J. CAPOGNA

DSP OF NEUROCARDIAC SIGNALS IN PATIENTS WITH CHF

M.A.Sc.

Digital Signal Processing of Neurocardiac Signals in Patients with Congestive Heart Failure

By
JOSHUA CAPOGNA, B.ENG

A Thesis Submitted to
the School of Graduate Studies
in Partial Fulfillment of the Requirements
for the Degree
Master of Applied Science

McMaster University

Master of Applied Science (2003) McMaster University

(Electrical and Computer Engineering) Hamilton, Ontario

TITLE:	Digital Signal Processing of Neurocardiac Signals in Patients with Congestive Heart Failure
AUTHOR:	Joshua Capogna, B.Eng. (McMaster University)
SUPERVISOR:	Dr. Markad V. Kamath, Ph.D., P.Eng.
NUMBER OF PAGES:	clxiv, 101

Dedication

This work is dedicated to my Uncle Tony Kopczewski. His untimely death due to heart disease was the inspiration for this academic pursuit

Abstract

Recent work has found that a frequency domain and time domain analysis of the heart rate variability signals can provide significant insights into function of the heart in healthy subjects and in patients with heart disease. Patients with congestive heart failure are an important clinical health issue and it is hoped that this work will contribute towards gaining knowledge of this debilitating pathological condition.

Our laboratory has recently acquired more than three thousand 24-hour ECG tapes recorded during called Study of Left Ventricular Dysfunction (SOLVD). The SOLVD trial was conducted between 1987-1990 to test the efficacy of a medication called, Enalapril, to treat patients with heart failure. There were an equal number of patients with (group A) and without overt heart failure (group B). The work reported in this thesis describes the development of a hardware and software framework used to analyze the ECG signals recorded on these tapes. Primary objective of this work was to develop and test a system which would assist in analyzing the above tapes so as to examine if there are differences between two groups using the HRV parameters from both frequency and time domain.

The research was conducted in three steps: Hardware design, software and algorithm development and finally the validation phase of the design, to test the usefulness of the overall system. The tapes were replayed on a tape recorder and the ECG was digitized at a rate corresponding to 500 samples/second. Labview software was invoked for this task. Secondly a set of algorithms were developed to perform QRS-

detection and QT- interval identification. The detection algorithms involved placing critical ECG fiducials onto the ECG waveform through the use of a trained model. The model construction used patient specific pre-annotated data coupled with statistical and genetic algorithm techniques. The beat-to-beat HRV signal was thus generated using the annotation data from the ECG.

Frequency domain indices were obtained using power spectral computation algorithms while time domain statistical indices were computed using standard methods. QT-interval algorithms were tested using a set of manually and automatically tagged set of beats from a sample of subjects. For the third part of this research, i.e. validation phase, we set up a test pool of 200 tapes each from patients with overt heart failure and with no heart failure, recorded at the baseline before the subjects entered the study. This phase of the study was conducted with the help of a statistician in a blinded fashion. Our results suggest that there is significant difference between frequency domain and time domain parameters computed from the HRV signals recorded from subjects belonging to group A and group B. The group A patients had a lot of ectopic beats and were challenging to analyze. These results provide a confirmation of our analytical procedures using real clinical data. The QT-analysis of the ECG signals suggest that automatic analysis of this interval is feasible using algorithms developed in this study.

Acknowledgements

I would like to thank my supervisor Dr.Markad V.Kamanth for his support and assistance. Dr. E.L. Fallen, M.D., Professor Emeritus, Cardiology Division, made the tapes generously available for the present work and his help in obtaining the research material is acknowledged. Dr.S.Yusuf who was the lead investigator for SOLVD study helped us find the clinical data necessary for analysis. I wish to thank Dr.Adrian Upton, M.D., Professor Medicine and Director of Neurology and DeGroote Foundation for making the facilities available for conducting the research presented in this thesis. Natural Sciences and Engineering Research Council provided the funding for Dr.Markad V.Kamath's laboratory and for my research scholarship and their help is gratefully acknowledged.

I would also like to thank Sangeeta Ullal, Karen Yuen and my family (Stella, Sebastian and Stephany Capogna) for their support. Without them this thesis certainly would not have come together in such a timely fashion. Finally I would like to show my appreciation to Janice Pogue for lending us her time to complete the analysis.

TABLE OF CONTENTS

List of Figures

List of Tables

GLOSSARY

ACE Inhibitor Angiotensin Converting Enzyme Inhibitor. ACE inhibitors (i.e., Enalapril, Captropril) reduce peripheral vascular resistance via blockage of the angiotensin converting enzyme. This action reduces the myocardial oxygen consumption, thereby improving cardiac output and moderating left ventricular and vascular hypertrophy.

Annotation: A label indicating the presence of a Fiducial on the ECG

Auto-Annotation: The process of using an unsupervised software tool to place fiducials on digitized ECG data.

Baseline wander: The low frequency drift of the average measured ECG voltage due to both physiological and environmental factors.

ECG waveform: A single cycle of cardiac polarization and depolarization.

Ectopic beat: A neurocardiac signal that is initiated by a nerve bundle other that the SA node

Ejection Fraction The proportion, or fraction, of blood pumped out of your heart with each beat.

Fiducial: A point of reference on the ECG.

Holter ECG: An ECG recording taking using a Holter monitor system.

LF:HF ratio Ratio of the powers in the low frequency and high frequency bands of a power spectrum of the heart rate variability.

Model (of and ECG): A template used to approximate a single ECG waveform.

N-N intervals: Normal sinus conducted beats.

Null Hypothesis: The hypothesis that two sets come from the same population and therefore have the same mean.

pNN50 Percentage of adjacent intervals between normal sinus conducted beats within a given heart rate data set that differ by more than 50 milliseconds.

QRS complex: the portion of the ECG waveform that starts from the onset of the Q wave and ends at the end of the S wave.

QT-interval: The time between the onset of the Q wave and the end of the T wave in an ECG.

Regime (of and ECG): A section of the ECG that is uniquely filtered to expose an ECG waveform feature.

Respiratory Sinus Arrhythmia (RSA) Fluctuations in heart rate that occur concurrently with the respiratory cycle. In general, heart rate increases on inspiration and decreases on expiration.

R-MSSD Root mean square of successive differences between normal sinus conducted beats in a given heart rate data set.

R-R interval: The time between the onset of sequential R waves in an ECG.

SDANN Standard Deviation of the mean interval between normal sinus conducted beats in adjacent 5 minute segments of a given heart rate data set.

SDNN index Mean of the standard deviation of intervals between normal sinus conducted beats in adjacent 5 minute segments of a given heart rate data set.

SDNN Standard Deviation of all intervals between normal sinus conducted beats in a 24 hour segments of a given heart rate data set.

Zone (of an ECG): A section of the ECG Model that is independently scaled to better match the model to the ECG waveform that is being examined.

List of Abbreviations

AMI	Acute Myocardial Infarction
AR	autoregressive
BPV	Blood Pressure Variability
CHF	Congestive Heart Failure
CONSENSUS	The Cooperative North Scandinavian Enalapril Survival Study
DCM	Dilated Cardiomyopathy
ECG	Electrocardiogram
FT	Fourier Transform
FWHM	Full Width Half Maximum
GA	Genetic Algorithm
GUI	Graphical User Interface
HF	Heat Failure
HRV	Heart Rate Variability
MED	Minimum Euclidian Distance
MF	Merit Function
NYHA	
PSD	Power Spectrum Density
SD	Standard Deviation
SOLVD	Study of left ventricular dysfunction.
TNF	Tumor Necrosis Factor
UI	User Interface
VEL	Virtual ECG Lab

Chapter 1: Introduction

1.1 Background and Motivation

A study of heart rate variability (HRV) and blood pressure variability (BPV), commonly known as hemodynamic variables in physiological literature, through digital signal processing techniques has led to clinically relevant discoveries which have improved lives of many patients with the heart disease. Central processing of physiological input signals in the mid-brain, where autonomic regulation takes place, has a major role in controlling the hemodynamic variables. Due to its relative inaccessibility, autonomic nervous system is often studied through its efferent variables, namely HRV and BPV. An analysis of ambulatory signals generated by the heart has helped researchers understand a variety of disorders where autonomic nervous system is involved (Kleiger et al.,1991; Akselrod et al., 1981). For example, it is now known that a low standard deviation of the 24 hour mean beat-to-beat interval (also called R-R interval time series) following a myocardial infarction can, with high probability, predict a significant cardiac event. The effects of medications such as beta blockers, vago-mimetic agents (Scopolamine), circadian variations, smoking and agents associated with depression are among the various issues being studied in this field. Both short term (<20-30 minutes) and long term (24 hours or more) HRV signals have been analyzed under various pathological conditions and experimental paradigms during the last 2 decades (Kamath and Fallen, 1993; Malik M, 1993). The analysis of the hemodynamic signals

has therefore gained wide acceptance in medical literature. Early on, frequency domain analysis of R-R intervals yielded information regarding how the brain exercised cardiac control on a beat-to-beat basis. As 24 hour ECG, clinically known as Holter ECG, became easily available throughout 1970s, computerized analytical techniques were developed to identify abnormal electrophysiological signatures of the heart through the identification and enumeration of the abnormal beats and beats that have not been initiated by SA node conduction, (known as ectopic beats). With the development of the HRV processing, the same instrumental procedure was extended to study the autonomic function through a computation of the time domain statistics and frequency domain indices of the HRV signal. Because of the ease of recording the ECG (i.e.non-invasively), an analysis HRV signals has become the signal of choice between BPV and HRV.

Congestive heart failure is a pathological condition, occurring primarily in patients with diseases of the coronary arteries, although other causes of its origin are known. Often a precipitating event such as a myocardial infarction triggers the process of congestion and a failure of the pumping capacity of the heart. Subsequent adaptation, called left ventricular remodeling results in a chronic condition perpetuated the onset of a diminished ability of the left ventricle to pump blood. Further, the process of congestion impairs the autonomic nervous system. In Canada, demographics suggest an ageing population with a large percentage of people occupying an age group called 'Boomer generation' (i.e. those born between 1945-1964) who are entering their fifth and sixth decades of life. The Boomer generation is widely believed to have a disproportionately high risk of developing heart disease. The increased risk is largely due to the tendency of

those in that demographic group to lead increasingly stressful lives, even as their age advances. Given this trend, there is also an increased likelihood that a significant portion of this demographic group shall suffer from the progressive form of congestive heart failure. Therefore, it is pertinent to leverage engineering expertise to study this disease category in general, and from the point of hemodynamic signal processing in our laboratory, in particular.

The research reported in this thesis presents rationale, methodologies and results of HRV analysis using modern signal processing algorithms applied to a sample tapes (n=400) of baseline Holter ECG. The Holter ECG was obtained from a pre-recorded set of tapes containing data from patients with and without overt congestive heart failure (CHF). These patients had enrolled in a muti-centre trial called a Study of Left Ventricular Dysfunction (SOLVD) to test the efficacy of a pharmacological/therapeutic agent called Angiotensin-Converting-Enzyme (ACE) inhibitor for treating congestive heart failure. The tapes were made available to our laboratory for research.

Chapter 2 provides an overview of the congestive heart failure as a pathological condition. A summary of the two pioneering clinical trials which initially identified and ascertained the role of ACE inhibitors in treating patients with CHF is presented. Previous work conducted by investigators in this field is reviewed in chapter 3, as pertaining to both time domain and frequency domain analysis. In chapter 4, an identification and formulation of various questions which led to this research is presented. In chapter 5 we examine the relevance of QT interval analysis in patients with CHF and some of the algorithms developed for this work. Chapter 6 outlines the practical

details of R-R interval and QT interval analysis and chapter 7 presents our experimental methodology used in this study. The results of our techniques and subsequent R-R interval analysis can be found in chapter 8. In chapter 9 a discussion of the relevance of these results in a large picture is presented with a few suggestions for future work. Chapter 10 presents a summary of the research in this thesis.

1.2 Summary

In this introductory chapter we have identified the research question in the larger context and have discussed how the subsequent chapters will present our work.

Chapter 2: Congestive Heart Failure and the role of ACE Inhibitors in treating Congestive heart failure

2.1 Introduction

In this chapter we present a brief epidemiology of congestive heart failure (CHF) and its classification. We also outline the process of neurhumoral activation following a myocardial infarction and the resulting left ventricular modeling. We elucidate the role of angiotensin-converting enzyme inhibitors in treating congestive heart failure and highlight two studies; CONSENSUS and Study of Left Ventricular Dysfunction (SOLVD) which tested the efficiency of ACE inhibitors in treating SOLVD. The SOLVD study is the source of clinical heart rate data for this thesis.

2.2 Epidemiology and New York classification of Congestive Heart Failure

Congestive heart failure is a complex clinical syndrome that is affecting an increasing number of people annually, especially those in the Western world. Approximately 4.6 million Americans are currently living with heart failure. It is estimated that 450,000 cases will be diagnosed this year alone (Heart Failure Society of America, 2002). In Canada, approximately, 330,000 patients have been diagnosed CHF and in addition, an estimated 500,000 patients have CHF but they may not be aware of it (Heart and stroke Foundation of Canada, 2002). There is a mortality rate of 25-40%

within one year of diagnosis of CHF and this rate increases significantly in patients over the age of 65 years (Townley & Howlett, 2002). The disease costs upwards of 1 Billion dollars for the Canadian Health Care system. The bulk of this expenditure (70-80%) is as a result of hospitalization costs. With ageing baby-boomers, it is anticipated that CHF will be a serious and debilitating disorder that will lay a major claim to increasing medical costs in both the government regulated health care system and in private supplementary care in the coming decade. In order to determine the best course of therapy, physicians often assess the stage of heart failure according to the New York Heart Association (NYHA) functional classification system. This system relates symptoms to everyday activities and the patient's quality of life and is shown in table 2.1.

Table 2.1 New York Heart Association Functional Classification System of Congestive Heart Failure

Class	**Patient Symptoms**
Class I (Mild)	No limitation of physical activity. Ordinary physical activity does not cause undue fatigue, palpitation, or dyspnea (shortness of breath).
Class II (Mild)	Slight limitation of physical activity. Comfortable at rest, but ordinary physical activity results in fatigue, palpitation, or dyspnea.
Class III (Moderate)	Marked limitation of physical activity. Comfortable at rest, but less than ordinary activity causes fatigue, palpitation, or dyspnea.
Class IV (Severe)	Unable to carry out any physical activity without discomfort. Symptoms of cardiac insufficiency at rest. If any physical activity is undertaken, discomfort is increased.

2.3 Neurohumoral Activation and Left Ventricular Remodelling

The function of the heart is to maintain a specific blood pressure volume relationship within the hemodynamic system so that nutrients can be supplied efficiently to the whole body. Heart failure may be viewed as a progressive disorder that is initiated after an index event. An index event, such as an acute myocardial infarction, damages the heart muscle. The result is a loss of functioning cardiac myocytes and leads to a disruption of the ability of the myocardium to generate contraction and the pumping action. The congestive heart failure may be precipitated by an event such as a myocardial infarction or gradual insidious hemodynamic overloading due to non-specific factors. In some patients, genetic factors may cause cardiomyopathies with a resulting heart failure (Mann, 1999). Such events will cause an initial decline in the pumping capacity of the heart. In most instances the patient remains asymptomatic or minimally symptomatic following the index event. During this period many compensatory mechanisms attempt to preserve the pump output and provide the subject with a functioning heart within the parameters of normal physiological function. For example, adaptive responses are invoked that preserve stroke volume by involving non-infarcted remote myocardium. Other compensatory mechanisms are early activation of the sympathetic nervous system and salt and water retaining system. A family of vasodilatory molecules such as natriuretic peptides, prostaglandins (PGE_2 and $PGEI_2$ and nitric oxide) (Mann) may also participate in such compounds. In addition to these chemical factors, other issues such as

gender, age, genetic background and may compound the disease process. However, with time the symptoms of pump failure will manifest and transition from asymptomatic to symptomatic heart failure is accompanied by further activation of neurohumoral and cytokine systems along with a series of adaptive changes in the myocardium, generally called LV remodeling. Left ventricular remodeling is the process by which ventricular size, shape and function are regulated by mechanical, neurohumoral and genetic factors. In this thesis we will be dealing with data from patients who have had an infarct and therefore we will focus on LV remodeling following an acute myocardial infarction.

A number of proteins such as Norepinephrine, angiotension II, endothelin, aldosterone and tumor necrosis factor (TNF) have been implicated as potentially active molecules whose biochemical toxic properties contribute to the disease progression in the failing heart. It is believed that these molecules may degenerate the heart function independent of the hemodynamic status of the patient. Animal studies have shown that high concentration of certain neurohormones can bring about heart failure like condition (Mann, 1999). Clinical studies have also shown that antagonizing neurohormones leads to clinical improvement for patients with heart failure (Mann, 1999). It is hypothesized that long-term activation of a variety of neurohormonal mechanisms produce direct end-organ damage within the heart and circulation and therefore, heart failure may develop insidiously many years after an acute myocardial infarction, despite the absence of ongoing ischemia.

2.4 Role of ACE inhibitors in Congestive Heart Failure and Cardioprotective Effects of ACE inhibitors in Congestive Heart Failure

The efficiency of ACE inhibitors in attenuating left ventricular dilatation after an infarction was first demonstrated in rats and its effect on remodeling was associated with improved survival. In a human study, Captopril (an ACE Inhibitor) resulted in increased stroke volume index and ejection fraction when compared to another medication or placebo. The mechanism of improvement with ACE inhibition is related in part to peripheral vasodilatation, ventricular unloading and ventricular dilatation.

A number of cardioprotective and vasculoprotective effects of ACE inhibitors have been observed in both animal models and in clinical trials. Table 2.2 lists some of the cardioprotective and vasculoprotective Effects of ACE inhibitors (Lonn et al.1994). Based on these observations, clinical practice of using ACE inhibitors in treating patients with Congestive heart failure has a rational basis. For the work reported in this thesis, the source of 24-Holter ECG data arose out of a sub-study which tested the efficacy of ACE inhibitors (called SOLVD) in treating patients with congestive heart failure. The objective of this thesis is to report the development of the HRV analysis technology and test a hypothesis regarding the relationship between various HRV measures and CHF. We will not be evaluating the neurocardiac related effects of ACE. However, these same algorithms developed in this study may be used for evaluating the effects of ACE inhibitors in patients at a future date.

Table 2.2 Effects of ACE inhibitors (Lonn et al. 1996)

Cardioprotetective Effects:

Restoring the balance between myocardial oxygen supply and demand

Reduction in left ventricular preload and afterload

Reduction in left ventricular mass

Reduction in sympathetic stimulation

Beneficial effects on reperfusion injury*

Vasculoprotective Effects:

Direct Atherogenic effect*

Antiproliferative and animigratory effects on smooth muscle cells, neurophils and mononuclear cells

Improvement and/or restoration of endothelial function

Protection from plaque rupture*

Antiplatelet effects

Enhancement of endogenous fibrnolysis

Antihypertensive effects

Improvement in arterial compliance and tone

---

* Not demonstrated conclusively in humans

2.5 Clinical Trials of ACE inhibitors on patients with congestive heart failure (CHF)

Beneficial effects of ACE inhibitors on patients with CHF was sufficiently important to warrant clinical trials initially in Europe (called CONSENSUS) and subsequently by through a multi-center trial called SOLVD (Study of Left Ventricular Dysfunction) in U.S, Canada and Belgium. We present an overview of these 2 key trials evaluating the effects of ACE inhibitors on humans.

2.5.1 The CONSENSUS (Cooperative North Scandinavian Enalapril Survival Study) Trial

To evaluate the influence of the ACE inhibitor enalapril (2.5 to 40 mg per day), a group of European investigators examined the prognosis of severe congestive heart failure (New York Heart Association [NYHA] functional class IV), the CONSENSUS trial study randomly assigned 253 patients in a double-blind study to receive either placebo (n = 126) or enalapril (n = 127). Follow-up averaged 188 days. The mortality at the end of six months (primary end point) was 26 % in the enalapril group and 44 % in the placebo group; a reduction of 40 % ($p = 0.002$). Mortality was reduced by 31 % at one year ($p = 0.001$). By the end of the study, there were 68 deaths in the placebo group vs. 50 in the enalapril group and thus a 27 % reduction ($p = 0.003$) of deaths was observed. A reduction of 50% in total mortality was found to be primarily among patients with progressive heart failure. There was a significant improvement in NYHA classification in the enalapril group, and a reduction in heart size and a reduced

requirement for other medication for heart failure. It was concluded that the addition of enalapril to conventional therapy in patients with severe congestive heart failure can reduce mortality and improve symptoms.

2.5.2 The SOLVD (Study of Left Ventricular Dysfunction) trial

The SOLVD Investigators recruited patients from North America (USA, Canada) and Europe. They studied the effect of an angiotensin-converting-enzyme inhibitor, enalapril, on mortality and hospitalization in patients with chronic heart failure and with an ejection fraction ≤ 0.35. Patients receiving conventional treatment for heart failure were randomly assigned to receive either placebo (n = 1284) or enalapril (n = 1285) at doses of 2.5 to 20 mg per day in a double-bind trial. Approximately 90 % of the patients were in NYHA functional classes II and III. The follow-up averaged 41.4 months. There were 510 deaths in the placebo group (39.7 %), as compared with 452 in the enalapril group (35.2 %) (reduction in risk, 16 %; P = 0.0036). Although reductions in mortality were observed in several categories of cardiac deaths, the largest reduction occurred among the deaths attributed to progressive heart failure (251 in the placebo group vs. 209 in the enalapril group; reduction in risk, 22 %;). Fewer patients died or were hospitalized for worsening heart failure (736 in the placebo group and 613 in the enalapril group; risk reduction, 26 %; $p < 0.0001$). It was concluded that an addition of enalapril to conventional therapy therefore significantly reduced mortality and hospitalizations for heart failure in patients with chronic CHF and reduced ejection fractions.

2.6 Summary

In this chapter we have highlighted the importance of congestive heart failure and its clinical classification based on New York Heart Association. The process of left ventricular remodeling following a myocardial infraction has been outlined. cardioprotective effects of ACE inhibitors in patients with CHF have been described and two pioneering trials of these medications(CONSENSUS and SOLVD) in patients with CHF have been summarized.

Chapter 3: Digital Processing of the HRV Signal and its Clinical Relevance

3.1 Introduction

This chapter will examine and summarize research conducted by leading researchers in the field of HRV analysis field. The clinical significance of HRV factors and their usefulness in the diagnosis of various conditions will also be touch upon. It is anticipated that such a study may provide a basis to analyze the HRV signal obtained from the Holter tapes recorded during the SOLVD trial.

3.2 Heart Rate Variability in Clinical And Physiological Research.

In this section, we will briefly review work done in laboratories by Kleiger (1987), Bigger (1991), Eckberg (1983), Bianchi (1993), Malik (1993), and Goldberger (2001). During the past 20 years, these researchers have been at the forefront of the emerging trends in the field. Each of these investigators chose a clinical issue to examine and used analytical methodologies derived from signal processing on the HRV as a key tool to solve the problem; a pursuit not unlike our own paradigm. All investigators chose a Holter to record their long term ECG data in an analogue format. Further, their laboratories have had experience in processing HRV signals from human subjects as well as animal models. We further note that all these investigators focused on the heart disease.

Kleiger et al. (1987) wanted to examine the survival profile of the HRV recorded from patients following an acute myocardial infarction. He studied heart rate variability in 808 survivors of acute myocardial infarction under the age of 70 years. Time domain indices were calculated from 24-hour ambulatory Holter monitoring and mean follow up time was 31 months following the initial event. It was found that an SDNN of <50 ms was associated with a 5.3 fold increased risk of mortality compared to a corresponding value of SDNN >100ms. Heart rate variability indices remained a significant marker of mortality risk following adjustment for left ventricular ejection fraction, ventricular premature complex frequency, New York Heart Association class and drug treatment regime. The work done by Kleiger et al has served as a foundation on which much of the subsequent time domain research of HRV has been developed.

Kleiger et al. (1991) examined the 24-hour heart rate power spectrum in 68 post-AMI patients at 0, 3, 6, 9 and 12 months after enrollment in the study (enrollment was 25 ± 17 days post-MI). Comparisons were made with 95 normal (unafflicted control) individuals of similar age and gender. Upon inception of the study, baseline data showed all frequency bands exhibited a significant power reduction in patients, compared to controls. There was a substantial increase in power across the entire spectrum in patients, between the baseline data and 3 months later, after which time the measured indices remained relatively stable. However, after twelve months all HRV measures were still only 50-66% of the values observed in the control subjects. It was apparent that the recovery of HRV appears to be completed during approximately the same time period in

which post-MI mortality rates drop to a stable value. However, the authors acknowledged that the association may be coincidental rather than causal.

Mortality in chronic heart failure (HF) remains high with a significant number of patients dying of the progressive disease. Identification of these patients is therefore important from a clinical point of view. Kearney et al (2002) explored the value of noninvasive predictors of death in ambulant outpatients with chronic heart failure (HF) with the goal of estabilsing a link between mortality and HRV factors. They recruited 553 ambulant outpatients age 63 ± 10 years with symptoms of chronic HF (New York Heart Association functional class, 2.3 ± 0.5) and objective evidence of left ventricular dysfunction (ejection fraction <45%, cardiothoracic ratio >0.55, or pulmonary edema on chest radiograph). After 2,365 patient-years of follow-up, 201 patients had died, with 76 events due to progressive CHF. They found SDNN as an independent predictor of all-cause mortality assessed with the Cox proportional hazards model, The SDNN was also an independent predictor of death due to progressive CHF. The study concluded that from 24 hours Holter one can identify increased risk of death due to progressive HF.

Bianchi et al. (1993) developed a time-variant algorithm of autoregressive (AR) identification and applied it to the HRV signal from patients with ischemia. The power spectrum is calculated from the AR coefficients derived from R-R interval epoch. Time-variant AR coefficients are determined through adaptive parametric identification with a forgetting factor. The forgetting factor obtains weighted values on a running temporal window of 50 preceding measurements. Power spectrum density (PSD) was obtained at each cardiac cycle, making it possible to follow the dynamics of the spectral parameters

and LF:HF ratio on a beat-by-beat basis. Bianchi et al. applied the above methodology to Holter tapes recorded from subjects suffering from transient myocardial ischemia. The time variant spectral parameters suggest an early activation of LF component in the HRV power spectrum prior to ischemia. The increased LF precedes by approximately 1.5-2 min the tachycardia and the ST displacement, generally indicative of the onset of an ischemic episode.

Goldberger (2001) and his team have studied complexity in patients with heart disease. This complexity arises from the interaction of a myriad of structural units and regulatory feedback loops that operate over a wide range of temporal and spatial scales, enabling the organism to adapt to the stresses of everyday life. Quantifying and modeling a variety of behaviors exhibited by living organisms is one of the major challenges of contemporary medical science (Goldberger, 2001). The combination of non-linearity and non-stationarity, more the rule than the exception in the output of physiologic systems, poses a major challenge to conventional biostatistical assessments and standard reductionist modeling methods. To describe and quantify the mechanisms of these "nonhomeostatic" behaviors, Goldberger et al (2001) have employed new techniques derived from nonlinear dynamics. The appropriate application and interpretation of such metrics, however, remains incompletely explored. After a review of the work conducted by Goldberger et al., it was concluded that complexity and nonlinear dynamics do not help correlate physiological observations with abstract mathematical descriptions.

3.3 Summary

Here we explored the subject of HRV analysis and it's application in the analysis of clinical patient data. Work relating to time, frequency and non-linear analysis was discussed as a background describing the state-of-the-art in the field. From this research it can be concluded that both time and frequency measures of the HRV have clinical relevance as well as potential for further HRV studies.

Chapter 4: QT-interval analysis and its implications in congestive heart failure

4.1 Introduction

This chapter outlines the significance of the QT-interval measurement in an ECG recording. It also discusses how patients with congestive heart failure can have abnormal QT-intervals. We will investigate current state-of-the-art, in terms of engineering methods, in QT-interval analysis in patients with CHF.

4.2 Definitions

The QT interval begins at the onset of the QRS complex and to the end of the T wave (See also Appendix A). It represents the time between the start of ventricular depolarization and the end of ventricular repolarization. It is useful as a measure of the duration of repolarization cycle. The QT interval will vary depending on the heart rate, age and gender. It increases with bradycardia and decreases with tachycardia. Men have shorter QT intervals (0.39 sec) than women (0.41 sec). The QT interval is influenced by electrolyte balance, drugs, and ischemia. Often, a normalized QT interval is obtained by dividing QT by the square root of the R-R interval. The calculation of the QT interval for interval analysis purposed is explored in more detail in chapter 6.

4.3 QT-dispersion and congestive heart failure

QT- interval as measured from the 12 lead electrocardiogram gives information about the spatial differences in myocardial recovery time. Homogeneity of recovery time has been postulated to protect against arrhythmias whereas dispersion of recovery time is arrhythmogenic. QT- interval dispersion calculated from the 12-lead ECG can provide a noninvasive measurement for quantifying the degree of myocardial repolarization and inhomogeneity. QT dispersion has been linked as a risk indicator for arrhythmic cardiac death in patient populations having coronary artery disease, chronic heart failure, myocardial infarction, sustained ventricular arrhythmias, ventricular fibrillation, peripheral vascular disease, unstable angina pectoris and ischemia. Anomalies have also been linked with reprofusion therapy, drug arrhythmogenicity and hypertrophic cardiomyopathy. Long -QT syndrome is a condition specifically attributed to this interval measurement. While the work presented in this thesis does not incorporate any 12 lead ECG data, the importance of QT-dispersion in the diagnosis of severe heart conditions cannot be ignored. Therefore, it is imperative that an algorithm for measuring the QT-interval be available for analyzing SOLVD Holter ECG data. Burger et al. (1997) have developed an algorithm which uses a correlation between a template and the next beat and obtains a difference between the two to compute QT- variability.

Increased QT-variability has been studied in patients with dilated cardiomyopathy (DCM). Berger et al (1997) found that DCM is associated with beat-to-beat fluctuations in QT-interval that are larger than normal and uncoupled from variations in heart rate.

QT-interval variability increases with worsening functional N.Y.H.A. class but is independent of ejection fraction.

4.4 Summary

QT-interval variability has found clinical utility and therefore algorithms developed in this research can be used for understanding QT-interval variation in patients with CHF. Research conducted by Berger et al. revealing a relationship between R-R and QT-interval variability in DCM patients was discussed as a justification for the need of a QT-interval measurement method for future analysis.

Chapter 5: Development of Ideas and Tools for the Signal processing R-R and QT-intervals

5.1 Introduction

This chapter explores the proposed goal of our study as well as the relevance of the goal to the field of cardiac care and diagnosis of congestive heart failure. We will also examine some of the technical challenges we face in achieving that goal such as patient physiology, processing requirements and user needs. We shall also review previous techniques used to solve similar signal processing tasks and formulate our own approach to the problem.

5.2 R-R, QT-interval and patients with CHF

Given the widespread occurrence of the disease and the high probability that an ever greater portion of an aging population will succumb to CHF, it is important to develop new diagnostic techniques so that a suitable therapy can be applied early. One such method of early detection could be based on HRV analysis as suggested by the study conduced by Eckberg et al. (1983), which was described in detail in chapter 3.

Our goal will be to use the 24 hour Holter data collected for the SOLVD trial to test the hypothesis that both time domain and frequency domain properties a patients' HRV can be used to separate CHF patients from patients from those that have not been diagnosed with the disease. The HRV measurements must be based entirely on sinus

rhythms, most easily characterized by the presence of the P wave. Therefore, a system shall be developed to pinpoint this ECG fiducial point. Also, as noted in chapter 4, Berger et al (1997) found a relationship between patients' with dilated cardiomyopathy (DCM) and their QT-interval. Given that DCM often accompanies CHF, a system for analyzing the QT intervals in the time and frequency domain shall also be constructed and tested.

5.3 Technical challenges

There are a plethora of technical challenges that make the study of HRV a non-trivial task. One such challenge is the collection of data. It is a difficult task to record artifact free, consistent HRV data using ECG recorded using a Holter monitor, especially from a group of older patients with a serious pathological condition. The major factors that limit the quality of measured ECG data are summarized in table 5.1.

Most of the above challenges could be mitigated against if the data recording process was carefully and consistently monitored. However, obtaining large amounts of ECG data in this way is exceedingly time consuming. Therefore, large amounts of data must be gathered from pre-recorded sources. In our case data is drawn from the from the SOLVD trial which, in many cases, is less than ideal for HRV analysis. Therefore, the software developed for this study must be constructed with enough flexibility to analyze less than ideal data.

Table 5.1: Technical challenges in collecting ECG data		
Challenge	Cause	Consequence
1. Improper electrode attachment	Patients skin, especially the elderly, is rough or loose, causing the electrode attachment to shift in place during motion	ECG recording may drop substantially in voltage or may even disappear all together due to varying conduction through the skin
2. Many physiology related artifacts in ECG such as ectopic beats	Often patients we wish to study the most (i.e. CHF patients), have an increased occurrence of artifacts and ectopic beats	Ectpic beats are not initiated by the sinus node (see appendix A for ECG anatomy). Therefore, they introduce noise into the HRV signal.
3. 24 hr ECG data contains many noise components	24 hr ECG data is often recorded while the patient is going about his or her daily life. Actions such as coughing, walking, electrical interference introduce noisy segments in the continuous ECG.	Noise caused by changing environments may cause spikes or baseline wandering in the ECG recording.
4. Poor QRS complex in CHF patients	Patients with CHF often have widened, attenuated and inconsistent QRS complexes	Changes in the QRS complex may make QRS detection especially difficult especially if the detection criteria is static.
5. Inconsistent lead placement	Often large chests or other physical impediments prevent the electrode from being placed in the optimal spot. Also, varying patient heart positions within the chest cavity may alter the morphology of the ECG.	A QRS detection algorithm and even signal filters that are designed specifically for V1 placement will likely not work for V2 or V3 data.

The second major technological challenge stems from the need for flexibility. If a QRS and QT detection system is to be useful as a clinical analysis tool it must be able to adapt to varying qualities and morphology of ECG with minimal intervention by the user. Any clinical QRST detection system should not require the user (nurse, technician

or cardiologist) to have any specialized knowledge outside of his or her professional skill set. Also, if large amounts of data are to be analyzed then it cannot take daunting amounts of time to re-tool the system to maximize its performance for multiple patients.

5.4 The approach

The bulk of ECG analysis is conducted through the use of various software tools operating on digitized data. Therefore, it seems only natural that we too develop software that analyzes digitized ECG data. In this section, we will outline an approach to develop signal processing software and hardware for analyzing the SOLVD data.

5.4.1 Data Collection and storage

In its current form, the SOLVD data set consists of more than 2000 ECG recordings stored on spooled magnetic tape. To make processing possible we must play back the analogue data on these tapes and digitize them so that a software application can conduct the analysis. A high-speed playback device coupled with a digitization hardware (A/D converter) capable of high sampling rates (~60 KHz) is needed to minimize the operator's digitization time. Finally, the data should be compressed for convenient storage.

5.4.2 Data Selection

Although we wish to build a system that can handle a variety of patients, we must establish the following minimal conditions for data quality so that HRV analysis is possible.

1. At least 20 minutes of continuous sinus beats of ECG/ per hour be available. (as indicated by the presence of a P wave)
2. Fewer than six ectopic beats per minute in the remaining data.
3. An absolute recorded start time.
4. A full 23-24 hours of data must have been recorded.

In order to conduct a blinded experiment, the selection of test data must be selected randomly from both the CHF group and the Non-CHF group. To make this selection, a third party experienced in the construction of such an experimental design, preferably a statistician is needed.

5.4.3 Placement of Fiducial Points

Once the 24 hour ECG recorded on the Holter has been digitized, a sample of the data must be annotated for use in the training of various algorithms. Annotation consists of identifying and placing critical fiducial points on the ECG wave from so that an interval analysis can be conducted (please see appendix A for an explanation of the morphology of the ECG waveform). The critical fiducials for R-R interval analysis are the start of the Q wave, the R wave peak and the end of the S wave. These three fiducials make up the QRS complex. Without verification that a QRS complex is present, the position of the R wave for R-R interval analysis should not be recorded.

The QT interval is measured from the beginning of the Q wave to the end of the T wave. Naturally, the end of T wave fiducial is also a necessary component. There is also

a requirement that only sinus rhythms should be included in the interval analysis. Thus, the P wave must also be identified.

Over the past 20 years there have been many approaches to the automated placement of fiducials onto ECG data. All such algorithms have incorporated a noise reduction procedure followed by a signal identification subroutine. Noise reduction techniques that been examined in the past have taken the form of wavelet filtering (Li et al, 1995;Martinez et al, 2000; Sahambi et al, 2000) , band pass filtering (Berger et al, 1997; Shankara et all, 1992; Riccio et al, 1992; Friesen et al, 1990) as well as adaptive filtering (Jane et al, 1992; Xue et al, 1992). Each of these techniques attempt to filter out different components of noise commonly found in ECG waveforms. A subset of these noise sources, with their effect on ECG's are listed in table 5.2(Friesen et al, 1990).

Table 5.2: common source of noise in ECG recordings

Name	Source	Appearance on ECG
Power line interference	Caused by ECG picking up fluctuations generated by nearby AC power.	60Hz with subsequent harmonics
Electrode contact noise	Caused by loss of contact between the electrode and the patient's skin	Rapid baseline transition to the maximum recorder input voltage followed by a slow decay.
Baseline drift	Variance in electrode to skin contact quality caused by the shifting of the electrode on the patients skin which occurs as the patient moves and/or breaths	Gradual motion upwards and then downwards of the baseline for a duration of approximately 100-500 ms. This can be modeled as a 0.15 – 0.3Hz signal.
Muscle activation	Electrical noise caused by muscle activation of the body	Noise signal can be modeled as zero-mean Gaussian noise that is band limited anywhere between 0- 10K Hz

Once the data has been filtered to eliminate noise, the signal is then examined to identify the fiducial(s) points of interest. The most common method of fiducial detection involves a threshold based technique that scans the data until a threshold criteria is met (Ruha, 1997; Algra, 1987; Pellegrini, 1987; Pan, 1985). Other techniques have included detection of local minima and maxima (Li et al, 1995; Martinez et al,2000; Sahambi et al, 2000) , correlation of the signal with a trained or calculated model (Berger et al,1997; Kluge et al, 1997; Sadeh, 1987) and also artificial intelligence methods such as neural networks (Hopkins et al, 2000;Yang et al, 1993) and syntactic based analyzers (Pietka, 1991; Skordalakis, 1986).

The difficulty with many of these techniques is that they have been geared towards a particular lead (usually V1) and do not provide a robust method for stand-alone adaptation. However, separating the various ECG fiducials using multi-resolution analysis via wavelet filtering as suggested by Sahambi et al (2000) has a great deal of promise for our application since it has the ability of emphasizing specific waveforms and suppressing others.

Another promising technique is the construction of a model using patient specific pre-annotated ECG data and then stretching or squeezing that model so that it best fits the ECG waveform in question. Such a technique has been successfully tested Berger et al (1997). Although Berger's algorithm does not provide a techniques for verifying the authenticity of it's placement, such checks can be incorporated using maximum likelihood methods. An adaptation of the method must also be made to identify the P wave.

Upon examination of previous work, we conclude that a combination of the systems proposed by Sahambi et al. (2000) and by Berger et al (1997) provide an excellent starting point for our software. Additional robustness can be achieved by including a technique for building a patient model based on multiple annotated beats prior to analysis and a post-analysis system of verifying the plausibility of the proposed fiducial position.

5.4.4 Interval Analysis

Once the data has been annotated the R-R and QT-intervals shall be measured over the 24 hour time period that the patient was examined. Using the R-R interval data, we shall distil the following HRV properties:

a. Time domain
    i. HR
    ii. NN
    iii. SDNN
    iv. SDNNind
    v. SDANN
    vi. R_MSSD
    vii pNN50
b. Frequency domain
    i. LF freq
    ii. HF freq
    iii. %LF area
    iv. %HF area
    v. LF:HF
    vi. ABS LF
    vii. ABS HF

Techniques for extracting these measures and correcting for ectopic beats were described in chapter 3. Work done by Anne Hu (1999) in our laboratory will provide some guidelines, especially for frequency domain analysis of the HRV data. It is in not

the intention of this study to complete a detailed QT-interval analysis of the patient data. However, the software for doing so shall be constructed and thoroughly tested to facilitate future work.

5.4.5 Statistical Analysis

In order to maintain objectivity, statistical analysis shall be conducted by a statistician employed by the McMaster Medical center. The Statistician will be requested to determine the probability that HRV characteristics can separate CHF and non-CHF patients.

5.5 Summary

In this chapter we outlined our intent to examine the HRV characteristics of SOLVD patient data in order to find a relationship between a patient's HRV properties and the onset of CHF.

Our approach will involve a software based automated ECG fiducial placement system. The software shall separate the cardiac signal from ambient noise by first applying one or more Band Pass filters (wavelet or otherwise). The fiducials shall then be placed on the filtered data using a trained model method. Once HRV data has been collected statistical analysis will be employed to examine the relationship between HRV characteristics and patients with and without CHF. Algorithms for measuring the QT interval will be developed but will not be used to analyze patient data.

Chapter 6: Algorithms for ECG Analysis and Fiducial Placement

6.1 Introduction

In this chapter we explore the various methods and algorithms used in this study to filter, annotate and analyze ECG data. An explanation of the clinical significance of the various HRV parameters along with details on how they are obtained from the HRV signal is given. Model optimization methods based on Genetic algorithms are also discussed. Finally, statistical methods for discriminating between CHF and non-CHF patients using their time and frequency domain indices of the HRV are introduced.

6.2 Filtering Methods

The common noise components that are often found on ECG signals was discussed in Chapter 5. The goal of filtering is often to reduce or minimize the effects of noise on the signal. In this study however, the reduction of noise is not necessarily the only goal of filtering. To insure the most effective model for ECG fiducial placement the filter must suppress only those noise components that could be confused as features as well as emphasize true features. To accomplish this it has been found multiple filters are necessary if a range of features are to be examined (Li et al, 1995; Shambi et al, 2000). Section 6.2 describes in detail the filtering techniques used in this study and their effect on ECG signals.

6.2.1 Band Pass Filtering

Band Pass filtering of ECG signals has been explored heavily by researchers in the field (Hu, 1999; Berger et al, 1997; Ruha et al, 1997; Shankara et al, 1992; Riccio et al, 1992; Friesen et al, 1990, Sadeh, 1987; Pan and Willis, 1985). Techniques have ranged from linear digital filtering (Ruha et al, 1997) to adaptive matched filtering (Xue, 1992). The primary goal of the band pass filter is to isolate the QRS complex with a secondary goal being the isolation of the T wave.

Given the large volume of data that must be processed it was determined that a linear filter would require the least amount of computational overhead for the task. However, in many threshold based feature detection methods, much work must be done to insure phase information for the signal is not lost (Shankara et al. 1992). By utilizing a trained model method, an offset generated by unmatched linear filtering can be corrected by the model. Therefore, a simple bandpass filtering techniques can be used to ease the detection of the QRS complex. Sadeh et al. (1987) used such a filter, applied in the frequency domain, for a correlation based QT detection algorithm with success. The details of the method, which was adapted for our study are as follows:

Step 1. Let the impulse response of the digital filter be h(t) in the time domain and the ECG signal be x(t).

Step 2. Compute the Fourier transform of both h(t) and x(t) resulting in H(s) and X(s)

Step 3. Compute the product of the filter and the ECG waveform such that:

$$Z(s) = H(s)X(s) \tag{6.1}$$

Step 4. Obtain the filleted ECG by taking the inverse transform of Z(s)

A modification to these steps used in this study was to start with H(s). H(s) was constructed using the following (Sadeh et al.1987):

$$H(s) = \left[\frac{1}{\sqrt{1+\left(S_L/s\right)^{2n}}} \right] \left[1 - \frac{1}{\sqrt{1+\left(S_H/s\right)^{2m}}} \right] \qquad (6.2)$$

Where $H(S_L)=1/\sqrt{2}$ and $H(S_H)=1-(1/\sqrt{2})$ and n, m are the orders of the filter, with higher values resulting in a sharper cut-on.

For this study, software was constructed the allowed the user to specify the centre frequency, full width half maximum (FWHM) and slope of the filter in the frequency domain. The Fourier transform algorithm that was used for the band pass filtering routine operates on a data set of size N in $O(N \log_2 N)$ time and can be found in Numerical Recipes in C (Press et al, 1997). Further details of this software are covered in chapter 7.

6.2.2 Wavelet Filtering

An alternative method to linear band pass filtering is filtering using wavelets. Wavelets have the advantage of being able to capture the frequency make up of a signal as well as produce and inference of where those frequencies occur in time (Dremin et al, 2001). This property allows the wavelet transform to be particularly well suited for non-stationary signals such as ECG (Li et at. 1995). Through scaling of the wavelets, the presence of high and low frequency components can be determined for a particular moment in time (Bryll, 2000). The basis of wavelet analysis stems from the proper

choice of a basis function or mother wavelet , Ψ(t). In most cases, Ψ(t) should be infinitely differentiable and vanish as $t \rightarrow \pm \infty$. The wavelet transform is defined as (Sahambi et al., 2000):

$$Wf(\alpha,\tau) = \langle f(t), \psi_{\alpha,\tau}(t) \rangle \tag{6.3}$$

Where <...> denotes the inner product and:

$$\psi_{\alpha,\tau}(t) = \frac{1}{\sqrt{\alpha}} \psi\left(\frac{t-\tau}{\alpha}\right) \tag{6.4}$$

α is called the scaling parameter and τ is the translation parameter. In practice, α is used to examine the frequency power at various resolution contained in the signal at time t. τ is used to examine the presence of a particular frequency band at various times within the signal. The mother wavelet used this study was the first derivative of the smoothing or Gaussian function, similar to the wavelet chosen by Sahambi et al. Thus, Ψ(t) can be expressed as follows:

$$\psi(t) = -te^{\frac{-t^2}{2}} \tag{6.5}$$

Scaled versions of this wavelet can were used as band pass filters with varying bandwidths and centre frequencies. The wavelet filter was applied to the ECG data in a similar manner as the band pass filter. However, in the case of the wavelet filter, we start with the time domain representation of the filter (h(t)) and subsequently take it's Fourier transform. This method of applying the wavelet filter has the advantage of only

computing the filtered ECG signal that we need for analysis of, as well as gives us the flexibility to adjust the scale to an optimal value for the particular waveform. Another method of computing the wavelet transform is through the use of Mallat's algorithm (Vidakovic and Mueller, 1991). However, this method forces one to use discrete values of α with may not be optimal for the particular ECG.

6.3 Construction of ECG Model for Fiducial Placement

Model based pattern matching methods have an advantage over straight filter and threshold methods because of their ability adapt to both varying patient physiologies and varying environmental conditions (Xue et al. 1992; Sadeh et al, 1987). In this study we have implemented a model similar to that developed by Berger et al (1997) for the detection of QT intervals. The details of our implementation are described below.

6.3.1 Zone Scaling

The first step in the construction of our model was to split the ECG waveform up into areas of similar frequency content. This was done for 2 reasons. The first was so that the model could be scaled non-uniformly across the potential ECG signal. The second reason was to allow multiple filters to be applied to the ECG so that only the feature on interest would be enhanced while all other features and noise would be suppressed.

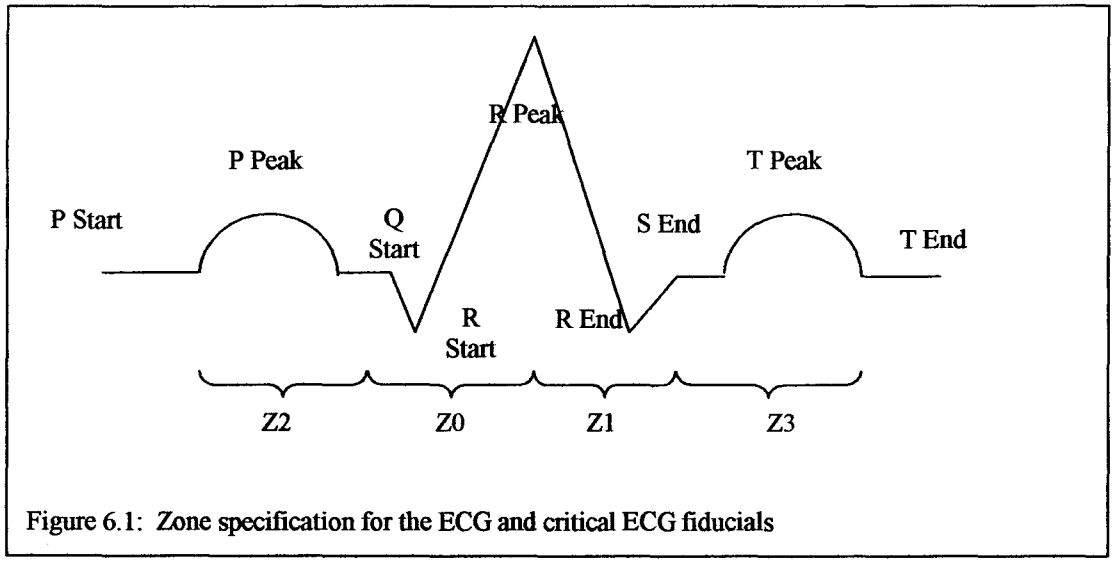

Figure 6.1: Zone specification for the ECG and critical ECG fiducials

Zones were specified based on the findings of Li et al (1995) Martinez et al. (2000) and Sahambi et al.(2000). Using wavelet filtering, these groups found that The QRS complex could be isolated using quadratic spline wavelets at scales of 1 or 2 (Martinez et al. 2000). P and T waves were similar in that they could be isolated using scales of 4 or 5 , although 5 exposed the waves to a great deal of baseline wander (Li et al, 1995). Using this information, the ECG waveform was split up into 4 zones. Zones 0 and 1 contained the QRS complex, Zone 2 was dedicated to the P wave and Zone 3 was created for the T wave. Figure 6.1 shows the start and end points of the Zones. Although Zones 0 and 1 represent the same filter regime, they are treated as separate because of the potential need to scale the QRS complex non-uniformly.

The use of zones allows the model to be constructed piecewise. This means that when the model segment for each zone can be compressed of stretched by a known amount. The result is a model that can obtain a better fit to the potential ECG waveform

than if the entire ECG model was scaled uniformly, as was the technique presented by Berger et al (1997).

6.3.2 Model Construction

A set of critical fiducials was chosen for the model. They are:

i. Start of the P wave (PwaveS)

ii. Peak of the P wave (Pwave)

iii. Start of the Q wave (QwaveS)

iv. Start of the R wave (RwaveS)

v. Peak of the R wave (Rwave)

vi. End of the R wave (RwaveE)

vii. End of the S wave (SwaveE)

viii. Peak of the T wave (Twave)

ix. End of the T wave (TwaveE)

Each of these model fiducials had additional information associate with them within the ECG model such as mean and variance of the fiducial location relative to R-wave.

The model was constructed by first selecting 30 continuous beats of the patient's ECG. These beats were in a section of ECG that contained very little ambient noise such as base-line wander. For each beat within the sample, all the critical fiducial points had to have been identifiable. Fiducials were placed on each of the 30 beats, forming an annotation file. The model was constructed by deconstructing the raw ECG data with 3 filters, each designed to emphasize the feature of interest for each zone. Then, using the

pre-defined annotations as a template, the region of interest was cut out from each of the filtered data set to create a series of model waves. These model waves were averaged to form a single most likely model wave. The wave was further deconstructed by approximating each zone with its cubic spline approximation. The number of points that are used for the cubic spline are zone specific and are specified by the user. The model constructor would then use the annotation file to generate maximum likelihood statistics for each of the fiducials. The local 30 beat R-R interval and average filtered R-wave height was also stored. Figure 6.2 shows this deconstruction/construction process in pictorial format.

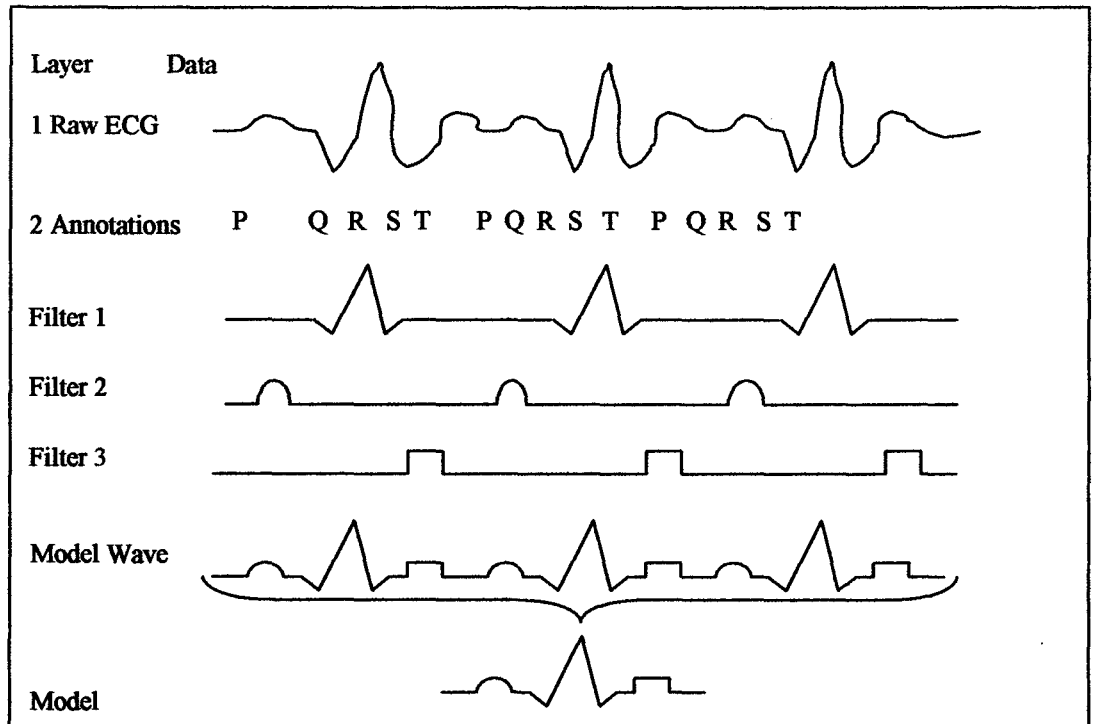

Figure 6.2: Model construction From Raw, annotated ECG

6.3.3 Model Optimization using Genetic Algorithms

A genetic algorithm (GA) is a very simple machine learning structure meant to simulate the most prevalent natural optimization method on earth, namely, the evolutionary process. The structure of the GA is general enough that it can solve many different types of problems with similar levels of efficiency through its inherently robust search strategies (Goldberg, 1989). One of the great attractions of genetic algorithms in optimization problems is their ability to find global solutions using only goal oriented merit function and a set of adjustable parameters (Back, 1996). This study uses a genetic algorithm to optimize the ECG auto-annotation model by searching for optimal model parameters based on fiducial placement accuracy. The software developed uses a library developed by Matthew Wall (1995) called GALib, which is distributed, free of charge, by Massachusetts Institute of Technology. GALib is based on work published by Goldberg (1989). The following subsections explain how genetic algorithms work and how they are applied in this study to the formulation of a patient specific model of the ECG.

6.3.3.1 How genetic algorithms work

Like most biological systems on earth, a genetic algorithm starts off with a set of genomes (Back, 1996; Goldberg, 1989) or chromosomes (Davis, 1991) and an ecology. Through processes of breeding, mutation and death the sets of genotypes are generated from genomes and are tested for fitness against the ecological goal (Back, 1996). Those genotypes that poorly meet the goal die off while those that satisfy it breed and populate the gene pool. Since the genotypes are generated from genomes spanning the entire

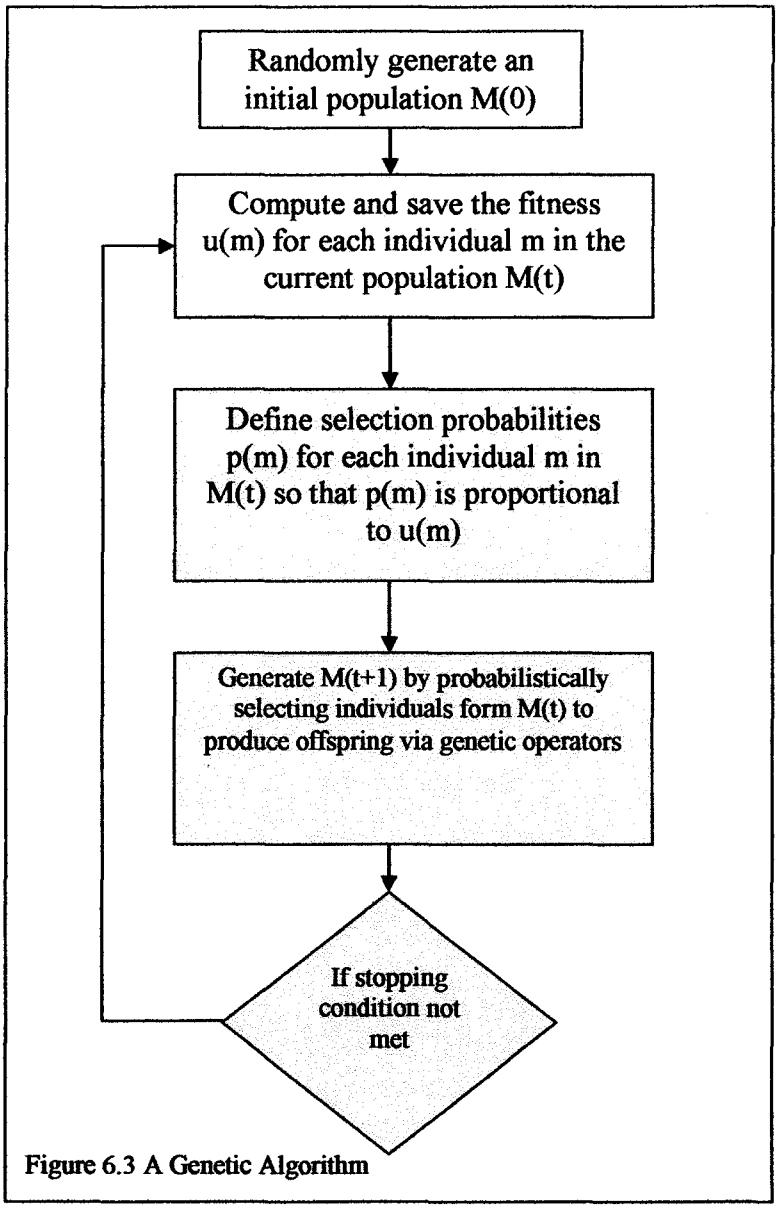

Figure 6.3 A Genetic Algorithm

solution space the likelihood that the global solution to the ecological goal will be found is quite high.

From a computational point of view, GA's use various statistical techniques to maximize or minimize a merit function given a set of variable parameters. A schematic diagram of the evolutionary process is given in figure 6.3 (Whitley, 1993).

6.3.3.2 Parameters affecting model Optimization

There are a number of different parameters that can be adjusted to have an effect on the performance of at GA. The key parameters are the merit function, scaling, population size, mating, mutation and crossover. For now, we will leave the design of the merit function for this study for later discussion.

Population size and selection techniques are closely linked. Population size is the maximum number of genotypes that can exist at any one time. To maintain the population size after mating of genotypes some older genotypes must die. To determine this, the merit function values obtained by each member of the population are first scaled using a scaling function. The most common of such a function is the linear static scaling function (Back, 1996):

$$\sigma(MF) = c_o F(MF) + c_1 \tag{6.6}$$

Where F(MF) is a function that converts the merit function result into a real number. This method is critical in establishing distance measures $\Phi(\sigma(MF_i))$ between members (i) of the population. Here again is another scaling factor in that the distance measure could be the minimum Euclidian distance (MED) or some other mapping. However, MED is the simplest and most common. The next step is to calculate survival probabilities of different individuals based on their distances from the goal $\Phi(\sigma(MF_{goal}))$. The chance of survival of a member (i) is governed by (Back, 1996):

$$p_{death}(\Phi_i) = \frac{\Phi_i}{\sum_{j=1}^{n} \Phi_j} \tag{6.7}$$

where n is the population size. Selection of the new population may take place before or after p_{death} is calculated. If this occurs, mating between genotypes occurs randomly and not all the children are expected to survive as only n of the least likely to die are selected for the next population. If mating is done afterwards, n-ν members of the population are allowed to survive and mate to create ν children.

Mutation and crossover dictates what those ν children look like after they have been mated. Crossover takes properties from both parents and combines them into the children. Crossover can either occur by selecting a parental data in either a fixed or random method. Fixed methods could mean a fixed amount of data is selected from a fixed point in each parent. Random selection is just the opposite. However, straight parameter swapping does not introduce new genetic material into the gene pool. Thus a system for slightly modifying individual parameters from both parents can be expressed as follows (Haupt and Haupt,1998):

$$P_{new} = \mathrm{B}P_{mother} + (1-\mathrm{B})P_{father} \tag{6.8}$$

where B is a random number between 0 and 1. P_{mother} and P_{father} are similar properties from each parent that are being mutated into the child. All or some of the swapped chromosomes can be transferred in this way.

Mutation is a way in which the solution can be prevented from converging too quickly. Mutation involves changing random chromosomes in random genotypes in the population after each generation. This is often done bit-wise by changing a random bit in a parameter from 1 to 0 or from 0 to 1 (Davis, 1991) although more complex methods are also possible and have been explored in Whitley (1993).

Finally, the number of iterations that a GA is allowed to run will affect the degree of convergence once the operation is complete. GA's are designed to run indefinitely. It is up to the user to decide through experimentation how many iterations are enough to achieve a satisfactory result.

6.3.3.3 GALib parameters for ECG Model building

In this study we selected the SimpleGA genome type available in the GAlib library for our optimization. The genome parameters, optimized through experimentation were as follows:

i. Goal: Minimization of merit function
ii. Scaling: Linear static scaling
iii. Population size: 30 (note larger was better but computationally expensive)
iv. Number of generations: 100
v. % chance of mutation: 0.1% (Maintained the software default)
vi. % chance of straight parental crossover: 90%

Genomes or variables were selected for each zone. In altering our ECG model we optimized each zone individually. The genomes were scale factor (if wavelet filter), centre frequency, FWHM and slope (if band pass filter), number of spline points and the R window detection threshold.

6.3.3.4 Design of the Merit Function for ECG Model optimization

The merit function has the greatest influence on GA performance as it is used to decide which genotypes live, die and mate. The design of the merit function varies according to the application with internal calculations being problem specific. In this

study a merit function was developed that measured the total prediction error of the zone in question. The calculation is based on the mean squared error of the fiducial placement relative to the user placed fiducial. The merit function value was calculated as follows:

$$MF_{Zone}(A) = \frac{1}{M}\sum_{i=0}^{M}(A_i - F_i)^2 \tag{6.9}$$

Where M is the total number of beats used in the training data, A is the location, relative to R, of the user placed fiducial and F is the location of the model placed fiducial.

6.4 Auto-Annotation of ECG Data

The auto annotation process occurs in 3 steps. The first step is a scanning step, which looks for R-waves. Once an R-wave is found, the model is fitted to the surrounding filtered data. Once an optimal fit is achieved, various annotations are placed on the ECG. Here we discuss the steps in detail.

6.4.1 Detection of the R-wave

The R-wave is found by first scanning the filtered data for a portion of the signal that is greater than a threshold. The threshold value is stored in the ECG Model and is determined using statistical properties (mean and variance) of the R-wave height. The algorithm then looks for a point where the data dips below the threshold. If the width of the entry and exit point matches the statistical model width with a high degree of

certainty the maximum value between the two points is found. This maximum value is recorded as an R-wave location.

The algorithm then used the R-R interval information stored in the model to determine the maximum likelihood of the next R-wave. If a spike is detected, the probability that it is an R-wave is determined based on the number of standard deviations it is from the most likely R-wave location. The merit function for determining if the detected spike is and R wave is as follows:

$$MF_{RR} = w\frac{p(R_x)+p(R_y)}{p(R_{x0})+p(R_{y0})} \tag{6.10}$$

Where p(Rx),p(Ry) is the probability of the R-wave being at location Rx and having a peak voltage of Ry. p(Rx0) and p(Ry0) are the probabilities that the position and voltage of the R-Wave will have the mean values Rx0 and Ry0, respectively. W is a weighting factor that is set during the model building phase.

If MF is > 0.5, in equation 6.10, the spike is deemed to be a true R-wave.

If the R-wave is not found within four standard deviations of the mean R-R interval the most likely location is marked with a "Unkwn" annotation and the algorithm looks ahead for the next beat. A flow chart of decision process can be found in Appendix B. It is important to note that the local mean R-R interval and mean R voltage is recorded for the last 60 minutes of data and is used to update the model parameters. In this way

the model learns and adapts to changing conditions but does not over react to sudden, short lived, variations.

6.4.2 Detection of the P wave, T wave and QRS complex

Once the R wave has been detected the Auto Annotation algorithm proceeds to identify the other fiducials in the wave. This is done by scaling portions of the model to obtain a best fit to the data. . The goodness of fit is determined using the merit function defined by Berger et al (1997).

$$MF_{Zone} = \frac{1}{n}\sum_{i=0}^{n}(X_i - Z_i)^2 \qquad (6.11)$$

Where n is the length (number of data points) of the Zone, Xi is the value of the filtered ECG data at point i from the R wave peak and Zi is the value of the scaled Zone. The best fit is determined by the scale at which the minimum MF in equation 6.11 is achieved. The MF is also compared to a threshold to determine is a minimum goodness of fit is met. If it is not met an "unknwn" fiducial is placed at the maximum likelihood position of the end fiducial. The search for the best scale is not exhaustive. Searching is achieved through the use of a binary search pattern. the procedure is as follows:

Step 1. Scale the Zone so that the Zone length is equal to the most likely position of the end fiducial as well as midway between ± three standard deviations from the mean. These points are the first pivot min and max points of our search.

Step 2. Determine the MF's for all three scales

Step 3. Select the Smallest MF scale as the new pivot point and the min and max points to be the scales midway in between the old points to the left and right of the new pivot.

Step 4. Obtain Scale for the Zone at the new pivot point as well as the new min and max points.

Step 5. Repeat steps 2 – 4 until the difference the min and max points drops below a threshold value.

The search and label procedure is carried out unit all four zones have been labeled. It must be noted that while the fiducials at the end points of the zone are scaled the fiducials inside the zone are not fit to the data. They are placed at fixed ratios between the beginning and end of the Zone.

6.5 R-R interval analysis

Once the data has been annotated, the measurement of the R-R intervals is straightforward. The system reads the annotation data which is saved as a separate layer beneath the raw ECG data. The algorithm scans though the annotation layer looking for instances of the "R-wave" fiducial. The time between "R-wave" fiducials are recoded in a separate data layer for viewing and analysis. This data is then inverted to convert R-R interval information (seconds per beat) to heart rate (HR) data (beats per second).

6.5.1 Ectopic correction

An ectopic beat is a neurocardiac signal that is initiated by a nerve bundle other that the SA node. Ectopic beats may occur due to the presence of a full or partial branch

block with causes internal pacemakers within the heart to generate a signal. It is important to remove ectopic beats from R-R interval data prior to power spectrum analysis because R-R interval analysis assumes that we are working with neuroly regulated beats and an extraneous beat is an electrical artifact.

The algorithm used for ectopic correction in this study is based on work published by Cheung et al (1981). The algorithm combines an error detection algorithm with an interpolation algorithm. The steps for the correction scheme are as follows:

Step 1. An R-R interval measurement is selected sequentially from the data

Step 2. It is deemed to be a autonomic controlled beat if its value is within 30% of the past four beats that have passed the test.

Step 3a. If it is a good beat it's value is recorded in a buffer (containing 4 beat values) and step 1 is repeated.

Step 3b. If it does not pass the, the program will find the first correct beat before and after the beat(s) in question.

Step 4. If there are more than one irregular interval in succession, a check is made of the number of erroneous beats and their total duration.

Step 5a. If this duration longer than the span of four recent correct beats, then the interval is tagged as an error that can not be corrected. The time series is then terminated at the last correct beat.

Step 5b. If the interval is less than four beats, the time duration between the two correct beats is now replaced with a combination of either the sum or two adjacent

beats, or divided into two beats or interpolated with a linear spline. The spline interpolation is calculated as follows:

If X(n) is the HRV signal for which an ectopic beat exists for an instant n, then estimate:

$$X(n) = X(n_0) + \left(X(n_1) - X(n_0) \frac{n - n_0}{n_1 - n_0} \right) \qquad (6.12)$$

where n(0) is the value of the signal at instant immediately prior to n and (n1) is the value of HR at an instant subsequent to n.

6.5.2 Time domain measurements

Of the various methods available for the investigation of HRV, perhaps the simplest to calculate and interpret are the time domain indices. These measures can be derived directly from the intervals between normal sinus conducted beats themselves (N-N intervals); and from the difference between adjacent cycles (Kleiger et al. 1992; Task Force 1996). The mean heart rate, standard deviation of all N-N intervals in a given time period, standard deviations of the mean heart rates of short recording segments, (commonly 5 min.) and the mean of the standard deviations of heart rate over similar short recording segments are completed from sequential N-N interval time series. Time domain indices are defined in the Glossary. These indices represent, not only short term beat-to-beat variations in heart rate but also oscillations of much longer time course, (such as diurnal variations). Time domain measures derived from the differences between adjacent N-N intervals include the percentage of adjacent variables that differ by more than a threshold value (commonly 50 ms or 6%) and the root mean square of

successive differences. These measures are therefore independent of secular or circadian trends and represent short term variability thought to be modulated solely by the parasympathetic nervous system (Kleiger et al. 1992).

The time domain parameters that were chosen for our study and there descriptions are as follows:

i. HR: Mean HR over the 24 hours of data
ii. NN: Mean R-R interval
iii. SDNN: Standard deviation of the 24 hour mean R-R interval
iv. SDNN Index: Mean of the standard deviation of intervals between normal sinus conducted beats in adjacent 5 minutes of the entire 24 hour period
v. SDANN: Standard deviation of SDNN index.
vi. R-MSSD: Root mean square of successive differences between normal sinus conducted beats in a the 24 hour period.
vii. pNN50: Percentage of beats where the time difference between them is exceeded by 50 ms or more.

6.5.3 Frequency domain measurements

The benefit of frequency domain analysis of the HRV signal lies in its ability to identify frequency specific oscillations of heart rate which correspond to distinct physiological mechanisms and thereby provide a comprehensive picture of neurocardiac regulatory system (Harvey, 1997).

Akselrod et al. (1981) identified peaks in three distinct frequency bands of the heart rate power spectrum of unconscious, anesthetized dogs. These same three frequency

specific oscillations had earlier been noted by Sayers (1973) in conscious human subjects. A high frequency peak (0.15Hz and 0.35Hz) is believed to correspond to parasympathetically mediated respiratory sinus arrhythmia (Akselrod et al. 1981 & 1985). In addition, two peaks are evident at lower frequencies (0.10Hz and 0.04Hz). The mid-frequency peak (0.10 Hz) is believed to correspond to baroreceptor modulation of heart rate and contains both a sympathetic and parasympathetic components (Akselrod et al. 198; Pagani et al. 1986; Task force 1996). Currently, many investigations focus on the power in well defined high frequency (HF) (0.15-0.40 Hz) and low frequency (LF) power bands (0.04-0.15 Hz).

Power spectral analysis utilizes either the fast Fourier transform (FT) or the autoregressive modeling (AR) algorithm to decompose the heart rate signal into its frequency components. The Fast Fourier method for computing the power spectrum can be expressed as:

$$P_x(f) = |X(f)|^2 \tag{6.13}$$

Where X(f) is the Fourier transform of the HRV signal. The FT method provides a simple and efficient measure the power spectrum. Unfortunately, the raw FT method without is not as accurate as other methods due to its strong ties with the Nyquist critical frequency (Press et al, 1997).

Hu (1999) investigated the use of the AR method for computing the power spectrum of HRV signals using the Yule-Walker equations. The method is summarized with the following equation (Hu):

$$P_{AR}(f) = \frac{\sigma^2}{\left(1 + \sum_{k=1}^{p} a[k] e^{-j2\pi fk}\right)^2} \tag{6.14}$$

Where a[k] are the coefficients that must be solved by the Yule-Walker equations and σ^2 is the white noise power. Details on how to solve for σ^2 the a[k]'s using Yule-Walker equations can be found in Hu (1999) and in Proakis and Manolakis (1996). Since (6.13) can be solved in O(N log N) time (Press et al , 1997) and (6.14) in $O(N^2)$ time, both methods were developed so that the computational time/accuracy trade-off could be made by the user.

The Frequency domain parameters the were chosen for our study and there descriptions are as follows:

i. % LF area: Total LF area as a percentage of the total power area

ii. % HF area: Total HF area as a percentage of the total power area

iii. LF:HF: Ratio of LF area to HF area

iv. ABS LF: absolute LF area

v. ABS HF: Absolute HF area

6.6 QT interval analysis

The QT interval begins at the onset of the QRS complex and terminates at the end of the T wave. The algorithm gathers QT interval data by searching for a "QwaveS" fiducial. Once it finds it the algorithm then looks for the TwaveE fiducial. If, instead it

finds another "QwaveS" it discards the last "QwaveS" and looks for the TwaveE relative to the new fiducial. The QT intervals are saved sequentially in a separate data layer. The QT interval is inversely related to heart rate. Therefore, the QT interval should be corrected for heart rate using Bazett's formula:

$$QTc(i) = \frac{QT(i)}{\sqrt{RR(i-1)}} \tag{6.15}$$

The QT interval was also transformed to the frequency domain using the same methods as was used for the HRV time domain data.

6.7 Statistical analysis of resultant data

In our study we have various time and frequency parameters that come from two sample distributions; those with CHF and those without. For of the parameters have an associated mean ($\bar{U}_{CHF}$,$\bar{U}_{NCHF}$) , variance (S_{CHF} ,S_{NCHF}) and Sample size (N_{CHF} ,N_{NCHF}). We wish to test the hypothesis that $\bar{U}_{CHF} = \bar{U}_{NCHF}$. If it turns out that the hypothesis is likely to be true we can say that the parameter is not likely to be a good discriminator between he two groups. If however it is likely to be false, then there is a chance that the parameter can be used to discriminate. Since our sample size is greater than 30, the hypothesis test is done using the z distribution. Our level of significance (two tailed) was taken to be 0.05. If the probability (p-value) that hypothesis is true was less than our level of significance, we can say that the means likely come from different populations. Our variables are calculated as (Mason, 1978):

$$z = \frac{\hat{U}_{CHF} - \hat{U}_{NCHF}}{\sqrt{\left(\frac{S_{CHF}}{\sqrt{N_{CHF}}}\right)^2 + \left(\frac{S_{NCHF}}{\sqrt{N_{NCHF}}}\right)^2}} \quad (6.16)$$

Degrees of freedom (v) = $N_{CHF} + N_{NCHF} - 2$ (6.17)

Using z and v the p-values is determined using a table or a software package. To keep the experiment unbiased, Ms.Janice Pogue, M.Sc. from the department of Medicine (cardiology division) was requested to perform the statistical analysis.

6.8 Summary

In this chapter, we introduced the use of filtering in the frequency domain using both band pass and wavelet techniques. Auto annotation of the raw ECG data will be accomplished by scaling various sections of a pre-constructed ECG model in order to obtain a best fit to the section of ECG data in question. Critical fiducials, needed for QT-interval and HRV analysis, will be placed based on the scaling of that model. The model construction will involve gathering statistical data from a number of pre-annotated beats within the ECG file. An optimization of the model using genetic algorithms is proposed to improve fiducial placement accuracy. Fourier and autoregressive techniques will be used to calculate the HRV and QT power spectra. The HF and LF frequencies of the HRV power spectra were introduced and their relationship the autonomic function of the heart was explained. The significance of various time domain measures was also explained. Once the time and frequency parameters have been collected the null hypothesis will be examined via the t-test for each variable to see if they are potential discriminating indicators of CHF or non-CHF patients.

Chapter 7: Experiment Methodology

7.1 Introduction

In this chapter we shall examine the methods used in this study to acquire and analyse 24 hour Holter data. An explanation of the methods used to test the various algorithms developed for this experiment shall also be provided. The GUI based software program called Vitural ECG Lab (VEL), which was developed for this study, shall be introduced and its capabilities discussed.

7.2 Data Acquisition: The SOLVD Tapes

In chapter 2 we identified the source of our data to be from the SOLVD study conducted from 1986 to 1990. The study collected pre and post treatment data on 4228 patients with known instances of heart disease. To qualify for the study, patients had to have been previously diagnosed with heart disease and exhibited ejection fractions of 0.35 or less. In addition, selected patients had not been prescribed diuretics, digoxin of vasodilators as a therapy for their heart failure. The breakdown on the physiology of the patients is presented in Table 7.1.

After acceptance into the trial, each patient was fitted with a magnetic tape Holter recorder. The recorder electrodes were placed as close to standard V1 position as possible (see appendix A for lead placement information). The recording rate of the tape

Table 7.1: Baseline patient demographics for SOLVD (The SOLVD Investigators, 1992)		
Characteristic	**Placebo (N = 2117)**	**Enalapril (N = 2111)**
	mean value	
Age (yr)	59.1	59.1
Ejection Fraction	0.28	0.28
Blood Pressure (mm Hg)		
Systolic	125.6	125.3
Diastolic	78.0	77.9
Heart Rate (beats/min)	75.2	74.6
	% of group	
Male Sex	88.6	88.5
NYHA Functional Class†		
I	67.1	66.3
II	32.7	4.1
History		
Ischemic Heart Disease	82.9	83.5
Myocardial Infarction	79.4	80.5
Hypertension	37.3	36.8
Diabetes Mellitus	15.1	15.4
Idiopathic Dilated Cardiomyopathy	10.1	8.6
†Five patients in NYHA class III were inadvertently enrolled in the Prevention Trial and have been retained in the analyses. No deaths or hospitalizations occurred among these five patients.		

was set to 1 inch per second. As SOLVD was a multi-national, multi-institution study a variety of Holter records were used with varying tolerances on recording speed, tape

distortion and recording quality. As the make and model of Holter used for each patient was not recorded, the data taken from the tapes must be assumed to be of uniform recording quality. The Holter tape contained enough magnetic material to record 24 hours of patient data.

A subset of SOLVD patients, consisting of 400 randomly selected samples was selected for this analysis. The samples were selected by an independent 3rd party who used a random selection technique to generate a list of patients from both the CHF and non-CHF category. The selection was blind. At no point, before or after the experiment, did we know which group the patients came from. Once the 400 patients were selected their ECG's were examined to insure the data was suitable for HRV analysis. The criteria for this selection were described in chapter 4. Of the 400 selected patients 222 were deemed acceptable for further study.

7.3 Data Digitization

The SOLVD Holter data tapes were digitized using a reel-to-reel play back device and an A/D converter. The reel-to-reel tape player was a Marquette Electronique model 22-2 professional grade, multi channel audio playback and recording system. Its settings were adjusted to play back the Holter tapes at a rate of 60 inches per second. Thus, the total time to playback and digitize a tape took approximately 24 minutes per tape. The outputs from the tapes were captured via the reel-to-reel's line-out jack.

The analogue signal was channelled into a National Instruments model PCI-6023E PCI data acquisition board which was hooked up to a National Instruments BNC-

2120 connector block. The data acquisition card had a resolution of 12 bits and a maximum sampling rate of 200 000 samples per second. However, such a high sampling rate was not needed. The card was set at 30000 samples per second so that we would obtain 500 samples per second on the Holter tape. A 500 Hz sampling rate was chosen based on recommendations published by Kamath et al (1993) and Merri et al. (1990).

During acquisition, data was saved in 1 hour intervals to hard disk. Each patient was assigned a directory, which was labelled with a unique identification marker. The 1 hour data sets were saved as floating point values (32 bit) in a binary file. The binary file name contained the identification marker and a numeric value representing the hour from which the data was taken. Software was developed in LabView™ to manage the digitization process. The front panel of the software in operation can be found in figure 7.1a. The wiring diagram for the LabView™ VI is shown in figure 7.1b.

Each 1 hour file was 7032 Kbytes in size and contained 1800000 data points. In total, a patient's 24 hour ECG record occupied 168Mbytes of disk space. Storage requirements were reduced considerable by compressing the data using WinZip™, a freely available data compression utility. Using the WinZip™ software each patient folder was compressed to approximately 20% of its original size.

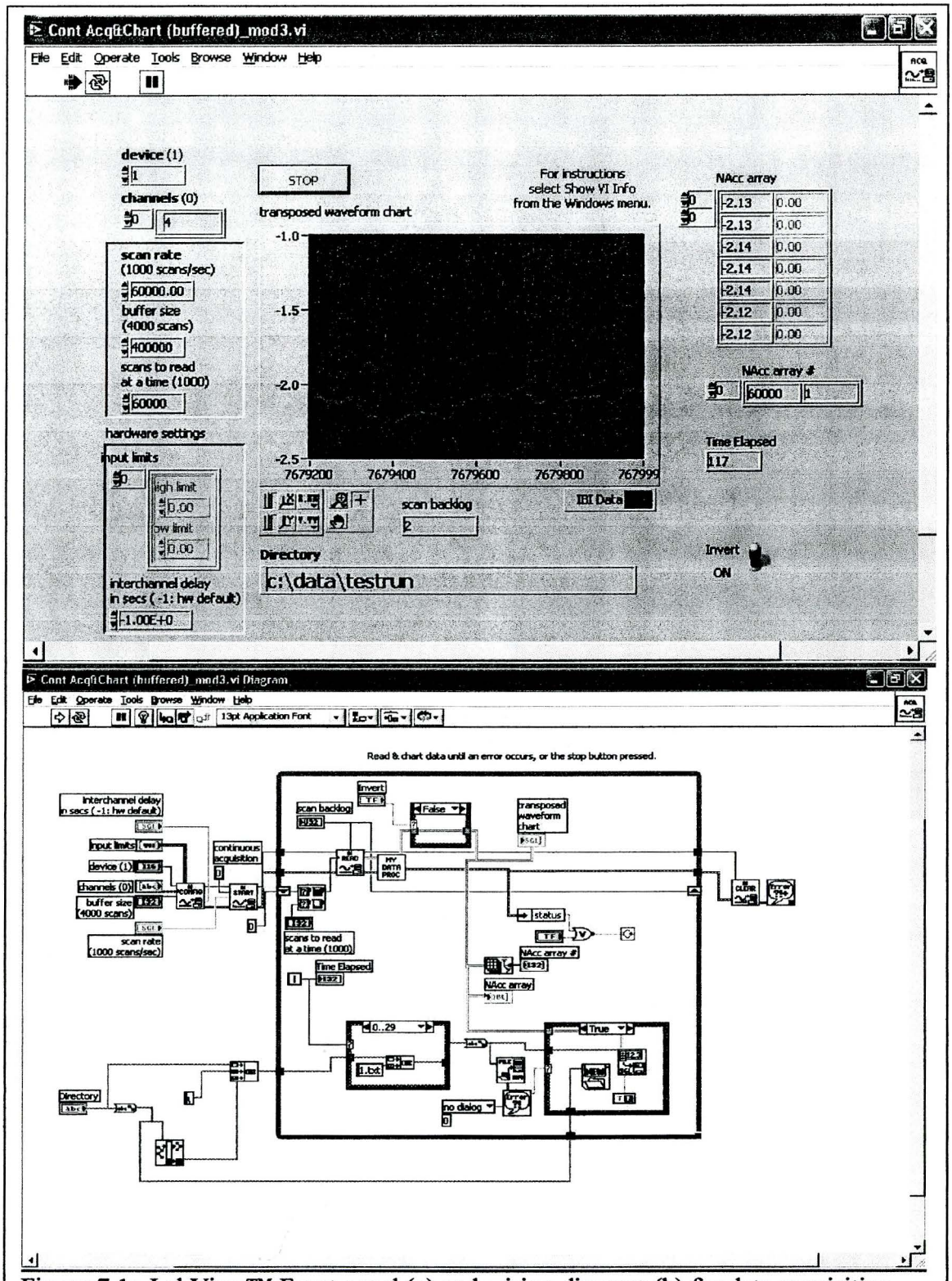

Figure 7.1: LabView™ Front panel (a) and wiring diagram (b) for data acquisition

7.4 Algorithm Testing

7.4.1 User interface

VEL (Virtual ECG Lab) was created to assist the user in annotating and analyzing ECG signals. The goal was to develop software simple enough for routine clinical use but would also have advanced features that would allow a researcher to adjust various low level functions to optimize its performance. Figure 7.2 contains a screen shot VEL being used to manually annotate a section of data.

The user interface was tested for ease of use by a technician with experience in using windows software and ECG analysis but little or no knowledge of signal processing or computer science. The testing consisted of asking the user to open a data set, manually annotate data, build a model, auto annotate the data and then perform HRV analysis. Success was determined based on the user's ability to operate the software successfully with little or no assistance once he or she has been shown how to use it a few times.

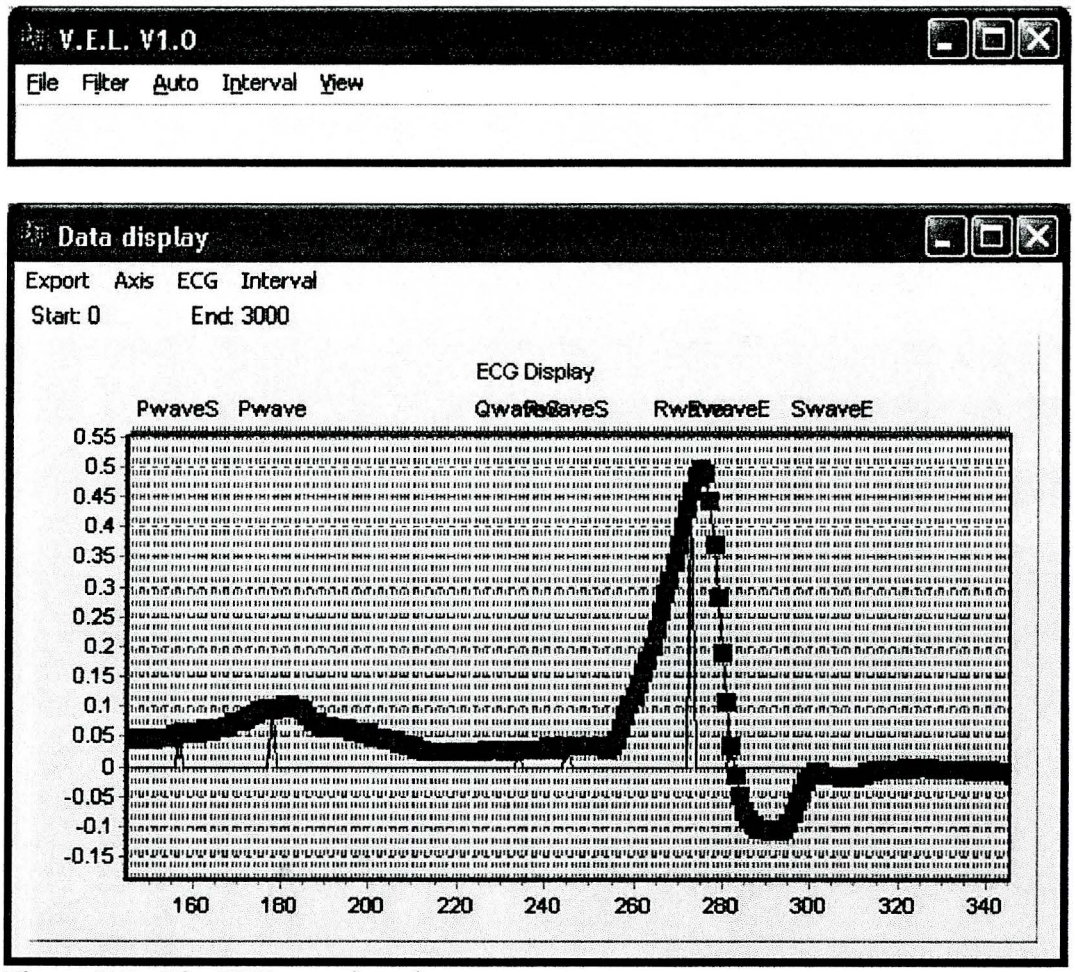

Figure 7.2: The VEL user interface

7.4.2 Filter Creation and ECG Filtering

FilterBuilder is filter creation software developed to assist the user in building and modifying filters to be used in VEL. It allows complex filters to be created by applying individual filters sequentially to ECG data. The filters available in FilterBuilder that can be stacked in this way are low pass, high pass, notch, band pass and quadratic wavelet filters. FilterBuilder allows the user to add, delete and modify any sequence of filters in a series as well as modify the properties of individual filters in the sequence. Figure 7.3

shows a screen shot to FilterBuilder being used to add a band-pass filter to a sequence

The program also allows the user to preview his of her filter before applying it to the data. The filtering algorithms in Filterbuilder were tested by applying various filters to data with known frequency components as well as white and coloured noise signals. The software was deemed successful if it could accurately and repeatably filter various known noise signals from a test signal. The performance of the wavelet filters on ECG data was tested against by comparing results in Filterbuilder with known results published by Martinez et al (2000), Sahambi et al (2000) and Li et al (1995).

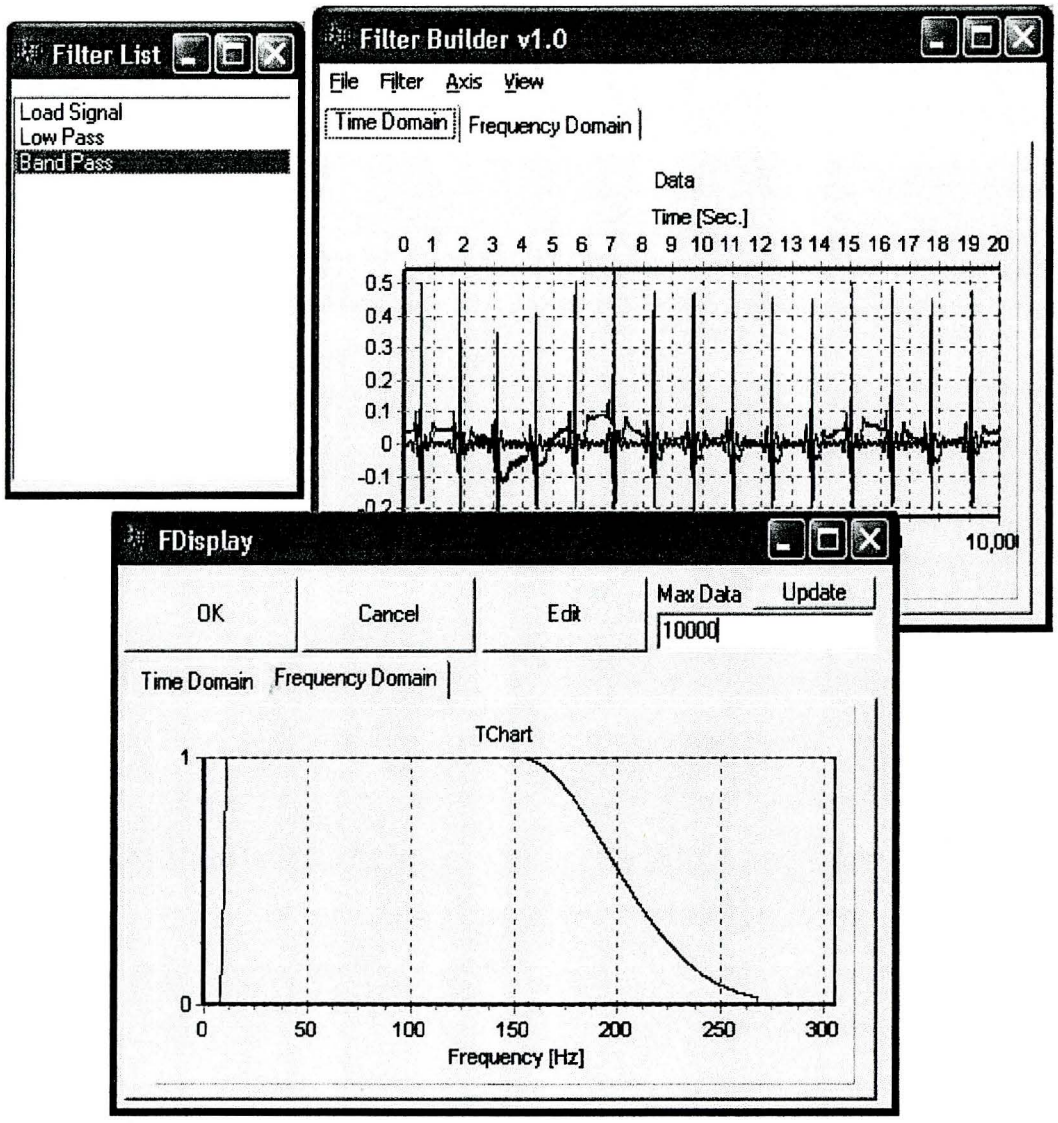

Figure 7.3: FilterBuilder being used to modify a Band Pass filter

7.4.3 Model Generation

VEL contains 3 different methods to generate a patient ECG model. The first is the default method that uses common ECG V1 parameters. The second model allows the user to specify three filters to be used for the three different regions of the ECG beat waveform as discussed in chapter 6. The third option allows the computer build an optimized model using a genetic search technique. Figure 7.4 shows the model builder dialogue in VEL. If the user selects the "Auto Generate" model option he or she also has

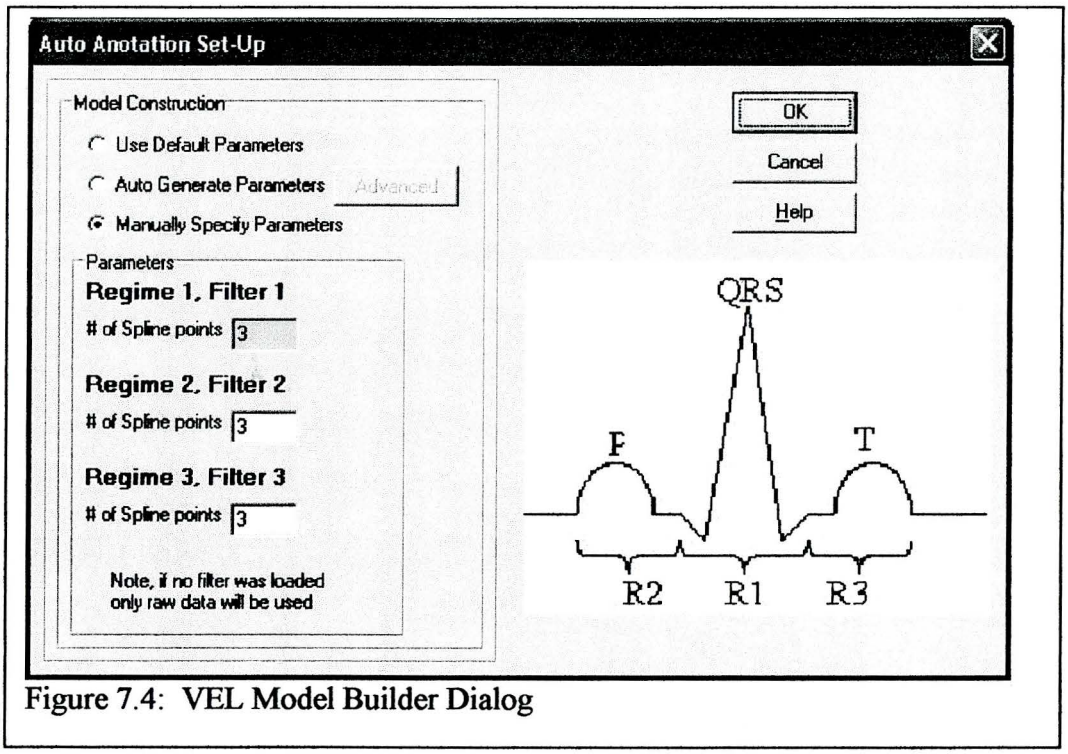

Figure 7.4: VEL Model Builder Dialog

the option of adjusting the genetic algorithm settings for optimal performance. Once the model has been built VEL generates a report showing the details of the model (Figure 7.5). The user has the option of adjusting the model for performance via the dialog. VEL also provides an error estimation dialog so that the user can test the performance of the model on trained or untrained data without analysing the entire data set. Figure 7.6

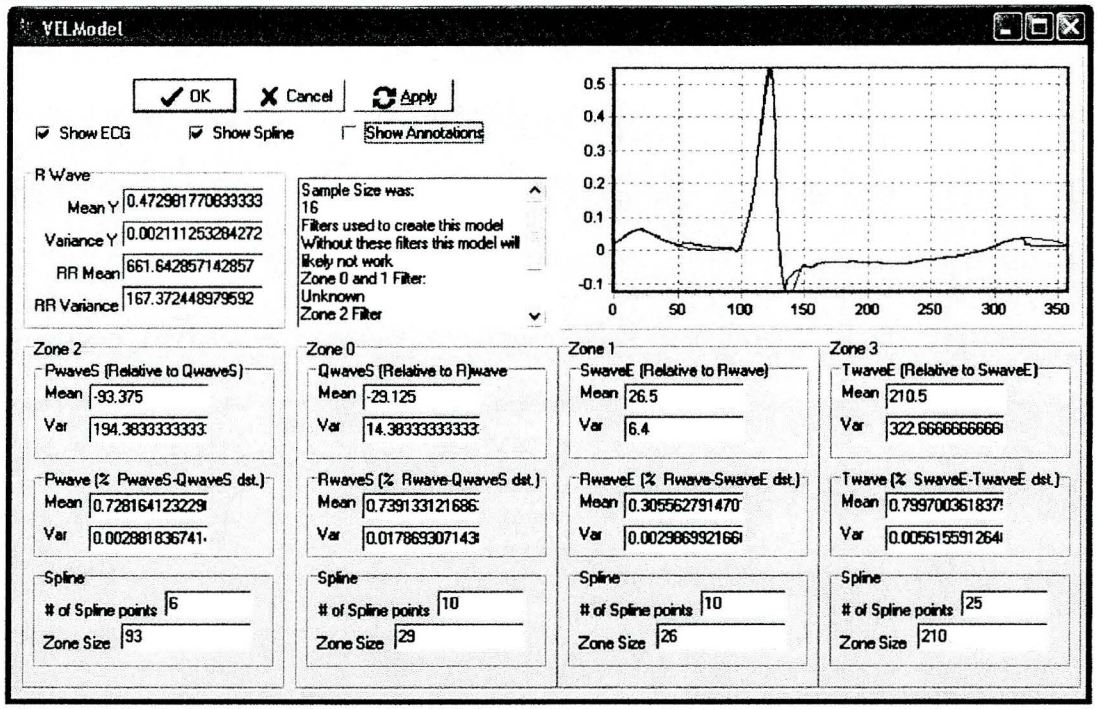

Figure 7.5: VEL Model Display Dialog

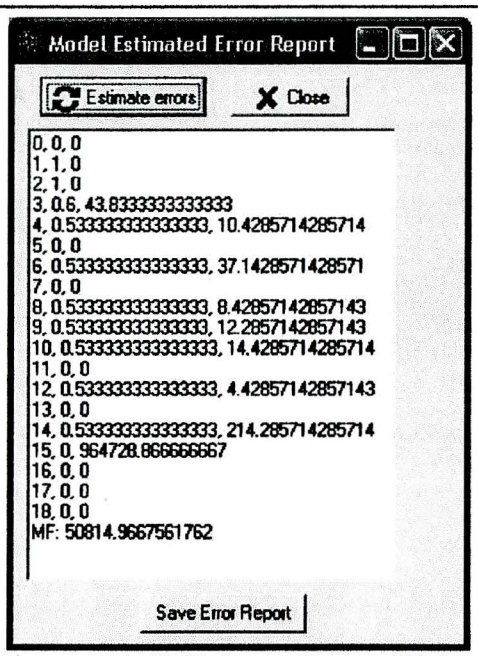

Figure 7.6: VEL Error Report Dialog

shows the VEL error analysis dialog. If the user loads a new set of pre-annotated data, VEL will test its results against the user's annotations.

Testing of the model generator was done in three parts. The first part was to test all three modeling techniques on 6 sets of pre-annotated data. The three models were compared with each other for accuracy using the pre-annotated ECG provided for the training. The default ECG was used as a baseline or control case. Then the models were tested for accuracy on untrained data using a second set of pre-annotated data taken from a different hour of ECG. Finally, the models were tested for repeatability by Auto-annotating the same portion of ECG data twice and examining any differences.

7.4.4 R-R Interval Analysis and Ectopic Correction

The R-R interval was taken for 1 hour of data after it had been auto-annotated. The time and power spectrum obtained were compared with results obtained by Hu (1999). Any unexpected data artefact was investigated and corrected if necessary. Ectopic correction of the HRV data was tested in the same way. The process was repeated for a number of 1 hour long data sets from different patients.

7.4.5 QT-interval analysis

The QT-interval was taken from the same hour of data that was used for the R-R interval testing. Time and power results were examined against trends discussed by Berger et al (1997). Inconsistencies were analysed and the software was adjusted if necessary.

7.5 Data annotation

For each of the 400 patients, 30 beats were manually annotated by members of the lab. Those 30 beats were then used by the genetic model builder to generate a model for each patient. The model was used to automatically annotate the data. Each of the 30 beats was labelled with the following fiducials:

i. AnaS (The start of the 30 beat set)
ii. AnaE (The end of the 30 beat set)
iii. PwaveS (The start of the P wave)
iv. Pwave (Peak of P wave)
v. QwaveS (Start of QRS complex)
vi. RwaveS (Start of R wave)
vii. Rwave (peak of R)
viii. RwaveE (end of R)
ix. SwaveE (end of QRS complex)
x. Twave (T peak)
xi. TwaveE (end of T)

The 30 beats that were used to make the patient model were selected from a section of the ECG that contained only sinus rhythms and clearly identifiable P waves, QRS complexes and T waves. The annotated data was then saved as a separate file to be used by the interval analysis software.

7.6 R-R interval analysis

Once an annotated data set has been loaded into VEL the software gives the user the option of examining the patient's HRV properties. VEL provides a separate dialog the displays a R-R interval graph (Figure 7.8a) and an average power spectrum graph (Figure 7.8b). The user is also given the option of adjusting the data to exclude ectopic beats. In this study all R-R interval data was corrected for ectopic beats before power spectrum data was analysed. However, corrected data was not used for the calculation of time domain properties. The power spectra was obtained taking average HRV power spectrum data over 128 millisecond intervals. This was done because ECG data cannot be assumed to be ergodic (Kamath et al.,1993).

VEL also allows the user to view and save all the HRV properties outlined in chapter 4. In this study the HRV properties were saved using the dialog provided by VEL. The properties were then compiled in an MS Excel™ spread sheet for further analysis.

7.7 QT-interval analysis

VEL also supplies a dialog for viewing QT time and power spectra although no QT properties were recorded. Figure 7.9a and 7.9b show the interval graph and power spectrum data for the QT interval of a sample patient.

7.8 Statistical analysis

Once all 222 patients had been analysed the data was sent back to the Biostatistician party who selected the data to determine if there was a significant difference between HRV properties exhibited by the CHF group verses the non-CHF group. The techniques used to do this are outlined in chapter 6. The resulting statistical data was then analyzed to determine if differences, if any, were of significance from a clinical point of view. In our case a p value of less than 0.05 was deemed adequate to declare the differences to be of clinical relevance.

7.9 Summary

In this chapter we discussed the method of selecting patients and digitizing Holter data tapes using a reel-to-reel audio playback device attached to a National Instruments data acquisition card. The data size and aspects of managing such large amounts of data was discussed. New signal analysis software developed in the lab called VEL (virtual ECG Lab) was introduced as a platform on which the analysis of the ECGs took place. The implementation of VEL for HRV and QT-interval analysis was demonstrated. The statistical techniques used to identify the relationships between HRV parameters and CHF were also discussed.

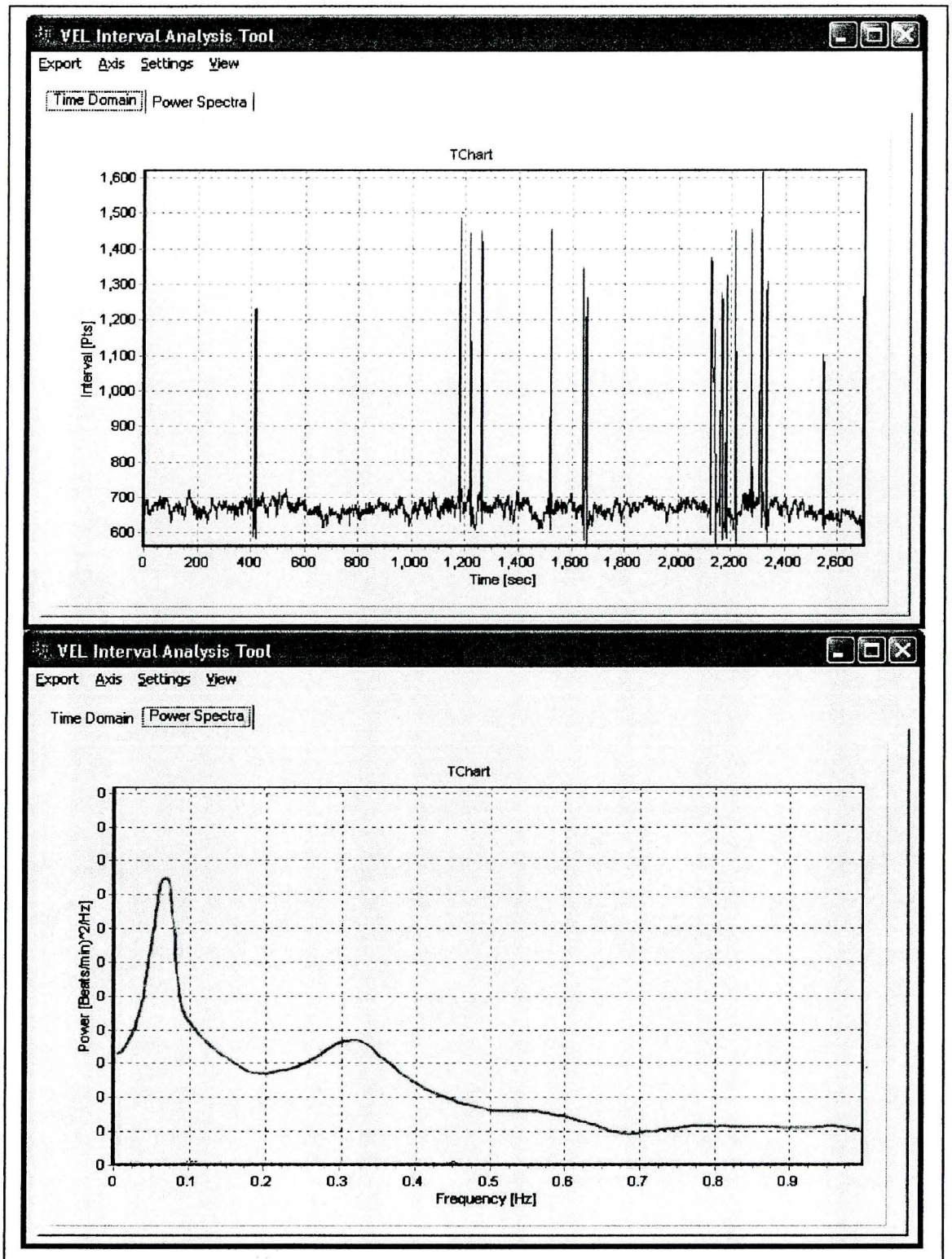

Figure 7.7: R-R interval graph (a) and power spectrum (b)

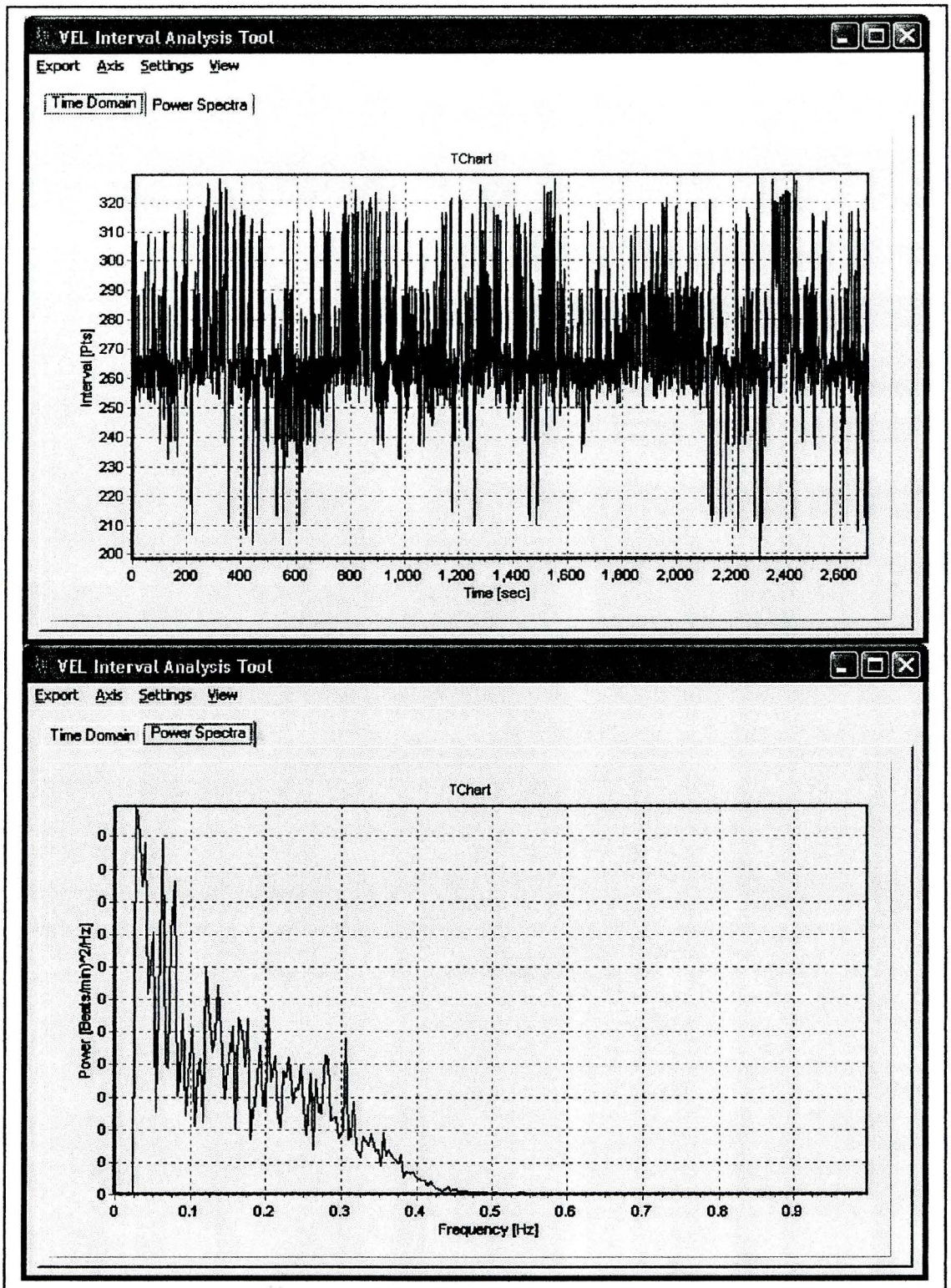

Figure 7.8: QT interval graph (a) and power spectrum (b)

Chapter 8: Experimental results

8.1 Introduction

In this chapter we delineate both engineering and physiological findings of our study. We present the statistical data which was generated to determine the relationship between HRV parameters and patients with and without overt CHF. Also provided are results which characterize the performance of filtering, model building and auto-annotation algorithms, which were used to build R-R interval and QT-interval libraries.

8.2 Algorithm testing

8.2.1 User Interface (UI) testing

The software usability test was preformed with the help 5 individuals. Each had varying amounts of education (from high school to PhD.) as well as varying amounts of experience using computers. After a brief demonstration each of the 5 subjects was able to operate the software with minimal supervision. Assistance was needed on few occasions when the software generated an error. Some suggestions were made as to how the tool's user interface could be altered to better reflect standard health care practice such as graph formatting and terminology. These changes are cosmetic and would not affect the performance of the overall software.

8.2.2 ECG Filter Testing

Through testing on various patients, it was found that straight-forward band pass filtering was more effective than wavelet filtering for isolating the QRS complex. Figure 8.1a shows an ECG recording with baseline wander in red. Figure 8.1b shows the result of band-pass filtering with a centre frequency of 20 Hz and full width half maximum bandwidth (FWHM) of 20 Hz. One can see that the filter is simple and effective. Not only is baseline wander in the signal removed but a good portion of the noise is eliminated as well. The drawback is that a number of ripples are created that mask the presence of the T-wave as well as P-wave, which has also been was reported by Hu (1999). However, the filter maintains enough QRS-signal information so that the data can be used as a unique identifier in our model.

Tests on various patients showed that shifting the centre frequency shifted between ±8 Hz improved QRS detection accuracy. Similar effects were also observed by altering the FWHM width of the filter by +5 or -2 Hz but the effects were less pronounced. These limits were used as guides when constructing parameters for the genetic algorithm.

Although the wavelet filter seemed not as effective in isolating the QRS, its effect on isolating P-waves and T-waves more than made up for its shortcoming. In contrast to the band pass filter, the wavelet filter suppressed multiple ripples and generated unique signatures for both smaller waves. Figure 8.2a shows the effect of wavelet filtering the previous ECG with a scale of 1.2. Figure 8.2b shows the filtered signal in isolation. One can see most of the extraneous ripples have been eliminated and a single relatively high

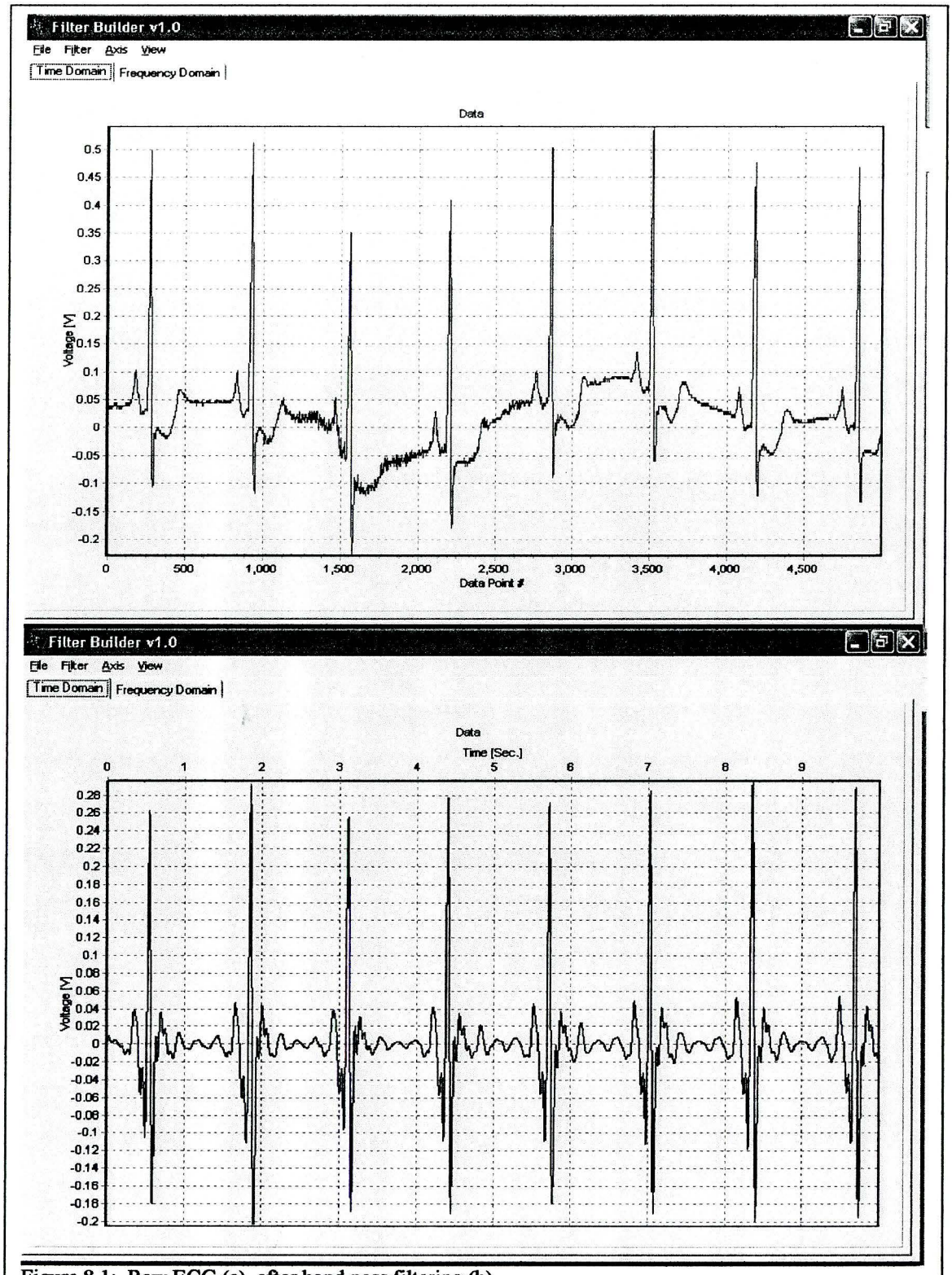

Figure 8.1: Raw ECG (a), after band pass filtering (b)

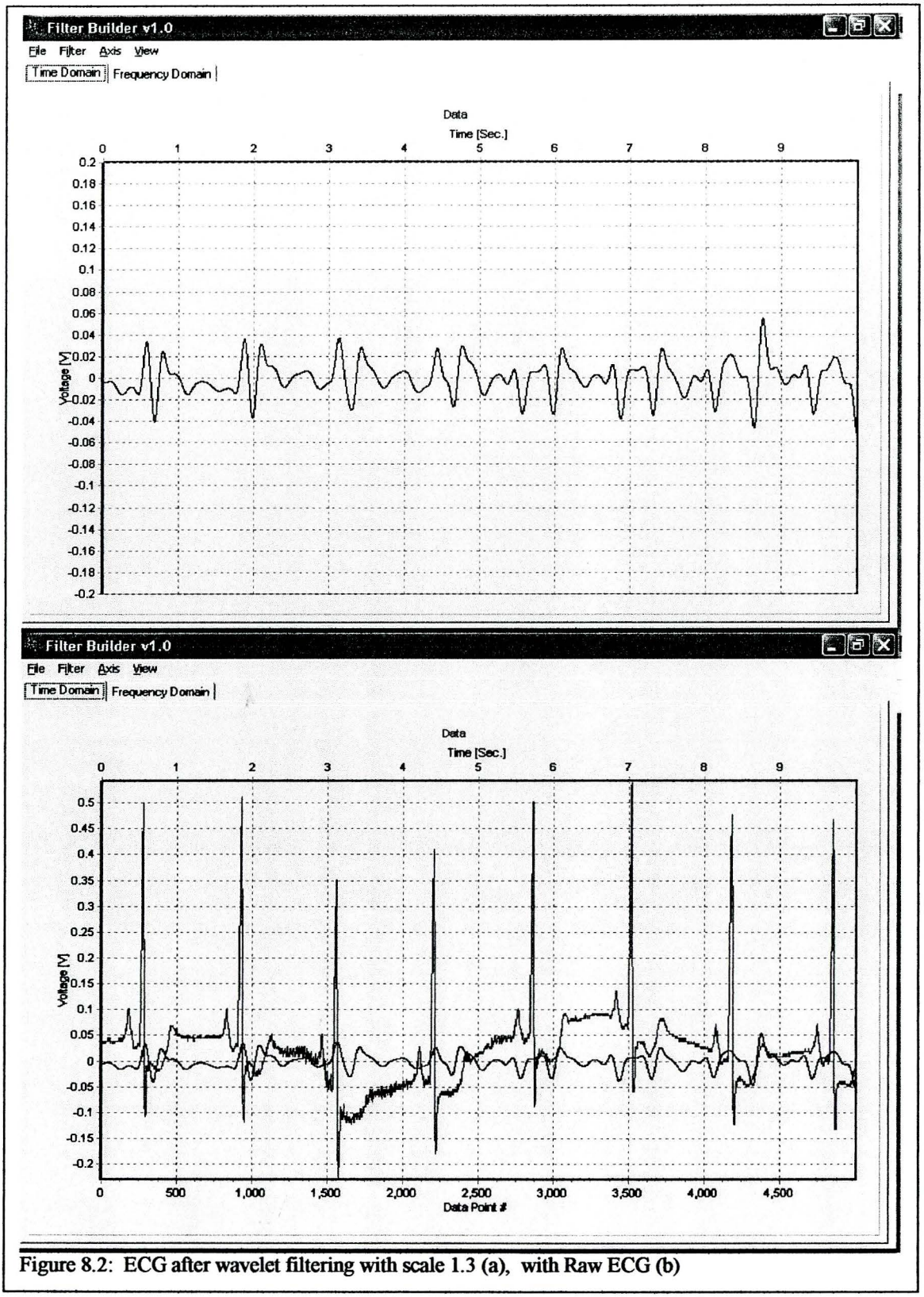

Figure 8.2: ECG after wavelet filtering with scale 1.3 (a), with Raw ECG (b)

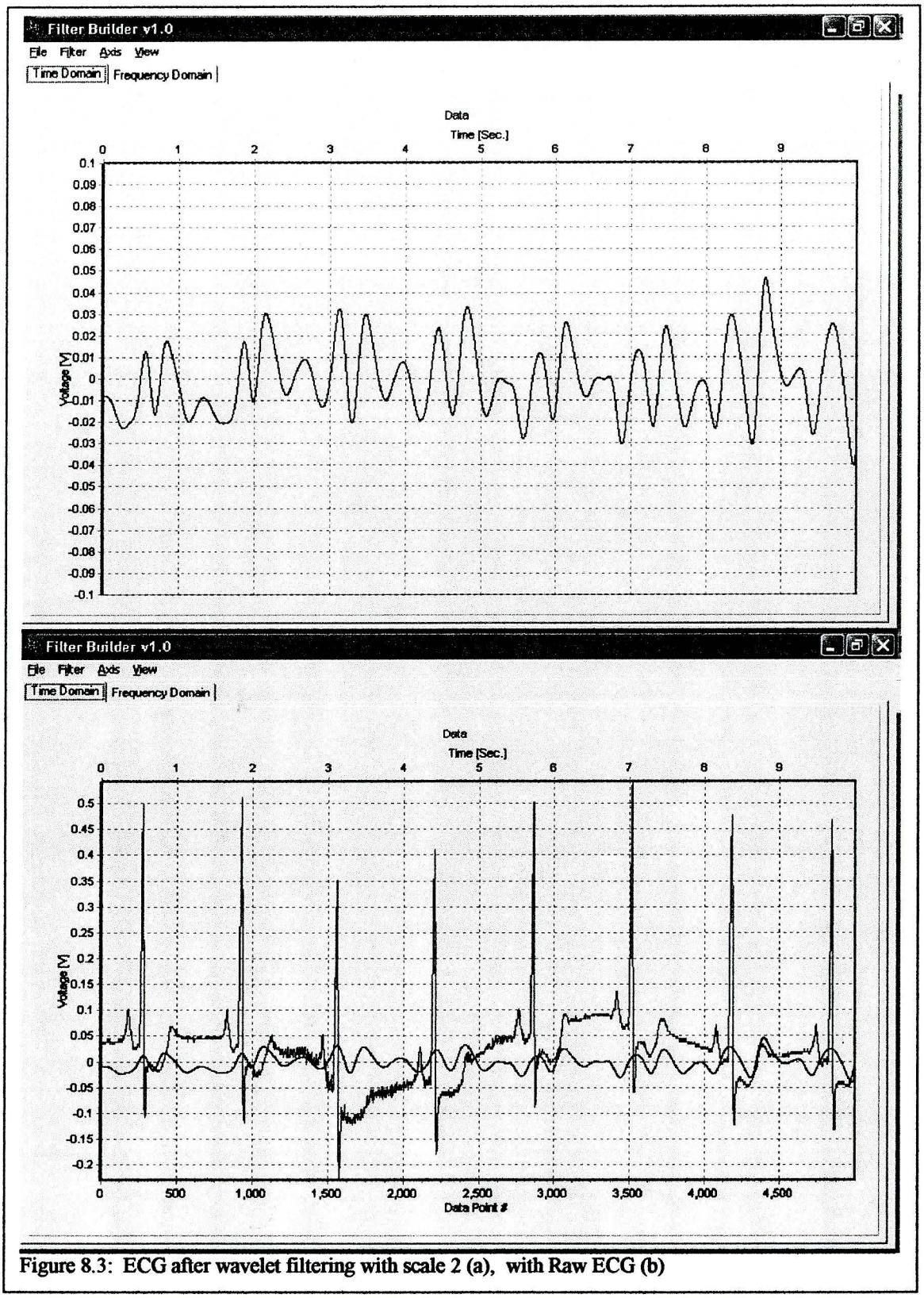

Figure 8.3: ECG after wavelet filtering with scale 2 (a), with Raw ECG (b)

However, if the scale is increased to 2 (figure 8.3a and b) then a consistent pattern begins to emerge for the P-wave as well. Experimentation showed that the scales for P and T-waves had to be between 1 and 3 in order to obtain acceptable detection accuracies. The passband at these scales (between 4 and 23 Hz) agreed with the range of wavelet passbands observed by Sahambi et al. (2000) in the detection of the T wave (between 7 and 22 Hz). The study by Sahambi et al. (2000) did not attempt to detect the P-wave.

8.2.3 Model Testing and comparison

8.2.3.1Performance of Training data

Table 8.1a shows the average mean squared errors obtained using various ECG model building techniques. For the performance test the same data that was used to train those models were used to test them. The model parameters were as follows:

Model 1: Model generated using no-patient specific data.

Model 2: Model generated using manually optimized filter and parameter settings.

Model 3: Model generated by the genetic algorithm method.

Table 8.1b shows the number of missed or incorrectly classified fiducials generated by each of the models.

Table 8.1a:Average Mean squared errors of fiducial placement on training data [sample points]						
Fiducial	Model 1		Model 2		Model 3	
	Mean	SD	Mean	SD	Mean	SD
PwaveS	26.91	6.45	8.32	5.50	5.68	2.57
Pwave	25.43	5.06	6.64	3.32	4.58	2.31
QwaveS	15.70	4.64	5.55	2.59	4.94	2.85
RwaveS	12.22	2.16	4.62	1.73	4.20	1.61
Rwave	3.04	0.88	1.76	0.78	1.40	0.83
RwaveE	12.12	3.38	4.48	4.23	3.75	2.74
SwaveE	21.17	10.28	9.44	11.80	7.63	7.29
Twave	28.80	9.21	11.27	10.26	7.25	5.70
TwaveE	32.72	8.41	14.41	9.80	9.93	6.68
Average	19.79	5.61	7.39	5.56	5.48	3.62

Table 8.1b: Average Number of Missed Fiducial placements on training data			
	Model 1	Model 2	Model3
PwaveS	0.33	0.00	0.00
Pwave	0.33	0.00	0.00
QwaveS	0.17	0.00	0.00
RwaveS	0.17	0.00	0.00
Rwave	0.17	0.00	0.00
RwaveE	0.17	0.00	0.00
SwaveE	0.17	0.00	0.00
Twave	0.40	0.07	0.07
TwaveE	0.40	0.07	0.07
average	0.08	0.01	0.01

8.2.3.2Performance on non-Training data

Table 8.2a shows the average mean squared errors obtained of various models in placing fiducials using data that was not used to train the models. The data was taken from a different segment of time from the same set as the training data.

Table 8.2a: Average Mean squared errors of fiducial placement on non-Training data [sample points]

Fiducial	Model 1		Model 2		Model 3	
	Mean	SD	Mean	SD	Mean	SD
PwaveS	26.00	7.59	11.08	7.53	5.91	5.68
Pwave	26.61	5.44	5.92	2.99	4.22	4.08
QwaveS	17.57	1.82	7.85	6.65	4.71	4.32
RwaveS	13.73	3.71	5.40	3.13	3.65	2.97
Rwave	2.44	0.66	0.76	0.55	1.39	0.92
RwaveE	10.13	1.66	2.80	0.87	2.25	1.71
SwaveE	20.01	5.37	8.81	6.27	2.87	2.11
Twave	24.29	6.64	8.47	5.13	2.86	2.54
TwaveE	28.30	5.06	12.75	6.18	5.89	5.82
Average	25.08	5.45	7.09	4.37	3.75	3.35

Table 8.2b: Average number of missed fiducial placements on non-training data

	Model 1	Model 2	Model3
PwaveS	0.40	0.03	0.03
Pwave	0.40	0.03	0.03
QwaveS	0.16	0.00	0.00
RwaveS	0.16	0.00	0.00
Rwave	0.16	0.00	0.00
RwaveE	0.16	0.00	0.00
SwaveE	0.16	0.00	0.00
Twave	0.33	0.07	0.03
TwaveE	0.33	0.07	0.03
average	0.08	0.02	0.01

One can see from these results that the genetic algorithms (GA) generated Model performed as well as the model that we optimized both on training and non-training data. The success is partially due to our narrowing of the possible genotype values based on our earlier investigation of the effects of filtering the ECG. These results are still very

encouraging since the GA model required no knowledge of signal processing from the user. Steps that can be taken to further improve the classification accuracy of the GA model are discussed in chapter 9.

Not surprisingly the performance of models 2 and 3 were the best. On average the mean squared error of fiducial placement with the most accurately placed fiducial was the R-wave (0.76 data points or 0.002 seconds) and the least accurate being the end of the T-wave (14 data points ,0.028 seconds). The error corresponded to an error in seconds of about 0.01 on average with less than 0.01 seconds being the best and 0.02 seconds being the worst.

In real life, ECG annotation is largely a manual process. Fiducial placement accuracy is greatly dependant on the experience and judgment of the annotator. The success of the model also depends on the ability of the annotator to accurately and repeatably place fiducials on the beats used for training. Given that much of the training data annotation was done by individuals with only an academic knowledge of ECG fiducials (i.e. the author and several volunteers from the lab), the placement accuracy and repeatability are subject to variation. It is very likely that the accuracies achieved by the models are comparable to the accuracies achieved by a trained but inexperienced technician.

8.2.3.3 Repeatability Performance

All models displayed no errors in repeatability. Fiducial placements did not vary between repeated annotations of the same data segment. This result was not unexpected as neither the data nor the model changed between iterations.

8.3 R-R interval analysis

Table 8.3 contains the mean and standard HRV parameter data collected over the 222 patients. 8.3a contains the time domain indices and 8.3b contains the power spectral indices. As it turns out, 171 of our patents from our random sample were from the non-CHF group and 51 were from patients who were diagnosed with CHF. As, the sampling was random it was not possible to insure the numbers of patients were equal for both groups. However, the volume of patient data was deemed high enough to give significant insight into the patients' physiology and confidence in the results obtained thereof. Again, mention is made of the fact that the samples were chosen blindly and results were analyzed by an independent statistician who was objective and unbiased.

Table 8.3a: Time domain parameters of HRV data

Parameter	Non- CHF Patients			CHF Patients			p-value
	Sample size	Mean	SD	Sample size	Mean	SD	
HR	171	70.7	10.1	51	75.3	10.8	0.0059
NN	171	866.0	124.3	51	812.4	109.4	0.0060
SDNN	171	115.0	35.3	51	95.0	35.9	0.0005
SDNN index	171	46.6	16.2	51	36.2	15.3	0.0001
SDANN	171	102.0	33.1	51	85.1	35.2	0.0023
R-MSSD	171	28.4	12.7	51	26.1	11.6	0.2494
PNN50	171	0.1	0.1	51	0.1	0.1	0.5455

Table 8.3b: Frequency domain parameters of HRV data							
Parameter	Non-CHF Patients			CHF Patients			p-value
	Sample size	Mean	SD	Sample size	Mean	SD	
%LF Area	171	57.5	7.1	51	52.5	8.5	<0.0001
%HF Area	171	42.5	7.1	51	47.5	8.5	<0.0001
LF:HF	171	1.6	0.4	51	1.4	0.4	<0.0001
ABS LF	171	6152.0	596.5	51	5693.5	794.4	0.0003
ABS HF	171	4661.0	929.4	51	5270.2	1115	0.0001

One can see that many of the parameters have p-values of less than, 0.05 (our level of significance). The low p-values indicate that most HRV measures in both the time and frequency domains are valid indicators for the probably presence of CHF in patients. Of these, the strongest indicators are potentially the LF:HF ratio as well as the SDNN index. The weakest potential indicators are R-MSSD and PNN50. These results suggest that LF:HF ratio, as a measure of sympathovagal balance is tilted towards being sympathetic in patients without CHF. This is confirmed by SDNN in the time domain because a high SDNN present in patients without CHF is associated with increased vagal modulation. Power spectral analysis of the HRV also disperses energy across the whole vagal band and thus some specific peaks at the breathing frequency are not very clear.

8.4 QT- interval analysis

We have shown that as a proof of concept, the software demonstrated that QT analysis could be done on patient's ECG data. As a demonstration, QT- interval information from both time and frequency domains were generated using VEL. These

graphs were presented previously in chapter 7. Methods detailing how VEL could be used for future QT-interval analysis on SOLVD data are presented in chapter 9.

8.5 Summary

The results of our research demonstrate two sets of findings. The first observation is that it patients with CHF have significantly differing HRV parameters in both the time and frequency domain. Strongest indicators in this category include LF:HF ratio and the SDNN index with p-values of <0.0001 and $p< 0.0001$ respectively. The worst potential HRV indicators were found to be R-MSSD and PNN50 with p-values of 0.249 and 0.545.

The second finding is that a robust, model based, ECG auto annotation system can be constructed that can place all of the critical fiducials of the ECG with a high degree of precision, based on user input. The flexibility of the system makes it possible for varying types of interval analysis to be conducted on the same platform. Further, possible for researchers with a variety of backgrounds to analyze the data.

Chapter 9: Discussion and Suggestions For Future Work

9.1 Introduction

In this chapter we discuss the significance of the research work contained in this thesis. We compare our results with those of researchers who have analyzed neurocardiac signals from patients with congestive heart failure. Limitations of our work are also delineated. Finally, we conclude with suggestions for future work.

9.2 Significance of the Research

To our knowledge our laboratory is the first to develop a hardware/software system to analyze the HRV signals recorded in patients with congestive heart failure from the SOLVD trial. Our results with regards to the hardware are clear: the hardware combination of a reel-to-reel tape recorder, a 16 bit A/d converter and a personal computer running at 2.05 GHz worked successfully and provided us digitized data. Four hundred tapes were digitized using the virtual instrument software developed using Labview 6.2 (National Instruments, www.ni.com) development system. The software developed for QRS-detection, P detection, QT detection, power spectral computation and time domain statistical computation have all performed well. Our results show that it is

possible to differentiate between patients with overt congestive heart failure and no CHF, with a high degree of confidence, for LF:HF ratio with $p<0.0001$. Such information may be helpful to stratify clinical risk associated with CHF who go on to therapy such as Enalapril.

9.3 Comparison with other researchers

The relationship between HRV characteristics and chronic heart failure has been recorded by a number of researchers. Kearney et al (2002) studied the relationship between an number of non-invasive diagnostic techniques to find a suitable predictor of death in patients with mild to moderate progressive heart failure. The study found that SDNN was an effective identifier of patients with an increased risk death.

La Rovere et al (2003) also found relationships between HRV characteristics and patients at risk with sudden death. Rovere et al. found that increased sympathetic discharge in CHF, often associated with patients at risk of sudden death, is correlated with a marked reduction in LF power. A reduction in the LF:HF ratio also proved to be a valid indicator of sudden death along with a reduced RR interval variability.

Lombardi et al (1992) used power spectral data to study changes in HRV in patients shortly after the onset of a myocardial infarction. A significant change in the LF:HF ratio was noted in these patients as they recovered from the event, strengthening the idea that the LF:HF is a measure of sympathovagal balance.

Lastly, Berger et al (1997) conducted research on QT interval parameter measures and their relationship to RR interval measures. Although the QT interval

parameters were not measured in this study, casual observation of the time and power spectrum graphs showed similar relationships between RR and QT intervals as were suggested by the Berger et al, study.

9.4 Limitations of our work

There are several limitations in our work which need to be addressed. Here outline a few of them.

9.4.1 Speed variations of the tape recorder

The variation of recording and play back speed of the various devices used were not measured. The SOLVD trial was conducted through the assistance of many international institutions. Each used a different type of Holter monitor with varying recording speed accuracies. The recording devices used were not logged and therefore it is impossible to trace which type was recorder was used for each patient. Therefore, even if the playback device was characterized exhaustively, variations in the ECG waveform due to tape speed could never be fully qualified. It is known that variations of recording speed do have an effect on HRV parameters (Maestri et al , 1994). However, the degree at which they manifest themselves in our data can never be fully investigated.

9.4.2 Quality of the training data

Due to the heavy workload of nurses at the McMaster medical centre, it was not possible to get the assistance of experienced nurses to annotate ECG data for the purposes of training our models. In order to obtain these annotations we had to resort to analysing

the data ourselves. Given that the members of the lab are not experienced health care workers and have only a rudimentary knowledge of ECG interpretation it is likely that there are inherent errors in the models that we developed. Luckily, the QRS complex is quite easy to identify. The T wave was far more difficult. It is unlikely that the annotated data sets generated in this study are appropriate for other sorts of interval analysis due to our inexperience in generating the training sets.

9.4.3 ECG annotation algorithms

One of the major draw backs of our mythology in auto-annotating ECG data is the time it takes to accomplish the task. Model building using the genetic algorithm technique takes an especially long time to complete. This is mostly due to the time it takes to filter the data set three times per iteration to obtain ECG features. Speed can be improve substantially if the filters are applied to only the training data and not the entire 24 hour data set. Currently, this requires someone to go in and clip out the data manually, which is also a time consuming task. A future program upgrade should probably included a way to automatically separate and filter the training data.

9.5 Suggestions for future work

9.5.1 Improvements to algorithms and software

The mechanics of stretching the zones to obtain an optimal fit to the data was discussed in chapter 7. Also mentioned was the fact that fiducials that rested between the Zone endpoints were not optimally positioned. They were placed based on statistical

probabilities of their position relative to the endpoint fiducials. To improve the placement accuracy of mid-fiducial points the Zone could be first stretched to an optimal size and then warped to better fit the midpoints. Warping could be accomplished by altering the distribution of the number of points on either side of the mid-fiducial while maintaining the total number of data points for each zone.

There is also potential for improving fiducial placement accuracy by optimizing the parameters for genetic evolution of the model. An obvious adjustment is to increase both the number of iterations of genetic optimization as well as the range of values the genotypes can assume. Other improvements could be made by allowing the genetic algorithm to weight the importance of each pre-annotated beat that is used to build the model.

The auto-annotation routines currently place an "unkwn" fiducial on any waveform that does not meet the minimum criteria. It would be more useful from a clinical research point of view if the system could identify what sort of error occurred and classify the beat accordingly. Such a system could be implemented using well know classification techniques such as neural networks (Hopkins and Suleman, 2000).

It is also possible that the auto-annotation system will encounter heavily corrupted ECG data in the middle of a good data set. It would be more efficient if an ECG quality measure could be developed so that the section of bad data could be marked off. The auto-annotation system could then skip that portion of data and avoid potential miss classifications as well as decrease analysis time. The ectopic correction routine would also have to be modified to take such areas of corrupted data into account.

9.5.2 Future directions for clinical based studies

The most immediate avenue of future research will likely be to analyze additional patients in out HRV data set to improve the statistical significance or our findings. However, the results of this study and the development of the VEL software have opened up a wide range of clinical analysis possibilities. As mentioned earlier, the effect of drug therapies such as Enalapril on HRV parameters have not been investigated using the SOLVD database. The effects of similar medications on other ECG interval measures such as QT have also not been explored. In addition we have not investigated other physiological variables such as severity of the disease (eg: ejection fraction, quality of life) and neuro-hormonal variables (eg: level of norepinephrine) on the neurocardiac signals. The tools developed for this study have made such avenues of future research easier and less time consuming to explore.

9.6 Summary

The findings and techniques developed in this study are significant to future work in the area of HRV analysis of patients with progressive heart failure. They also strengthen relationships suggested by various other researchers in the field. Despite some short comings we were also able to make significant progress towards the study of a variety of other interval measures related to ECG recordings. Future improvements to the technology to improve speed and accuracy will likely result in more research being done in this area.

Chapter 10: Summary

10.1 Introduction

The research presented in this thesis represents a possible solution towards analyzing several thousand 24 hour Holter ECG tapes presented to our laboratory by the SOLVD investigation into the effects of ACE inhibitors on patients with congestive heart failure. This chapter summarizes our efforts and the results obtained thereof.

10.2 Approach

In chapters 2 and 3 we identified the problem of congestive heart failure in a population growing older in both Canada and United States and evaluated the magnitude of the effects due to this pathological condition. As an example, it is estimated that about C$ 1Billion is spent in hospitalization costs alone in Canada during 2002. Also the mortality rate of 50% in patients with CHF is unacceptably high. It is also known that Angiotensin Converting Enzyme (ACE) inhibitors can help patients with CHF and prolong their life. The importance of QT-interval analysis is explained in chapter 4. Chapter 5 presented some of the key questions we would attempt to answer.

The key premise underlying these questions is that hidden signals are embedded within the ECG signals in the form of variations of R-R intervals and QT-intervals. Modern digital signal processing methods can provide insights into their underlying physiological mechanisms. A fundamental question that can be asked is if statistical

parameters computed from these hidden signals provide discriminatory power to separate patient groups who have overt congestive heart failure compared to those who do not have a failing ventricle. Statistical questions were set for testing through the help of an unbiased statistician who determined minimum sampling size and who performed the statistical analysis independently. We were fully blinded from identifying the data or its type.

10.3 Digitization Hardware and Software for Digitization of the Data

Digitization is currently handled by what are called virtual instruments programmed through National Instruments software development package. The software was flexible enough to select our sampling rate (corresponding to 500 Hz). The tapes were played at 60 times the recording speed and digitized at a corresponding sampling rate at 16 bit accuracy.

10.4 Algorithms for analysis of the signals embedded in the ECG signal

In chapter 6 we provided implementation details of various algorithms used to obtain the computed R-R and QT intervals. Statistical variations in a 24 hour ECG in mean values of certain indices can identify the differences in various pathologies. These include standard deviation of 24-hour R-R intervals and relative powers within certain specific bands (low frequency(LF): 0.015-0.15 Hz; high frequency (HF): 0.15-0.4 Hz) of the power spectra of the HRV signals.

Initial analysis of the ECG was done by first filtering the ECG signal using 3 different linear filters. These included one band pass filter and two wavelet filters. Each filter emphasized a particular critical waveform of the ECG (P wave, QRS complex, or T wave).

Once the filtering was completed a portion of the data was manually annotated so that an ECG waveform model could be created. This model was constructed in one of three ways. The first was using only literature-based estimates of the ECG waveform. The second model was created using user specified filters and parameters. The third model type was generated using parameters specified by a trained classifier based on genetic algorithm techniques.

Using the model 24 hour ECG data was automatically analyzed and critical ECG fiducials were placed on the data. These fiducials were then used as the basis for both RR and QT interval analysis routines.

HRV characteristics were generated using both statistical measures in the time domain (mean, variance etc.) and power spectral measures of the LF and HF areas. The power spectrum measured using both Fourier and auto-regressive techniques.

Finally the HRV parameter data was compiled and tested against the null hypothesis using a t test with a level of significance of 0.05.

10.5 Results and Discussion

Our results (chapter 8) suggest that after examining 220 patient tapes many of the indices computed for this work are sensitive enough for characterizing the patient

conditions objectively. Our results also show that both time domain and frequency domain parameters can be used to critically identify a patient population. QT interval data was successfully extracted form the data. QT interval time and frequency graphs were displayed for the user but not in-depth QT parameter analysis was done to the data. Limitations of our work and suggestions for future work are listed in chapter 9.

10.6 Summary

This chapter summarizes the research presented in this thesis through a systematic discussion of various issues such as the key questions, origin of the data from patients with congestive heart failure, instrumentation and software tools employed for objective analysis of the hidden signals embedded in 24 hour Holter ECG and finally an evaluation of the results obtained from our investigation.

Appendix A: Anatomy of the ECG

A.1 Introduction

The ECG (electrocardiogram) is a recording of the heart's electrical activity. One cardiac contraction of the heart is called a systole, and a relaxation of the heart is called a diastole.

A.2 Parts of the ECG

The heart consists of pacemaker cells, which are the electrical power sources of the heart. These cells are approximately 5-10 micrometers long. The dominant

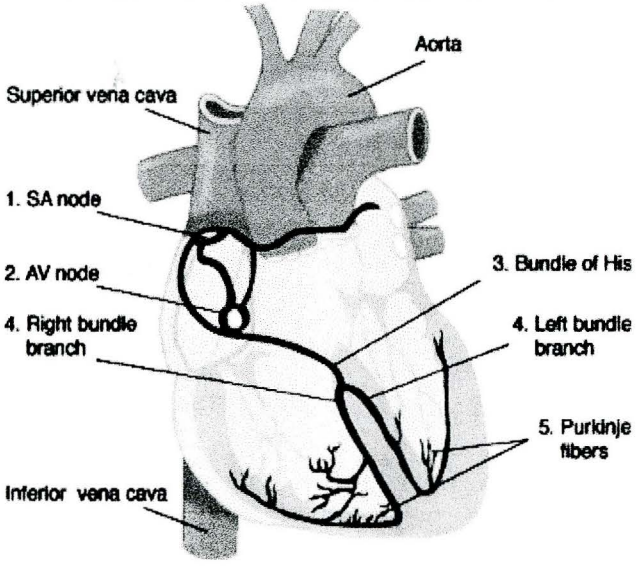

Figure A.1: Conduction anatomy of the heart (Bianco, 2003)

pacemaker cells are located in the right atrium of the heart and are called sinoatrial nodes or sinus nodes (SA node). Pacing signals travel from the SA node to the AV node (see

figure A.1) where they are transferred to the Bundle of His. From here they are distributed to the left and right bundle branches.

P wave

As the SA node fires a wave of depolarization spreads outward into the atrial myocardium. This results in atrial contraction which is denoted by the P wave (see figure A.2). The first part of the P wave predominantly represents right atrial depolarization and the second part shows the left atrial depolarization. (Thaler, 1999). The P wave may be upright, inverted, or isoelectric (where there is an equal amount of deflection above and below the baseline of the ECG tracing.)

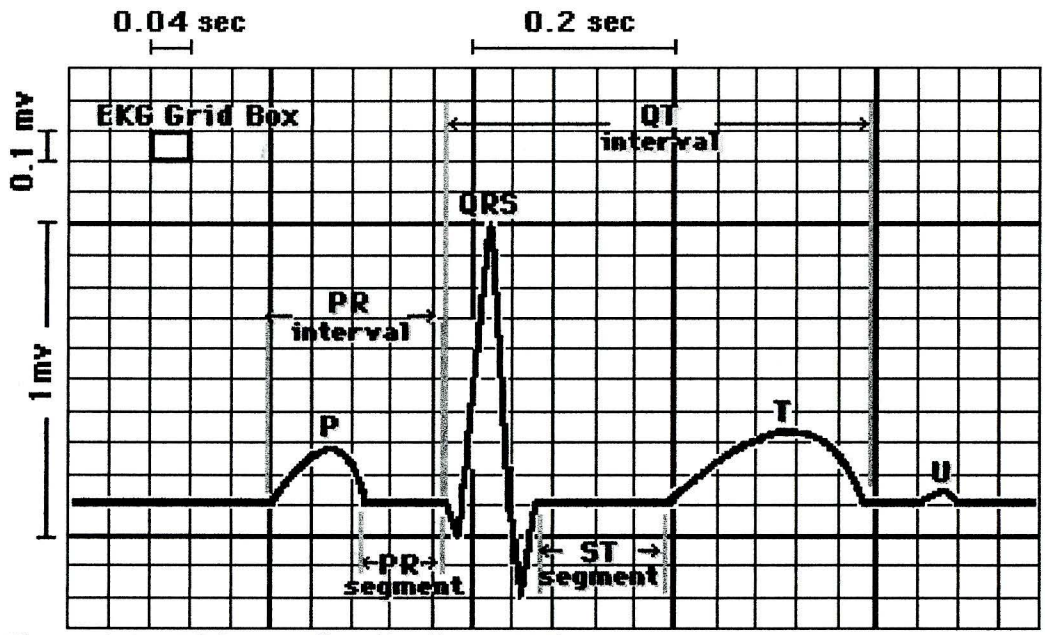

Figure A.2: ECG recording (Seeling, 1992)

QRS complex

Following the P wave is the QRS complex. This is usually the dominant wave used for ECG tracing. It is produced by the depolarization of the ventricles and is normally less than 0.12 seconds in duration. (Seelig, 1992). The QRS complex may start either with a downward deflection (Q wave) or an upward deflection (R wave).

T wave

The T wave immediately follows the QRS complex and is produced by the depolarization of the ventricles. It may be upright, inverted, biphasic (both up and down) or sometimes absent (flat).

A.3 Common Hook-up methods – Leads

The electrical currents generated by the heart are commonly measured by an array of electrodes placed on the body surface. These electrode leads are connected to a device that measures potential differences between selected electrodes in order to produce the characteristic electrocardiographic tracings. The standard ECG consists of 12 leads where each lead is determined by the placement and orientation of the various electrodes on the body. Two electrodes are placed on the wrists and two on the ankles. Six are placed across the chest.

Bibliography

Akselrod S, Gordon D, Madwed JB, Snidman NC, Shannon DC, Cohen RJ. "Hemodynamic regulation: investigation by spectral analysis." Am J Physiol 1985; 249:H867-H875.

Akselrod S, Gordon D, Ubel FA, Shannon DC, Barger AC, Cohen RJ. "Power spectrum analysis of heart rate fluctuations: a quantitative probe of beat-to-beat cardiovascular control." Science 1981;213:220-2.

Akselrod S, Karin J, Hirsch M. "Computerized detection of fetal ECG from maternal abdominal signal." IEEE Comp Cardiol 1987:261-4.

Algra A, Brun L, and Zeelenberg C , "An Algorithm for Computer Measurement of QT Intervals in the 24 Hour ECG", Computer in Cardiology, 1987: 117-119

Algra A, Tijssen JGP, Roelandt JRTC, Pool J, Lubsen J. "Heart rate variability from 24-hour electrocardiography and the 2-year risk for sudden death. Circulation" 1993;88:180-5.

Back, Thomas (1996), Evolutionary Algorithms in Theory and Practice. New York: Oxford Press.

Bevington, Pillip R., and Robinson D. Keith (1992), Data Reduction and Error Analysis for the Physical Sciences. 2nd Ed, NewYork: WCG McGraw-Hill.

Bianchi A.M., Mianardi L., Petrucci E., et al. "Time-variant power spectrum analysis for the detection of transient episodes in HRV signal." IEEE Transactions on Biomedical Engineering 1993: 40(2); 136-44

Bianco C, "How Your Heart works:, on-line, internet. Available: http://science.howstuffworks.com/heart4.htm. last updated: Jan 23, 2002

Bryll R (2000). "Introduction to Wavelets Presentation". On-line. Internet. Available: http://vislab.cs.wright.edu/~rbryll/wavelet-present.html. Last updated: Sept 19, 2000.

Davis, Lawrence (1991), Handbook of Genetic Algorithms. New York: Van Nostand Reinhold

Dremin IM, Ivanov OV, and Nechitailo VA (2001). "Wavelets and their use". On-line. Internet. Available http://arxiv.org/abs/hep-ph/0101182. Last updated: Jan 16, 2001.

Eckberg DL, Nerhed C, Wallin BG. "Respiratory modulation of muscle sympathetic and vagal cardiac outflow in man." J Physiol 1985;365:181-96.

Friesen G, Jannett T, Jadallay MA, Yates SL, Quint SR, and HT Nagle (1990), "A Comparison of Noise Sensitivity of Nine QRS Detection Algorithms", IEEE Trans Bio Eng, 37(1): 85-97.

Goldberg, David E. (1989), Genetic Algorithms in Search, Optimization & Machine Learning. Toronto: Addison Wesley Longman, Inc.

Haupt, Randy L., and Haupt, Sue E. (1998), Practical Genetic Algorithms. Toronto: John Wiley & Sons, Inc.

"What is the Impact of Heart Failure?" , Heart Failure Society of America. On-Line. Internet. Available: http://www.abouthf.org/questions_impact.htm. Last Modified, March 28, 2002.

"General Info - Incidence of Cardiovascular Disease", Heart and Stroke Foundation of Canada. On-Line, Internet. Available: www.heartandstroke.ca . Last Modified June 12, 2002

Hopkins CB, and Suleman J (2000), "An Artificial Neural Network for the Electrocardiographic Diagnosis of Left Ventricular Hypertrophy", Crit Rev Bio Eng, 28(3&4): 435-438.

Hu A (1990), "Digital Signal Processing of Hemodynamic Signals", Master Thesis, dept. of Engineering, McMaster University.

Jane T, Laguna P, Thakor NV, and Caminal P (1992), "Adaptive Baseline Wander Removal in the ECG: Comparative Analysis With Cubic Spline Technique", Computer in Cardiology, 143-146.

Kamath M, and Fallen E (1993). "Power Spectral Analysis of Heart Rate Variability: A Noninvasive Siganture of Cardiac Autonomic Function", Crit Rev Bio Eng, 21(3): 245-311.

Kearney MT. Fox KA. Lee AJ et al., "Predicting death due to progressive heart failure in patients with mild-to-moderate chronic heart failure", Am J Physiol: 2002; 40(10):1801-8

Kleiger RE, Bigger JT, Bosner MS et al. "Stability over time of variables measuring heart rate variability in normal subjects." Am J Cardiol 1991;68:626-30.

Kleiger RE, Miller JP, Ab J, Bigger T, Moss AJ. "Decreased heart rate variability and its association with increased mortality after acute myocardial infarction." Am J Cardiol 1987;59:256-62.

Kleiger RE, Stein PK, Bosner MS, Rattman JN, "Time domain measurements of heart rate variability." Cardiol Clin 1992;10:487-98.

Kluge P, Walter T, and Neugebauer A (1997), "Comparison of QT/RR Relationship Using Two Algorithms of QT Interval Analysis for Identification of High Risk Patients for Life-Threatening Ventricular Arrhythmias", ANE 2(1): 3-8.

La Rovere MT., Pinna GD., Maestri R., et al. "Short-term heart rate variability strongly predicts sudden cardiac death in chronic heart failure patients." Circulation 2003; 107(4):565-70.

Li C, Zheng C, and Tai C (1995), "Detection of ECG Characteristic Points Using Wavelet Transform", IEEE Trans Biomed Eng, 42(1):21-28.

Lombardi F, Sandrone G, Mortara A et al. "Cricadian variation of spectral indices of heart rate variability after myocardial infarction." Am Heart J 1992;123:1521-9.

Lonn EM, Yusus S, Jha P, Montague TJ, Teo KK, Benedict C, Pitt B. "Emerging role of angiotenssion-converting inhibitors in cardiac and vascular protection." Circulation, 90:2056-202069, 1994.

Madson, Robert D. (1978), Statistical Techniques in Business and Economics. 4th Ed, Illinois: Richard D. Irwin, Inc.

Malik M, Camm JA. "Components of heart rate variability - what they really mean and what we really measure." Am J Cardiol 1993;72:821-2.

Mann DL. Mechanisms and Models in Heart Failure. Circulation, 100:999-1008, 1999.

Martinea JP, Olmos S, and Laguna P (2000), "Evaluation of a Wavelet-Based ECG Waveform Detector on the QT Database", Computer in Cardiology, 27:81-84

Merri M, Farden D, Mottley JG, and Titlebaum EL (1990). "Sampling Frequency of the Electrocardiogram for Spectral Analysis of the Heart Rate Variability", IEEE Trans Bio Eng, 37(1): 99-106.

Pagani M, Lombardi F, Guzzetti S et al. "Power spectral analysis of heart rate and arterial pressure variabilities as a marker of sympatho-vagal interaction in man and concious dog." Circ Res 1986;59:178-93.

Pan J, and Tompkins WJ (1985), "A Real-Time QRS Detection Algorithm", IEEE Trans Bio Eng, BME-32(3): 230-236.

Pellegrinin F, Ferretti GF, and Nervoso S (1987), "Use of a Computerized Automation in Studying the Circadian Variability of the Q-T Interval by Means of the Dynamic Egg", Computer in Cardiology, 121-124.

Pietka E, "Feature Extraction in Computerized Approach to the ECG Analysis" Pattern Recognition 1991; 24(2): 139-146

Pinna GD, Maestri R, Di Ceare A, et al, "The accuracy of power-spectrum analysis of heart-rate variability form annotated R-R lists generated by Holter systems" Physiol Meas: 1994; 15(2):163-79

Press WH, Teukolsky SA, Vetterling WT, and Flannery BP (1999). Numerical Recipes in C. Cambridge: Cambridge University Press.

Proakis, John G., and Manolakis, Dimitris G. (1996), Digital Signal Processing Principles, Algorithms, and Applications. 3rd Ed, New Jersey: Prentice Hall.

Reddy BRS, Elko PE, Christenson DW, and Rowlandson GI (1992), "Detection of P Wave in Clinical ECG A Preliminary Study", Computer in Cardiology, 87-90.

Riccio ML, and Belina JC (1992), "A versatile Design Method of Fast, Linear-Phae FIR Filtering Systems for Electrocardiogram Acquisition and Analysis Systems", Computer in Cardiology, 147-150.

Ruha A, Sallinen S, and Nissila S (1997), "A Real-Time Microprocessor QRST Detector System with a 1-ms Timing Accuracy for the Measurement of Ambulatory HRV", IEEE Trans Bio Eng, 44(3): 159-167.

Sadeh D., Shannon DC, Abboud S, Akselrod S, and Cohen R (1987), "A New Technique to Determine the Correlation Between the QT Interval and Hear-Rate for Control and SIDS Babies", Computer in Cardiology, 125-127.

Sahambi JS, Tandon SN, and Bhatt RKP (2000), "An Automated Approach to Beat-by-Beat QT-Interval Analysis", IEEE Eng Med Bio Mag, pp. 97-101, May/June 2000.

Sayers BM. Analysis of heart rate variability. Ergonomics 1973;16:17-32.

Seelig CB. Simplified EKG Analysis – A Sequential Guide to Interpretation and Diagnosis, Hanley & Belfus, Inc.: Scarborough, 1992

Skordalakis E. "Syntactic ECG Processing: A Review" Pattern Recognition 1986; 19(4): 305-313

SOLVD Investigators. "Effect of Enalapril on Mortality and the Development of Heart Failure in Asymptomatic Patients with Reduced Left Ventricular Ejection Fractions." *NEMJ:*1992; 327: 685-91.

SOLVD Investigators. "Effect of Enalapril on Survival in Patients with Reduced Left Ventricular Ejection Fractions and Congestive Heart Failure." *NEMJ*: 1991; 325: 293-302.

SOLVD Investigators. "Effect of Acute Angiotensis Converting Enzyme Inhibition on Left Ventricular Filling in Patients With Congestive Heart Failure." *NEJM:* 1991.

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 1996;93:1043-65.

Thaler MS. The Only EKG Book You'll Ever Need (3rd ed): Ch. 1. Lippincott Williams & Wilkins: New York, 1999.

The CONSENSUS Trial Study Group. "Effects of Enalapril on Mortality in Severe Congestive Heart Failure." *NEMJ*: 1987; 316: 1429-35.

Townley R, Howlett J. "What's new in heart failure?" Perspectives in Cardiology, Nov/Dec.2002; pp. 33-41

Vidakovic B, and Mueller P (1994). "Wavelets for Kids A Tutorial Introduction". On-line. Internet. Available: http://www.isye.gatech.edu/~brani/wavelet.html.

Weiss IW. Essentials of Heart Rhythm Analysis: Ch. 1. F.A. Davis Company: Philadelphia, 1973.

Whitley, L. Darrell (1993), Foundations of Genetic Algorithms 2. San Mateo: Morgan Kaufmann Publishers.

Xue Q, Hu YH, and Tompkins WJ (1992), "Neural-Network-Based Adaptive Matched Filtering for QRS Detection", IEEE Trans Bio Eng, 39(4): 317-329.

Yang TF, Devin B, and Macfarlane PW (1994) "Artificial Neural Networks for the Diagnosis of Atrial fibrillation", Med. & Biol. Eng. & Comput., 32: 615-649.