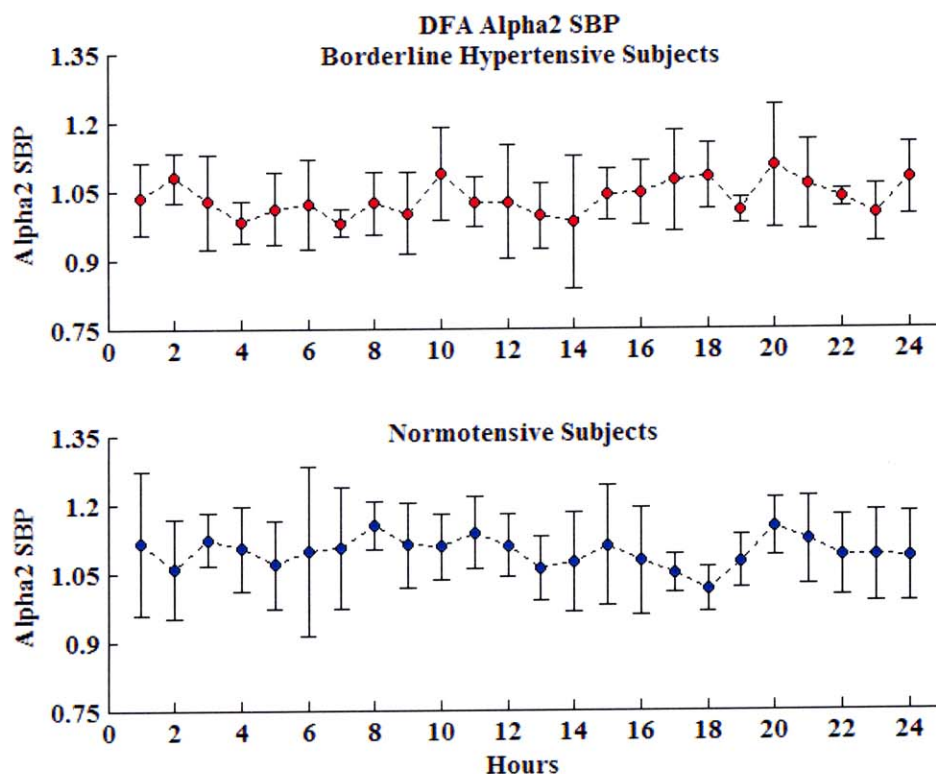
Figure 8.19 Detrended fluctuation analysis (DFA) long-term scaling exponent, α_2, of systolic blood pressure data for both groups. Results are shown as mean +/- the standard deviation. Results are statistically significant to $p < 0.05$.

An overlay plot of the data is shown in Figure 8.20 to provide an improved graphic of the differences in the α_2 values between groups. It is apparent in Figure 8.20, that the normotensive group maintains a higher long-term scaling exponent value ($\alpha2$) over the course of the 24-hour recording. This may be an indication that the scaling behavior is beginning to deteriorate in borderline hypertension leading to more random dynamics. This view is validated by elevated approximate entropy values, which indicates random behavior, in borderline hypertension over the normotensive group.

The divergence between groups in the α_2 values versus the α_1 values is greater. This is easily verified by comparing Figures 8.18 and 8.20. Although the α_2 values in the borderline hypertensive group would not necessarily be considered abnormal they are lower, than in the normotensive group. This may indicate a premature deterioration of the long-term scaling exponent in borderline hypertension. Application of detrended fluctuation analysis in a more physiologically diverse cohort may provide additional information regarding loss of long-term scaling properties in hypertension.
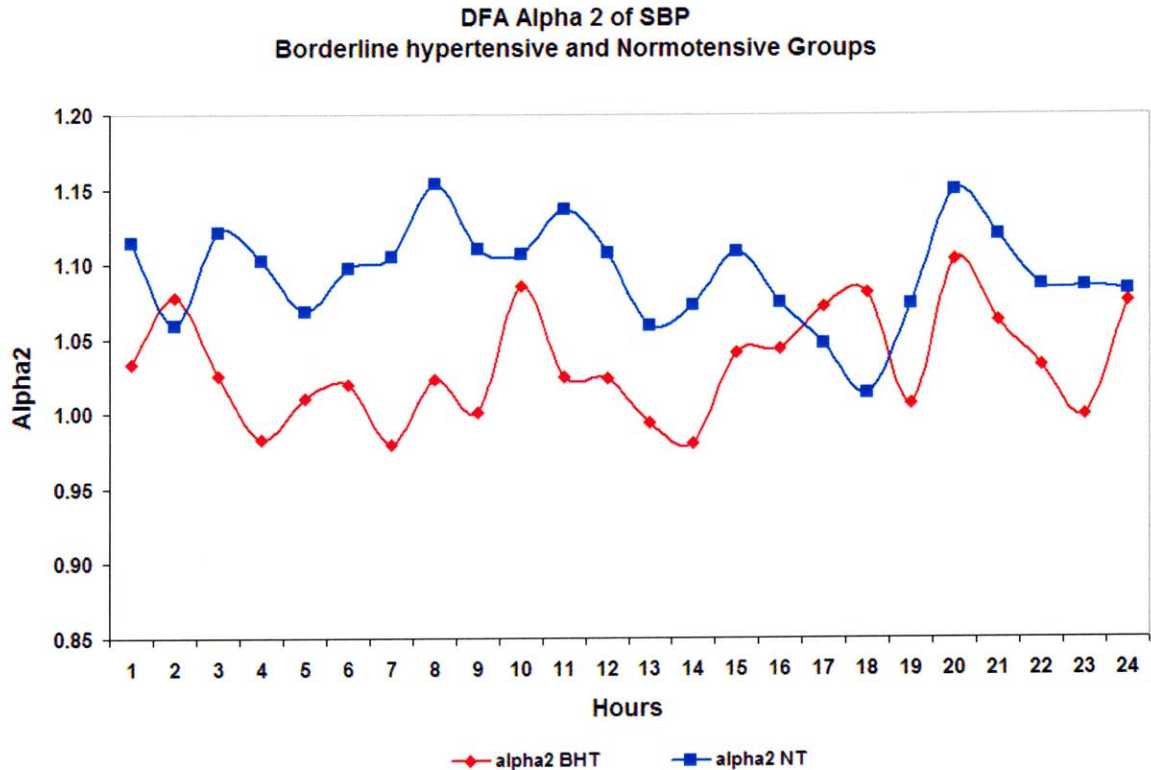
Figure 8.20 Overlay plot of the long-term scaling exponent α_2 from analysis of systolic blood pressure data. Statistically these results are significant to $p < 0.05$.

A scatter plot of $\alpha 1$ versus $\alpha 2$ from systolic blood pressure data was created. This is shown in Figure 8.21. This plot was created to investigate the clustering behavior of these two variables for both groups. It can be seen in Figure 8.21 that the two groups cluster together in the same general area of the scatter plot. There is, however, a difference in cluster pattern between the two groups. The spread of the normotensive group is more aligned to an imaginary line of identity, $y = x$. The spread of the borderline hypertensive group is near orthogonal to an imaginary line of identity and thus, to the normotensive group.
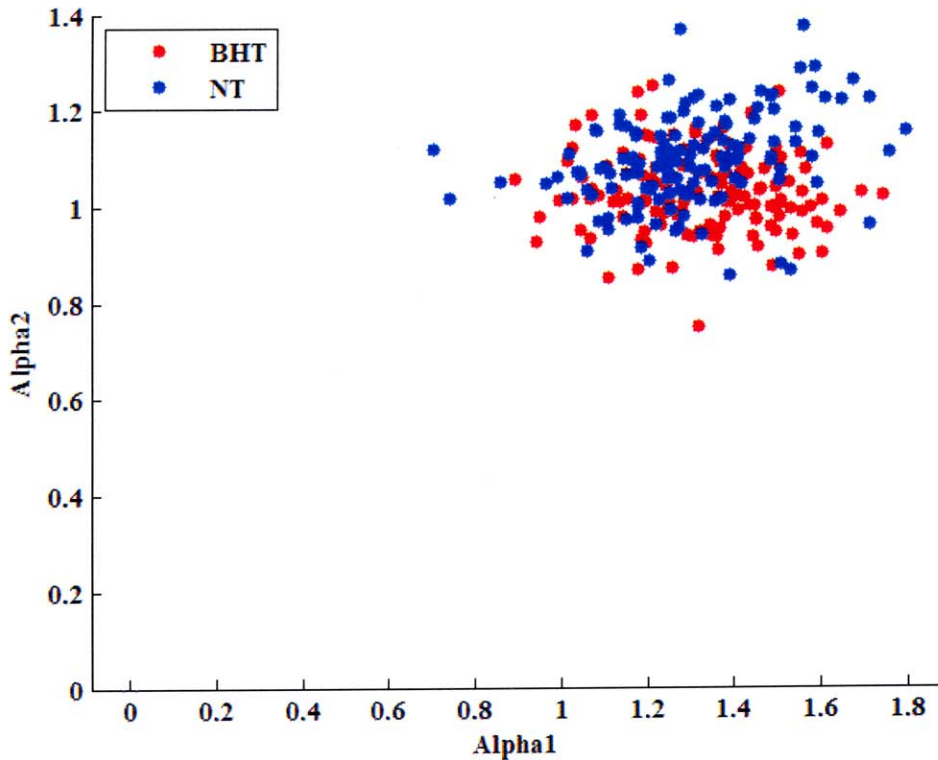
Figure 8.21 Scatter plot of scaling exponents α_1 versus α_2 from SBP data for the normotensive group (blue circle) and borderline hypertensive group (red circle).

Scatter plots for each group were developed to highlight individual group scattering. Figure 8.22 is a scatter plot of the scaling exponents $\alpha1$ versus $\alpha2$ for the borderline hypertensive group. As viewed in Figure 8.22, the cluster direction is in a near orthogonal direction to an imaginary line of identity. An ellipse has been added that encompasses the main cluster for this group. Figure 8.23 is an identical scatter plot for the normotensive group. In the normotensive group, the direction of clustering more closely follows an imaginary line of identity.
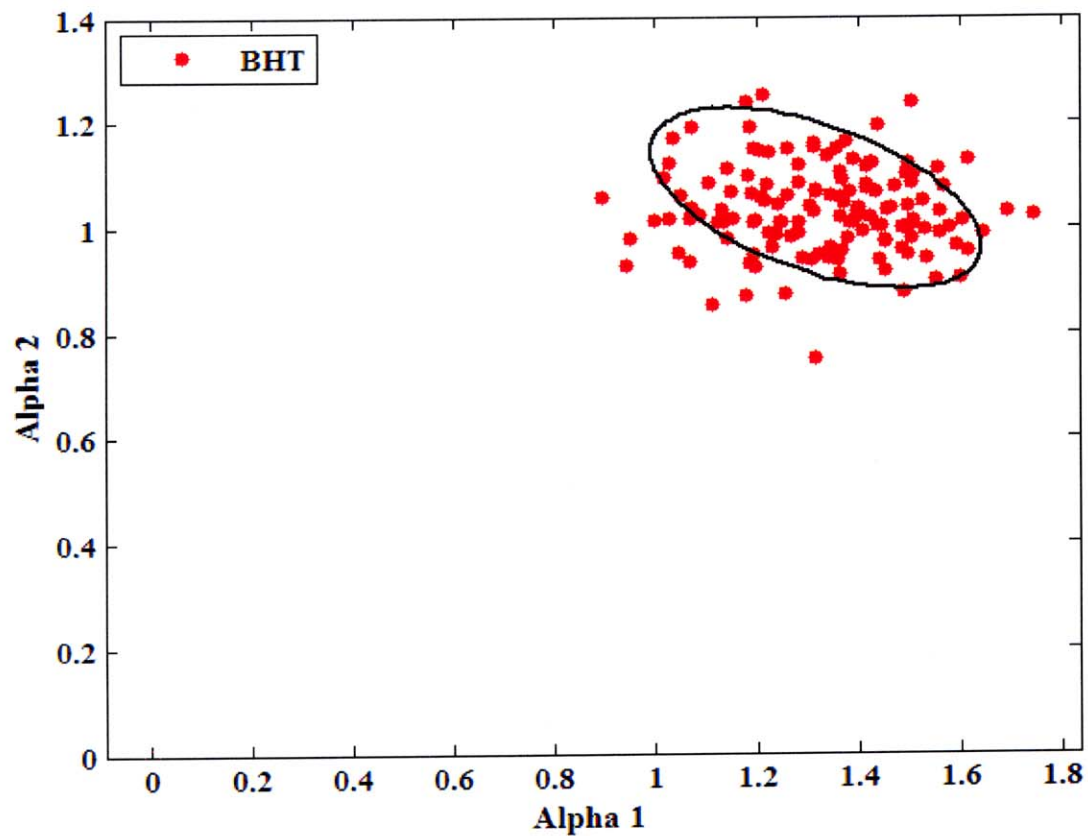
Figure 8.22 Scatter plot of scaling exponents α_1 versus α_2 for the borderline hypertensive group with an ellipse added. The ellipse emphasizes cluster direction.

Figure 8.23 Scatter plot of scaling exponents α_1 versus α_2 for the normotensive group with an ellipse added. The ellipse emphasizes cluster direction.

While it may not be possible to assign group membership from the clusters in these scatter plots, there does appear to be different behavioral patterns forming within each group. In a hypertensive cohort that is more physiologically diverse, scatter plots of the scaling exponents α_1 and α_2 may provide more conclusive evidence as to subject grouping and aid in early detection of hypertension. It is concluded that detrended fluctuation analysis of systolic blood pressure data, has the potential to aid in the characterization of hypertension.

Detrended fluctuation analysis (DFA) results for α_1 from analysis of pulse interval data for both groups are shown in Figure 8.24. Statistically the results are not significant. It can be seen in Figure 8.24 that the results are very similar in trend and closely matched in data dispersion about the mean between the two groups.
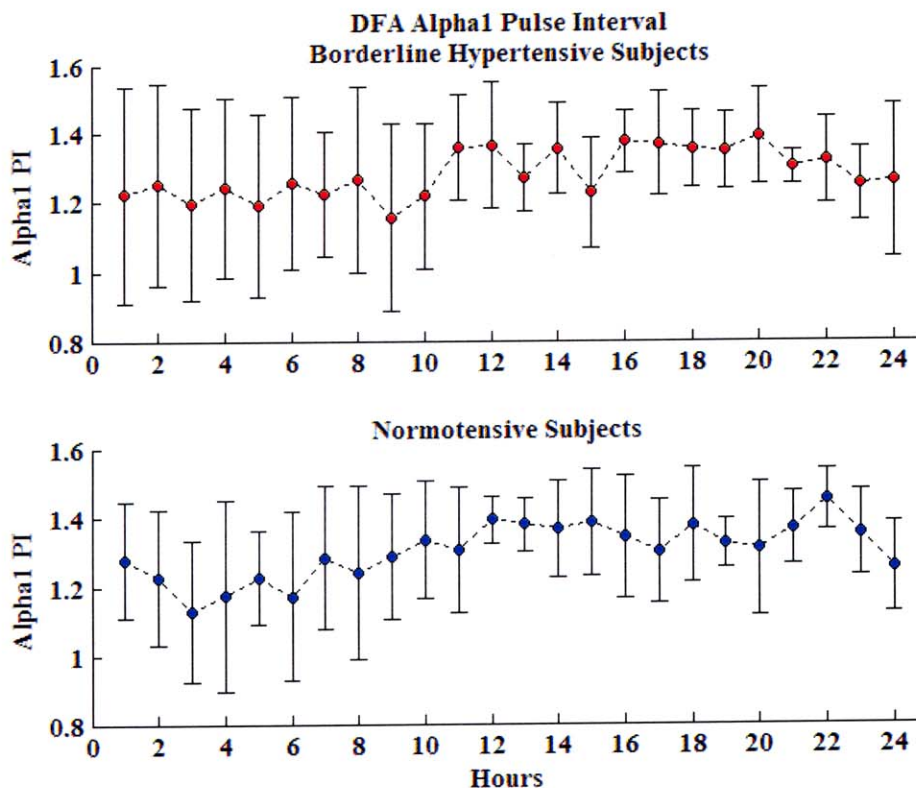
Figure 8.24 Short-term scaling exponent (α_1) from detrended fluctuation analysis of pulse interval data for both groups. Results are shown hourly as the mean +/- the standard deviation. Statistically these results are not significant.

In Figure 8.25, the results of the long-term scaling exponent for the pulse interval data are shown for both groups. Statistically these results are not significant over the full 24-hour period. They are nonetheless, statistically significant to $p < 0.05$ for the first twelve hours and again from hour 16 to 23. During the first twelve hours, the

normotensive group trends to higher long-term scaling exponent values than the borderline hypertensive group. This behavior changes after the first twelve hours with a four hour transition period where behavior between the groups is similar. The results diverge again from hour 16 through 23 at which time they are again statistically significant to $p < 0.05$. An overlay plot created in Figure 8.26 demonstrates more clearly the gross behavior between groups.
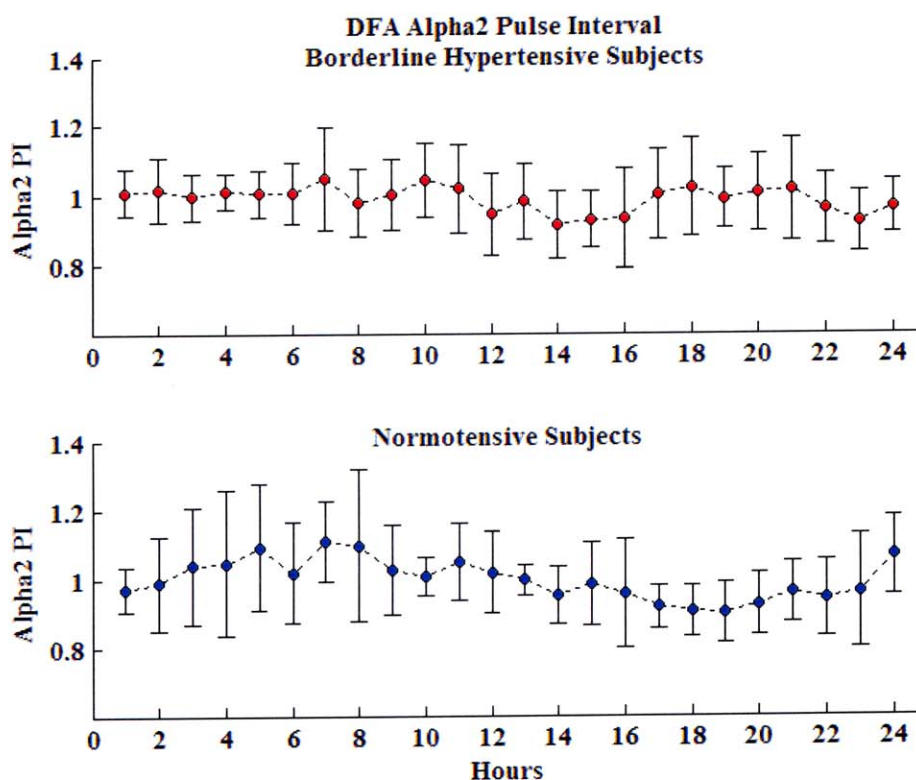
Figure 8.25 The long-term scaling exponent from DFA analysis of pulse interval data for both groups. Results are shown as the mean +/- the standard deviation. There is no statistical difference between the groups when analyzed over the full 24-hour period. The first twelve hours are statistically significant to $p < 0.05$.
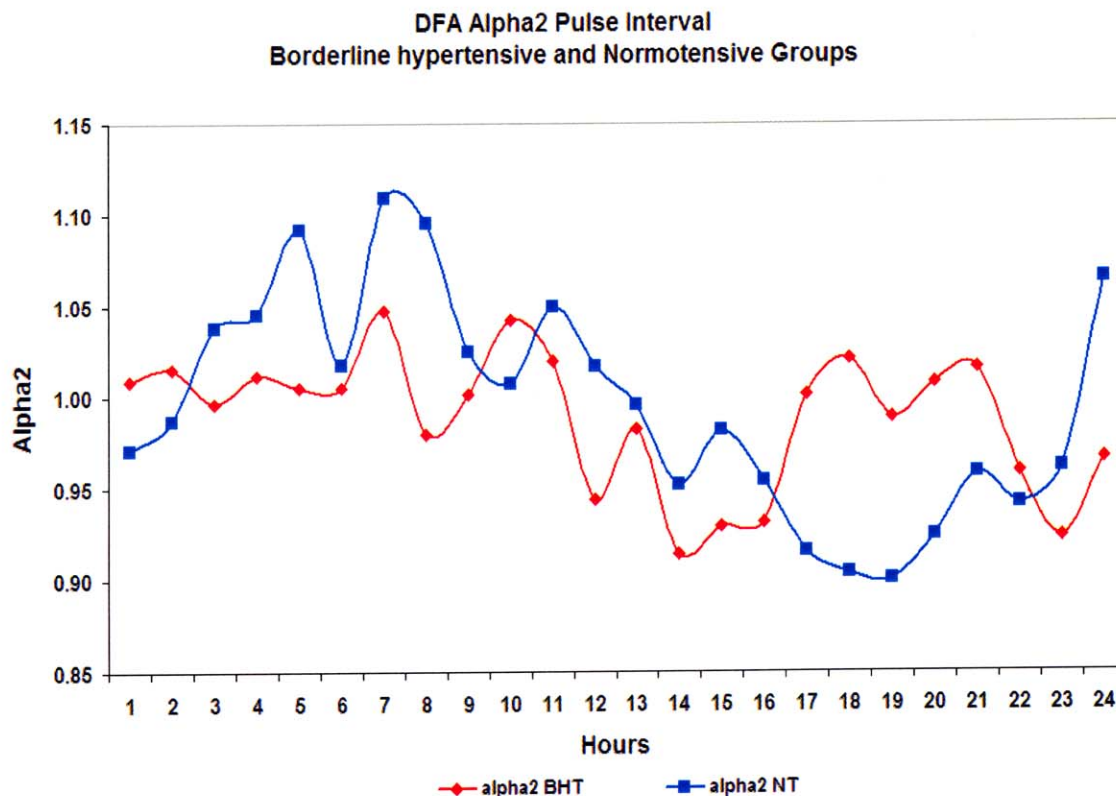
Figure 8.26 An overlay plot of the long-term scaling exponent from detrended fluctuation analysis of pulse interval data for both groups. Results are statistically significant from hours one through twelve, and again from hours 16 through 23 to $p < 0.05$.

In Figure 8.26, note the change in the y-axis to highlight the statistically different areas in the data. Again the long-term scaling exponent obtained through detrended fluctuation analysis has indicated statistically significant differences between the groups. The α_2 values from systolic blood pressure indicated statistically significant differences between subject groups and were highly instrumental in a-priori subject separation. While α_2 derived from pulse interval data did not aid in a-priori subject separation, given a more physiologically diverse cohort, it appears to have the potential to aid in characterization of hypertension.

8.7 Poincaré Plots

The predominant use of Poincaré plots in medical science has been in the study of heart rate variability. Poincaré analysis has not been identified as having previously been used in the analysis of blood pressure data or in a cohort of normotensive and borderline hypertensive subjects. The Poincaré map, also called the return map, was discussed in greater detail in Section 3.3.

Figure 8.27 is a plot of SD1 from Poincaré analysis of systolic blood pressure for both groups. Results are shown as the mean +/- the standard deviation. The parameter SD1 is an indication of the instantaneous beat-to-beat variability (see Figure 3.2) of the data. The results are not statistically different between groups over the full 24-hour period. They are however, significantly significant to $p < 0.05$ for the first twelve hours. During this time the subjects are asleep, waking up at 7AM, eating breakfast followed by watching television, reading a book or working with a computer. Individual subject logs are located in Appendix C.

In comparing the first twelve hours of Figure 8.27, the overall behavior between groups is similar with the normotensive group exhibiting greater dispersion about the mean during only a few hours over the borderline hypertensive group. In the last twelve hours the opposite is true, yet only the first twelve hours are statistically different to $p < 0.05$. On average, the borderline hypertensive group has a higher SD1 value than the normotensive group. These results did not weigh heavily in the decision as to subject grouping.
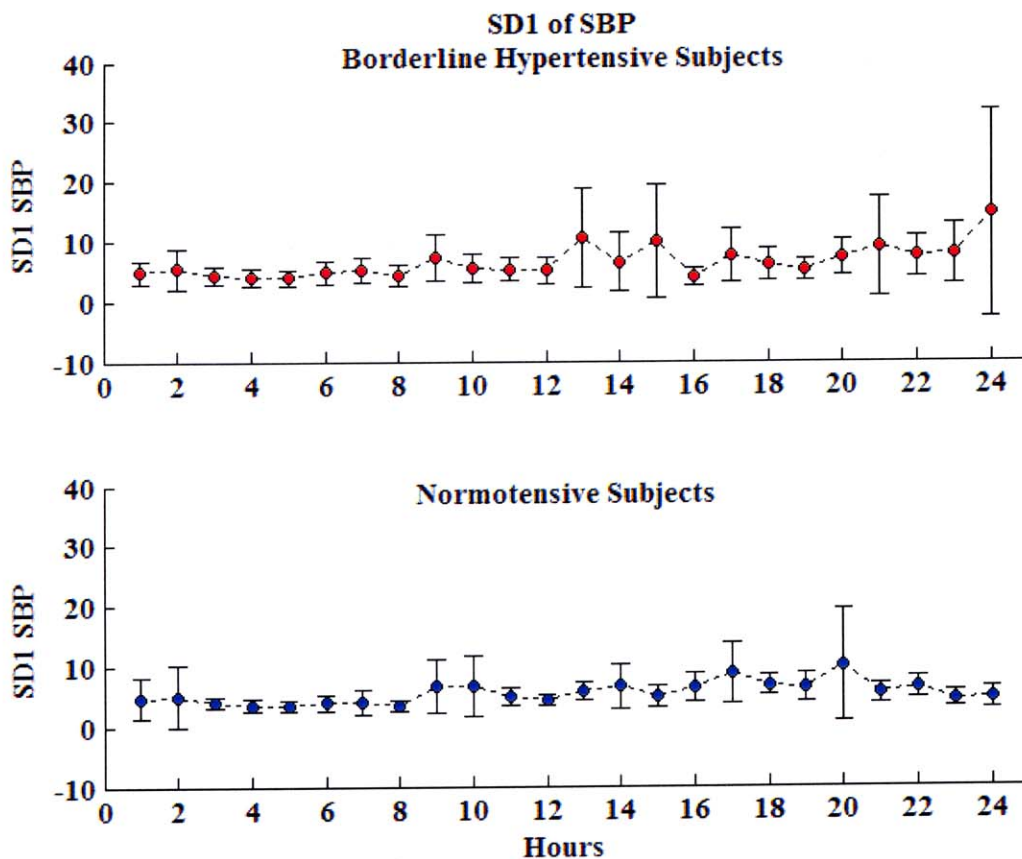
Figure 8.27 Parameter SD1 from Poincaré analysis of systolic blood pressure. Results are shown as the mean +/- the standard deviation. The first twelve hours are statistically significant to p < 0.05.

Figure 8.28 shows the parameter SD2 from Poincaré analysis of systolic blood pressure for both groups. This parameter is an indication of long-term variability (see Figure 3.2) of the systolic blood pressure data. These results along with the long-term scaling exponent α_2 from detrended fluctuation analysis were the two variables that were crucial in determining subject grouping. Subject grouping is discussed in detail in Chapter 7. These results are statistically significant to p < 0.05 as determined by paired t-testing over the full 24-hour recording period.
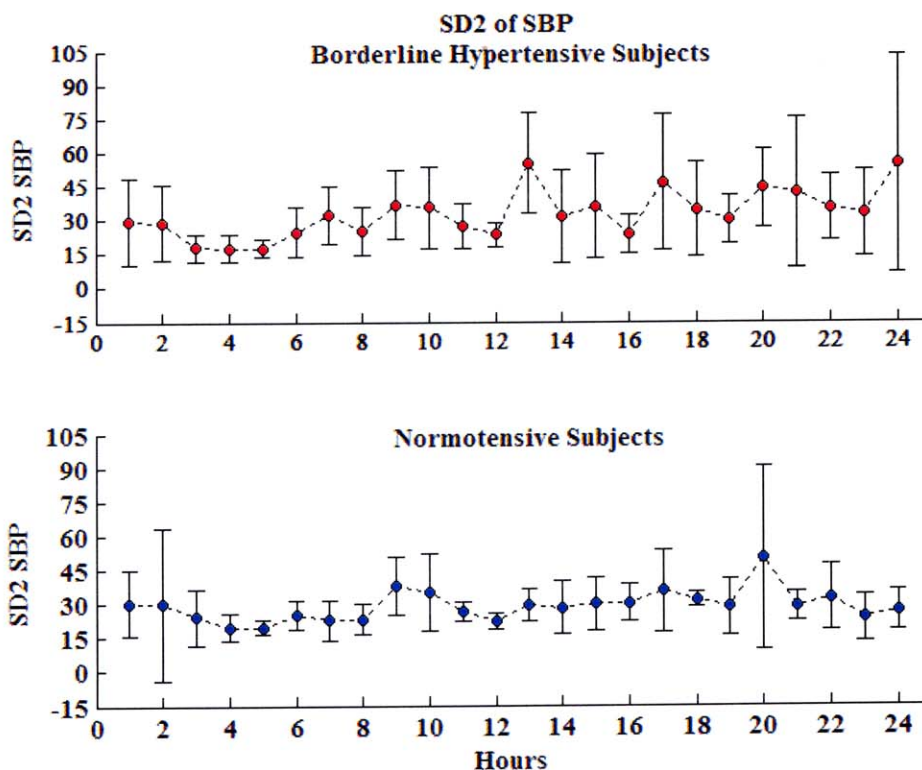
Figure 8.28 SD2 results from Poincaré analysis of systolic blood pressure data. Results are shown as the mean +/- the standard deviation. Results are statistically significant to $p < 0.05$ between groups as determined by paired t-testing.

In Figure 8.28, there is a greater dispersion of data about the mean in the borderline hypertensive group than in the normotensive group. Dispersion in this group increases beyond hour 12 and continues for the remaining hours. This would indicate an increased long-term variability in the borderline hypertensive group which supports previous research of elevated blood pressure variability in hypertension. Poincaré analysis, specifically the long-term variability parameter SD2 was paramount in its contribution to the decision making process towards subject separation.

Poincaré analysis of pulse interval data was also performed. Figure 8.29 is a plot of SD1 of the pulse interval data for both groups. Results are shown as the mean +/- the standard deviation. There is no statistical significance between the two groups for this variable.
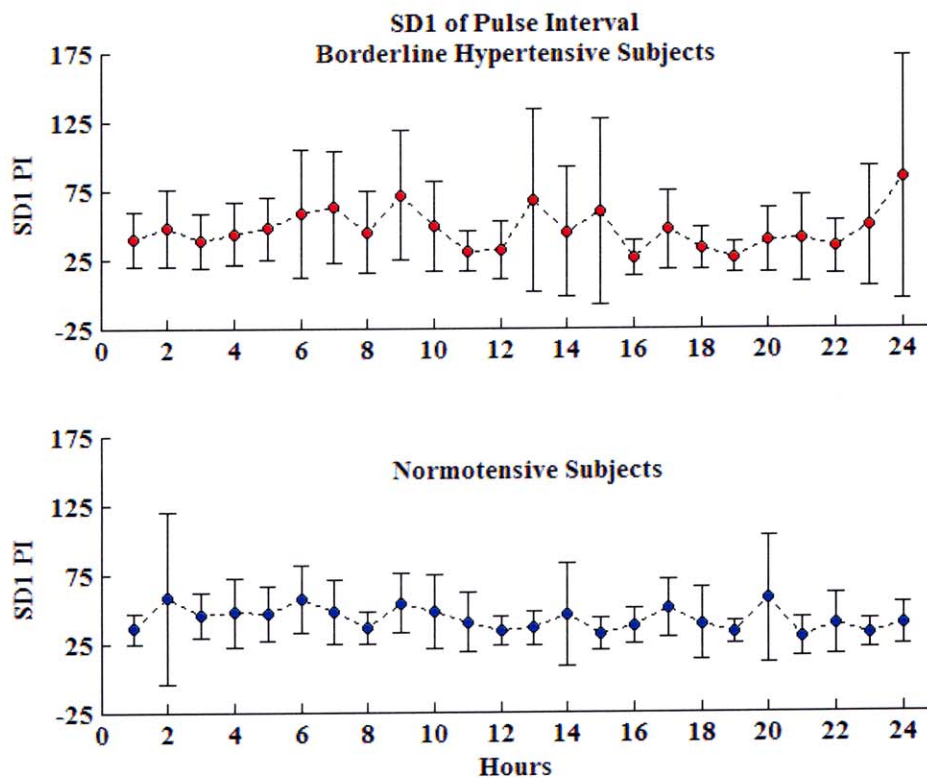
Figure 8.29 SD1 results from Poincaré analysis of pulse interval data. Results are shown as the mean +/- the standard deviation. There is no statistical significance between groups.

Figure 8.30 is a plot of SD2 from Poincaré analysis of pulse interval data for both groups. Results are shown as the mean +/- the standard deviation. There is no statistical significance between groups for this parameter.
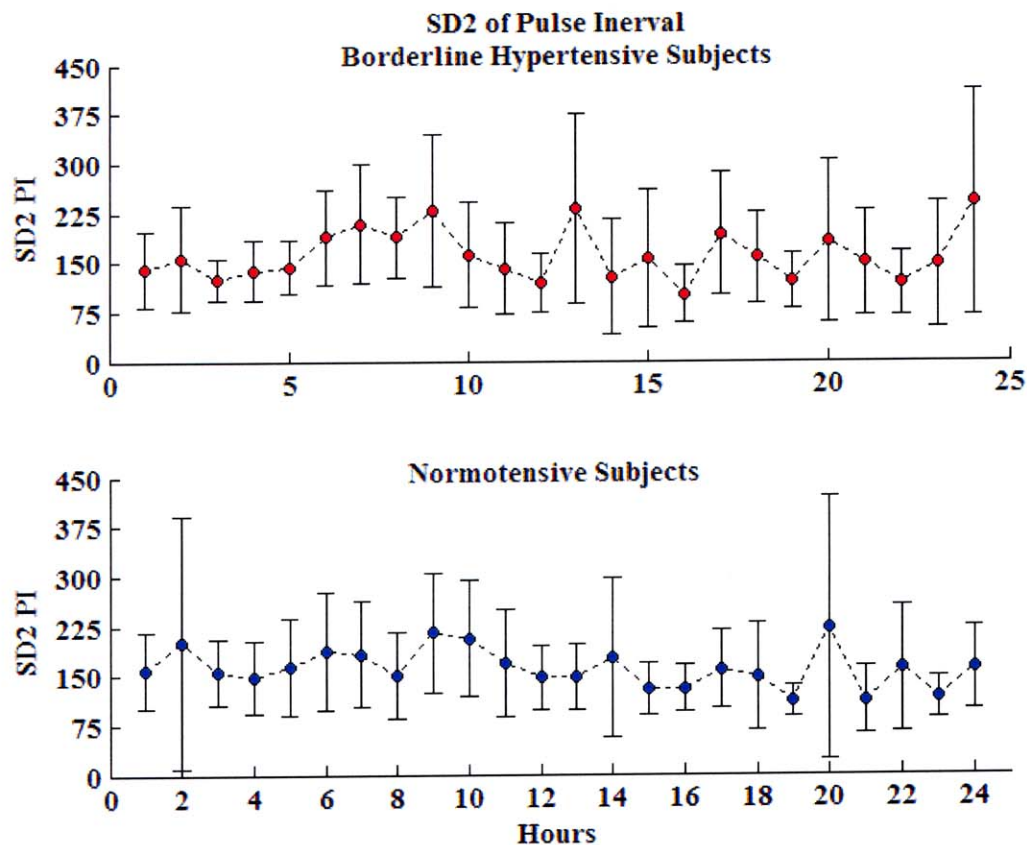
Figure 8.30 SD2 results from Poincaré analysis of pulse interval data for both groups. Results are shown as the mean +/- the standard deviation. There is no statistical significance between groups.

Poincaré plots of the systolic blood pressure data are shown for the borderline hypertensive and the normotensive groups in Figure 8.31 and 8.32, respectively. The parameters SD1 and SD2 are indicated on each plot. These parameters represent the instantaneous beat-to-beat systolic blood pressure variability and the long-term continuous variability, respectively, for each subject over the 24-hour recording period.

POINCARÉ PLOTS OF BORDERLINE HYPERTENSIVE GROUP

SYSTOLIC BLOOD PRESSURE
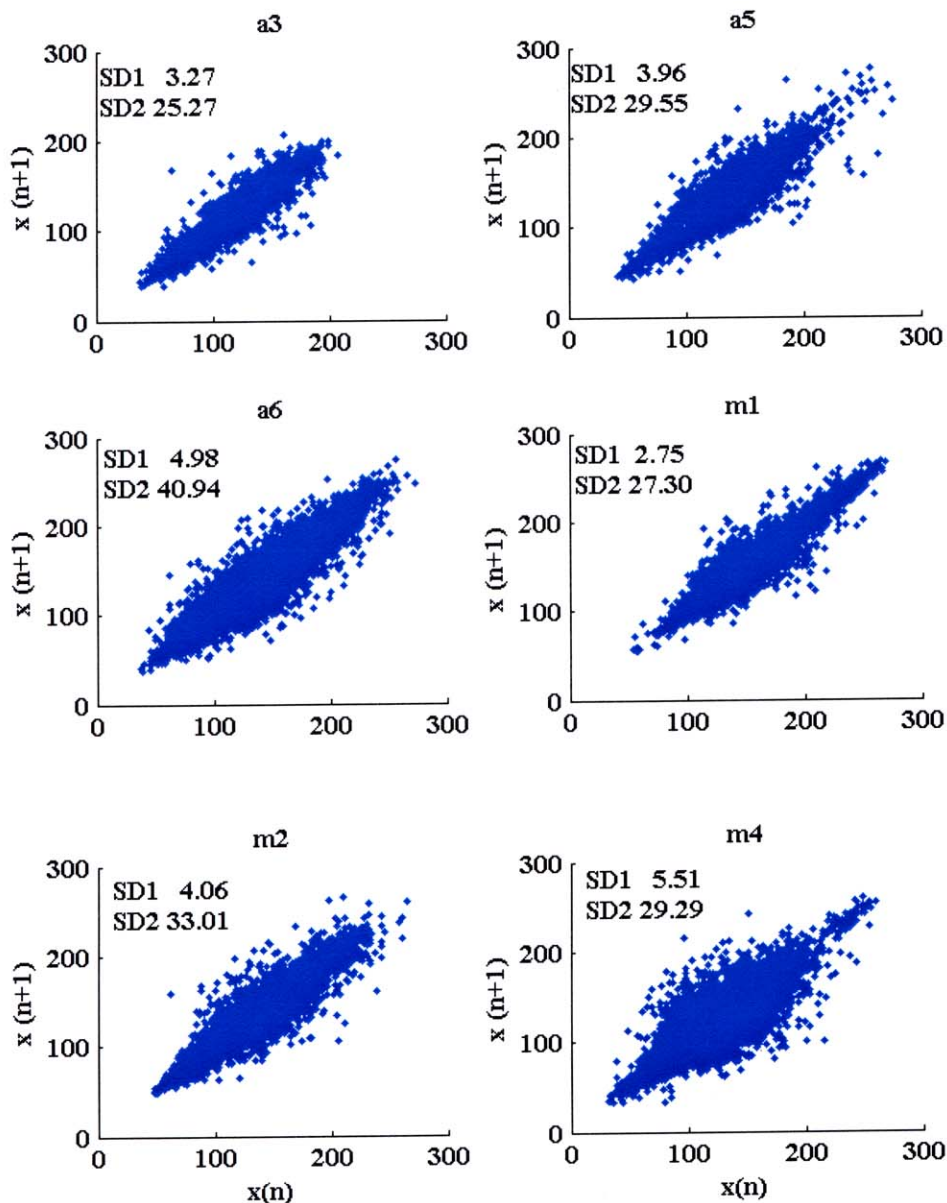
Figure 8.31 Individual Poincaré plots of systolic blood pressure data in the borderline hypertensive group. Each plot indicates the instantaneous beat-to-beat systolic blood pressure variability (SD1) along with the long-term continuous systolic blood pressure variability (SD2) over the full 24-hour recording period. In all plots, the x-axis is x(n); subject identifiers appear above each plot.

POINCARÉ PLOTS OF NORMOTENSIVE GROUP

SYSTOLIC BLOOD PRESSURE
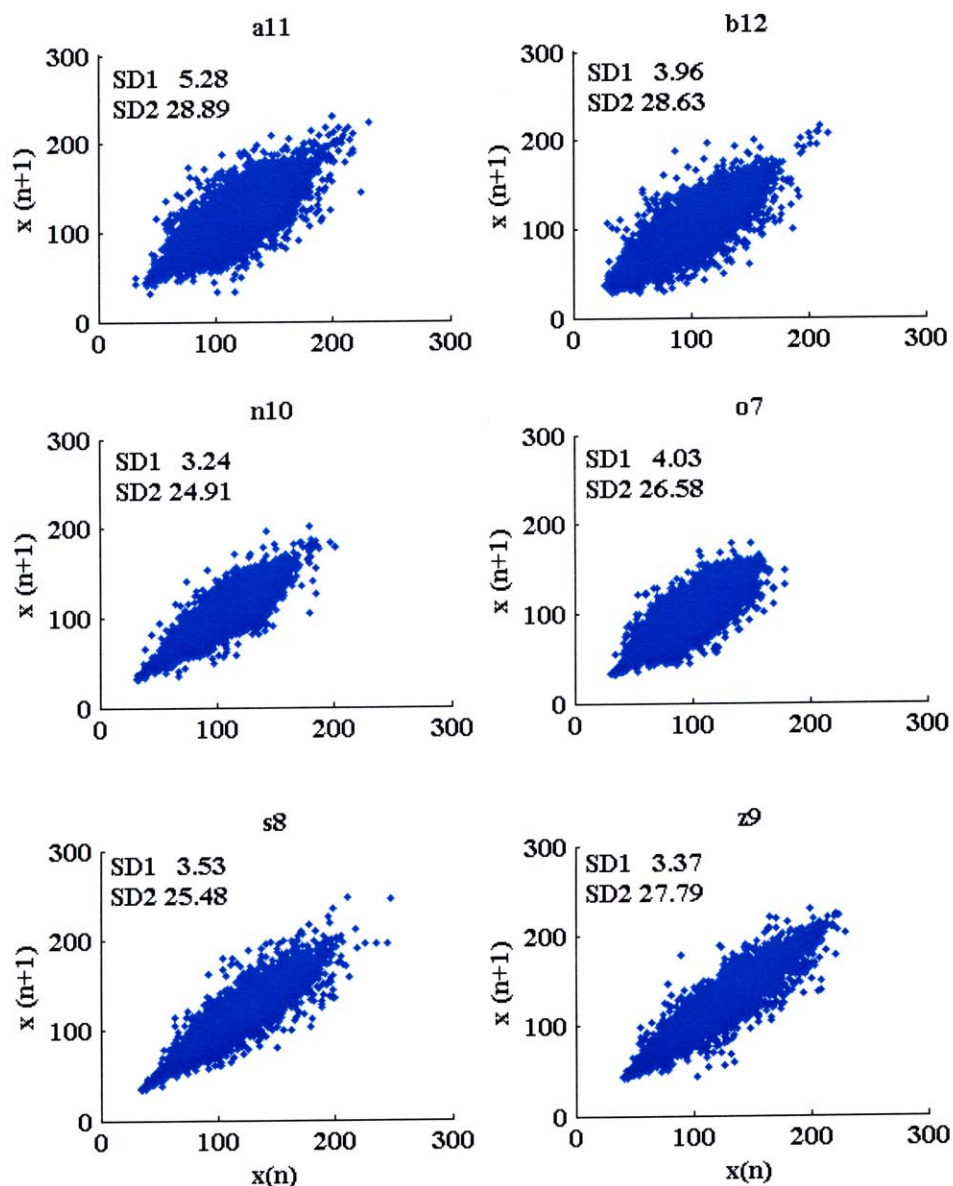
Figure 8.32 Individual Poincaré plots of systolic blood pressure in the normotensive group. Each plot indicates the instantaneous beat-to-beat systolic blood pressure variability (SD1) along with the long-term continuous systolic blood pressure variability (SD2) over the full 24-hour recording period. In all plots, the x-axis is x(n); subject identifiers appear above each plot.

The Poincaré plots of systolic blood pressure from both groups can be characterized as fan or torpedo shaped. In Figure 8.31, subject m4 is the only one in this group with an increased center width. The ranges of systolic blood pressure values in the borderline hypertensive plots are larger than the normotensive plots. In Figure 8.32, subjects n10, s8 and z9 in the normotensive group, appear very close in shape to the borderline hypertensive group of Figure 8.31. The other three normotensive subjects all have more condensed plots with widened centers. The shape of the Poincaré plots, while interesting, do not discriminate adequately between subject groups. Physiologically this cohort are very close, this may help explain why the Poincaré plots of systolic blood pressure do not elucidate subject differences.

Poincaré plots of the pulse interval data are shown for the borderline hypertensive group in Figure 8.33. Identical plots for the normotensive group are shown in Figure 8.34. The parameters SD1 and SD2 are indicated on each plot. They represent the instantaneous beat-to-beat pulse interval variability and the long-term continuous variability, respectively, for each subject over the 24-hour recording period. Although the Poincaré plots of the pulse interval are different within each group, it does not seem possible to classify shape or determine subject grouping from these plots. Poincaré plots of the systolic blood pressure were more consistent within each group than the pulse interval data.

In the case of more severely hypertensive subjects, Poincaré plots may provide complementary information to characterize hypertension. While the Poincaré plots themselves do not discriminate between subjects, the parameter SD2 from systolic blood pressure was extremely valuable in determining a-priori subject grouping.

POINCARÉ PLOTS OF BORDERLINE HYPERTENSIVE GROUP

PULSE INTERVAL
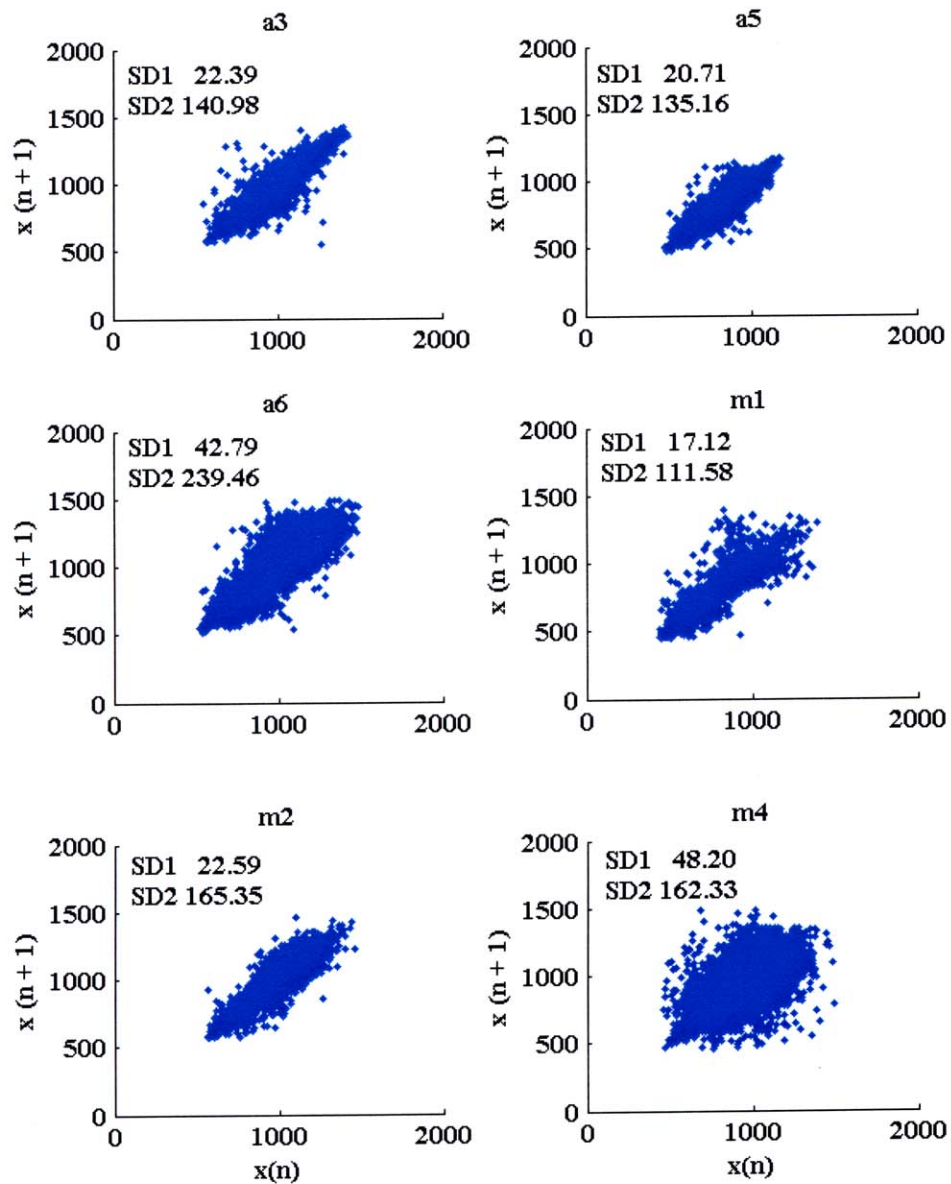
Figure 8.33 Individual Poincaré plots of pulse interval data in the borderline hypertensive group. Each plot indicates the instantaneous beat-to-beat systolic blood pressure variability (SD1) along with the long-term continuous systolic blood pressure variability (SD2) over the full 24-hour recording period. In all plots, the x-axis is x(n); subject identifiers appear above each plot.

POINCARÉ PLOTS OF NORMOTENSIVE GROUP

PULSE INTERVAL
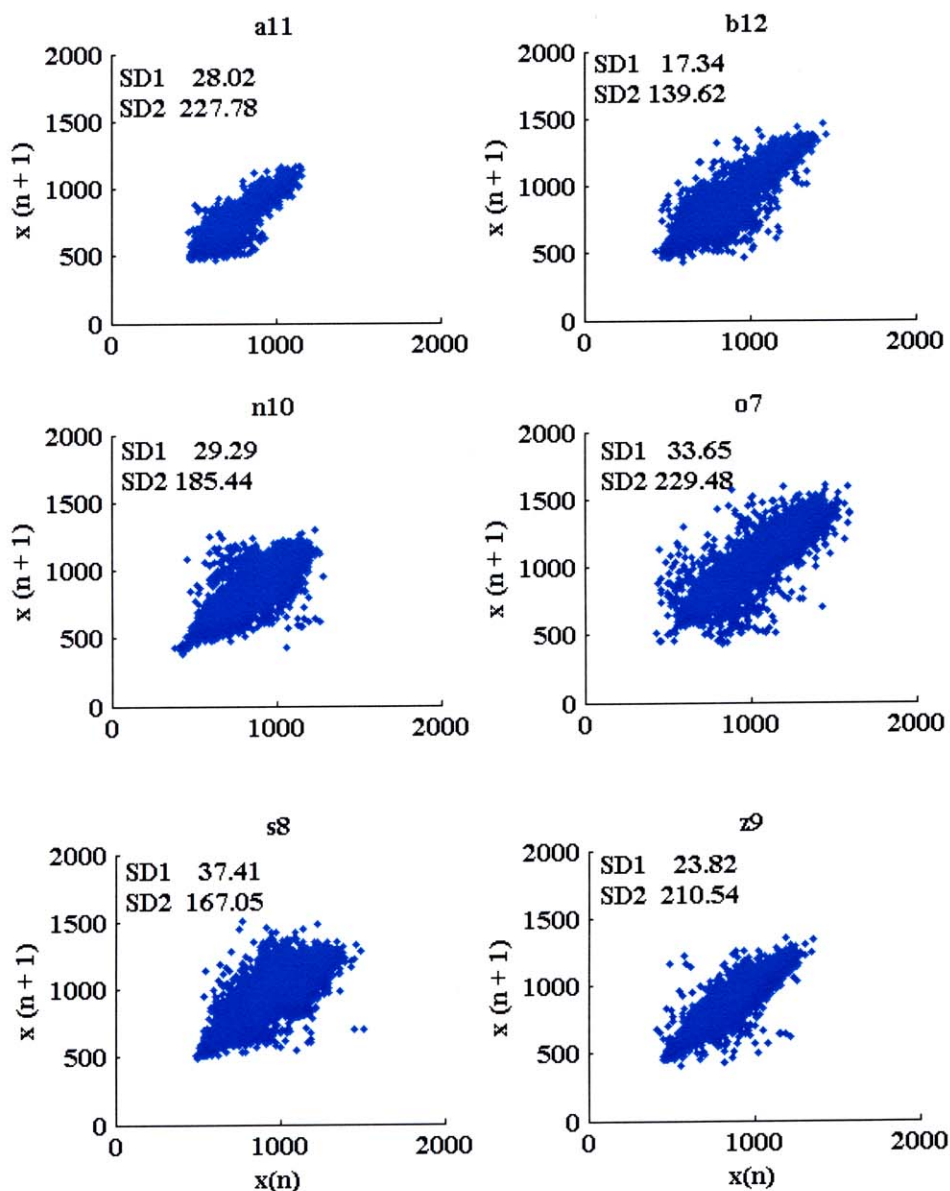
Figure 8.34 Individual Poincaré plots of pulse interval data in the normotensive group. Each plot indicates the instantaneous beat-to-beat pulse interval variability (SD1) along with the long-term continuous systolic blood pressure variability (SD2) over the full 24-hour recording period. In all plots, the x-axis is x(n); subject identifiers appear above each plot.

CHAPTER 9

CONCLUSIONS

Several hypotheses were put forward. One was that the use of nonlinear methods would provide valuable information to aid in the characterization of hypertension. Another was that the use of nonlinear methods would reveal information in the data that could not be revealed by linear methods. The third was that the combination of linear and nonlinear analysis methods applied to blood pressure data, may lead to advancements in the characterization of hypertension, and elucidate autonomic dysfunction. The overall goals of this thesis were then outlined as follows:

- Utilize linear and nonlinear analysis methods with blood pressure and pulse interval data.

- Evaluate results of nonlinear analyses to determine their efficacy as an aid in the characterization of hypertension.

- Assess the significance and clinical utility of these methods.

The first goal was accomplished by processing the data, as explained in detail in Chapter 5, and analyzing the data with linear and nonlinear methods. The linear and nonlinear methods that were used are outlined in Chapter 1, with detailed explanations in Chapter 3. This was followed by a-priori subject separation into two evenly sized groups based solely on the analysis results. The task of subject separation began by applying principal component and cluster analyses to the various results. This was followed by an exhaustive process requiring very careful consideration and evaluation of all cluster configurations. The two variables that lead to the determination of subject grouping

were detrended fluctuation analysis scaling exponent α_2, and SD2 derived from Poincaré mapping, α_2-SBP and SD2-SBP, respectively were analyses of hourly results from 24-hour systolic blood pressure data over all subjects. The variable that provided the strongest separation was α_2-SBP; while this variable exhibited strength of separation the two groups were uneven, with seven in one group and five in the other. The variable SD2-SBP separated the subjects with near identical groupings and comparable strength with the exception that subject division for this variable was even with six subjects in each group. The linear analyses were not instrumental in subject separation during any phase of this research. Had this cohort been more physiologically diverse, a-priori subject separation may have been more transparent and exhibited greater accuracy. Once subjects were separated into two groups, it was possible to fulfill the last two goals, to evaluate individual results, and determine their efficacy as an aid in the characterization of hypertension.

In addition to elevated sympathetic activity, previous literature has noted that blood pressure variability is elevated and heart rate variability is decreased in hypertension [76-78]. In terms of blood pressure variability, the borderline hypertensive group had, on average, a higher low frequency and lower high frequency component than the normotensive group. In low frequency power, the borderline hypertensive group exhibited elevated sympathetic activity as compared to the normotensive group. The parasympathetic response by the normotensive group to a lower magnitude sympathetic outflow was greater than that of the borderline hypertensive group with a higher sympathetic outflow. Although, the effects of respiration on the high frequency power for either group are impossible to define, the parasympathetic response in borderline

hypertension may be too slow, too weak or both to counteract the strong sympathetic activity evident in hypertension. Elevated blood pressure variability and increased sympathetic outflow were determined by the increased low frequency component of blood pressure variability. Statistically, the low frequency power from blood pressure variability analysis was significant to $p < 0.05$ for the full 24-hour recording period. The high frequency power was not significantly different for the full 24-hour period. However, during hours 16 through 21 the difference between groups was significant to $p < 0.05$. During those hours it was the normotensive group that exhibited elevated parasympathetic activity which led to the difference between the groups. Parasympathetic dominance in the normotensive group during these hours is evident in the LF/HF ratio of blood pressure variability. The BPV LF/HF ratio was significantly different to $p < 0.05$ between the hours 13 through 24.

In terms of heart rate variability, the borderline hypertensive group was found to have, on average, a lower low frequency component with a higher high frequency component. This behavior is the exact opposite of the findings in blood pressure variability. Using low frequency power as a determinant in heart rate variability, these findings indicate decreased heart rate variability in borderline hypertension. This is in contrast to one study where similar heart rate variability was reported in a similar cohort with decreased variability not evidenced until a more severe hypertensive state was reached [71].

As in blood pressure variability, the low frequency component of heart rate variability (HRV) was statistically significant to $p < 0.05$ over the full 24-hours. The high frequency power in HRV was significantly different between hours one through 18,

to p < 0.05 and the sympathovagal ratio, LF/HF, was significantly different to p < 0.05 over the full 24-hour recording period.

Baroreflex function was assessed by the alpha method, a spectral analysis of the baroreflex response. The low frequency alpha index, α_{LF}, was statistically significant over the full 24-hour period to p < 0.05, with the borderline hypertensive subjects exhibiting a lower baroreflex response than their normotensive counterparts. With the exception of a few hours, the normotensive group exhibited greater dispersion about the mean, which is interpreted as heightened responsiveness of the baroreceptor reflex in the normotensive group over the borderline hypertensive group. This was further illustrated in Figure 8.12 and agrees with previous studies that have shown baroreflex sensitivity to be negatively correlated with rising blood pressure [73-75].

The high frequency alpha index, α_{HF} was not statistically significant between groups. It is possible that this is due to respiration and/or the broadband approach to detecting coherence in the high frequency band. Respiration in this cohort may have frequently been in the low frequency range which would reduce the high frequency alpha index for both groups, rendering it insignificant.

Given previous research findings that blood pressure variability increases with increasing blood pressure [4], it was hypothesized that approximate entropy (ApEn) would be a valuable aid in the characterization of hypertension. Further, it was anticipated that ApEn would be elevated in borderline hypertensive versus normotensive subjects. The reasoning behind this assumption was due to the definition of approximate entropy. Approximate entropy is a regularity statistic that returns low values for regular behavior and high values for more random behavior. Therefore, given previous reports

of elevated blood pressure variability in hypertension, and the belief that these analyses would agree with previous research, a decrease in regularity was expected which would produce an elevated ApEn value. It was found that approximate entropy analysis of systolic blood pressure data was statistically significant to $p < 0.05$, with the borderline hypertensive group exhibiting higher ApEn values. The two groups are very closely matched physiologically. Approximate entropy confirmed this narrow physiologic difference by returning higher entropy values indicating increasing irregular behavior. The approximate entropy results compliment the blood pressure variability results. There was no statistical difference between groups when applying approximate entropy to pulse interval data. However, approximate entropy analysis of systolic blood pressure data has been shown here to be a valuable tool for use in the characterization of hypertension.

It was stated that due to the reported elevated blood pressure variability in hypertension, that it was believed the scaling exponents from detrended fluctuation analysis will be, on average, lower in the borderline hypertensive group than those of the normotensive group. It was postulated that this is due to the break down of the scaling behavior in hypertension as the system tends towards more random dynamics.

Detrended fluctuation analysis of systolic blood pressure data revealed statistically significant results in the short-term scaling exponent between the hours 13 through 24 with $p < 0.05$. It was clear in Figures 8.17 and 8.18 that the short-term scaling behavior was similar in the first half of the day and divergent in the last half. This observation was confirmed statistically. This may indicate that the short-term scaling exponent deteriorates as the day progresses. It may also be due to the time-scale over which the short-term scaling behavior was calculated. It is entirely possible that the

time-scale used to calculate α_1 was too short to fully capture short-term scaling behavior. The range used here was $4 <= n < 11$, which may need to be increased to capture short-term scaling behavior.

The long-term scaling exponent was statistically significant to $p < 0.05$ over the full 24-hour recording period. As stated earlier, the long-term scaling exponent was highly influential in a-priori subject separation. It was apparent in Figure 8.20, that the normotensive group maintained a higher long-term scaling exponent value over the course of the 24-hour recording. This may be an indication that the scaling behavior is beginning to deteriorate in borderline hypertension leading to more random dynamics. This view is validated by elevated approximate entropy values, which indicates greater random behavior, in borderline hypertension over the normotensive group.

Detrended fluctuation analysis of the pulse interval data for the short-term scaling exponent did not reveal statistically significant results. Analysis of the long-term scaling exponent however, revealed statistically significant results between hours one through 12 and again between hours 16 through 23 to $p < 0.05$. The only hours that were not statistically different were hours 13 through 15 and hour 24. Why these hours are similar to the normotensive group is unknown, nevertheless it may be due to the close physiological status of this cohort. It is also possible that the borderline hypertensive group may still present "normal" scaling behavior periodically through out the day. As the disease progresses, this oscillatory behavior may disappear altogether. If this is the case, detrended fluctuation analysis has shown that it can determine these differences and aid in the characterization of hypertension. This is especially true for the long-term scaling exponent.

Interestingly, scatter plots of α_1 versus α_2, revealed slightly different clustering patterns. These results indicate that in the borderline hypertensive group, a high α_1 value is more likely followed by a low α_2 value and conversely. Therefore, the cluster trend moves downward. The opposite is true in the normotensive group. A higher α_1 value is likely to be followed by an increasing α_2 value. The clustering trend in the normotensive group tends to move upward. This is highlighted in Figures 8.22 and 8.23. This may be an indication of the beginning of the deterioration of scaling properties in hypertension towards lower values and more random behavior. If these trends persist as the disease worsens, group membership could potentially be determined on a subject by subject basis by evaluation of these scatter plots. Detrended fluctuation analysis, particularly the long-term scaling exponent, has been shown to be a useful tool in the characterization of hypertension. Although the short-term scaling exponent from systolic blood pressure was not as statistically significant over the full 24-hours, its efficacy is demonstrated in the scatter plots. As noted earlier, it may be necessary to broaden the time scale, window size, over which the short-term scaling exponent is calculated. In this research window size for the short-term scaling exponent was $4 \leq n < 11$, this may be inadequate to fully capture short-term scaling behavior in this physiologically close cohort.

It was hypothesized that application of two dimensional Poincaré plots, would provide additional insight into the dynamics of blood pressure, and support the discrimination of subjects. It was further hypothesized that quantification of these plots would aid in the characterization of hypertension. Poincaré plots are shown in Figures 8.31 through 8.34. While the dynamics of systolic blood pressure and pulse interval data

are interesting, the shape of the Poincaré plots did not provide adequate information to sufficiently determine group membership.

Quantification of the Poincaré plots was extremely valuable in determining group membership. The parameter SD2, from analysis of systolic blood pressure data, provided strong support in determining group membership and was statistically significant to $p < 0.05$ over the full 24-hour period. The short-term parameter SD1 was significant over the first half of the day from hour one through 12.

Poincaré analysis of pulse interval data did not reveal statistically significant results between groups. Although it was believed that analysis of the pulse interval data would be significant between groups, this was not the case. This may be due to the manner in which the parameter was calculated. Heart rate and blood pressure variability were determined by Fourier analysis of small blocks of data that were averaged to provide hourly values for each subject. Poincaré results were determined from the full hour. The short-term variability parameter SD1, may not be as robust as the long-term variability parameter SD2, and may require smaller blocks of data to adequately determine short-term differences in pulse interval data.

Of the nonlinear methods used here, detrended fluctuation analysis and Poincaré maps have not been identified as ever having been applied in the study of blood pressure, certainly not in a cohort of borderline hypertensive and normotensive subjects. Approximate entropy was the only nonlinear method that was identified as having been previously used, in one study. That study used a pharmacological agent to study blood pressure in healthy individuals, not an actual disease state, and not on 24-hour blood pressure data. After utilization of six methods, three linear and three nonlinear, it was

the nonlinear methods that were crucial in determining group membership. Given that the linear results here concur with previously reported results, they were helpful in elucidating the nonlinear relationships, but they were not instrumental in a-priori subject separation. Significant differences were identified by all of the nonlinear measures used. These differences were predominantly determined from analysis of systolic blood pressure data versus pulse interval data. A summary of results and their significance are presented in Appendix B.

In addition to the novel use of nonlinear methods in blood pressure analysis, the presentation of results is also unique. Activity plots of approximate entropy, detrended fluctuation analysis and Poincaré parameters have not been seen in literature researched to date. Generally, presentation of these parameters is one gross measure versus temporal changes over the course of 24-hour recordings.

It was demonstrated that detrended fluctuation analysis is an extremely robust measure. Many physical and biological signals are noisy or rife with gaps which inevitably create problems during analysis. It was demonstrated here that detrended fluctuation analysis is an extremely robust measure capable of correctly detecting scaling behavior with large amounts of data removed. When discontinuities were introduced the data were rejoined with no nearest neighbor averaging. Detrended fluctuation analysis identified scaling behavior with 99% correlation compared with results from the full data set. Testing the effect of nonstationarities on detrended fluctuation analysis has been identified as having been done in only one previous study [67]. Although the testing performed here was not as in-depth as that performed previously [67], the findings here concur with that study.

The main theme of this thesis was the application of nonlinear analysis methods to systolic blood pressure data in an effort to improve characterization of hypertension. These analyses were applied to a cohort whose physiologic separation was very narrow, and still the nonlinear methods adequately characterized hypertension in its very early stages. Conversely, linear methods were not helpful in characterization of hypertension during any phase of this research. Application of these same nonlinear analysis methods in a cohort whose physiologic state is broader, may further advance the utility of nonlinear analysis toward an improved understanding of the pathophysiology of hypertension.

CHAPTER 10

FUTURE WORK

The results here provide some insight into the dynamics between a closely matched cohort of borderline hypertensive and normotensive subjects. Because analysis of blood pressure data by nonlinear methods is vastly underutilized, expansion of this work is almost endless. Several of the analysis possibilities are described.

Analysis of diastolic blood pressure, with the techniques utilized here, may expand the scope of knowledge and provide additional information regarding hypertension. It is possible, that diastolic blood pressure may well exhibit slightly different dynamics than systolic blood pressure. If these differences do exist, understanding how and why would provide a broader range of information and perhaps reveal additional characteristics of the pathophysiology of hypertension. As such combined analysis of systolic and diastolic blood pressure could provide additional information to aid in the determination of the onset of hypertension.

Pulse pressure is increasingly being recognized as a risk factor for cardiovascular disease [79-82]. Stroke volume and the properties of arterial circulation directly affect pulse pressure. Because blood pressure has a direct affect on arterial wall elasticity, analysis of pulse pressure data in a hypertensive cohort with nonlinear methods, might provide additional information to determine borderline hypertensive individuals from their normotensive counterparts.

While Poincaré plots of systolic blood pressure did nothing to clarify the nature of hypertension in a very closely matched cohort, similar plots of diastolic blood pressure or pulse pressure may. Further these data have never been identified as having been

analyzed by any nonlinear method. Given the positive results here it is reasonable to believe inclusion of these data would also provide additional information.

Quantitatively, Poincaré maps proved very useful in a very closely matched cohort. Poincaré plots created from a more severe hypertensive state may reveal dynamics that are not yet present in the borderline hypertensive state. Additionally, information may be gleaned by comparing the overall shape of the maps between cohorts that are more physiologically diverse.

The Poincaré parameter SD2 from analysis of systolic blood pressure was extremely useful while the parameter SD1 was not. Decreasing the data size for which the SD1 parameter is calculated may improve its ability to characterize hypertension and aid in understanding the dynamics of hypertension.

Scatter plots of SD1 versus SD2, similar to those created with detrended fluctuation analysis $\alpha 1$ versus $\alpha 2$, are worth investigating and may lead to additional information as to clustering of data between groups.

A more comprehensive investigation of sleep versus wake of all of the analysis methods used here through the use of fixed or sliding time windows may improve characterization of hypertension. Further this type of analysis may define timeframes over which data collection is paramount, thus avoiding the need for 24-hour recordings. Transitions between sleep/wake and wake/sleep have been shown to improve classification. Analysis by fixed or sliding windows may further elucidate these differences and provide insight into hypertension pathophysiology.

Comparison of the total area under each frequency curve between subjects may provide information that will aid in classification of hypertension and provide additional information as to the pathophysiology of the disease.

Evaluation of data during specific activities with both linear and nonlinear analyses may aid in determining characteristics of the system and how they differ between hypertension and controls.

R-WAVE AND BLOOD PRESSURE PEAK DETECTOR
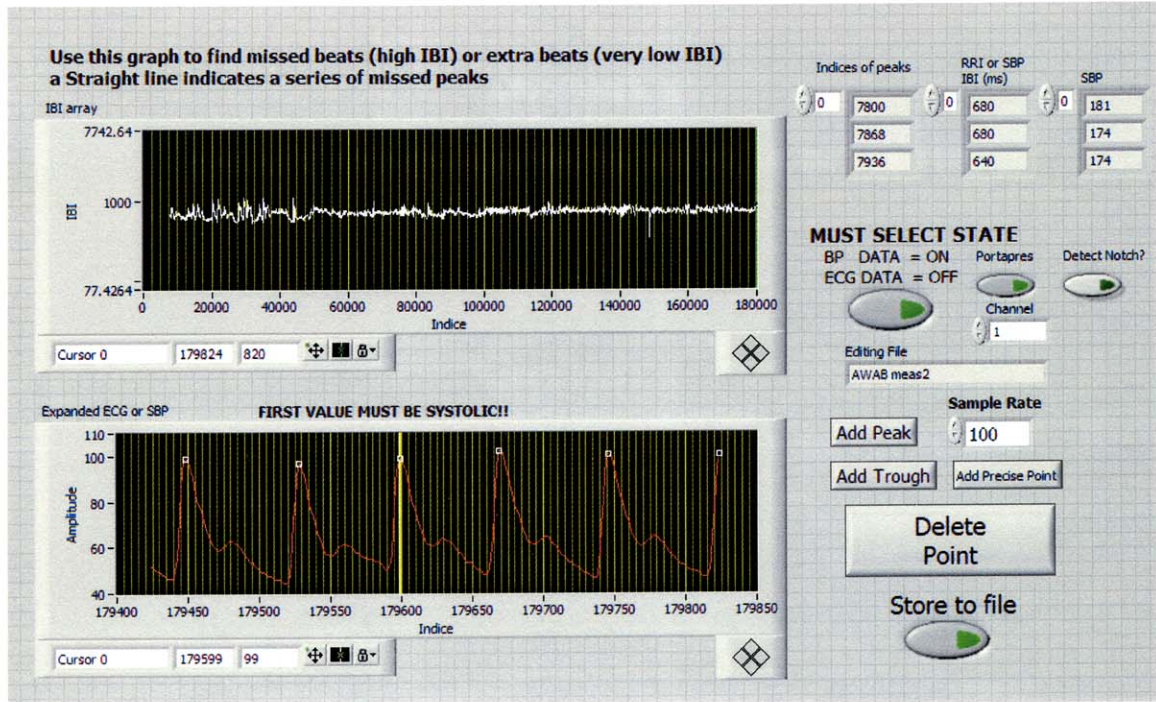
Figure A.1 Partial front panel view of R-Wave and Blood Pressure Analyzer.

A partial front panel view of the custom R-wave and blood pressure LabVIEW® program is shown in Figure A.1. The program requires initialization to correctly process input data. The program must be told what type of data is being analyzed; the sampling rate, the channel that contains the data, and whether or not the data was recorded by a Portapres. In order for proper output these buttons must be selected prior to running the program.

The top graph is the interbeat interval from the detected peaks. This aids in identifying any errors in peak detection. If there are missed beats there will be a spike in the top graph indicating possible missed beats or calibrations. If a downward spike is

present in the IBI this may indicate premature or dual peak detection. The bottom graph is an expanded window of the input data. This provides a better view of the peak detection and a window for editing. There are cursors on each graph. The cursor on the top IBI graph is tied to both graphs thus when it is moved both graphs scroll. This enables quick alignment of the graphs for inspection and editing if necessary. The cursor on the bottom expanded view graph is associated with that graph only. This cursor is used to move around in the current expanded window. If a peak is missed moving the bottom cursor downstream of the missed peak and hitting the button "Add Peak" inserts the missed point. If a peak was identified twice or noise in the data caused erroneous peak detection positioning the cursor upstream of the point to be removed and hitting the button "Delete Point" removes the offending point. Deletion of a point is also possible by positioning the cursor directly on the point to be removed then hitting the "Delete Point" button. A precise point can also be added if needed by positioning the cursor and using the "Add Precise Point" button.

Following any necessary editing and activation of the "Store to File" button the program writes the data to file. For this work blood pressure data are written to file which are the systolic blood pressure values in millimeters of mercury (mm Hg), an interbeat interval array in milliseconds (ms) and header information.. The header information includes the name of the raw data file that generated the data arrays, the sample rate as well as titles for each data column written to file.

It is impossible to illustrate the block diagram for the main program given its size. It is as impossible to list and describe all the supporting subroutines written for the same program. Therefore in order to provide an indication of the scope of work necessary to originally create to program for R-to-R peak detection and to modify it to include systolic blood pressure peak detection the program hierarchy is shown in Figure 9.2.
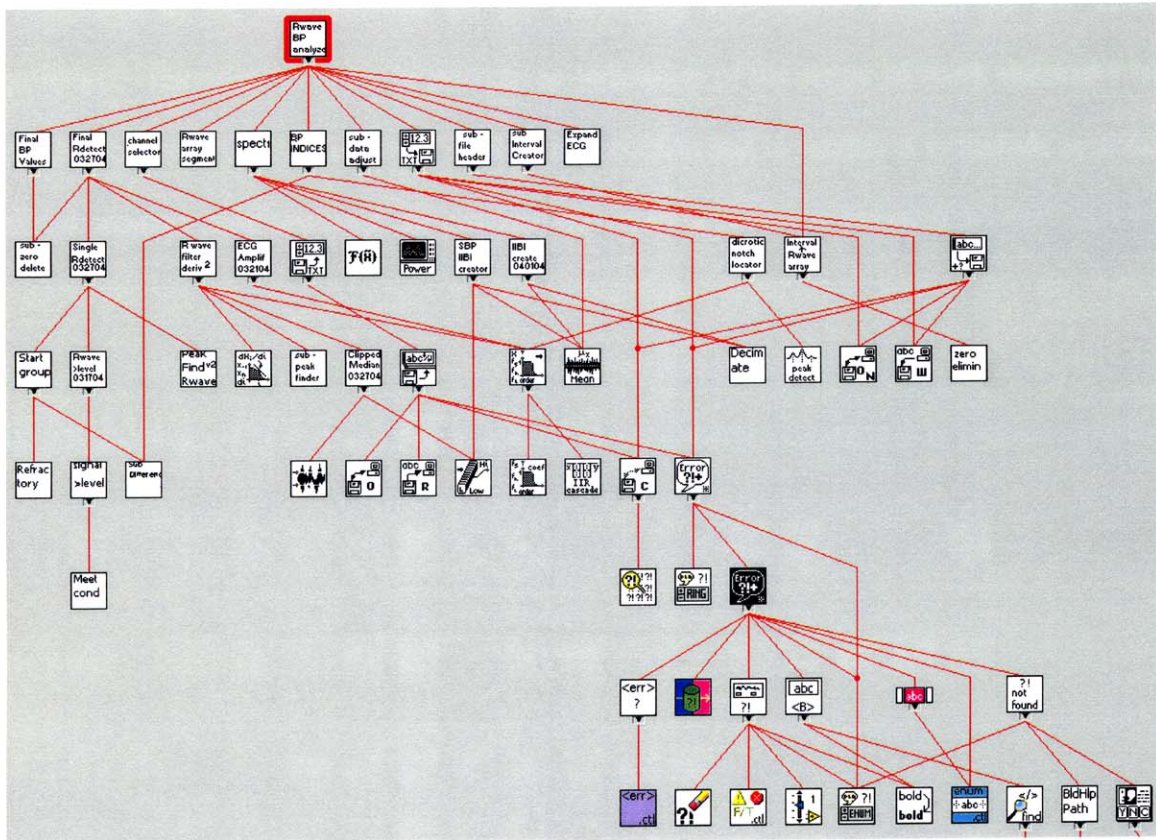
Figure A.2 Hierarchy of R-Wave and Blood Pressure peak detection program.

As noted earlier, the personnel working in and with the Signals and Systems lab developed the original program to detect R-waves in electrocardiogram data. This program was heavily modified to include the detection of systolic blood pressure peaks.

APPENDIX B

ANALYSIS BREAKDOWN

The various analysis methods, and the parameters associated with each, are summarized. A "yes" in column three indicates the analysis results for that parameter were significant over the full 24-hour period. If a parameter was not statistically significant over the full 24-hour period, but was significant for a block or blocks of hours, column four indicates the hours for which those parameter results were statistically significant. An "ns" indicates not significant. Statistical significance is defined as $p < 0.05$.

Table B.1 Analysis Breakdown by Parameter and Significance

Analysis Method	Parameter	24-hours	Hours Significant
Blood Pressure Variability	LF	yes	
	HF		16 → 21
	LF/HF		13 → 24
Heart Rate Variability	LF	yes	
	HF		1 → 18
	LF/HF	yes	
Alpha Index	LF	yes	
	HF	ns	ns
Approximate Entropy	SBP	yes	
	PI	ns	ns
Detrended Fluctuation	α_1 SBP		13 → 24
	α_2 SBP	yes	
	α_1 PI	ns	ns
	α_2 PI		1 →12 & 16 → 23
Poincaré	SD1-SBP		1 → 12
	SD2-SBP	yes	
	SD1-PI	ns	ns
	SD2-PI	ns	ns

APPENDIX C

SUBJECT LOG FILES

The log files indicating the daily activities for all twelve subjects as provided by Columbia researchers are shown. Each log file is marked with subject identification and number. Subject group membership was added for clarity. The units of time are hours and are shown in military notation.

Subject: a3 #1   Borderline Hypertensive

Time	Activities
1	Sleep
2	Sleep
3	Sleep
4	Sleep
5	Sleep
6	Wake up go to bathroom
7	Back to sleep
8	Wake up, sitting, eating breakfast
9	Sitting, watching TV.
10	Same as 9
11	Same as 9
12	Same as 9
13	Same as 9 +snacks
14	Watching TV
15	Phone calls
16	Phone calls
17	Lunch
18	Reading
19	Working on PC
20	Same as 19
21	Watching TV +Snacks
22	Go to sleep
23	sleep
24	sleep

Subject: a5 # 2   Borderline Hypertensive

Time	Activities
1	Sleep
2	Sleep
3	Sleep
4	Sleep
5	Sleep
6	Wake up go to bathroom
7	Back to sleep
8	Wake up, sitting, eating breakfast
9	Sitting, watching TV
10	Same as 9
11	Same as 9
12	Same as 9
13	Same as 9 +snacks
14	Chatting
15	Chatting
16	Chatting
17	Lunch
18	Reading
19	Working on PC
20	Same as 19
21	Watching TV +Snacks
22	Go to sleep
23	sleep
24	sleep

Subject: a6 # 3   Borderline Hypertensive

Time	Activities
1	Sleep
2	Sleep
3	Sleep
4	Sleep
5	Sleep
6	Wake up go to bathroom
7	Back to sleep
8	Wake up, sitting, eating breakfast
9	Sitting, PC work
10	Same as 9
11	Same as 9
12	Same as 9
13	Same as 9 +snacks
14	Chatting
15	Chatting
16	Chatting
17	Lunch
18	Reading
19	Working on PC
20	Same as 19
21	Watching TV
22	Go to sleep
23	sleep
24	sleep

Subject: m1 # 6   Borderline Hypertensive

Time	Activities
1	Sleep
2	Sleep
3	Sleep
4	Sleep
5	Sleep
6	Wake up go to bathroom
7	Back to sleep
8	Wake up, sitting, eating breakfast
9	Sitting, studying M.D. book.
10	Same as 9
11	Same as 9
12	Same as 9
13	Same as 9 +snacks
14	Watching TV
15	Chatting ,Watching TV
16	Chatting, Watching TV
17	Lunch
18	Reading
19	Working on PC
20	Same as 19
21	Watching TV +Snacks
22	Go to sleep
23	sleep
24	sleep

Subject: m2 # 7   Borderline Hypertensive

Time	Activities
1	Sleep
2	Sleep
3	Sleep
4	Sleep
5	Sleep
6	Wake up go to bathroom
7	Back to sleep
8	Wake up, sitting, eating breakfast
9	Sitting, watching TV
10	Same as 9
11	Same as 9
12	Same as 9
13	Same as 9 +snacks
14	Chatting
15	Chatting
16	Chatting
17	Lunch
18	Reading
19	Working on PC
20	Same as 19
21	Watching TV +Snacks
22	Go to sleep
23	sleep
24	sleep

Subject: m4 # 8   Borderline Hypertensive

Time	Activities
1	Sleep
2	Sleep
3	Sleep
4	Sleep
5	Sleep
6	Sleep
7	Wake up go to bathroom
8	Wake up, sitting, eating breakfast
9	Sitting, PC work
10	Same as 9
11	Same as 9
12	Same as 9
13	Same as 9 +snacks
14	chatting
15	chatting
16	Lunch
17	Reading
18	Reading
19	Working on PC
20	Chatting
21	Watching TV+ eating banana
22	Go to sleep
23	sleep
24	sleep

Subject: a11 # 4   Normotensive

Time	Activities
1	Sleep
2	Sleep
3	Sleep
4	Sleep
5	Sleep
6	Wake up go to bathroom
7	Back to sleep
8	Wake up, sitting, eating breakfast
9	Sitting, PC work
10	Same as 9
11	Same as 9
12	Same as 9
13	Same as 9 +snacks
14	Chatting
15	Chatting
16	Chatting
17	Lunch
18	Reading
19	Working on PC
20	Same as 19
21	Watching TV
22	Go to sleep
23	sleep
24	sleep

Subject: b12 # 5   Normotensive

Time	Activities
1	Sleep
2	Sleep
3	Sleep
4	Sleep
5	Sleep
6	Wake up go to bathroom
7	Back to sleep
8	Wake up, sitting, eating breakfast
9	Sitting, watching TV
10	Same as 9
11	Same as 9
12	Same as 9
13	Same as 9 +snacks
14	Chatting
15	Chatting
16	Chatting
17	Lunch
18	Reading
19	Working on PC
20	Same as 19
21	Watching TV +Snacks
22	Go to sleep
23	sleep
24	sleep

Subject: n10 # 9   Normotensive

Time	Activities
1	Sleep
2	Sleep
3	Sleep
4	Sleep
5	Sleep
6	Wake up go to bathroom
7	Back to sleep
8	Wake up, sitting, eating breakfast
9	Sitting, watching TV
10	Same as 9
11	Same as 9
12	Same as 9
13	Same as 9 +snacks
14	Watching TV
15	Chatting ,Watching TV
16	Chatting, Watching TV
17	Lunch
18	Reading
19	Working on PC
20	Same as 19
21	Watching TV +Snacks
22	Go to sleep
23	sleep
24	sleep

Subject: o7   # 10   Normotensive

Time	Activities
1	Sleep
2	Sleep
3	Sleep
4	Sleep
5	Sleep
6	Wake up go to bathroom
7	Back to sleep
8	Wake up, sitting, eating breakfast
9	Sitting, watching TV
10	Same as 9
11	Same as 9
12	Same as 9
13	Same as 9 +snacks
14	Chatting
15	Chatting
16	Chatting
17	Lunch
18	Reading
19	Working on PC
20	Same as 19
21	Watching TV +Snacks
22	Go to sleep
23	sleep
24	sleep

Subject: s8 # 11   Normotensive

Time	Activities
1	Sleep
2	Sleep
3	Sleep
4	Sleep
5	Sleep
6	Wake up go to bathroom
7	Back to sleep
8	Wake up, sitting, eating breakfast
9	Sitting, reading a book
10	Same as 9
11	Same as 9
12	Same as 9
13	Same as 9 +snacks
14	chating
15	lunch
16	Phone calls
17	Phone calls
18	Reading
19	Working on PC
20	Same as 19
21	Chewing gum
22	Go to sleep
23	sleep
24	sleep

Subject: z9 # 12   Normotensive

Time	Activities
1	Sleep
2	Sleep
3	Sleep
4	Sleep
5	Sleep
6	Wake up go to bathroom
7	Back to sleep
8	Wake up, sitting, eating breakfast
9	Sitting, PC work
10	Same as 9
11	Same as 9
12	Same as 9
13	Same as 9 +snacks
14	Watching a final soccer game(he was very excited)
15	Same as 14
16	Same as 14
17	Phone calls
18	Lunch
19	Working on PC
20	Same as 19
21	Phone calls+ snacks
22	Go to sleep
23	sleep
24	sleep

APPENDIX D

A-PRIORI CLUSTER RESULTS

Cluster results from variables α_2 SBP and SD2-SBP from detrended fluctuation and Poincaré analyses, respectively, are shown. These variables and the process of determining group membership are discussed in Chapter 7, Section 7.3.

There are three columns for each variable. The first column indicates cluster membership, the second column specifies subject identification and the third column illustrates the strength of separation. Subject identification is discussed in Chapter 6 and illustrated in Table 6.1. Cluster analysis is discussed in Chapter 6, Section 6.3. Detrended fluctuation and Poincaré analyses are discussed in Chapter 3, Sections 3.2 and 3.3, respectively.

α_2-SBP			SD2 SBP		
Cluster	Subject	Strength	Cluster	Subject	Strength
1	1	0.7452	1	1	0.8145
1	2	0.9023	1	2	0.6418
1	4	0.5309	1	4	0.8395
1	5	0.883	1	5	0.8586
1	10	0.8009	1	10	0.895
1	11	0.8671	1	11	0.8353
1	12	0.4458	2	3	0.4691
2	3	0.5307	2	6	0.5729
2	6	0.5626	2	7	0.5127
2	7	0.8399	2	8	0.5849
2	8	0.6708	2	9	0.733
2	9	0.7661	2	12	0.7666

Figure D.1 a-priori cluster results of variables α_2-SBP and SD2-SBP from detrended fluctuation and Poincaré analysis, respectively.

APPENDIX E

A-POSTERIORI CLUSTER RESULTS

A-posteriori analysis is discussed in Chapter 7, Section 7.4. All of the results shown here represent cluster results from a-posteriori analysis of blocks of hours from the combined matrix. The combined matrix and all data configurations are discussed in Chapter 6. All data for a-posteriori analysis was analyzed consistent with a-priori analyses. The data were first analyzed with principal component analysis followed by cluster analysis. The only cluster results that were considered during a-posteriori analysis were those that separated the cohort into two evenly sized groups of six each. All of the results shown improved the accuracy of subject separation from 67%, or four of six subjects correctly identified per group, to 83%, five of six subjects correctly identified per group.

The first result that improved upon subject separation is shown in Figure E.1. These results represent analysis of the first eight hours of the combined matrix. Subject activity during this time was sleep with transition to waking in the 7th and 8th hours. The only variable that improved upon subject separation during this period was $\alpha 2$-SBP from detrended fluctuation analysis. This variable was instrumental in the original determination of group membership based solely on the results. Cluster results for this variable during the first eight hours (Figure E.1) do not exhibit the strength of clustering shown in Appendix D (Figure D.1) for the full 24-hour period for the same variable. Cluster results which originally determined group membership are shown in Appendix D and discussed in Chapter 7.

α_2-SBP		
Cluster	Subject	Strength
1	1	0.5925
1	2	0.182
1	6	0.7296
1	7	0.5001
1	8	0.3384
1	9	0.6709
2	3	0.7124
2	4	0.8622
2	5	0.8335
2	10	0.541
2	11	0.5413
2	12	0.8453

Figure E.1 A-posteriori cluster results of variable α_2-SBP from detrended fluctuation analysis as applied to the first eight hours of the combined matrix. Subject separation determined by this variable during this time was 83% accurate.

Figure E.2 shows cluster results from analysis of the last eight hours of the combined matrix. The only variable that improved upon subject separation during this time was approximate entropy. Although this variable did improve separation accuracy, the strength of separation is generally low. Approximate entropy (ApEn) was not instrumental in the original determination of subject grouping. Subject activity during the last eight hours is indicated in subject logs in Appendix C.

Cluster	ApEn-SBP Subject	Strength
1	1	-0.0385
1	2	0.5617
1	3	0.5841
1	7	0.5551
1	8	0.0209
1	9	0.197
2	4	0.3355
2	5	0.4908
2	6	0.4078
2	10	0.262
2	11	0.6953
2	12	0.7306

Figure E.2 A-posteriori cluster results of variable ApEn-SBP from approximate entropy analysis as applied to the last eight hours of the combined matrix. Subject separation determined by this variable during this time was 83% accurate.

Figure E.3 shows cluster results from analysis of the final four hours of the combined matrix. The only variable that improved upon subject separation during the final four hour analysis was SD2-IBI from Poincaré quantification. Six four hour blocks of data were analyzed. Only the final four hour block improved upon subject separation. The remaining four hour blocks did not separate subjects beyond the original 67% accuracy rate. The only clusters that were considered when determining improvement in subject separation were those that successfully separated subjects evenly, six in each group.

Cluster	SD2 IBI Subject	Strength
1	5	0.5913
1	6	0.5362
1	9	0.3534
1	10	0.5451
1	11	0.2439
1	12	0.3375
2	1	0.3252
2	2	0.8206
2	3	0.6749
2	4	0.8189
2	7	0.721
2	8	0.83

Figure E.3 A-posteriori cluster results of variable SD2-SBP from Poincaré analysis as applied to the last four hours of the combined matrix. Subject separation determined by this variable during this time was 83% accurate.

Subsequent analysis of the final five hours of the combined matrix did not produce any variables that improved subject separation. During analysis of the final six hours of the combined matrix variable α_1-IBI from detrended fluctuation analysis indicated an improved separation. During this time this was the only variable that improved the accuracy of subject separation. These results are shown in Figure E.4.

| | α₁-PI | |
Cluster	Subject	Strength
1	1	0.824
1	2	0.7465
1	3	0.2621
1	4	0.7843
1	7	0.7544
1	8	0.5621
2	5	0.8508
2	6	0.8289
2	9	0.4448
2	10	0.8484
2	11	0.4665
2	12	0.7343

Figure E.4 A-posteriori cluster results of variable α_1-IBI from detrended fluctuation analysis as applied to the last six hours of the combined matrix. Subject separation determined by this variable during this time was 83% accurate

Additional analyses were performed a-posteriori and are discussed in Chapter 7, Section 7.4. Only those variables that improved upon subject separation are shown in Appendix E.

REFERENCES

[1]   S. Eyal, Y. Almog, O. Oz, S. Eliash, S. Akselrod S., "Nonlinear dynamics applies to blood pressure control," *Autonomic Neuroscience: Basic & Clinical*, vol. 89, pp. 24-30, 2001.

[2]   T.H. Makikallio, J.M. Tapanainen, M.P. Tulppo, H.V. Huikuri, "Clinical Applicability of Heart Rate Variability Analysis by Methods Based on Nonlinear Dynamics," *Cardiac Electrophysiology Review*, vol. 6, pp. 250-255, 2002.

[3]   G. Parati, A Frattola, M. Di Rienzo, G. Mancia G., "Blood Pressure Variability Importance in Research and in Clinical Hypertension," *BPA research and clinical hypertension*, vol. 67, no. 2, pp. 131-133, 1996.

[4]   G. Parati, "Blood pressure variability: its measurement and significance in hypertension," *Journal of Hypertension*, vol. 23, pp. S19-S25, 2005.

[5]   C.K. Peng, S. Havlin, H.E. Stanley, A.L. Goldberger, "Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series," *Chaos*, vol. 5, no. 1, pp. 82-87, 1995.

[6]   Joint National Committee on Prevention Detection, Evaluation, and Treatment of High Blood Pressure, "The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure," *JAMA* vol. 289, pp. 2560-2572, 2003.

[7]   O. Oz, S. Eliash, S. Cohen, S. Akselrod, "The Effect of Changes in Blood Volume on Low Frequency Blood Pressure Fluctuations in Spontaneously Hypertensive Rats," *IEEE Computers in Cardiology, 1989 Proceedings*, pp. 61-64, 1989.

[8]   S. Eliash, O. Oz, S. Cohen, S. Akselrod, "The role of renin-angiotensin and alpha control in the regulation of blood pressure in a normotensive versus a hypertensive system," *IEEE Computers in Cardiology, 1990 Proceedings*, pp. 155-158, 1990.

[9]   S. Akselrod, O. Oz, M. Greenberg, L. Kesselbrener, "Autonomic response to change of posture among normal and mild-hypertensive adults: Investigation by time-dependent spectral analysis, "*Journal of the Autonomic Nervous System*, vol. 64, pp. 33-43, 1997.

[10] S. C. Malpas, B. L. Leonard, S. J. Guild, J.V. Ringwood, M. Navakatikyan, P. C. Austin, G. A. Head, D. E. Burgess, "The Sympathetic Nervous System's Role in Regulating Blood Pressure Variability," *IEEE Engineering in Medicine and Biology*, pp. 17-24, 2001.

[11] R. L. Davrath, Y. Goren, I. Pinhas, E. Toledo, S. Akselrod, "Early autonomic malfunction in normotensive individuals with a genetic predisposition to essential hypertension," Am. J. Physiol. Heart Circ. Physiol., vol. 285, pp. H1697-H1704, 2003.

[12] S. M. Pincus, "Approximate entropy as a measure of system complexity," *Proc. Natl. Acad. Sci. USA*, vol. 88, pp. 2297-2391, 1991.

[13] S. M. Pincus, A. L. Goldberger, "Physiological time-series analysis: what does regularity quantify?," *Am. J. Physiol. Heart Circ. Physiol.*, vol. 266, pp. H1643-H1656, 1994.

[14] A. Voss, J. Kurths, H. J. Kleiner, A. Witt, N. Wessel, P. Saparin, K. J. Osterziel, R. Schurath, R. Dietz, "The application of methods of non-linear dynamics for the improved and predictive recognition of patients threatened by sudden cardiac death," *Cardiovascular Research*, vol. 31, pp. 419-433, 1996.

[15] S. Vickman, T. H. Makikallio, S. Yli-Mayry, S. Pikkujamsa, A. Koivisto, P. Reinikainen, K. E. Airaksinen Juhani, H. Huikuri H., "Altered Complexity and Correlation Properties of R-R Interval Dynamics Before the Spontaneous Onset of Paroxysmal Atrial Fibrillation," *Circulation*, vol. 100, pp. 2079-2084, 1999.

[16] T. H. Makikallio, H. V. Huikuri, A. Makikallio, L. F. Sourander, R. D. Mitrani, A. Castellanos, R. J. Myerburg, "Prediction of Sudden Cardiac Death by Fractal Analysis of Heart Rate Variability in Elderly Subjects," *J. Am. Coll. Cardiol.*, vol. 37, no. 5, pp. 1395-1402, 2001.

[17] J. K. Kanters, M. V. Hojgaard, E. Agner, N. H. Holstein-Rathlou N. H., "Short-and long-term variations in non-linear dynamics of heart rate variability," *Cardiovascular Research*, vol. 31, pp. 400-409, 1996.

[18] P. Mansier, J. Clairambault, N. Charlotte, C. Medigue, Ch. Vermeiren, G. LePape, F. Carre, A. Gounaropoulou, B. Swynghedauw, "Linear and non-linear analyses of heart rate variability: a minireview," *Cardiovascular Research*, vol. 31, pp. 371-379, 1996.

[19] T. H. Makikallio, T. Seppanen, K. E. Juhani Airaksinen, J. Koistinen, M. Tulppo, C. K. Peng, A. L. Goldberger, H. V. Huikuri H. V., "Dynamic Analysis of Heart Rate May predict Subsequent Ventricular Tachycardia After Myocardial Infarction," *The American Journal of Cardiology*, vol. 80, pp. 779-783, 1997.

[20] S. Kagiyama, A. Tsukashima, I. Abe, S. Fujishima, S. Ohmori, U. Onaka, Y. Ohya, K. Fujii, T. Tsuchihashi, M. Fujishima M., "Chaos and spectral analyses of heart rate variability during head-up tilting in essential hypertension," *Journal of the Autonomic Nervous System*, vol. 76, pp. 153-158, 1999.

[21] F. Lombardi, "Chaos Theory, Heart Rate Variability, and Arrhythmic Mortality," *Circulation*, pp. 8-10, 2000.

[22] P. van Leeuwen, H. Betterman, "The status of nonlinear dynamics in the analysis of heart rate variability," *Herzschr Elektrophys.*, vol. 11, pp. 127-130, 2000.

[23] M. Di Rienzo, P. Castiglioni, G. Mancia, G. Parati, A. Pedotti, "Critical Appraisal of Indices for the Assessment of Baroreflex Sensitivity," *Methods of Information in Medicine*, vol. 36, pp. 246-249, 1997.

[24] M. Di Rienzo, P. Castiglioni, G. Mancia, A. Pedotti, G. Parati, "Advancements in Estimating Baroreflex Function," *IEEE Engineering in Medicine and Biology*, pp. 25-32, 2001.

[25] G. Parati, M. Di Rienzo, G. Mancia, "How to measure baroreflex sensitivity: From the cardiovascular laboratory to daily life," *Journal of Hypertension*, vol. 18; no. 1, pp. 7-19, 2000.

[26] G. Parati, M. Di Rienzo, G. Mancia, "Dynamic Modulation of Baroreflex Sensitivity in Health and Disease," *Annals New York Academy of Sciences*, pp. 469-487.

[27] C. K. Peng, S. Buldyrev, S. Havlin, M. Simons, H. E. Stanley, A. L. Goldberger, "Mosaic organization of DNA nucleotides," *Physical Review E*, vol. 49, no. 2, pp. 1685-1689, 1994.

[28] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, PCh. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C. K. Peng, H. E. Stanley H. E., "PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals," *Circulation,* vol. 101, no. 23, pp. 215-220, 2000.

[29] G. A. Head, E. V. Lukoshkova, S. L. Burke, S. C. Malpas, E. A. Lambert, B. J. A. Janssen, "Comparing Spectral and Invasive Estimates of Baroreflex Gain," *IEEE Engineering in Medicine and Biology*, pp. 43-52, 2001.

[30] R. L. Hughson, L. Quintin, G. Annat, Y. Yamamoto, C. Gharib C., "Spontaneous baroreflex by sequence and power spectral methods in humans," *Clin. Physiol.*, vol. 13, pp. 663-676, 1993.

[31] G. Beevers, G. Lip, E. O'Brien E., "The pathophysiology of hypertension," *BMJ,* vol. 322, pp. 912-916, 2001.

[32] G. C. Butler, Y. Yamamoto, R. Hughson, "Fractal nature of short-term systolic BP and HR variability during lower body negative pressure," *Am. J. Physiol.*, vol. 267, pp. R26-R33, 1994.

[33] C. D. Wagner, R. Mrowka, B. Nafz, P. B. Persson, "Complexity and "chaos" in blood pressure after baroreceptor denervation of conscious dogs," *Am. J. Physiol. Heart Circ. Physiol.*, vol. 269, pp. H1760-H1766, 1995.

[34] C. D. Wagner, B. Nafz, P. B. Persson, "Chaos in blood pressure control," *Cardiovascular Research*, vol. 31, pp. 380-387, 1996.

[35] Y. Almog, S. Eliash, O. Oz, S. Akselrod, "Nonlinear analysis of BP signals Can it detect malfunctions in BP control?," *Am. J. Physio.*, vol. 271, pp. H396-H403, 1996.

[36] D. Mestivier, N. Phong Chau, S. Chanudet, B. Bauduceau, P. Larroque, "Relationship between diabetic autonomic dysfunction and heart rate variability assessed by recurrence plot," *Am. J. Physiol.*, vol. 272, pp. H1094-H1099, 1997.

[37] D. Mestivier, H. Dabire, N. Phong Chau, "Effects of autonomic blockers on linear and nonlinear indexes of blood pressure and heart rate in SHR," *Am. J. Physiol. Heart Circ. Physiol.*, vol. 281, pp. H1113-H1121, 2001.

[38] J. J. Gonzalez, J. J. Cordero, M. Feria, E. Pereda, "Detection and sources of nonlinearity in the variability of cardiac R-R intervals and blood pressure in rats," *Am. J. Physiol. Heart Circ. Physiol.*, vol. 279, pp. H3040-H3046, 2000.

[39] H. Dabire, D. Mestivier, J. Jarnet, M. Safar, N. Phong Chau, "Quantification of sympathetic and parasympathetic tones by nonlinear indexes in normotensive rats," *AJP – Heart*, vol. 275, pp. 1290-1297, 1998.

[40] T. T. Jartti, T. A. Kuusela, T. J. Kaila, K. U. O. Tahvanainen, I. A. T. Valimaki, "The dose-response effects of terbutaline on the variability, approximate entropy and fractal dimension of heart rate and blood pressure," *Br. J. Clin. Pharmacol.*, vol. 45, pp. 277-285, 1998.

[41] T. A. Kuusela, T. T. Jartti, K. U. O. Tahvanainen, T. J. Kaila, "Nonlinear methods of biosignal analysis in assessing terbutaline-induced heart rate and blood pressure changes," *Am. J. Physiol. Heart Circ. Physiol.*, vol. 282, pp. H773-H781, 2002.

[42] Steve Pincus, "Approximate entropy (ApEn) as a complexity measure," *Chaos*, vol. 5, pp. 110-117, 1995.

[43] S. M. Pincus, "Approximate entropy in cardiology," *Herzschr Elektrophys.*, vol. 11, pp. 139-150, 2000.

[44] D. Cysarz, H. Betterman, P. van Leeuwen, " Entropies of short binary sequences in heart period dynamics," *Am. J. Physiol. Heart Circ. Physiol.*, vol. 278, pp. H2163-H2172, 2000.

[45] T. Penzel, J. W. Kantelhardt, L. Grote, J. H. Peter, A. Bunde, "Comparison of Detrended Fluctuation Analysis and Spectral Analysis for Heart Rate Variability in Sleep and Sleep Apnea," *IEEE Transactions on Biomedical Engineering*, vol. 50, pp. 1143-1151, 2003.

[46] J. Piskorski, P. Guzik, "Filtering Poincaré plots," *Computational Methods in Science and Technology*, vol. 11, no. 1, pp. 39-48, 2005.

[47] M. Meesmann, B. Holbach, M. Koller, V. Lingg, C. Braun, P. Kowallik, "Heart rate variability and nonlinear dynamics," *Herzschr Elektrophys.*, vol. 11, pp. 151-158, 2000.

[48] H. V. Huikuri, T. Seppanen, M. J. Koistinen, K. E. Airaksinen, M. J. Ikaheimo, A. Castellanos, R. J. Myerburg, "Abnormalities in Beat-to-Beat Dynamics of Heart Rate Before the Spontaneous Onset of Life-Threatening Ventricular Tachyarrhythmia in Patients with Prior Myocardial Infarction," *Circulation*, vol. 93, pp. 1836-1844, 1996.

[49] H. V. Huikuri, T. H. Makikallio, J. Perkiomaki, "Measurement of Heart Rate Variability by Methods of Nonlinear Dynamics," *Journal of Electrocardiology*, vol. 36, pp. 95-99, 2003.

[50] M. Brennan, M. Palaniswami, P. Kamen, "Do Existing Measures of Poincaré Plot Geometry Reflect Nonlinear Features of Heart Rate Variability?," *IEEE Transactions on Biomedical Engineering*, vol. 48, no. 11, pp. 1342-1347, 2001.

[51] P. W. Kamen, H. Krum, A. M. Tonkin, "Poincaré plot of heart rate variability allows quantitative display of parasympathetic nervous activity in humans," *Clinical Science*, vol. 91, pp. 201-208, 1996.

[52] M. P. Tulppo, T. H. Makikallio, T. E. Takala, T. Seppanen, H. V. Huikuri, "Quantitative beat-to-beat analysis of heart rate dynamics during exercise," *Am. J. Physiol.*, vol. 271, no. 40, pp. H244-H252, 1996.

[53] M. Brennan, M. Palaniswami, P. Kamen, "Poincaré plot interpretation using a physiological model of HRV based on a network of oscillators," *Am. J. Physiol. Heart Circ. Physiol.*, vol. 283, pp. H1873-H1886, 2002.

[54] http://www.cbi.dongnocchi.it/glossary/Glossary.html, August 19, 2002.

[55] Lauralee Sherwood, *Human Physiology From Cells to Systems*, 4th ed. Pacific Grove, CA: Brooks/Cole Publishing, pp. 126; 306; 330; 355, 2001.

[56] G. J. Tortora and S. R. Grabowski, *Principles of Anatomy and Physiology*, 7th ed. New York: Harper Collins College Publishers, pp. 630, 1993.

[57] http://www.12leadecg.com/4044GarciaCh08.pdf, June 2, 2006.

[58] P. S. McKinley, P. A. Shapiro, E. Bagiella, M. M. Myers, R. E. DeMeersman, I. Grant, R. P. Sloan, "Deriving heart period variability from blood pressure waveforms," *J. Appl. Physiol.*, vol. 95, pp. 1431-1438, 2003.

[59] P. B. Persson, "Spectrum analysis of cardiovascular time series," *AJP – Regulatory, Integrative and Comparative Physiology*, vol. 273, no. 4, pp. R1201-R1210, 1997.

[60] A. B. Ritter, S. Reisman, B. B. Michniak, Bozena B., *Biomedical Engineering Principles*. Boca Raton, Florida: CRC Press, Taylor & Francis Group, pp. 307, 2005

[61] Harald M. Strauss, "Heart rate variability," *Am. J. Physiol. regul. Integr. Comp. Physiol.*, vol. 285, pp. R927-R931, 2003.

[62] 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*, vol. 923, pp. 1043-1065, 1996.

[63] http://www.bmi-tno.nl/products/portapres.htm, September 3, 2005.

[64] B. P. M. Imholz, W. Weiling, G. A. van Montfrans, K. H. Wesseling, "Fifteen years experience with finger arterial pressure monitoring: assessment of the technology," *Cardiovascular Research*, vol. 38, pp. 605-616, 1998.

[65] P. K. Stein, A. Reddy, "Non-Linear Heart Rate Variability and Risk Stratification in Cardiovascular Disease," *Indian Pacing and Electrophysiology Journal*, vol. 5, no. 3, pp. 210-220, 2005.

[66] M. P. Tulppo, T. H. Makikallio, T. Seppanen, K. Shoemaker, E. Tutungi, R. L. Hughson, H. V. Huikuri, "Effects of pharmacological adrenergic and vagal modulation on fractal heart rate dynamics," *Clinical Physiology*, vol. 21, no. 5, pp. 515-523, 2001.

[67] Z. Chen, Ch. P. Ivanov, K. Hu, E. H. Stanley, "Effect of nonstationarities on detrended fluctuation analysis," *Physical Review E*, vol. 65, pp. 65-79, 2002.

[68] *MatLAB®*, Computer software. The MathWorks, Inc., ver. 7.2.0.232, R2006a, 2006.

[69] John L. Semmlow, *Biosignal and Biomedical Image Processing*. New York: Marcel Dekker, pp. 243-269, 2004.

[70] A. Lanfranchi, D. Spaziani, G. Seravalle, C. Turri, R. Dell'Oro, G. Grassi, G. Mancia, "Sympathetic Control of Circulation in Hypertension and Congestive Heart Failure," *Blood Pressure*, vol. 7, pp. 40-45, 1998.

[71] G. Parati, P. Saul, M. Di Rienzo, G. Mancia, "Spectral Analysis of Blood Pressure and Heart Rate Variability in Evaluating Cardiovascular Regulation," *Hypertension*, vol. 25, pp. 1276-1286, 1995.

[72] S. Akselrod, D. Gordon, F. A. Ubel, D. C. Shannon, C. A. Barger, R. J. Cohen, "Power Spectrum Analysis of Heart Rate Fluctuation: A Quantitative Probe of Beat-To-Beat Cardiovascular Control," *Science*, vol. 213, no. 4504, pp. 220-222, 1981.

[73] L. L. Watkins, P. Grossman, A. Sherwood, "Noninvasive Assessment of Baroreflex Control in Borderline Hypertension," *Hypertension*, vol. 28, pp. 238-243, 1996.

[74] C. H. Davos, L. C. Davies, M. Piepoli, "The Effect of Baroreceptor Activity on Cardiovascular Regulation," *Hellenic Journal of Cardiology*, vol. 43, pp. 145-155, 2002.

[75] G. Grassi, B. M. Cattaneo, G. Seravalle, A. Lanfranchi, G. Mancia, "Baroreflex Control of Sympathetic Nerve Activity in Essential and Secondary Hypertension," *Hypertension*, vol. 31, pp. 68-72, 1997.

[76] M. Kikuya, A. Hozawa, T. Ohokubo, I. Tsuji, M. Michimata, M. Matsubara, M. Ota, K. Nagai, T. Araki, K. H. Satoh, S. Ito, S. Hisamichi, Y. Imai, "Prognostic Significance of Blood Pressure and Heart Rate Variabilities The Ohasama Study," *Hypertension*, vol. 36, pp. 901-906, 2000.

[77] G. Mancia, G. Grassi, C. Giannattasio, G. Seravalle, "Sympathetic Activation in the Pathogenesis of Hypertension and Progression of Organ Damage," *Hypertension*, vol. 34, pp. 724-728, 1999.

[78] M. P. Schlaich, E. Lambert, D. M. Kaye, Z. Krozowski, D. J. Campbell, G. Lambert, J. Hastings, A. Aggarwal, M. D. Esler, "Sympathetic Augmentation in Hypertension Role of Nerve Firing, Norepinephrine Reuptake, and Angiotensin Neuromodulation," *Hypertension*, vol. 43, pp. 169-175, 2004.

[79] A. M. Dart, B. A. Kingwell, "Pulse Pressure A Review of Mechanisms and Clinical Relevance," *Journal of the American College of Cardiology*, vol. 37, no. 4, pp. 975-984, 2001.

[80] Luc M. A. B. Van Bortel, Harry A. J. Struijker-Boudier, Michel E. Safar, "Pulse Pressure, Arterial Stiffness, and Drug Treatment in Hypertension," *Hypertension*, vol. 38, pp. 914-921, 2001.

[81] R. Pastor-Barriuso, J. R. Banegas, J. Damian, L. J. Appel, "Systolic Blood Pressure, Diastolic Blood Pressure, and Pulse Pressure: An Evaluation of Their Joint Effect on Mortality," *Ann. Intern. Med.*, vol. 139, pp. 731-739, 2003.

[82] V. Rizzo, F. di Maio, F. Petretto, M. Marziali, G. Bianco, F. Barilla, V. Paravati, D. Pignata, S. V. Campbel, G. Donato, V. Bernardo, D. Tallarico, "Ambulatory Pulse Pressure, Left Ventricular Hypertrophy and Function in Arterial Hypertension," *Echocardiography*, vol. 21, no. 1, pp. 11-16, 2004.