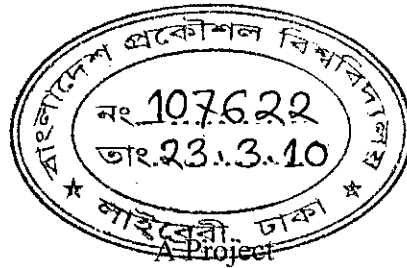


Detection of Cardiac Abnormality Using Entropy Measure of ECG Attractor.

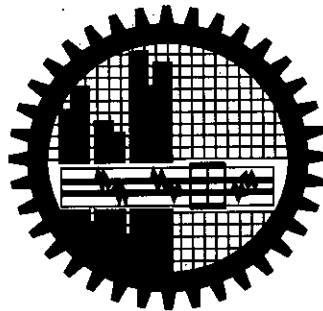
by

A. H. M. Zadidul Karim



Submitted to the Department of
Electrical and Electronic Engineering, BUET,
in partial fulfilment of the requirements for the degree of

**Master of
Electrical and Electronic Engineering**



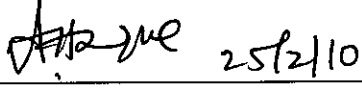
BUET


**DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING
BANGLADESH UNIVERSITY OF ENGINEERING AND TECHNOLOGY**


February, 2010

The project titled, “**Detection of Cardiac Abnormality Using Entropy Measure of ECG Attractor**” by **A. H. M. Zadidul Karim**, Roll No: 040506224 (F), Session: **April 2005**, has been accepted as satisfactory in partial fulfillment of the requirement for the degree of **Master of Electrical and Electronic Engineering** on **25 February 2010**.

BOARD OF EXAMINERS

1. 

Dr. Md. Aynal Haque
Professor
Dept. of EEE, BUET, Dhaka
Chairman
2. 

Dr. Mohammed Imamul Hassan Bhuiyan
Assistant Professor
Dept. of EEE, BUET, Dhaka
Member
3. 

Dr. Shaikh Anowarul Fattah
Assistant Professor
Dept. of EEE, BUET, Dhaka
Member

DECLARATION

It is hereby declared that the work presented here in this project or any part it has not been or being submitted elsewhere for the award of any degree or diploma.

Zadidul Karim

A. H. M. Zadidul Karim

Acknowledgements

I would like to express my gratitude to my supervisor Dr. Md. Aynal Haque for his kind supervision, patience and valuable support in making a difficult task to a pleasant one.

I wish to express my thanks and regards to the Head of the Department of Electrical and Electronic Engineering, BUET, for his support and cooperation.

Sincere thanks to my parents and wife for their encouragement, advice and support to complete my project.

Table of Contents

	ACKNOWLEDGEMENTS	iv
	ABSTRACT	viii
	LIST OF ABBREVIATIONS	ix
	LIST OF TABLES	v
	LIST OF FIGURES	vi
Chapter 1	INTRODUCTION	1
1.1	Background and Motivation	1
1.2	Objective of the Study	5
1.3	Organization of the Dissertation	5
Chapter 2	HEART AND ECG	6
2.1	The Heart	6
2.1.1	Anatomy and Physiology of the Heart	6
2.1.2	Electrical Conduction System of the Heart	8
2.2	Electrocardiogram	10
2.2.1	Electrical Basis of the ECG	10
2.2.2	Components of the ECG	11
2.2.3	ECG Leads	12
Chapter 3	SOME ECG ANALYSIS TECHNIQUES	14
3.1	Introduction	14

3.2	Analysis methods for HRV	15
3.3	Time-domain methods	16
3.3.1	Standard deviation of the NN intervals	16
3.3.2	The standard deviation of successive RR interval differences(SDSD)	17
3.3.3	The root mean square of successive differences (RMSSD)	17
3.3.4	Proportion- pNN50	18
3.4	Frequency-domain methods	18
3.4.1	Wavelet transform	19
3.5	Nonlinear methods	19
3.5.1	Central tendency measure(CTM)	20
3.5.2	Poincaré HRV plot	20
3.5.3	Approximate Entropy	22
3.5.4	Sample Entropy	23
3.6	Conclusion	25
Chapter 4	PROPOSED METHOD	26
4.1	Introduction	26
4.2	Conventional HRV indices	26
4.2.1	Mean IHR	27
4.2.2	Variance IHR	27
4.2.3	Standard deviation of the NN intervals (SDNN)	27
4.2.4	The standard deviation of successive RR interval differences (SDSD)	28
4.2.5	The root mean square of successive differences (RMSSD)	28
4.3	Poincaré plot analysis	29
4.4	Approximate Entropy	30
4.5	Sample Entropy	34

4.6	Conclusion	36
Chapter 5	EXPERIMENTAL RESULTS	37
5.1	Introduction	37
5.2	Poincaré plot analysis	38
5.3	Time domain indexes	64
5.4	Poincaré plot indexes	64
5.5	Approximate entropy(ApEn) values of IHR between Healthy groups and the groups with	65
5.6	Sample Entropy(SampEn) values of IHR between Healthy groups and the groups with abnormal rhythm	65
5.7	Conclusion	66
Chapter 6	CONCLUSIONS	67
6.1	Discussions	67
6.2	Future work	68
Appendix-A	QRS Detection algorythm	69
	REFERENCES	73

Abstract

Chaotic analysis and entropy measurement has been shown to be useful in a variety of medical applications, particularly in cardiology. Chaotic parameters like Poincare plot indexes have shown potential in the identification of diseases, especially in the analysis of biomedical signals like electrocardiogram (ECG).

In this work, entropy measurement, Poincare plot indexes and time domain parameters in ECG signals have been analyzed. First, the ECG signal is processed through a series of steps to extract the QRS complex. From this extracted feature, bit-to-bit interval (BBI) and instantaneous heart rate (IHR) have been calculated. We quantified several time domain HRV parameters: mean IHR, variance, standard deviation of normal IHR data (SDNN) and the square root of the mean squared difference of the successive IHR data (RMSSD), and nonlinear techniques like approximate entropy, sample entropy, and Poincare plot indexes have been determined from the IHR. Standard database of MIT-BIH is used as the reference data where each ECG record contains 650000 samples. Approximate entropy (ApEn) and sample entropy (SampEn) are calculated for IHR for each ECG record of the database. A much higher value of ApEn and SampEn for IHR is observed for eight patients with abnormal beats like T, AFIB, VT. On the contrary, the ApEn and SampEn for IHR of eight patients with normal rhythm shows lower value. The IHR time series and their corresponding Poincaré plots taken from the HRV patterns of patients with normal heart beat and abnormal heart beat are presented in Figure . Poincare plot indexes for IHR of eight normal rhythm records show lower values with a SD1 of 5.6087 and SD2 of 7.1364. On the other hand Poincare plot indexes for IHR of eight abnormal rhythm records show higher values with a SD1 of 23.2093 and SD2 of 22.6107. Time domain parameter are found lower for eight normal rhythm records with variance of 52.6808, SD of 6.5875, SDSD of 7.9260 and SMSSD of 77.5388. . Time domain parameter are found lower for eight abnormal rhythm records with variance of 562.3727, SD of 23.2972, SDSD of 32.8103 and SMSSD of 86.6614. These results indicate that ECG can be classified based on this chaotic modelling which works on the nonlinear dynamics of the system.

List of Abbreviations

AFIB-	Atrial fibrillation
ApEn -	Approximate Entropy
BBI -	Bit-to-Bit Interval
B -	Ventricular bigeminy
CoV -	Coefficient of Variation
CTM -	Central Tendency Measure
DFA -	Detrended Fluctuation Analysis
ECG -	Electrocardiogram
HRV -	Heart Rate Variability
IHR -	Instantaneous Heart Rate
RMSSD-	Root Mean Square of Successive Difference
SD -	Standard Deviation
SDNN-	Standard Deviation of Normal-to-Normal R-Peak interval
T-	Ventricular trigeminy
SampEn-	Sample Entropy
VFIB-	Ventricular Fibrillation
VT-	Ventricular Tachycardia

List of Tables

5.1	Beat annotation appear in the records	37
5.2	Sample Entropy (SampEn) , Approximate entropy (ApEn) obtained from IHR of normal rhythm (continued)	56
5.2	Sample Entropy (SampEn) , Approximate entropy (ApEn) obtained from IHR of normal rhythm (continued)	57
5.3	Sample Entropy(SampEn) , Approximate entropy (ApEn) obtained from IHR of abnormal rhythm (continued)	58
5.3	Sample Entropy(SampEn) , Approximate entropy (ApEn) obtained from IHR of abnormal rhythm (continued)	59
5.4	Sample entropy values of Healthy groups and abnormal rhythm	60
5.5	Approximate entropy values of Healthy groups and abnormal rhythm	60
5.6	HRV parameters in the time domain for Healthy subjects	61
5.7	HRV parameters in the time domain for the patient have the following symptoms: T, AFIB, VT.	61
5.7	HRV parameters in the time domain for the patient have the following symptoms: T, AFIB, VT.	62
5.8	Average HRV parameters in the time domain.	62
5.9	HRV parameters in the Poincare plot for Healthy subjects.	62
5.9	HRV parameters in the Poincare plot for Healthy subjects.	63
5.10	HRV parameters in the Poincare plot for the patient has the following symptoms: T, AFIB, VT.	63
5.11	Average values of HRV parameters in the Poincare plot.	64

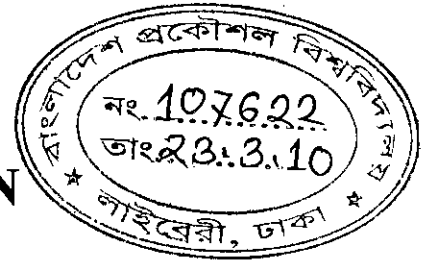
List of Figures

2.1	Anatomy of the heart	07
2.2	Electrical conduction system of heart	09
2.3	Electrical basis of the ECG	11
2.4	Components of the ECG	11
2.5	The standard (bipolar) limb leads I, II, and III.	12
2.6	The augmented (unipolar) leads aVR, aVL, and aVF.	13
2.7	Precordial (unipolar) leads	13
3.1	ECG signal. R peaks are marked with the crosses.	15
3.2	Derivation of two HRV signals from ECG	15
3.3	Ellipse fitting technique	21
4.1	Flowchart for ApEn Analysis	33
4.2	Flowchart for SampEn Analysis	35
5.1	MIT-BIH Arrhythmia Record_100: (a) Poincare plot (RR_i vs. RR_{i+1}); (b) IHR (n) vs. Sample (n). The patient has normal heart beat.	38
5.2	MIT-BIH Arrhythmia Record_105: (a) Poincare plot (RR_i vs. RR_{i+1}); (b) IHR (n) vs. Sample (n). The patient has normal heart beat.	39
5.3	MIT-BIH Arrhythmia Record_111: (a) Poincare plot (RR_i vs. RR_{i+1}); (b) IHR (n) vs. Sample (n). The patient has normal heart beat.	40
5.4	MIT-BIH Arrhythmia Record_112: (a) Poincare plot (RR_i vs. RR_{i+1}); (b) IHR (n) vs. Sample (n). The patient has normal heart beat.	41
5.5	MIT-BIH Arrhythmia Record_116: (a) Poincare plot (RR_i vs. RR_{i+1}); (b) IHR (n) vs. Sample (n). The patient has normal heart beat.	42
5.6	MIT-BIH Arrhythmia Record_118: (a) Poincare plot (RR_i vs. RR_{i+1}); (b) IHR (n) vs. Sample (n). The patient has normal heart beat.	43
5.7	MIT-BIH Arrhythmia Record_121: (a) Poincare plot (RR_i vs. RR_{i+1}); (b) IHR (n) vs. Sample (n). The patient has normal heart beat.	44
5.8	MIT-BIH Arrhythmia Record_122: (a) Poincare plot (RR_i vs. RR_{i+1}); (b) IHR (n) vs. Sample (n). The patient has normal heart beat.	45
5.9	MIT-BIH Arrhythmia Record_106: (a) Poincare plot (RR_i vs. RR_{i+1}); (b) IHR (n) vs. Sample (n). The patient has following symptoms: T,B	46

5.10	MIT-BIH Arrhythmia Record_119: (a) Poincare plot (RR_i vs. RR_{i+1}); (b) IHR (n) vs. Sample (n). The patient has following symptoms: T, B	47
5.11	MIT-BIH Arrhythmia Record_201: (a) Poincare plot (RR_i vs. RR_{i+1}); (b) IHR (n) vs. Sample (n). The patient has following symptoms: T, SVTA, AFIB, and NOD.	48
5.12	MIT-BIH Arrhythmia Record_208: (a) Poincare plot (RR_i vs. RR_{i+1}); (b) IHR (n) vs. Sample (n). The patient has following symptoms: T.	49
5.13	MIT-BIH Arrhythmia Record_210: (a) Poincare plot (RR_i vs. RR_{i+1}); (b) IHR (n) vs. Sample (n). The patient has following symptoms: T, B, AFIB, VT.	50
5.14	MIT-BIH Arrhythmia Record_219: (a) Poincare plot (RR_i vs. RR_{i+1}); (b) IHR (n) vs. Sample (n). The patient has following symptoms: T, B, and AFIB.	51
5.15	MIT-BIH Arrhythmia Record_221: (a) Poincare plot (RR_i vs. RR_{i+1}); (b) IHR (n) vs. Sample (n). The patient has following symptoms: T, B, and AFIB.	52
5.16	MIT-BIH Arrhythmia Record_222: (a) Poincare plot (RR_i vs. RR_{i+1}); (b) IHR (n) vs. Sample (n). The patient has following symptoms: AB, AFIB, SVTA, and NOD.	53
5.17	ApEn between the patient has normal heart rhythm vs. the patient has following symptoms: T, B, SVTA, AFIB, and NOD.	54
5.18	SampEn between the patient has normal heart beat vs. the patient has following symptoms: T, B, SVTA, AFIB, and NOD.	55

Chapter 1

INTRODUCTION



1.1 Background and Motivation

The electrical signals described are measured by the ECG where each heart beat is displayed as a series of electrical waves characterized by peaks and valleys. Electrocardiogram (ECG) is a wave that represents an electrical event in the heart, such as atrial depolarization, atrial repolarization, ventricular depolarization, ventricular repolarization, or transmission, and so on. An ECG gives two major kinds of information. First, by measuring time intervals on the ECG, the duration of the electrical wave crossing the heart can be determined and consequently we can determine whether the electrical activity is normal or slow, fast or irregular. Second, by measuring the amount of electrical activity passing through the heart muscle, a pediatric cardiologist may be able to find out if parts of the heart are too large or are overworked. The ECG signal is characterized by five peaks and valleys labeled by successive letters of the alphabet P, Q, R, S and T. A good performance of an ECG analyzing system depends heavily upon the accurate and reliable detection of the QRS complex, as well as the T and P waves. The P wave represents the both atria activation (depolarization). The first half of the P wave is the activation of the right atrium, whereas the second half of the P wave is the activation of the atria septum and the left atrium, the atria while the QRS wave (or complex) and T wave represent the excitation of the ventricles or the lower chambers of the heart. The detection of the QRS complex is the most important task in automatic ECG signal analysis. Once the QRS complex has been identified, a more detailed examination of ECG signal, including the heart rate, the ST segment, etc., can be performed.

Analyzing the ECG is a useful tool for diagnosing heart diseases. Based on the ECG signal shape and distance between fiducial points and other parameters, physicians diagnose heart diseases. Recognition of the fiducial points and calculation of

parameters is a tedious routine for the physician; 100000 cardiac cycles per patient are recorded in a day and a physician has to interpret this large amount of ECG data to search for a few abnormal cardiac cycles in the ECG. Therefore there is an urgent need for an automatic ECG interpreting system to help reduce the burden of interpreting the ECG.

Conventional methods of monitoring and diagnosing arrhythmia rely on detecting the presence of particular signal features by a human observer. Due to the large number of patients in intensive care units and the need for continuous observation of such conditions, several techniques for automated arrhythmia detection have been developed in the past two decades to attempt to solve this problem. Such techniques work by transforming the mostly qualitative diagnostic criteria into a more objective quantitative signal feature classification problem.

So an efficient mean is required to analyze the ECG signal and diagnose the heart diseases. For this the QRS complex in the ECG is used in many methods to distinguish between the healthy person and the ailing one. Classical techniques have been used to address this problem such as the analysis of electrocardiogram (ECG) signals for arrhythmia detection using the frequency domain features [1], using time domain analysis [2], wavelet transform and non linear analysis such as chaotic analysis, sample entropy , approximate entropy measurement[3]. Other techniques used adaptive filtering [4], sequential hypothesis testing [5].

Even though comparatively good results have been achieved using classical techniques, they seem to provide only a limited amount of information about the signal because they ignore the underlying nonlinear signal dynamics. Recently, there has been an increasing interest in applying techniques from the domains of nonlinear analysis and chaos theory in studying biological systems.

The analysis of heart rate variability is based mainly on analysis of RR intervals [1]. RR intervals are the series of time intervals between heartbeats [2]. We can observe RR

intervals in electrocardiogram, which is simply graphic representation of the electrical forces produced by the heart [3].

Central tendency measure (CTM) has also become very useful in describing the chaotic behavior of the system. When used with clinical parameters [11], CTM can become a powerful indicator of the absence of congestive heart failure. The CTM method for determining variability in nonlinear time series has been shown to be effective in both time series with inherent patterns, such as ECGs and in timer series without a pattern such as hemodynamic studies. CTM can be used along with other techniques like CD, and ApEn to study the underlying chaos in the ECG [12].

The phase space of a dynamical system is a mathematical space with orthogonal coordinate directions representing each of the vectors needed to specify the instantaneous state of the system. Usually taken method of delays is used to construct an attractor of dynamical system in a multidimensional state space from only the knowledge of a one-dimensional time sequence that describes the system behavior.

Studies have been performed to obtain the phase space density plot by mapping the distribution of points in the phase space of ECG signals and the phase space density values within a predefined window were used for arrhythmia detection [13]. Classification was performed using a multilayer back propagation neural network and very good result was obtained for classifying cardiac arrhythmias like premature ventricular contraction, atrial fibrillation, ventricular tachycardia and ventricular fibrillation. Phase space portrait of a single cycle ECG wave has also been derived to study the chaos of the system.

The measurement of heart rate variability (HVR) is a valuable tool in both clinical practice and physiological research. The Poincaré plot of RR intervals is one of the recent methods of HRV analysis [6]. It has also been used to measure the autonomic modulation and randomness of heart. The Poincare plot is a graphical representation of sequential correlations within the RR interval. Poincaré HRV plot is a graph in which each RR interval is plotted against next RR interval (a type of delay map). Various descriptors are associated with this plot, some of which have a convincing

physiological interpretation. Minor axis (SD1), major axis (SD2) and the SD1/SD2 ratio were compared against standard HRV indexes in time domain, in a group of normal subjects and in a group of abnormal subjects [7, 14].

Recent approaches to the study of nonlinearity in biological systems have found a powerful tool in the Approximate Entropy (ApEn) estimation [3], and [8]. Approximate entropy (ApEn) describes the complexity and irregularity of the signal [9, 10]. ApEn is low in regular time series and high in complex irregular ones. It can be applied to both deterministic and stochastic signals and their combinations. In those papers its ability to distinguish, in the short period, different physiological conditions in which the cardiovascular control system can influence the heart rate variability signal (HRV) were analyzed. The results confirm the ability of the regularity estimation to separate different patterns in the HRV time series giving at the same time a strong indication for a possible clinical application of the calculated parameters. In particular by using ApEn it has been possible to classify Myocardial Infarction patients who had different performances of the cardiac pump with a higher connected risk for sudden cardiac death.

The differences between ApEn and SampEn result from 1) defining the distance between two vectors as the maximum absolute difference between their components; 2) excluding self-matches, i.e., vectors are not compared to themselves; and 3) given a time series with L data points, only the first $L-m$ vectors of length m are considered. SampEn is precisely equal to the negative of the natural logarithm of the conditional probability that sequences (epochs) close to each other for m consecutive data points will also be close to each other when one more point is added to each sequence. Having all these features makes SampEn to be a useful tool for investigating the dynamics of heart rate and other time series.

1.2 Objective of the Study

The aim of this work is to detection of abnormality in ECG using Entropy of Chaotic Attractor. Chaos may be defined as the pattern that lies between the determinism and randomness of a system. Poincaré plot indexes as well as traditional time domain

analysis, approximate entropy (ApEn) and the sample entropy (SampEn) measure were used for analyzing variability and complexity of HRV. These non-linear techniques are applied instantaneous heart rate (IHR) that are derived from the sample ECG records from MIT-BIH database [28]. The results found from this work are analyzed to see if there is any abnormality present in the examined ECG record.

1.3 Organization of the Dissertation

Chapter 1 is an introductory chapter. It contains the background and motivation of analysis of ECG, objective and outline of the proposed algorithm and organization of this project. A discussion about the human heart, electrical conduction system of the heart, and electrocardiogram is presented in the chapter 2. Chapter 3 reviews some of the classical techniques that have been used to analyze the ECG signals for arrhythmia detection. These techniques include frequency domain features, time domain analysis, and wavelet transform. Other techniques used adaptive filtering, sequential hypothesis testing. This chapter also includes some of the non-linear methods that are applied to analyze ECG signals. The proposed techniques for analyzing ECG signal using non-linear techniques is described in the chapter 4. Chapter 5 includes the results of the proposed methods. Chapter 6 contains the summary, conclusions, and recommendations for continuation.

Chapter 2

HEART AND ECG

2.1 The Heart

2.1.1 Anatomy and Physiology of the Heart

The heart is a muscular cone-shaped organ about the size of a clenched fist of the same person. It is located in the upper body (chest area) between the lungs, and with its pointed end (called the apex) downwards, forwards, and pointing towards the left. The essential function of the heart is to pump blood to various parts of the body [16]. The heart has four chambers: right and left atria and right and left ventricles. The two atria act as collecting reservoirs for blood returning to the heart while the two ventricles act as pumps to eject the blood to the body. As in any pumping system, the heart comes complete with valves to prevent the back flow of blood. Deoxygenated blood returns to the heart via the major veins (superior and inferior vena cava), enters the right atrium, passes into the right ventricle, and from there is ejected to the pulmonary artery on the way to the lungs. Oxygenated blood returning from the lungs enters the left atrium via the pulmonary veins, passes into the left ventricle, and is then ejected to the aorta. The walls of the ventricles are composed of three layers of tissue: the innermost thin layer is called the endocardium; the middle thick, muscular layer, the myocardium; and the outermost thin layer, the epicardium. The walls of the left ventricle are more muscular and about three times thicker than those of the right ventricle. The atrial walls are also composed of three layers of tissue like those of the ventricles, but the middle muscular layer is much thinner. The two atria form the base of the heart; the ventricles form the apex of the heart. The inter-atrial septum (a thin membranous wall) separates the two atria, and a thicker, more muscular wall, the interventricular septum, separates the two ventricles. The two septa, in effect, divide the heart into two pumping systems, the right heart and the left heart, each one consisting of an atrium and a ventricle.

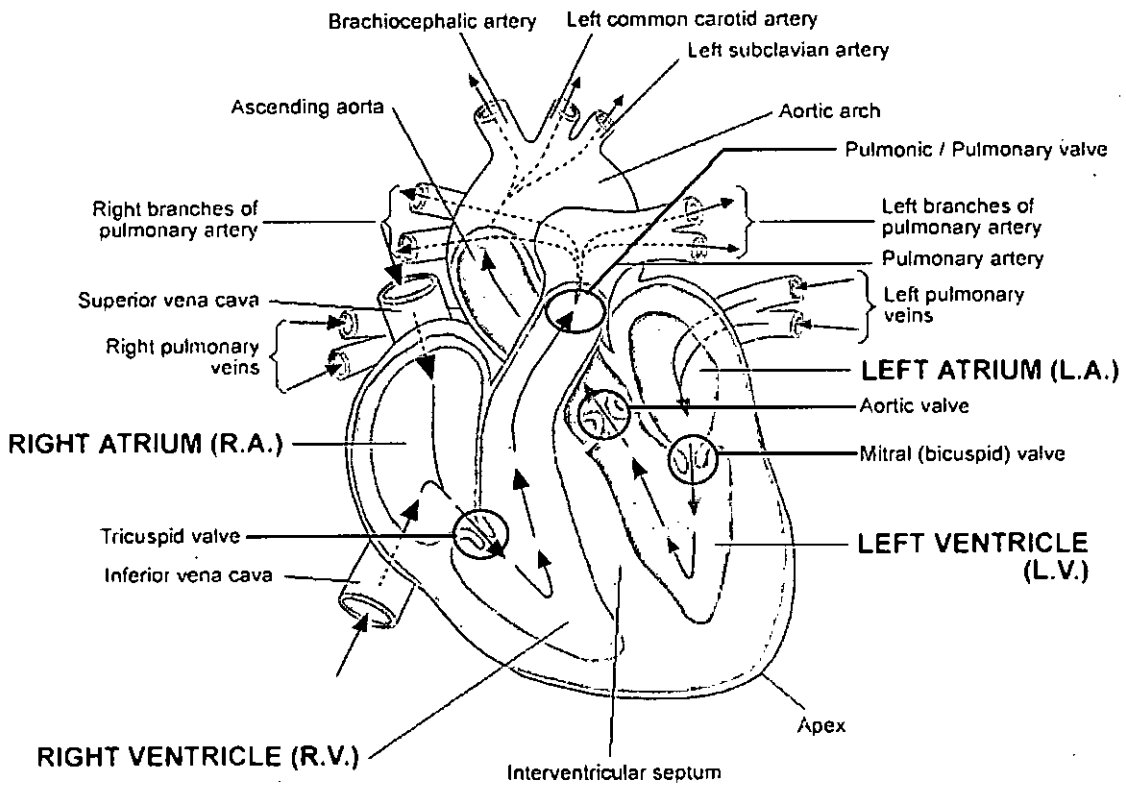


Fig 2.1: Anatomy of the heart

The heart performs its pumping action over and over in a rhythmic sequence. First, the atria relax (atrial diastole), allowing the blood to pour in from the body and lungs. As the atria fill with blood, the atrial pressure rises above that in the ventricles, forcing the tricuspid and mitral valves to open and allowing the blood to empty rapidly into the relaxed ventricles. Then the atria contract (atrial systole), filling the ventricles to capacity.

Following the contraction of the atria, the pressures in the atria and ventricles equalize, and the tricuspid and mitral valves begin to close. Then, the ventricles contract vigorously, causing the ventricular pressure to rise sharply. The tricuspid and mitral valves close completely, and the aortic and pulmonic valves snap open, allowing the blood to be ejected forcefully into the pulmonary and systemic circulations. Meanwhile, the atria are again relaxing and filling with blood. As soon as the ventricles empty of blood and begin to relax, the ventricular pressure falls, the aortic and pulmonic valves shut tightly, the tricuspid and mitral valves open, and the rhythmic cardiac sequence begins anew.

The period from the opening of the aortic and pulmonic valves to their closing, during which the ventricles contract and empty of blood, is called ventricular systole. The following period from the closure of the aortic and pulmonic valves to their reopening, during which the ventricles relax and fill with blood, is called ventricular diastole. The sequence of one ventricular systole followed by a ventricular diastole is called the cardiac cycle, commonly defined as the period from the beginning of one heart beat to the beginning of the next.

2.1.2 Electrical Conduction System of the Heart

The electrical conduction system of the heart (Figure 2.2) is composed of the following structures:

- Sinoatrial (SA) node
- Internodal atrial conduction tracts and the interatrial conduction tract (Bachmann's bundle)
- Atrioventricular (AV) junction consisting of the atrioventricular (AV) node and bundle of His
- Right bundle branch, left bundle branch, and left anterior and posterior fascicles
- Purkinje network

The prime function of the electrical conduction system of the heart is to transmit minute electrical impulses from the SA node (where they are normally generated) to the atria and ventricles, causing them to contract (Figure 2.2). The SA node lies in the wall of the right atrium near the inlet of the superior vena cava. It consists of pacemaker cells that generate electrical impulses automatically and regularly. The three internodal atrial conduction tracts, running through the walls of the right atrium between the SA node and the AV node, conduct the electrical impulses rapidly from the SA node to the AV node in about 0.03 second. The interatrial conduction tract (Bachmann's bundle), a branch of one of the internodal atrial conduction tracts, extends across the atria, conducting the electrical impulses from the SA node to the left atrium. The AV node lies partly in the right side of the interatrial septum in front of the opening of the coronary sinus and partly in the upper part of the interventricular septum above the base of the tricuspid valve. The primary function of the AV node is to relay the electrical impulses from the atria into the ventricles in an orderly and timely way. A ring of fibrous tissue insulates the remainder of the atria

from the ventricles, preventing electrical impulses from entering the ventricles except through the AV node.

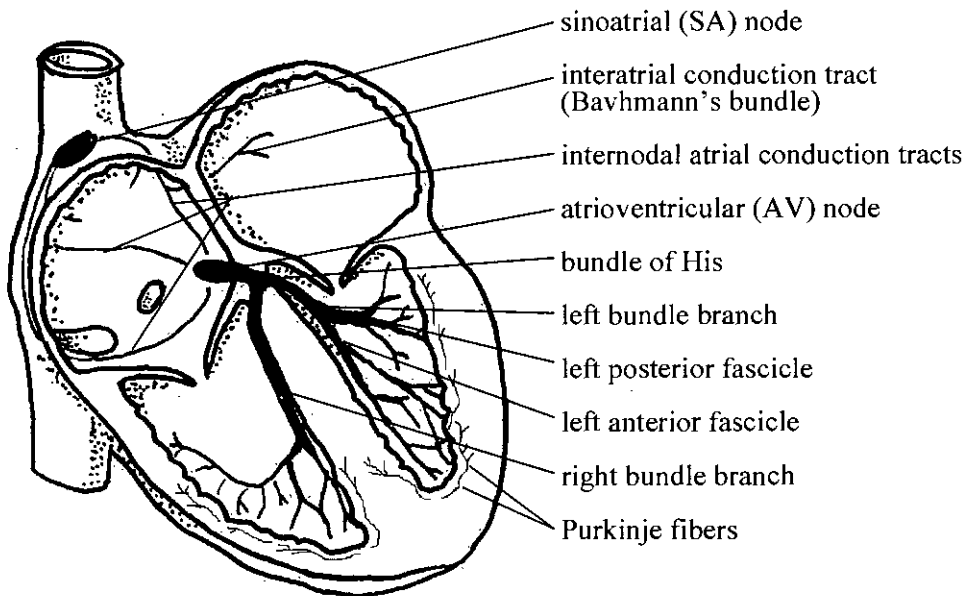


Fig 2.2: Electrical conduction system of heart

The electrical impulses slow as they travel through the AV node, taking about 0.06 to 0.12 second to reach the bundle of His. The delay is such that the atria can contract and empty, and the ventricles fill before they are stimulated to contract. The bundle of His lies in the upper part of the interventricular septum and connects the AV node with the two bundle branches. Once the electrical impulses enter the bundle of His, they travel more rapidly on their way to the bundle branches, taking 0.03 to 0.05 second. The right bundle branch and the left common bundle branch arise from the bundle of His, straddle the interventricular septum, and continue down both sides of the septum. The left common bundle branch further divides into two major divisions: the left anterior fascicle and the left posterior fascicle. The bundle branches and their fascicles subdivide into smaller and smaller branches, the smallest ones connecting with the Purkinje network, an intricate web of tiny Purkinje fibers spread widely throughout the ventricles beneath the endocardium. The ends of the Purkinje fibers finally terminate at the myocardial cells. The bundle of His, the right and left bundle branches, and the Purkinje network are also known as the His-Purkinje system of the ventricles. The electrical impulses travel very rapidly to the Purkinje network through the bundle branches in less than 0.01 second. All in all, it normally takes

the electrical impulses less than 0.2 second to travel from the SA node to the Purkinje network in the ventricles.

2.2 Electrocardiogram

The electrocardiogram (ECG) is a graphic record of the changes in magnitude and direction of the electrical activity, or, more specifically, the electric current, that is generated by the depolarization and repolarization of the atria and ventricles (Figure 2.3). This electrical activity is readily detected by electrodes attached to the skin. But neither the electrical activity that results from the generation and transmission of electrical impulses which are too feeble to be detected by skin electrodes nor the mechanical contractions and relaxations of the atria and ventricles (which do not generate electrical activity) appear in the electrocardiogram.

2.2.1 Components of the ECG

After the electric current generated by depolarization and repolarization of the atria and ventricles is detected by electrodes, it is amplified, displayed on an oscilloscope, recorded on ECG paper, or stored in memory. The electric current generated by atrial depolarization is recorded as the P wave, and that generated by ventricular depolarization is recorded as the Q, R, and S waves: the QRS complex. Atrial repolarization is recorded as the atrial T wave (Ta), and ventricular repolarization, as the ventricular T wave, or simply, the T wave. Because atrial repolarization normally occurs during ventricular depolarization, the atrial T wave is buried or hidden in the QRS complex. In a normal cardiac cycle, the P wave occurs first, followed by the QRS complex and the T wave (Figure 2.4).

The sections of the ECG between the waves and complexes are called segments and intervals: the PR segment, the ST segment, the TP segment, the PR interval, the QT interval, and the R-R interval. Intervals include waves and complexes, whereas segments do not. When electrical activity of the heart is not being detected, the ECG is a straight, flat line – the isoelectric line or baseline [30].

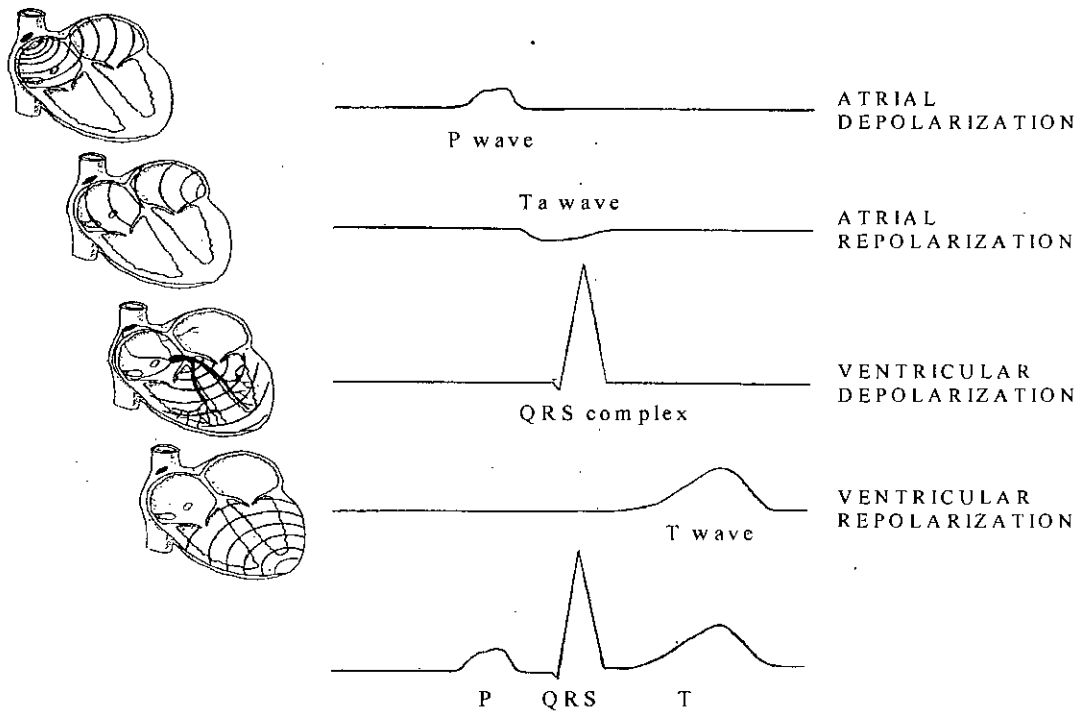


Fig 2.3: Electrical basis of the ECG

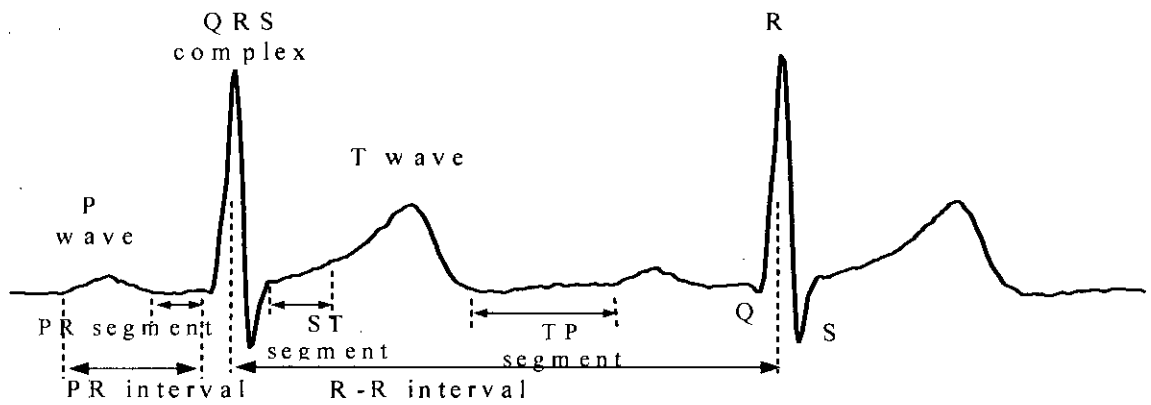


Fig 2.4: Components of the ECG

2.2.2 ECG Leads

An ECG lead is a record (spatial sampling) of the electrical activity generated by the heart that is sensed by either one of two ways: (1) two discrete electrodes of opposite polarity or (2) one discrete positive electrode and an “indifferent,” zero reference point. A lead composed of two discrete electrodes of opposite polarity is called a bipolar lead; a lead composed of a single discrete positive electrode and a zero reference point is a unipolar lead.

Depending on the ECG lead being recorded, the positive electrode may be attached to the right or left arm, the left leg, or one of several locations on the anterior chest wall. The negative electrode is usually attached to an opposite arm or leg or to a reference point made by connecting the limb electrodes together.

For a detailed analysis of the heart's electrical activity, usually in the hospital setting, an ECG recorded from 12 separate leads (the 12-lead ECG) is used. The 12-lead ECG is also used in the prehospital phase of emergency care in certain advanced life support services to diagnose acute myocardial infarction and to help in the identification of certain arrhythmias. A 12-lead ECG consists of three standard (bipolar) limb leads (leads I, II, and III) (Figure 2.5), Three augmented (unipolar) leads (leads aVR, aVL, and aVF) (figure 2.6), and six precordial (unipolar) leads (V_1 , V_2 , V_3 , V_4 , V_5 , and V_6) (Figure 2.7).

When monitoring the heart solely for arrhythmias, a single ECG lead, such as the standard limb lead II, is commonly used, especially in the prehospital phase of emergency care.

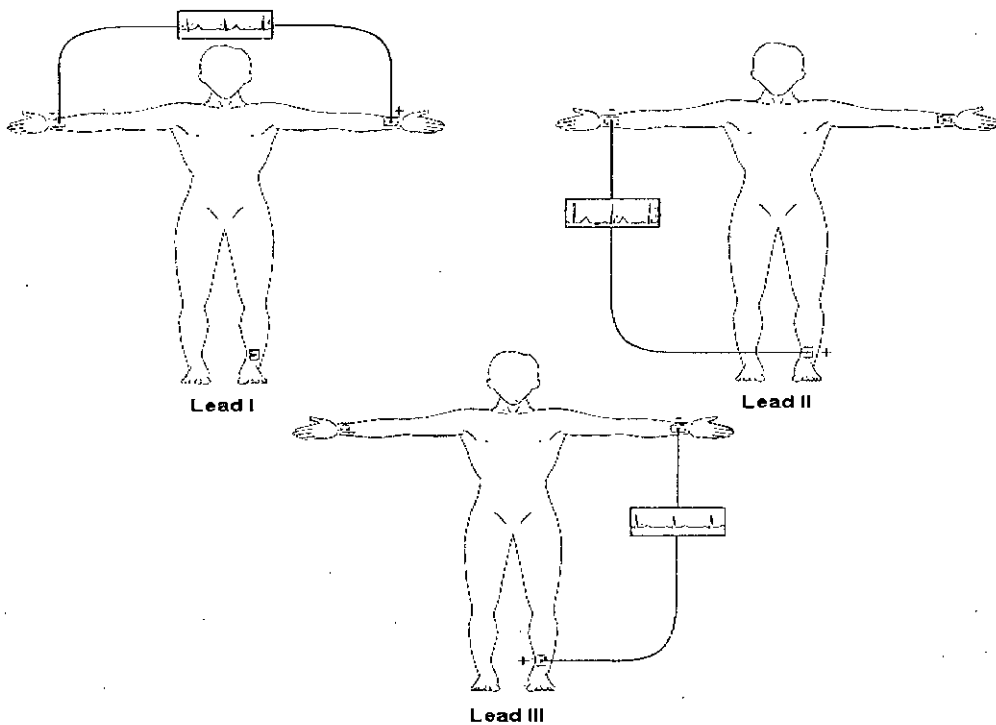


Fig 2.5: The standard (bipolar) limb leads I, II, and III.

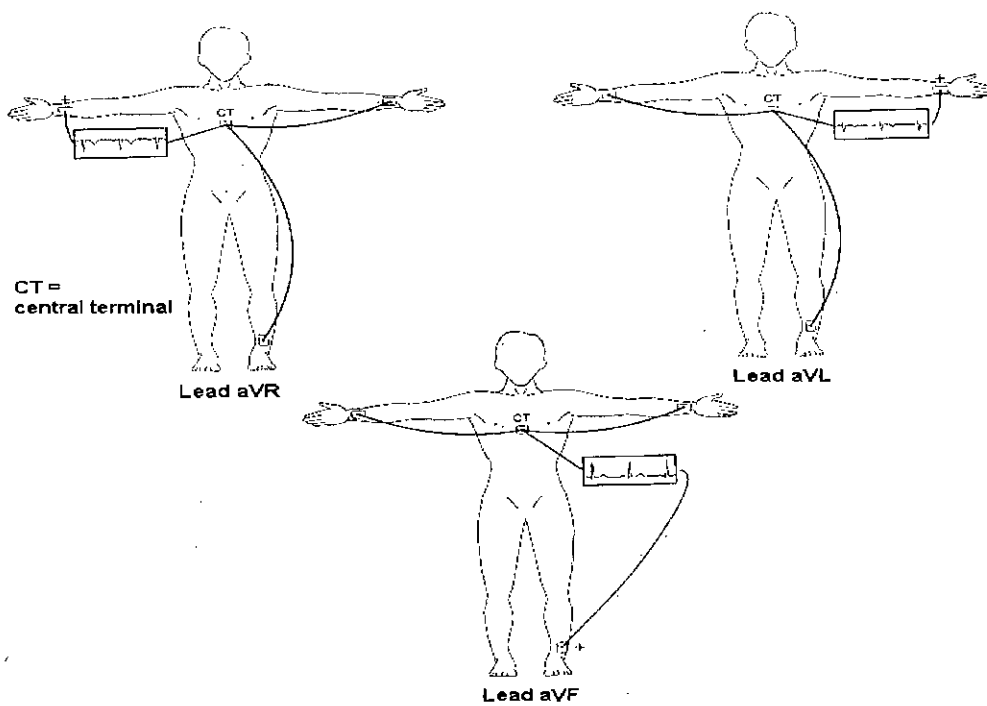


Fig 2.6: The augmented (unipolar) leads aVR, aVL, and aVF.

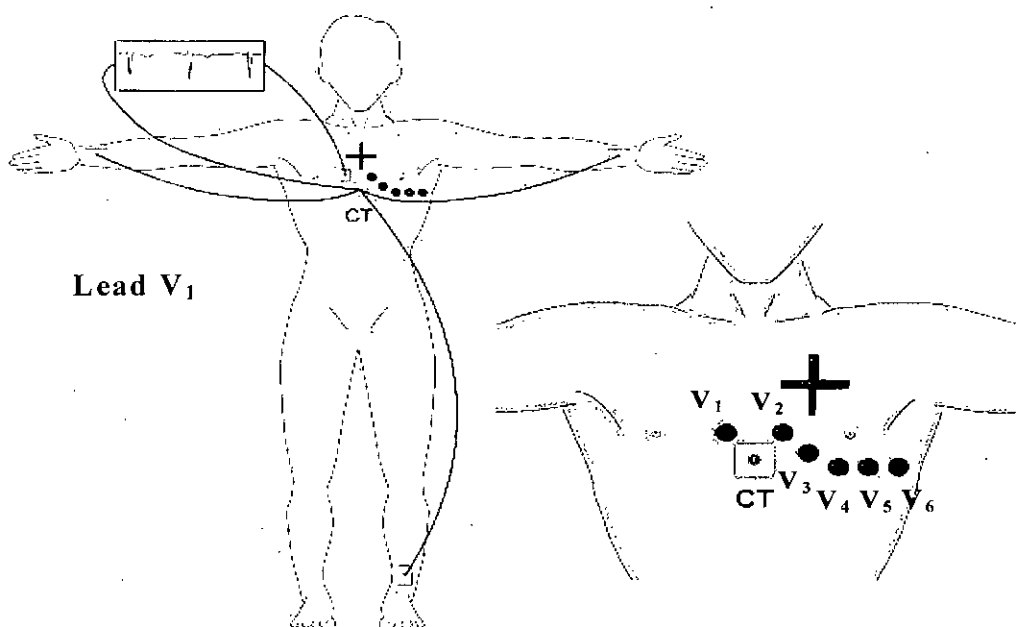


Fig 2.7: Precordial (unipolar) leads

Chapter 3

ECG ANALYSIS TECHNIQUES

3.1 Introduction:

The analysis of the ECG is widely used for diagnosing many cardiac diseases, which are the main cause of mortality in developed countries. A large variety of techniques for ECG analysis have been proposed and published over the last two decades. These techniques have become essential in a large variety of applications, from diagnosis through supervision and monitoring applications. Time domain measures, standard deviation of NN interval (SDNN), root mean square of successive NN interval differences (RMSSD), successive NN intervals differing more than 50 ms (NN50 Count), percentage value of NN50 count (pNN50), HRV triangular index and triangular interpolation of NN intervals (TINN) also show significant difference between the healthy patients and patients with different heart diseases[15,17]. Frequency domain analysis of extracted normal to normal interval (NN interval) data indicates significant difference in very low frequency (VLF) power, low frequency (LF) power and high frequency (HF) power, between the healthy patients and patients with different heart diseases. Other techniques used adaptive filtering, sequential hypothesis testing[17,18]. Recently, there has been an increasing interest in applying techniques from the domains of nonlinear analysis and chaos theory in studying biological systems. The Poincaré plot analysis is a geometrical and non-linear method to assess the dynamics of heart rate variability (HRV). the Poincaré plot gives a useful visual contact to the R-R data by representing both short- and long-term variations[14,19]. Analysis of Poincaré plots can be performed by a simple visual inspection of the shape of the attractor which has been used to classify the signal. Lack of regularity in physiological time series is often cardiology. Chaotic parameters have shown potential in the identification of disease, especially in the analysis of biomedical signals. This index can be efficiently evaluated even over relatively short time series, making it particularly suitable for the analysis of physiological signals. The objective of this chapter is to present some of the techniques for ECG signal analysis. Quantified by

computing the approximate entropy (ApEn) and sample entropy (SampEn) [8,10,21]. Chaotic analysis has been shown to be useful in a variety of medical applications especially in ECG.

3.2 Analysis methods for HRV

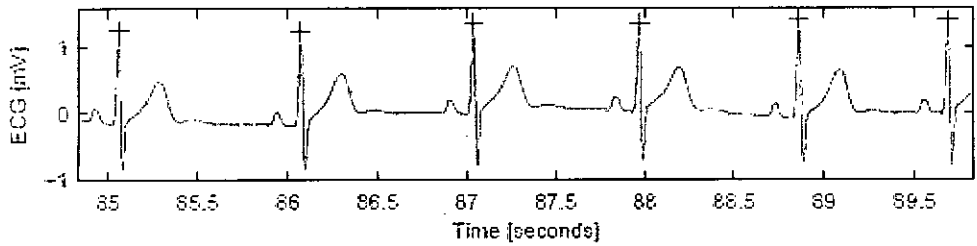


Fig 3.1: ECG signal. R peaks are marked with the crosses.

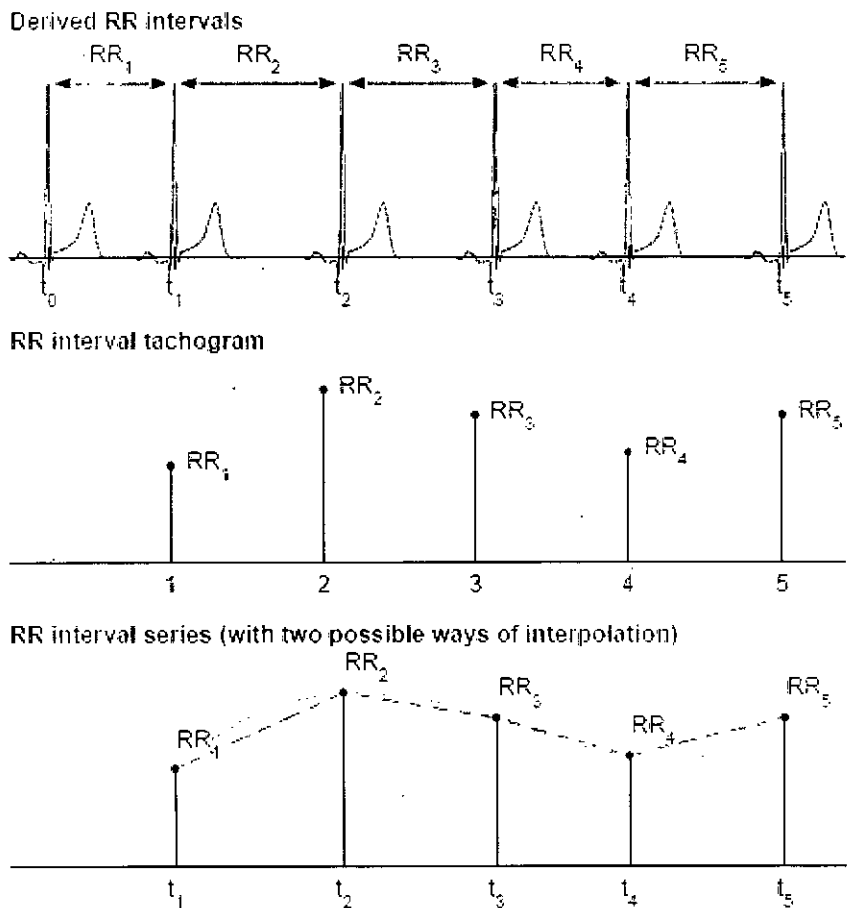


Fig 3.2: Derivation of two HRV signals from ECG:

The base of the HRV analysis is the electrocardiography (ECG) from which the QRS complexes are determined, see Fig. 2.3. Usually the QRS complex is detected at the R wave and the RR time series is used in the HRV analysis. Because the R wave has a better signal-to-noise ratio than the P wave, it can be more easily and reliably detected. As the time interval between P and R detections is constant, the R waves can be used reliably for recognition of QRS complexes [24]. Sometimes the term normal-to-normal (NN) is used instead of RR to emphasize that the intervals are between adjacent QRS complexes resulting from normal sinus node depolarization.

3.3 Time-domain methods

3.3.1 Standard deviation of the NN intervals

The time-domain methods are the simplest to perform since they are applied straight to the series of successive RR interval values. The simplest variable to calculate is the SDNN that is the square root of variance. Since variance is mathematically equal to total power of spectral analysis, SDNN reflects all the repeated components responsible for variability in the period of recording.

Mean: The sum of a list of numbers, divided by the total number of numbers in the list, Where $(RR_1, RR_2, RR_3, \dots, RR_N)$ is the sample here total number of sample is N.

$$\overline{RR} = \frac{1}{N} \sum_{i=1}^N RR_i \quad (3.1)$$

Variance: The variance is a measure of the dispersion of a set of values. The variance is the mean of the sum of the squares of the differences between the values and the mean of the sample [15].

$$\sigma^2 = \frac{1}{N-1} \sum_{i=1}^N (RR_i - \overline{RR})^2 \quad (3.2)$$

The most obvious such measure is the mean value of RR intervals \overline{RR} . In addition, several variables that measure the variability within the RR series exist. The standard deviation of RR

intervals (SDNN) is defined as

$$SDNN = \sqrt{\frac{1}{N-1} \sum_{i=1}^n (RR_i - \overline{RR})^2} \quad (3.3)$$

Where RR_i denotes the value of i 'th RR interval and N is the total number of successive intervals. Where $(RR_1, RR_2, RR_3, \dots, RR_N)$ is the sample and \overline{RR} is the mean of the sample. The denominator $N-1$ is the number of degrees of freedom in the vector $\{(RR_1 - \overline{RR}), \dots, (RR_N - \overline{RR})\}$. The SDNN reflects the overall (both short-term and long-term) variation within the RR interval series [15].

3.3.2 The standard deviation of successive RR interval differences (SDSD)

The standard deviation of successive RR interval differences (SDSD) given by

$$SDSD = \sqrt{Var (RR_N - RR_{N+1})} \quad (3.4)$$

SDSD can be used as a measure of the short-term variability.

3.3.3 The root mean square of successive differences (RMSSD)

The most commonly used measures derived from interval differences include the square root of the mean squared differences of successive NN intervals [15]. Calculation of root mean square is show in equation 5.

$$RMSSD = \sqrt{\frac{1}{N} \sum_{i=1}^N RR_i^2} \quad (3.5)$$

3.3.4 Proportion- pNN50

Another measure calculated from successive RR interval differences is the NN50 which is the number of successive intervals differing more than 50 ms or the corresponding relative amount [15]

$$pNN\ 50 = \frac{NN\ 50}{N - 1} \times 100\ \%$$

3.4 Frequency-domain methods

In the frequency-domain methods, a power spectrum density (PSD) estimate is calculated for the RR interval series. In HRV analysis, the PSD estimation is generally carried out using either FFT based methods or parametric AR modeling based methods [1, 22].

The advantage of FFT based methods is the simplicity of implementation, while the AR spectrum yields improved resolution especially for short samples. Another property of AR spectrum that has made it popular in HRV analysis is that it can be factorized into separate spectral components. The disadvantages of the AR spectrum are the complexity of model order selection and the contingency of negative components in the spectral factorization. Nevertheless, it may be advantageous to calculate the spectrum with both methods to have comparable results.

The generalized frequency bands in case of short-term HRV recordings are the very low frequency (VLF, 0–0.04 Hz), low frequency (LF, 0.04–0.15 Hz), and high frequency (HF, 0.15–0.4 Hz)[9]. The frequency-domain measures extracted from the PSD estimate for each frequency band include absolute and relative powers of VLF, LF, and HF bands, LF and HF band powers in normalized units, the LF/HF power ratio, and peak frequencies for each Band[17]. In the case of FFT spectrum, absolute power values for each frequency band are obtained by simply integrating the spectrum over the band limits. In the case of AR spectrum, on the other hand, if factorization is enabled distinct spectral components emerge for each



frequency band with a proper selection of the model order and the absolute power values are obtained directly as the powers of these components. If factorization is disabled the AR spectrum powers are calculated as for the FFT spectrum.

3.4.1 Wavelet transform

The wavelet transform has been used in the biomedical signal processing and has an important role in the ECG characterization and QRS detection [21]. These class of algorithms for ECG signal processing normally uses discrete wavelet transforms (DWT), dyadic wavelet transform (DyWT) and continuous wavelet transforms (CWT), and all of them with real wavelet functions to decompose, analyze, and compress the ECG signals[23].

The Fourier Transform (FT) decomposes signals in infinite extended sinusoidal functions, requiring the stationary hypothesis; so all the information localized in time, such QRS complex, are spread over the entire frequency axis. The Short Time Fourier Transform (STFT) is useful if the signal is non-stationary over the whole interval, but maintains its frequency characteristic during the short time interval. The STFT provides time-frequency information: a local spectrum is performed by windowing the signal through fixed dimension windows where it may be considered approximately stationary. The window dimension fixes both time and frequency resolutions. The Wavelet Transform (WT) is a tool for non-stationary signal processing that is alternative both to the classical spectral analysis in the frequency domain and to the STFT. The WT overcomes the fixed resolution analysis of the STFT. It satisfies the uncertainty principle or Heisenberg inequality, so time resolution increasing lowers frequency resolution.

3.5 Geometrical methods

Considering the complex control systems of the heart it is reasonable to assume that nonlinear mechanisms are involved in the genesis of HRV. The nonlinear properties of HRV have been analyzed using measures such as Central tendency measure (CTM), Poincar'e plot [19, 25, 26], approximate and sample entropy [3, 27], detrended fluctuation analysis, correlation dimension [16, 18], and recurrence plots. During the last years, the number of studies utilizing

such methods has increased substantially.

3.5.1 Central tendency measure (CTM)

Central tendency measure (CTM) is used to quantify the degree of variability in the second order difference plot. It is calculated by selecting a circular region of radius r , around the origin, counting the number of points that fall within the radius, and dividing by the total number of points. If N = total number of points, and r = radius of central area. Then

$$n = \left[\sum_{i=1}^{N-2} \delta (d_i) \right] / (N - 2) \quad (3.6)$$

Where,

$$\begin{aligned} \delta (d_i) &= 1, \text{ if } [(RR_{i+2} - RR_{i+1})^2 + (RR_{i+1} - RR_i)^2]^{0.5} < r \\ &= 0, \text{ otherwise} \end{aligned} \quad (3.7)$$

3.5.2 Poincaré HRV plot

Poincaré HRV plot is a graph in which each RR interval is plotted against next RR interval (a type of delay map). Poincaré plot has many synonyms like Scatter plot; scattergram Return map; phase delay map, Lorenz plot.

The Poincaré plot (return map) is a scattergram, which is constructed by plotting each RR interval against the previous one [25, 26]. The Poincaré plot may be analyzed quantitatively by fitting an ellipse to the plotted shape. The center of the ellipse is determined by average RR interval. $SD1$ means the standard deviation of the distances of points from $y = x$ axis, $SD2$ means the standard deviation of the distances of points from $y = -x + \overline{RR}$ axis, where \overline{RR} is the average R-R interval. $SD1$ (instantaneous beat-to-beat variability of the data) determines the width of the ellipse, $SD2$ (continuous beat-to-beat variability) determines the length of the ellipse. The ratio $SD1/SD2$ is the measure of heart activity [6].

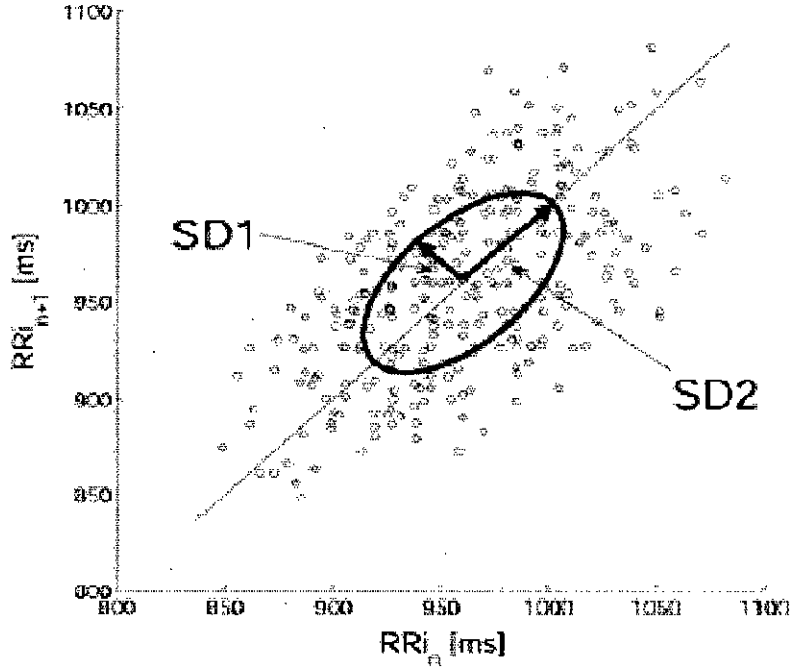


Fig 3.3: Ellipse fitting technique

SD1: dispersion (standard deviation) of points perpendicular to the axis of line-of-identity

$$\begin{aligned}
 SD_1^2 &= Var(x_1) = Var\left(\frac{1}{\sqrt{2}} RR_N - \frac{1}{\sqrt{2}} RR_{N+1}\right) \\
 &= \frac{1}{2} Var(RR_N - RR_{N+1}) = \frac{1}{2} SDSD^2
 \end{aligned} \tag{3.8}$$

SD2: dispersion (standard deviation) of points along the axis of line-of-identity.

$$SD_2^2 = 2SDNN^2 - \frac{1}{2} SDSD^2 \tag{3.9}$$

The ratio SD1/SD2 is the measure of heart activity.

The advantages of Poincare plot

- (1) Simple visualization tool
- (2) Outlier (ectopic beat or artifact) identifier
- (3) Possible insights into short-term and long-term variability

The limitations of Poincare plot

Derived statistics not independent of other time domain measures

3.5.3 Approximate Entropy

Approximate entropy (ApEn) describes the complexity and irregularity of the signal. ApEn is low in regular time series and high in complex irregular ones. It can be applied to both deterministic and stochastic signals and their combinations. Lack of regularity in physiological time series is often quantified by computing the approximate entropy, ApEn[27]. ApEn is related to the probability that segments of “m” data samples which are similar (i.e., closer each other than a given distance “r”) remain similar when the segment length increases to “m+1” [3]. Lower is this probability (and thus the predictability of the time series), greater is ApEn.

For a given r value, ApEn(r) is calculated as follows. First we set $m=2$ and from the RR series of $N = \text{Total number of beats}$, $RR(i) \quad i=1, \dots, N$, we created the series of $N-m+1$ vectors of m components $RR_m(i) = [RR(i), RR(i+1), \dots, RR(i+m)]^T$. The vector $RR_m(i)$ represents the sequence of m consecutive RR values starting at the beat i. Two vectors $RR_m(i)$ and $RR_m(j)$ are similar if the absolute difference between each couple of corresponding scalar components is less than the distance $r \times SD$. Calling $n_i^m(r)$ the number of $N-m+1$ vectors $RR_m(j)$ [9] which are similar to $RR_m(i)$, then

$$C_i^m(r) = \frac{n_i^m(r)}{N - m + 1} \quad (3.10)$$

is the probability to find a sequence of m beats similar to the sequence represented by $RR_m(i)$, and $C_i^m(r)$, defined as the mean of all $C_i^m(r)$, quantifies the prevalence of similar strings of m beats. $ApEn(r)$ is calculated as:

$$ApEn(N, m, r) = \ln \left[\frac{C_i^m(r)}{C_i^{m+1}(r)} \right] \quad (3.11)$$

A high degree of regularity means that sequences which are similar for m points are likely to be similar also for the next $m+1$ point, while this is unlikely to occur for irregular time series. Thus low values of $ApEn$ reflect high regularity.

3.5.4 Sample Entropy

Sample entropy (SampEn) values of IHR signals from all subjects were calculated. SampEn was developed to reduce the bias caused by the self matching in approximate entropy which is a mathematical approach to quantifying the complexity and regularity of a system [9]. SampEn is defined as the logarithmic likelihood that the patterns of the data that are close to each other will remain close for the next comparison within a longer pattern. SampEn does not use a template-wise approach when estimating conditional probabilities. It only requires that one template find a match of length $m+1$, then it computes the logarithm of a probability associated with the time series as a whole. Parameter m which specifies the pattern length was set at 2 and r defining the criterion of similarity was set at 10%-90% of the standard deviation of IHR data. Length of data (N) was varied from 1400-2000 beat numbers [9, 10].

Sample entropy (SampEn) values of IHR signals from all subjects were calculated. SampEn was developed to reduce the bias caused by the self matching in approximate entropy which

is a mathematical approach to quantifying the complexity and regularity of a system. SampEn is defined as the logarithmic likelihood that the patterns of the data that are close to each other will remain close for the next comparison within a longer pattern. SampEn does not use a template-wise approach when estimating conditional probabilities. It only requires that one template find a match of length $m+1$, then it computes the logarithm of a probability associated with the time series as a whole. A brief description is as follows.

Given a sequence of total N numbers of RR such as $RR(1), RR(2), \dots, RR(N)$. To compute SampEn of each RR data set, m -dimensional vector sequences $p_m(i)$ were constructed from the RR series $[p_m(1), p_m(2), \dots, p_m(N-m+1)]$, where the index i can take values ranging from 1 to $N-m+1$. If the distance between two vectors $p_m(i)$ and $p_m(j)$ is defined as $|p_m(i) - p_m(j)|$,

$$C_i^m(r) = \frac{1}{N-m+1} [\text{Number of vectors such that } p_m(j) - p_m(i) < r \text{ and } i \neq j] \quad (3.12)$$

Where m specifies the pattern length which is 2 in this study, r defines the criterion of similarity which was varied from 10~90% of the standard deviation of IHR data ($N=1400-2000$ beats). $C_i^m(r)$ is considered as the mean of the fraction of patterns of length m that resemble the pattern of the same length that begins at index i [3]. Then the SampEn is computed by using the following equation:

$$SampEn(N, m, r) = \ln \left[\frac{\sum_{i=1}^{N-m+1} C_i^m(r)}{\sum_{i=1}^{N-m} C_i^{m+1}(r)} \right] \quad (3.13)$$

We divided the data set into smaller sets of length, i.e., $m=2$. The next step is to

determine the number of subsets that are within the criterion of similarity r , excluding the self match. Then we repeat the same process for the second subset until each subset is compared with the rest of the data set[3]. This process computes $\sum_{i=1}^{N-m+1} C_i^m(r)$ part of equation (13) and $N-m+1$.

3.6 Conclusion:

In this chapter we discussed about various methods of analysis ECG signals. In time domain analysis methods we discussed about mean, variance, Standard deviation, standard deviation of successive difference (SDSD), root mean square of successive difference (RMSSD). In Poincare plot indexes we discussed about SD1, SD2, ratio of SD1 and SD2, and the area of the ellipse. Poincare plot gives us a visual observation of the ECG signals. By using Entropy measure methods like approximate entropy method and sample entropy we can easily separate normal rhythms from abnormal rhythms. Other methods also discussed like frequency domain analysis of ECG signals, Wavelet transform, Central tendency measure (CTM).

Chapter 4

PROPOSED METHOD

4.1 Introduction

By using some robust statistical tools decision making algorithm are designed in order to clarify normal and abnormal rhythm. This project work is based on time domain analysis and some non-linear methods. Heart rate variability (HRV) dataset are collected from MIT-BIH Arrhythmia data bank. The ECG signal to be analyzed is first processed [18] to extract the QRS complex. From that bit-to-bit interval (BBI) is calculated. From the BBI we get the instantaneous heart rate (IHR). On this dataset of IHR, various time domain parameters like mean, variance, standard deviation (SD), the standard deviation of successive RR interval differences (SDSD), the root mean square of successive differences (RMSSD) and some non-linear parameters like Poincaré HRV plot, approximate entropy (ApEn), and Sample entropy (SampEn) are determined. The result obtained from the application of these techniques is analyzed to distinguish the ECG signals between the healthy person and that of the ailing person. The HRV analysis described in the following sections was performed on IHR of 1400 to 2000 beats.

4.2 Conventional HRV indices

Variation in heart rate may be evaluated by a number of methods. Possibly the simplest to perform are the time domain measures. With these methods either the heart rate at any point in time or the intervals between successive normal complexes are determined. The statistical analysis of the calculated features indicate that they differ significantly between normal heart rhythm and the different arrhythmia types and hence, can be rather useful in ECG arrhythmia detection. The five statistical parameters considered for cardiac arrhythmia classification of the ECG signals are the mean of RR intervals, the variance of RR intervals, the standard deviation of the RR intervals (SDNN), the standard deviation of differences between adjacent RR intervals (SDSD) and the root mean square successive difference of intervals which are

extracted from heart rate signals (RMSSD). Time domain methods are easy to program in Matlab because of its simple mathematical expression. There is a wide range of variation found between normal and abnormal rhythms. We can easily separate normal and abnormal heart beat by using time domain method. For this reason I choose time domain analysis for separating normal and abnormal rhythms.

4.2.1 Mean IHR

In this project the mean is obtaining as

$$Mean = \frac{RR_1 + RR_2 + \dots + RR_N}{N} \quad (4.1)$$

The data saved in these files can be loaded into the MATLAB workspace by 'dload()' command. We can find mean by using `m = mean ()` command [29].

4.2.2 Variance IHR

Variance is the square of the standard deviation. A measure of the degree of spread among a set of values; a measure of the tendency of individual values to vary from the mean value. In this project we found variance by using the var command of MATLAB . We can find variance by using `v = var ()` command. Variance is symbolized by σ^2 .

$$Variance = (Standard\ deviation)^2$$

4.2.3 Standard deviation of the NN intervals (SDNN)

Standard deviation is a statistical measure of spread or variability. The standard deviation is the root mean square (RMS) deviation of the values from their arithmetic mean. The standard deviation of RR intervals (SDNN) is defined as

$$SDNN = \sqrt{\frac{1}{N-1} \sum_{i=1}^n (RR_i - \overline{RR})^2} \quad (4.2)$$

In this project we find SDNN by using MATLAB programming which is given below

$$SDNN = \sqrt{\frac{1}{n} \sum ((y-m)^2)}$$

y is the RR data and m is the mean of total RR

4.2.4 The standard deviation of successive RR interval differences (SDSD)

In this project we find SDSD from Poincare plot. We can find SDSD by using the given equation

$$SDSD = \sqrt{Var(RR_N - RR_{N+1})} \quad (4.3)$$

$$SDSD = \sqrt{2 \times SDI^2} \quad (4.4)$$

SDSD is calculated by using matlab program.

$$SDSD = \sqrt{2 * sd1^2}$$

SDI is the standard deviation of projection of the Poincare plot on the line perpendicular to the line of identity.

4.2.5 The root mean square of successive differences (RMSSD)

$$RMSSD = \sqrt{\frac{1}{N} \sum_{i=1}^N RR_i^2} \quad (4.5)$$

Where $RR(j) = RR(1), RR(2), \dots, RR(N)$. N = Total number of IHR data. In this project MATLAB program helps to find RMSSD. We use matlab command which is given below.

$$RMSSD = \sqrt{(\sum(RR_i.^2))/n}$$

4.3 Poincaré plot analysis

The Poincaré plot is a popular two-dimensional visualization tool for dynamic systems due to its intuitive display of the dynamic properties of a system from a time series. The length (SD2) and the width (SD1) of the Poincaré plot images represent short and long-term variability of any nonlinear dynamic system. The Poincaré plot was generated as a scatter plot of current instantaneous heart rate (IHR) against the IHR immediately preceding it. In this project we define the Poincaré plot for a data vector $RR_i = (RR_1, RR_2, \dots, RR_N)$ of length N . First, we define two auxiliary vectors:

$$RR_i^+ = (RR_1, RR_2, \dots, RR_{N-1}) \quad (4.6)$$

$$RR_i^- = (RR_2, RR_3, \dots, RR_N) \quad (4.7)$$

The Poincaré plot consists of all the ordered pairs:

$$(RR_i^+, RR_i^-), \quad i = 1 \dots N - 1. \quad (4.8)$$

We use matlab program to find out the (RR_i^+, RR_i^-) . Matlab command is given below

$$Xp = RR_i; Xp(\text{end}) = [];$$

$$Xm = RR_i; Xm(1) = [];$$

$SD1$ and $SD2$ are two standard Poincaré plot descriptors. $SD2$ is defined as the standard deviation of the projection of the Poincaré plot on the line of identity ($y = x$), and $SD1$ is the standard deviation of projection of the PP on the line perpendicular to the line of identity ($y = -x$) [1]. In this project we define $SD1$ and $SD2$ as:

$$SD1 = \sqrt{\text{Var}(X1)} \quad (4.9)$$

$$SD2 = \sqrt{\text{Var}(X2)} \quad (4.10)$$

Where,

$$X1 = \left(\frac{RR_i^+ - RR_i^-}{\sqrt{2}} \right) \quad (4.11)$$

$$X2 = \left(\frac{RR_i^+ + RR_i^-}{\sqrt{2}} \right) \quad (4.12)$$

We find SD1 and SD2 by using matlab command as:

$$SD1 = \text{std}(Xp - Xm) / \text{sqrt}(2);$$

$$SD2 = \text{std}(Xp + Xm) / \text{sqrt}(2);$$

We define a parameter which reflects the total variability as measured by the Poincaré plot:

$$S = \pi \times SD1 \times SD2 \quad (4.13)$$

Which is the area of the ellipse. S can be expressed in matlab as:

$$S = \text{pi} * SD1 * SD2;$$

4.4 Approximate Entropy

ApEn is a "regularity statistic" that quantifies the unpredictability of fluctuations in a time series. Intuitively, one may reason that the presence of repetitive patterns of fluctuation in a time series makes it more predictable than a time series in which such patterns are absent. ApEn reflects the likelihood that "similar" patterns of observations will not be followed by additional "similar" observations. A time series containing many repetitive patterns has a relatively small ApEn; a less predictable (i.e., more complex) process has a higher ApEn.

A brief summary of the calculations, as applied to a time series of heart rate measurements, $RR(i)$. Given a sequence N , consisting of N instantaneous heart rate measurements $RR(1), RR(2), \dots, RR(N)$, we must choose values for two input parameters, m and r , to compute the approximate entropy, $ApEn(N, m, r)$, of the sequence. The second of these parameters, m , specifies the pattern length, and the third, r , defines the criterion of similarity. We denote a subsequence (or pattern) of heart rate measurements, beginning at measurement i within N , by the vector $p_m(i)$. Two patterns, $p_m(i)$ and $p_m(j)$, are similar if the difference between any pair of corresponding measurements in the patterns is less than r , i.e., if

$$|p_m(i) - p_m(j)| < r \text{ for } 0 \leq k < m$$

Now consider the set $p_m(i)$ of all patterns of length m [i.e., $p_m(1), p_m(2), p_m(3), \dots, p_m(N-m+1)$], within N . For each subject and for each of the two experimental conditions, we considered a time series of $N=2000$ consecutive RR values; we calculated the standard deviation, SD , of RR and evaluated $ApEn$ setting $m=2$ and with r increasing from 0.1 to 0.9. For a given r value, $ApEn(r)$ was calculated.

We may now define:

$$C_i^m(r) = \frac{n_i^m(r)}{N - m + 1} \quad (4.14)$$

Where $n_i^m(r)$ is the number of patterns in P_m that are similar to $p_m(i)$ (given the similarity criterion r). The quantity $C_i^m(r)$ is the fraction of patterns of length m that resemble the pattern of the same length that begins at interval i . We can calculate $C_i^m(r)$ for each pattern in P_m , and we define $C_i^m(r)$ as the mean of these $C_i^m(r)$ values. The quantity $C_i^m(r)$ expresses the prevalence of repetitive patterns of length m in N . Finally, we define the approximate entropy of N , for patterns of length m and similarity criterion r , as

$$ApEn(N, m, r) = \ln \left[\frac{C_i^m(r)}{C_i^{m+1}(r)} \right] \quad (4.15)$$

i.e., as the natural logarithm of the relative prevalence of repetitive patterns of length m compared with those of length $m+1$.

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

In this project, we have estimated ApEn using the command of java program. For this program, we have taken instantaneous heart rate (IHR) values which are taken from MIT-BIH Arrhythmia Record. We use java vector class for intermediate data storage. However, after storing pertinent data, we constructed matrix of length 2 by taking two values from the data storage which are represented as $RR_1, RR_2, RR_3, RR_4, RR_5, RR_6, \dots, RR_N$. According to the algorithm of the ApEn we derived different values of approximate entropy. The flow chart for this algorithm is shown in fig (4.1). The comparison between two matrixes follows a threshold value which varies from 10% to 90% of the standard deviation for the data taken as input from IHR values. The comparison follows the following formula,

$$|\text{vector}(i+k) - \text{vector}(j+k)| < \text{Per} * \text{standard deviation for } 0 \leq k < m$$

Where per varies from 0.1 to 0.9 and vector is the object of Vector class in which we stored IHR values for our program. The value of m represents the pattern length and the right side of the above inequality means the criterion of similarity.

In our program, to calculate ApEn we evaluate the mean of the fraction of patterns of length m that resemble the pattern of the same length whose value depends on the distance between two vector. We attribute eight data set of healthy people as input and eight data set of abnormal patient and after analyzing by our program, we come across the discernible differences in output between healthy people and the people with abnormal rhythm for the value of the criterion of similarity which is greater than 0.2.

In our java program we use the following conspicuous classes

1. JFileChooser
2. File
3. BufferedReader
4. Vector
5. Math
6. StringToke

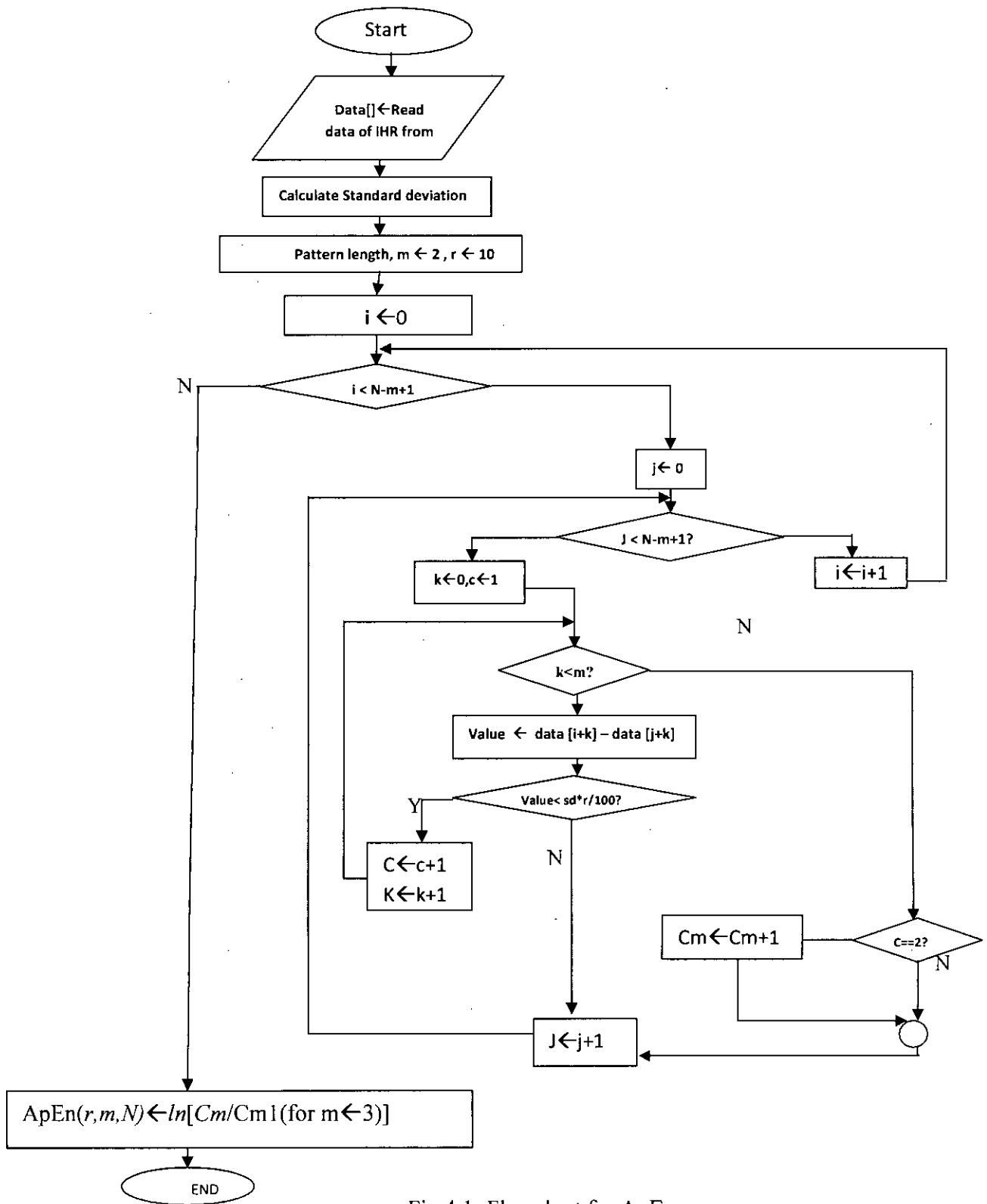


Fig 4.1: Flowchart for ApEn

4.5 Sample Entropy:

Entropy is related to dynamical systems which is the rate of information production. The algorithm for calculating ApEn and SampEn is almost similar. Pincus ascribe approximate entropy (ApEn) as a set of measures of system complexity which is closely associated to entropy and it is frequently applied to clinical cardiovascular and other time series analysis of sundry types of patients. It should be notified that ApEn statistics produce incompatible results. For this, We have developed a new and related complexity measure, sample entropy (SampEn), and have compared ApEn and SampEn by using them to analyze sets of random data with known probabilistic character. The most important thing is that ApEn counts each sequence as matching itself. We can attribute this self bias as the source of inconsistent result of Apen. In contrast, SampEn abates this self matching. The SampEn statistics is free of the bias grounded by self-matching. The name refers to the applicability to time series data which is sampled from a continuous process. In addition, the algorithm advocates ways to make use of sample statistics to appraise the results. However, we notice better result using sampan statistics. Quantification of the irregularity and difficulty of the heart rate using sample entropy [12] are increasingly used because they can be computed from shorter HRV records that are used in community screening. The conspicuous difference of sampan is shown in figure.... The SampEn is calculated by the following equation,

$$SampEn (N, m, r) = \ln \left[\frac{\sum_{i=1}^{N-m+1} C_i^m(r)}{\sum_{i=1}^{N-m} C_i^{m+1}(r)} \right] \quad (4.16)$$

Where m is the pattern length and $C_i^m(r)$ is the mean fraction of pattern length.

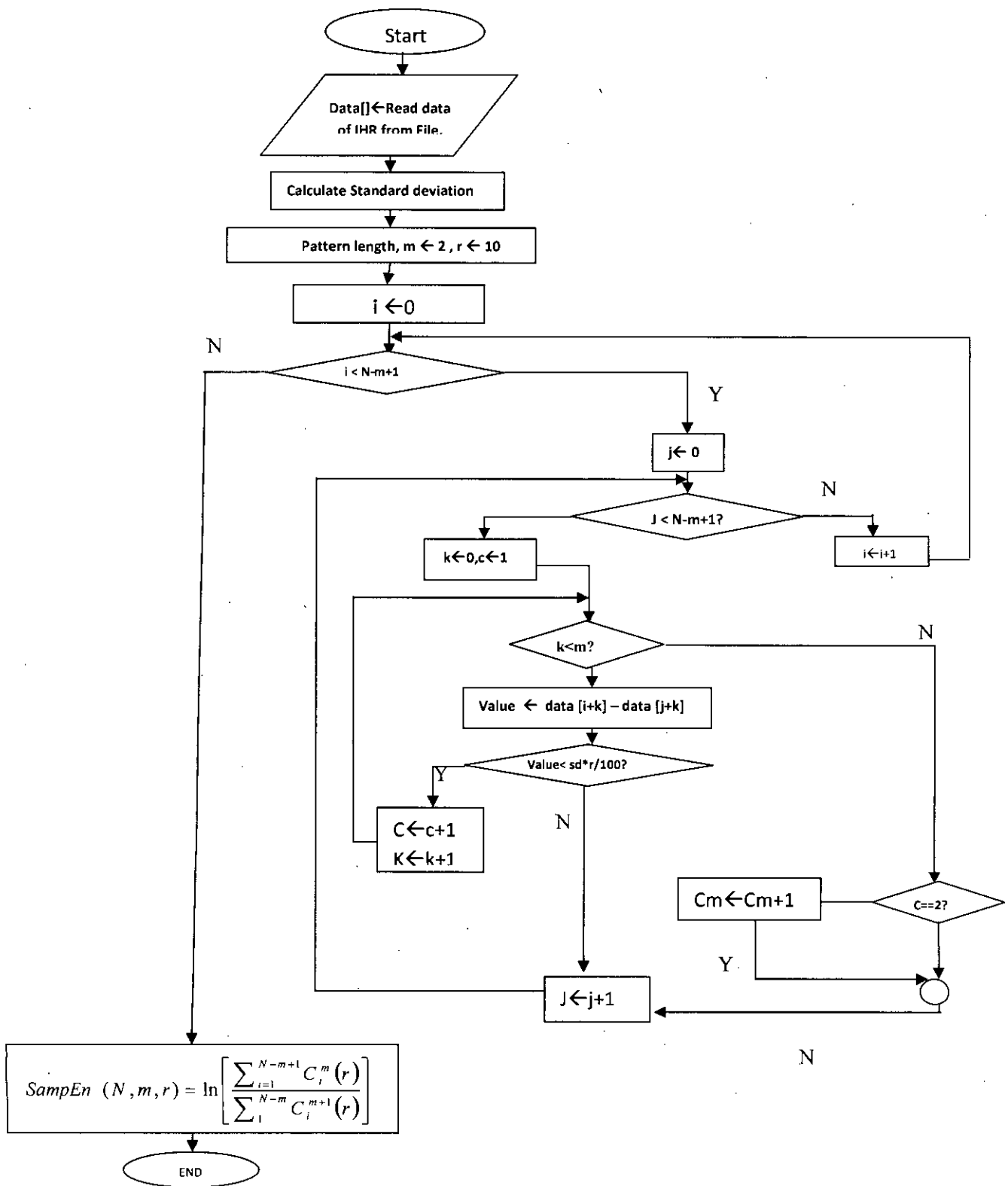


Fig 4.2: Flowchart for SampEn Analysis

Conclusion:

In this chapter some robust statistical tools decision making algorithm are designed in order to clarify normal and abnormal rhythm. All equations and Matlab command for time domain analysis, Poincare plot analysis and entropy analysis are discussed in this chapter.

Chapter 5

EXPERIMENTAL RESULTS

5.1 Introduction

To evaluate the act of the proposed method, we used the MIT-BIH Arrhythmia Database directory. The ECG signals in this directory were sampled at 360 Hz and with a quantization resolution of 11bits/sample. From these binary records, the QRS complex is determined using the algorithm described on APPENDIX-A. From these QRS complex locations, we determined IHR. From these IHR, the time domain parameter (like mean, variance, SD, SDSD, RMSSD), the Poincaré plot indexes (like SD1,SD2, SD1/SD2, area of ellipse), sample entropy(SampEn), approximate entropy(ApEn) are determined for healthy persons as well as ailing persons with abnormal rhythm like Atrial fibrillation(AFIB), Ventricular bigeminy(B), Ventricular trigeminy(T) and the results are compared.

Table 5.1: Beat annotation appear in the records

Symbol	Meaning
AB	Atrial bigeminy
AFIB	Atrial fibrillation
AFL	Atrial flutter
B	Ventricular bigeminy
BII	2° heart block
IVR	Idioventricular rhythm
N	Normal sinus rhythm
NOD	Nodal (A-V junctional) rhythm
P	Paced rhythm
SBR	Sinus bradycardia
SVTA	Supraventricular tachyarrhythmia
T	Ventricular trigeminy

5.2 Poincaré plot analysis

Sixteen ECG recordings were analyzed. Table 8 and table 9 summarizes the results from Poincaré indexes of the two groups.. Here in this project, Poincaré plot analysis has been used on IHR time series and the results are analyzed to see if any significant difference is found between normal and abnormal data series. Following are the portraits obtained using Poincaré plot analysis on IHR.

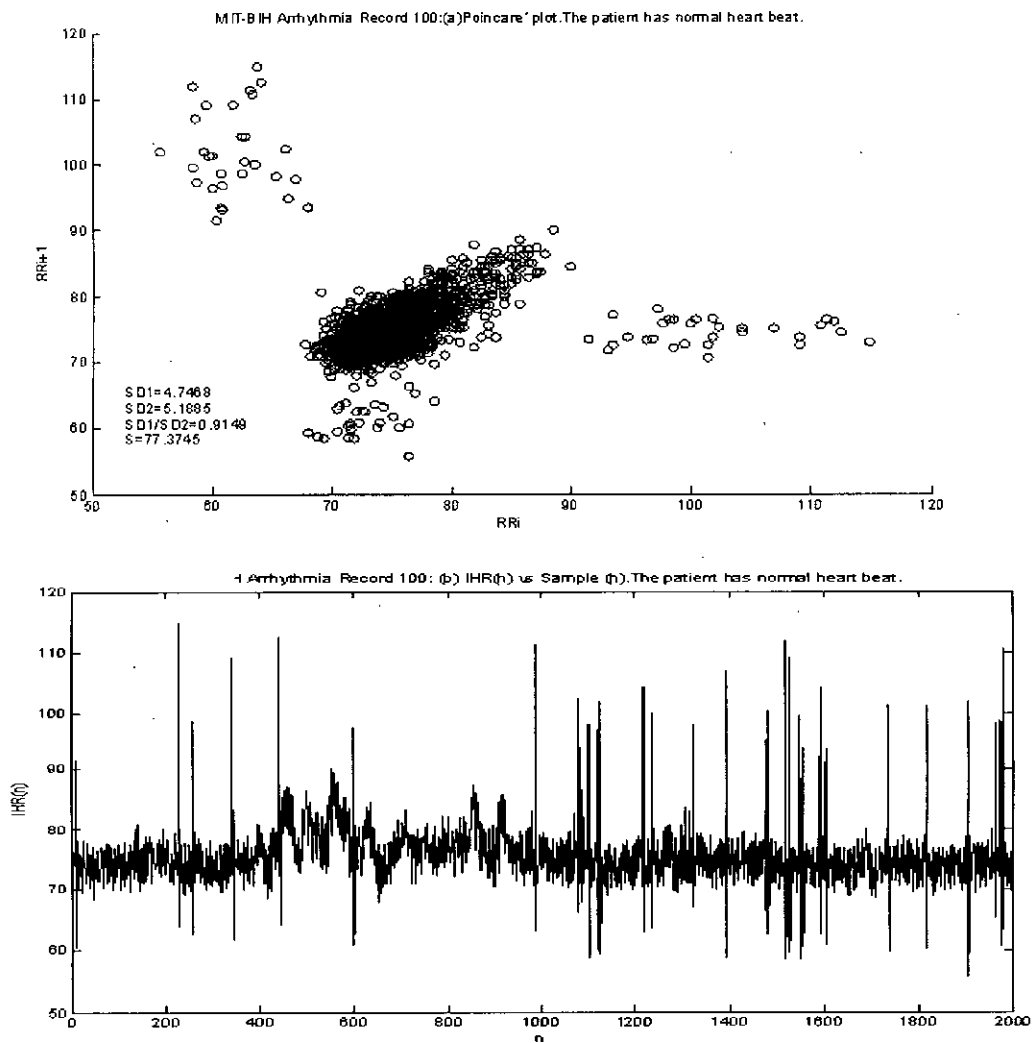


Fig 5.1: MIT-BIH Arrhythmia Record_100: (a) Poincaré plot (RR_i vs. RR_{i+1}); (b) IHR (n) vs. Sample (n). The patient has normal heart beat.

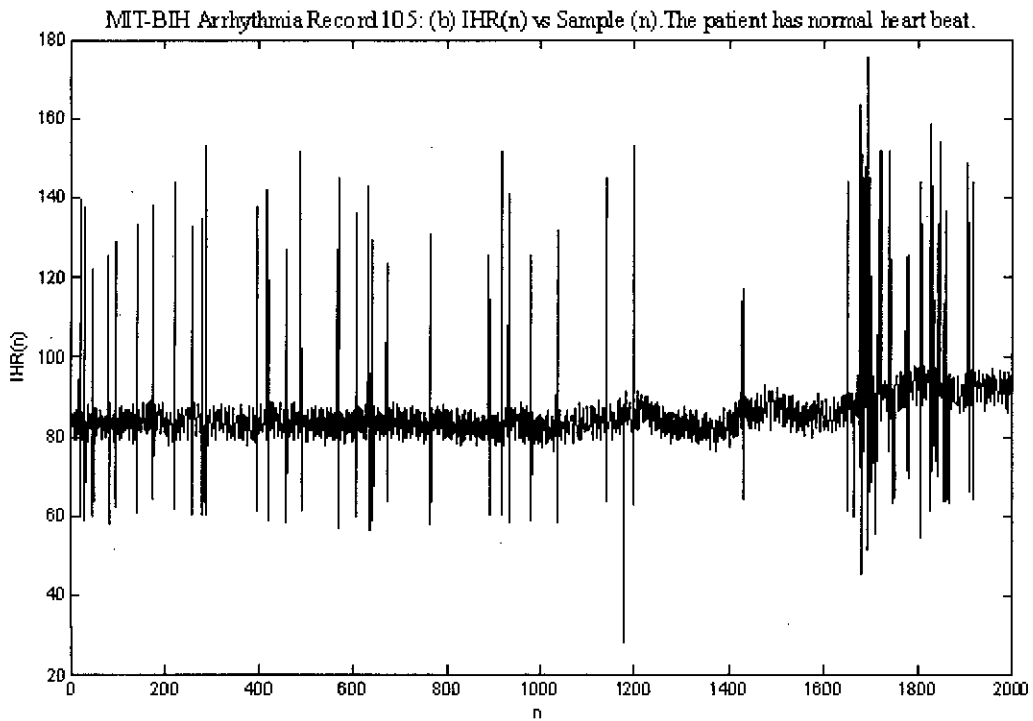
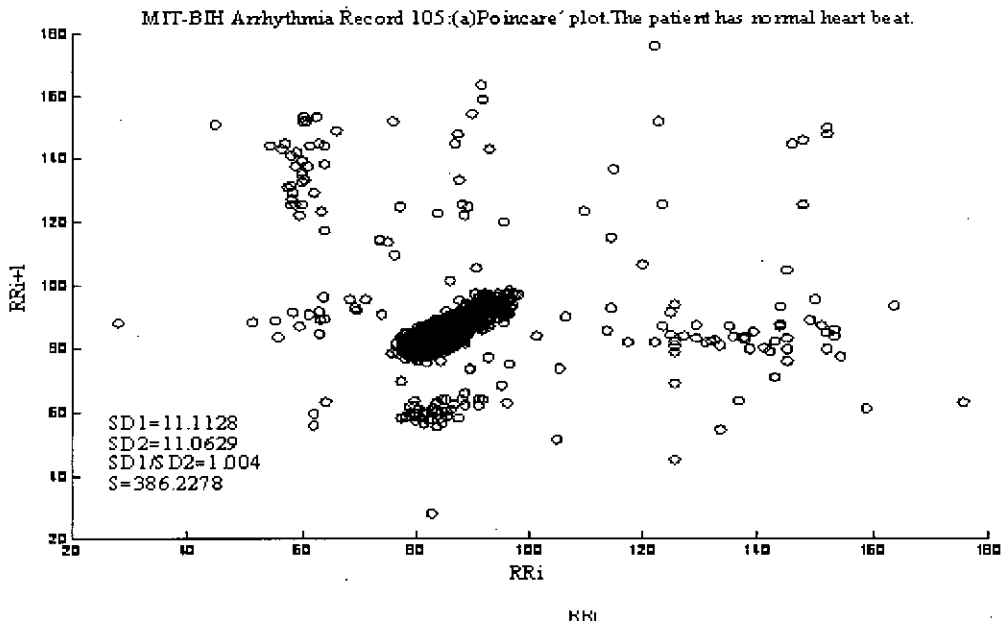


Fig 5.2: MIT-BIH Arrhythmia Record_105: (a) Poincare plot (RR_i vs. RR_{i+1}); (b) IHR (n) vs. Sample (n). The patient has normal heart beat.

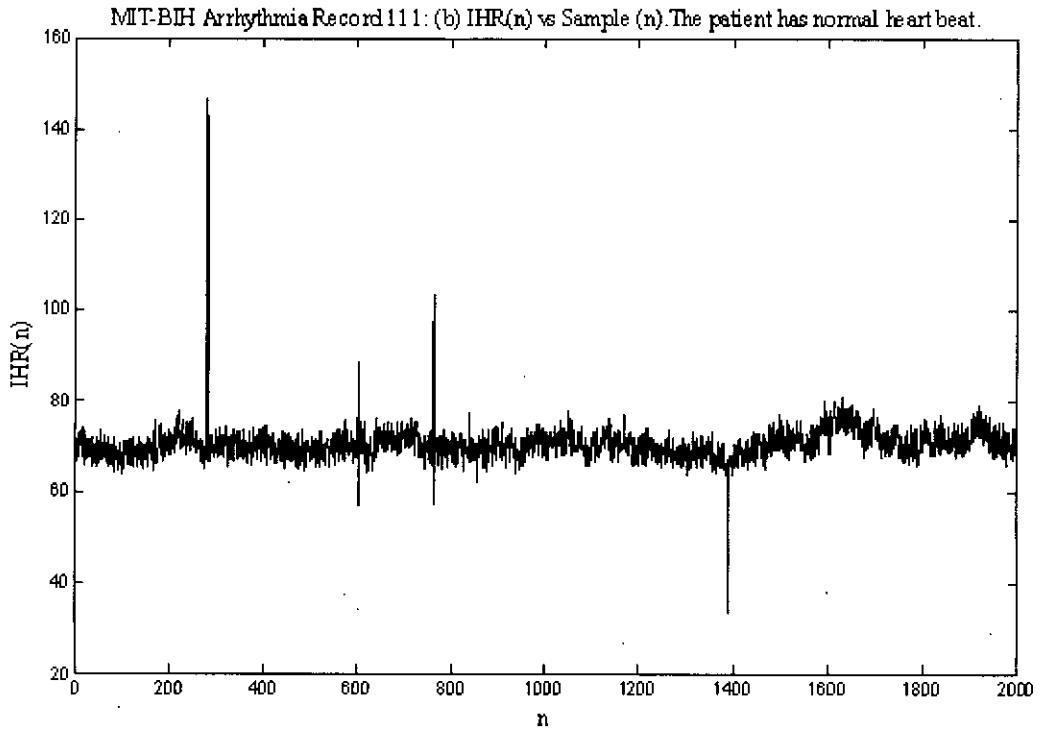
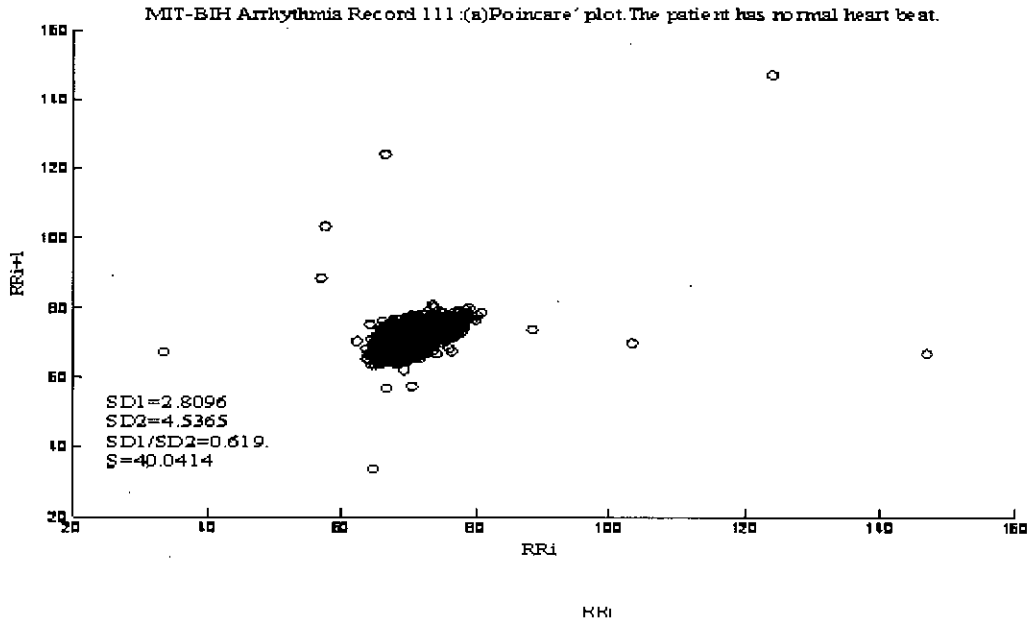


Fig 5.3: MIT-BIH Arrhythmia Record_111: (a) Poincare plot (RR_i vs. RR_{i+1}); (b) IHR (n) vs. Sample (n). The patient has normal heart beat.

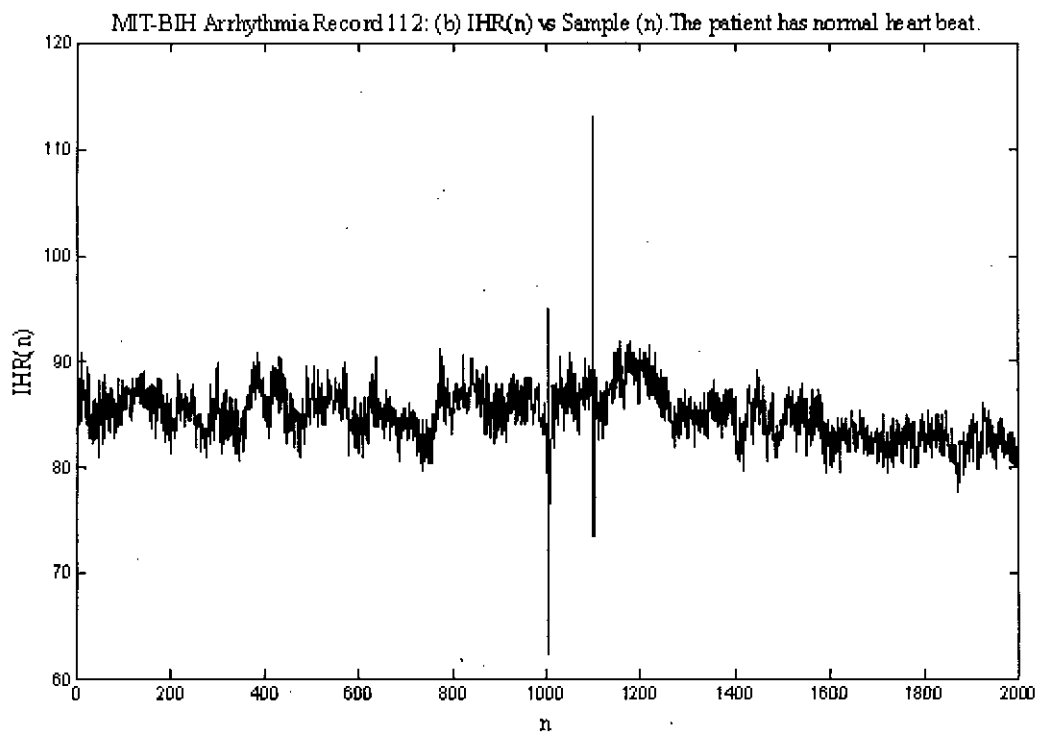
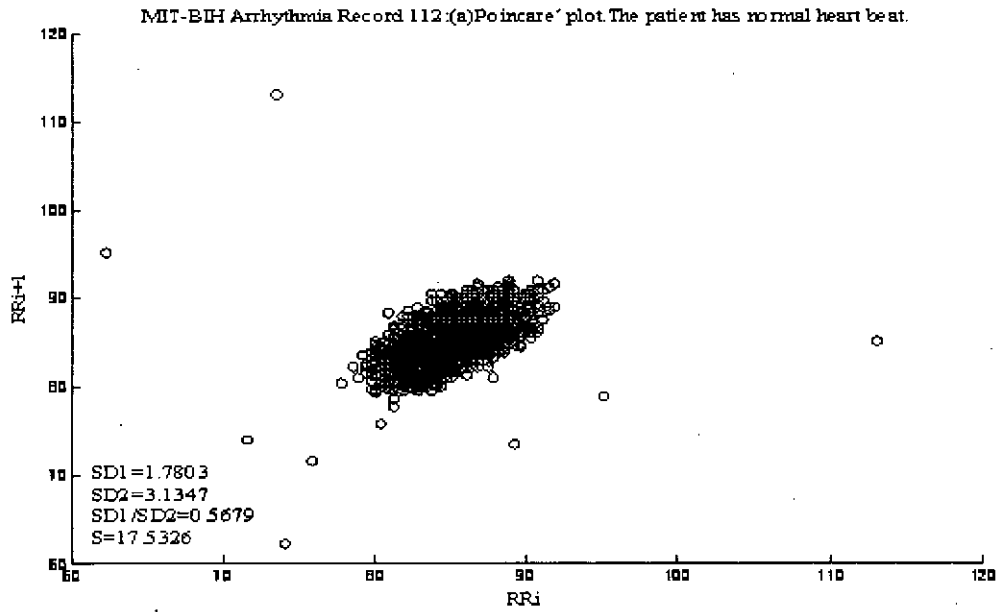


Fig 5.4: MIT-BIH Arrhythmia Record_112: (a) Poincare plot (RR_i vs. RR_{i+1}); (b) IHR (n) vs. Sample (n). The patient has normal heart beat.

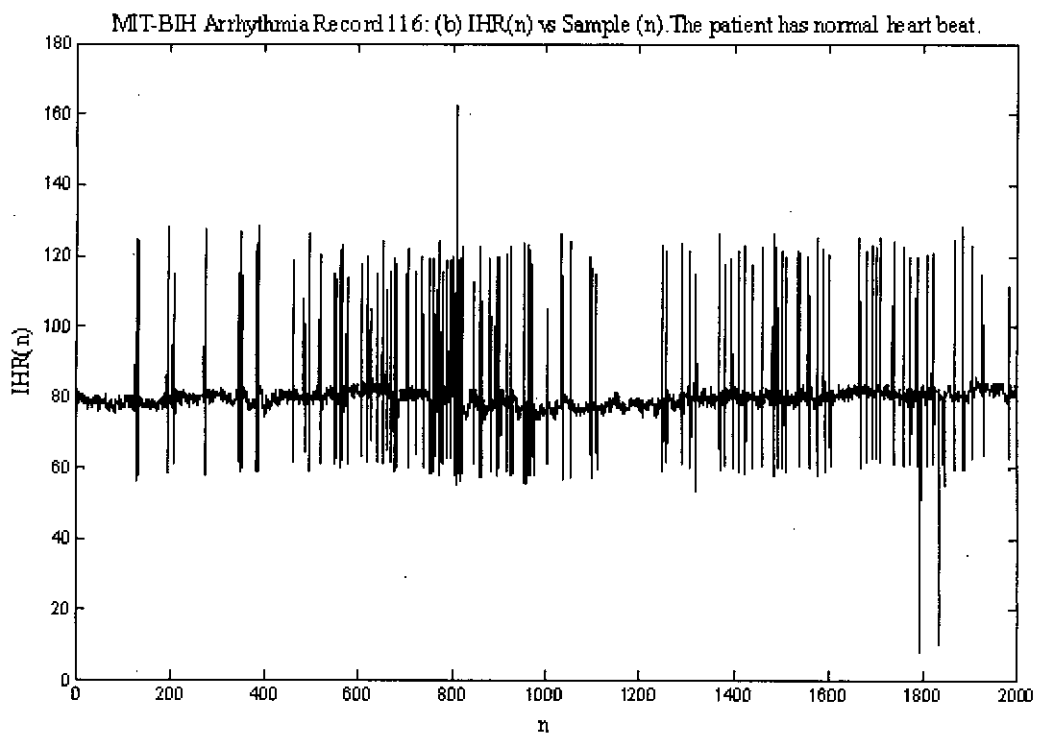
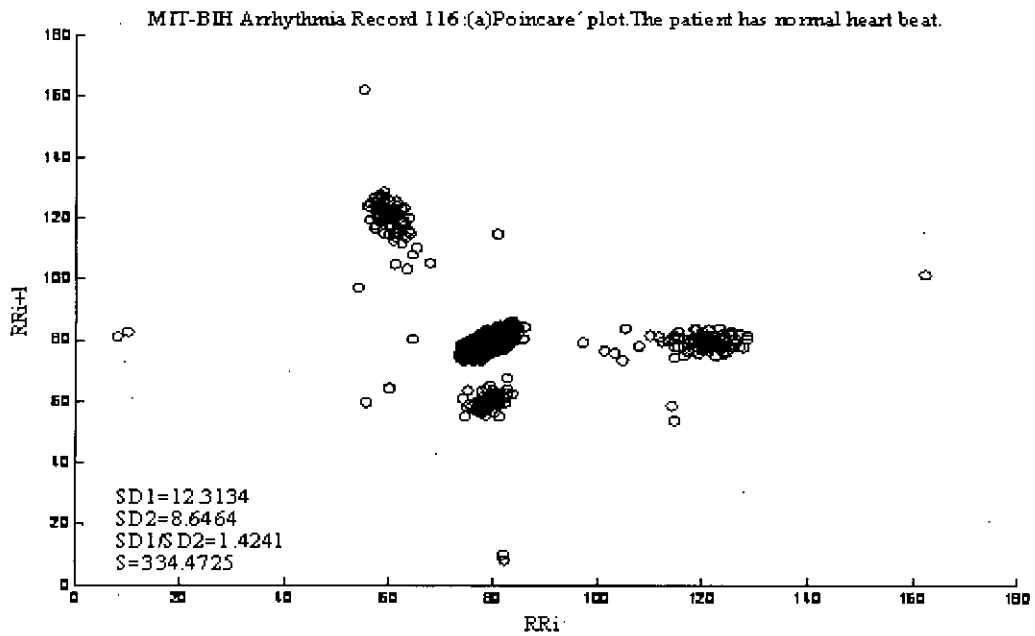


Fig 5.5: MIT-BIH Arrhythmia Record_116: (a) Poincare plot (RR_i vs. RR_{i+1}); (b) IHR (n) vs. Sample (n). The patient has normal heart beat.

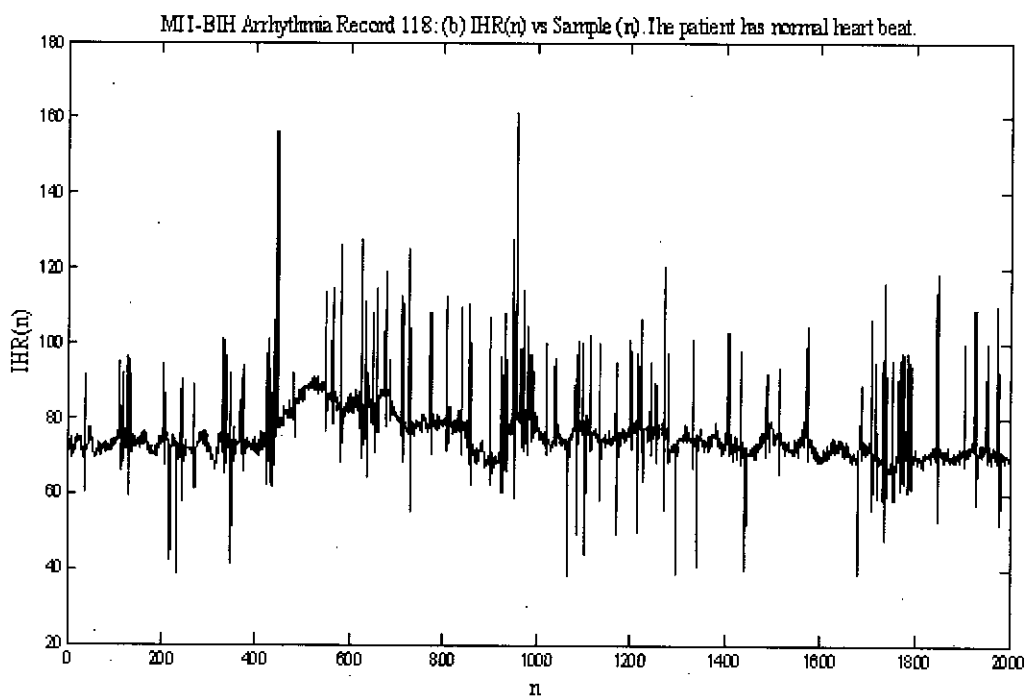
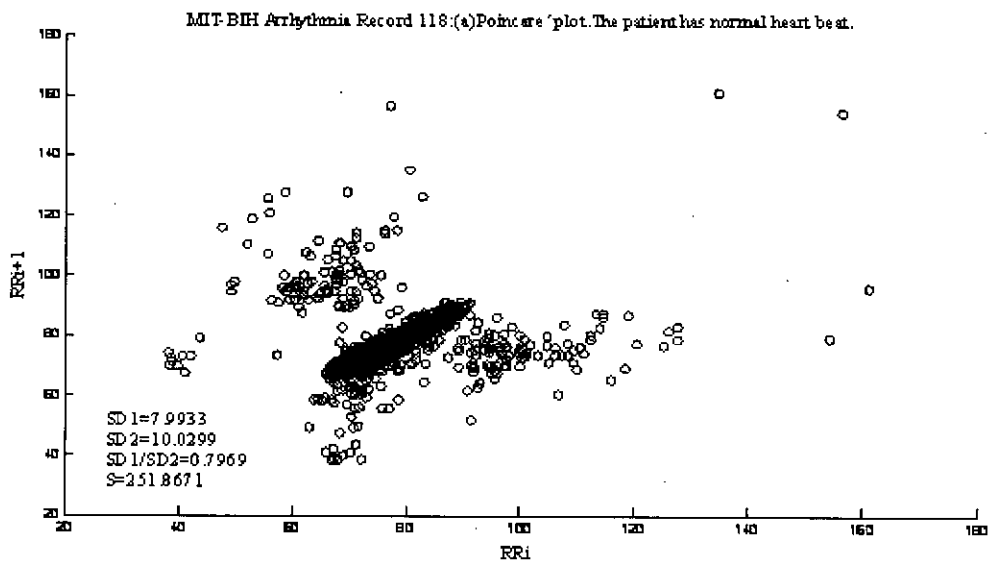


Fig 5.6: MIT-BIH Arrhythmia Record_118: (a) Poincare plot (RR_i vs. RR_{i+1}); (b) IHR (n) vs. Sample (n). The patient has normal heart beat.

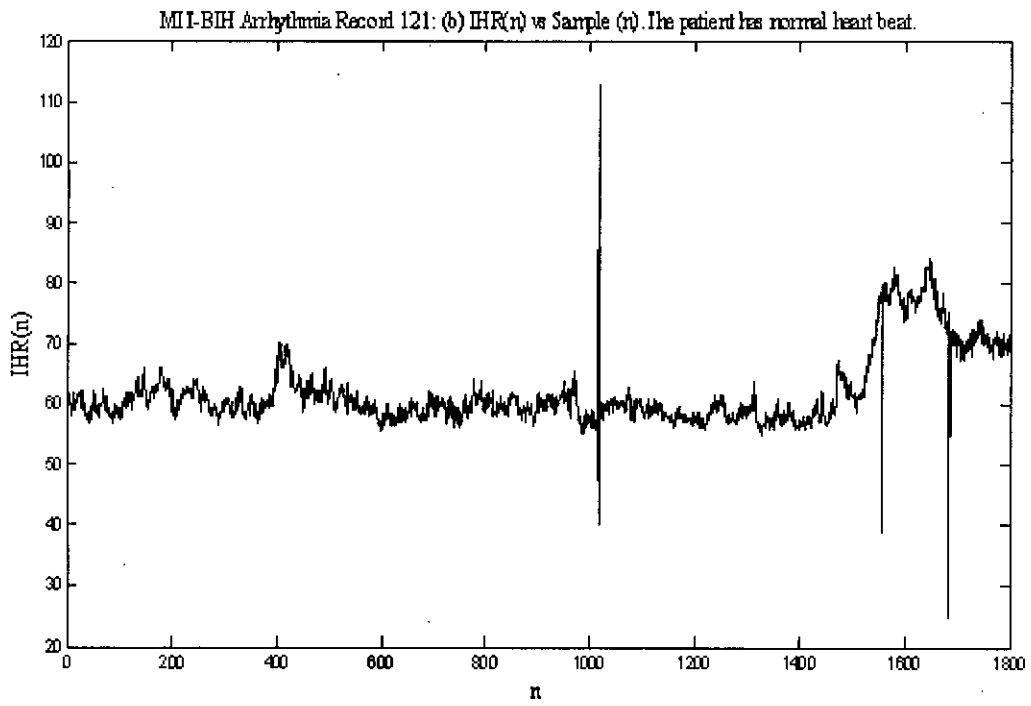
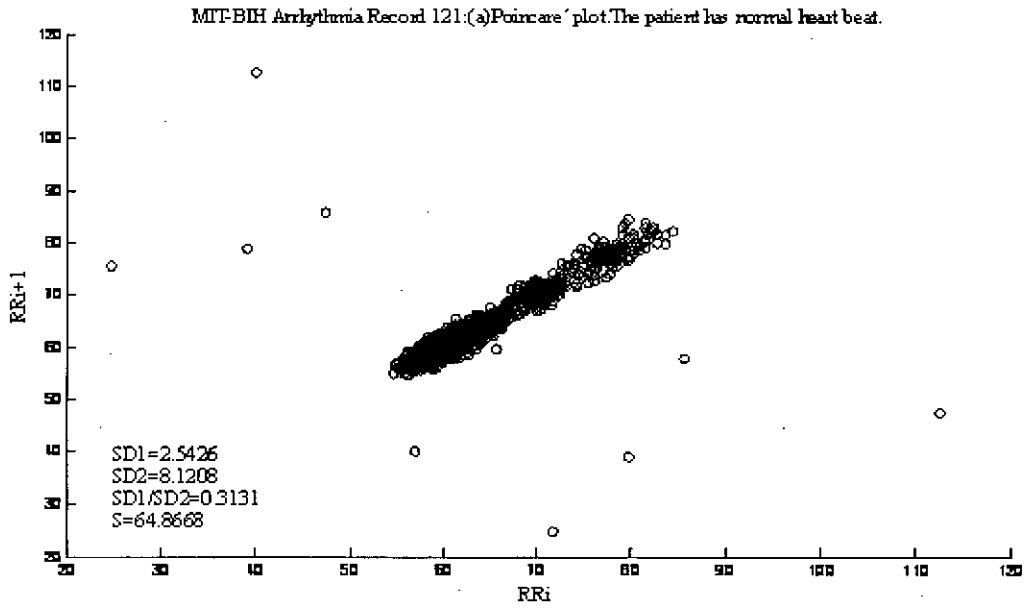


Fig 5.7: MIT-BIH Arrhythmia Record_121: (a) Poincare plot (RR_i vs. RR_{i+1}); (b) IHR (n) vs. Sample (n). The patient has normal heart beat.

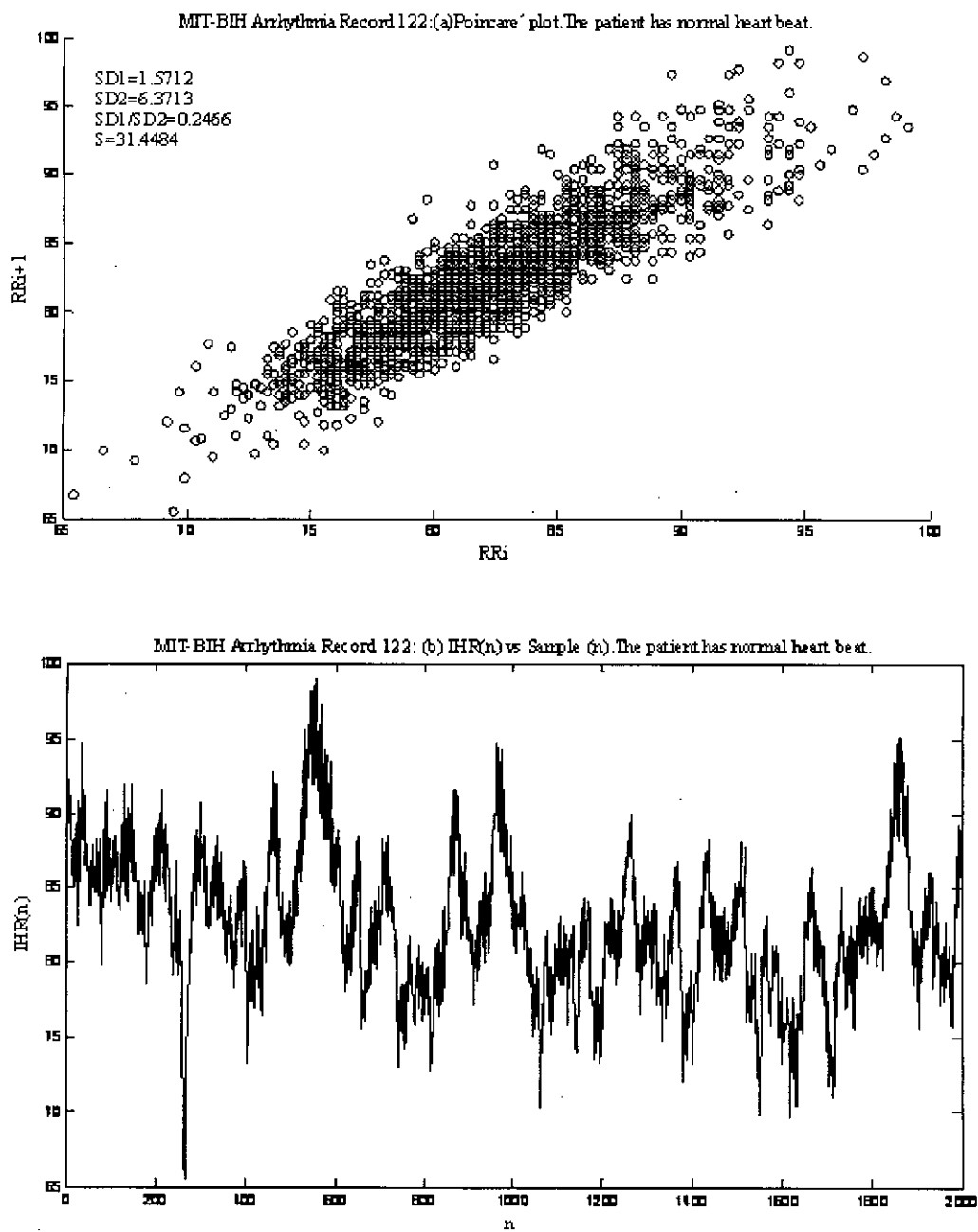


Fig 5.8: MIT-BIH Arrhythmia Record_122: (a) Poincare plot (RR_i vs. RR_{i+1}); (b) IHR (n) vs. Sample (n). The patient has normal heart beat.

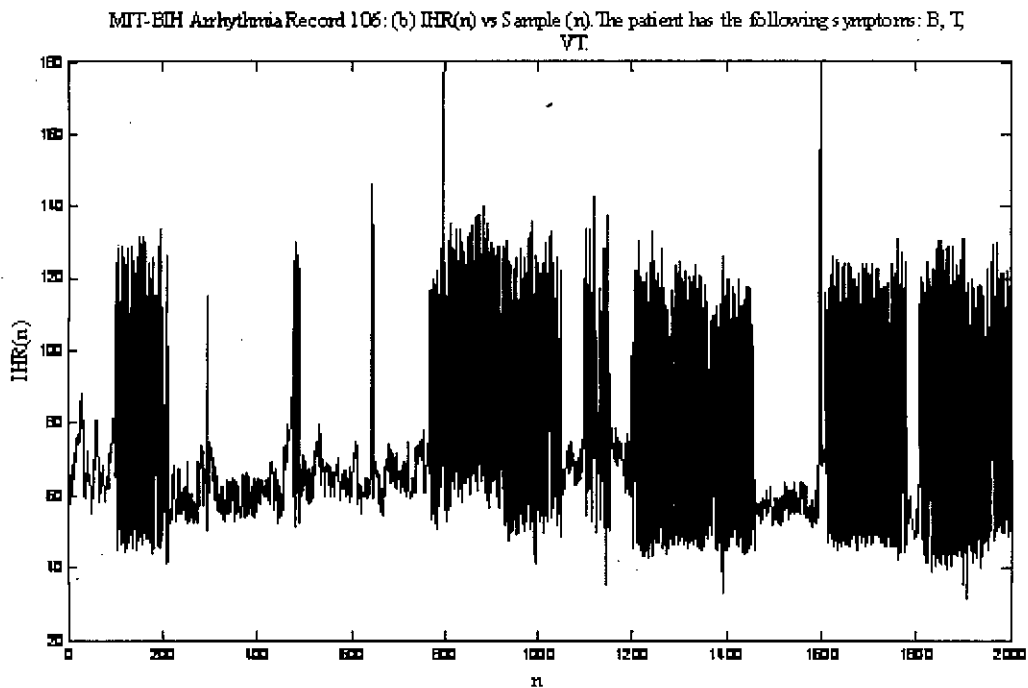
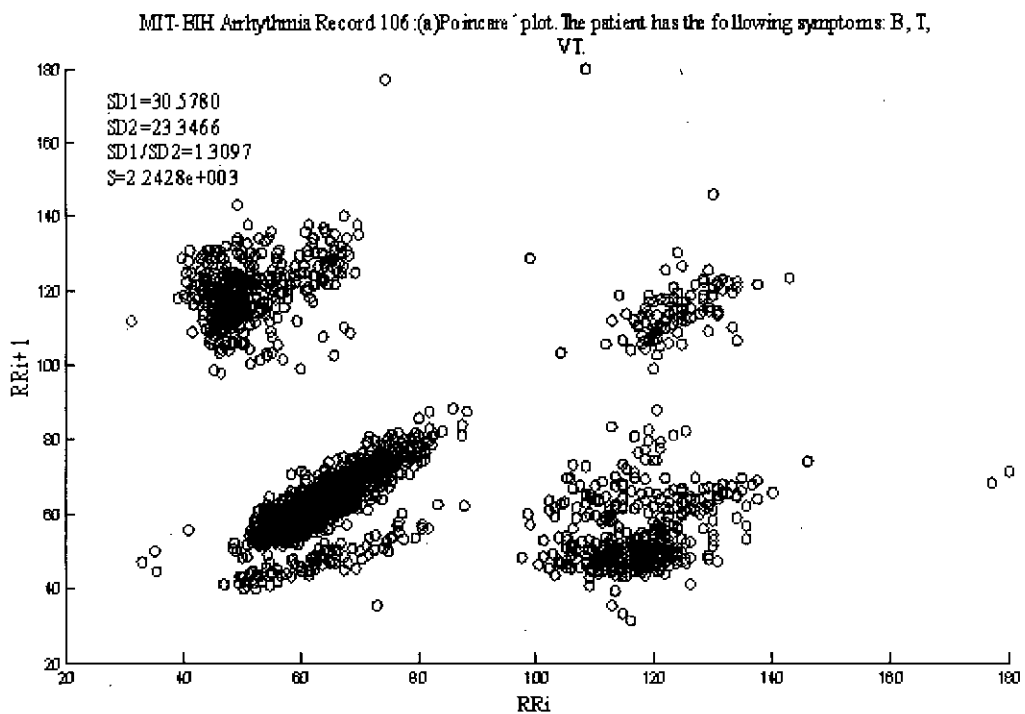


Fig 5.09: MIT-BIH Arrhythmia Record_106: (a) Poincare plot (RR_i vs. RR_{i+1}); (b) IHR (n) vs. Sample (n). The patient has following symptoms: T, B.

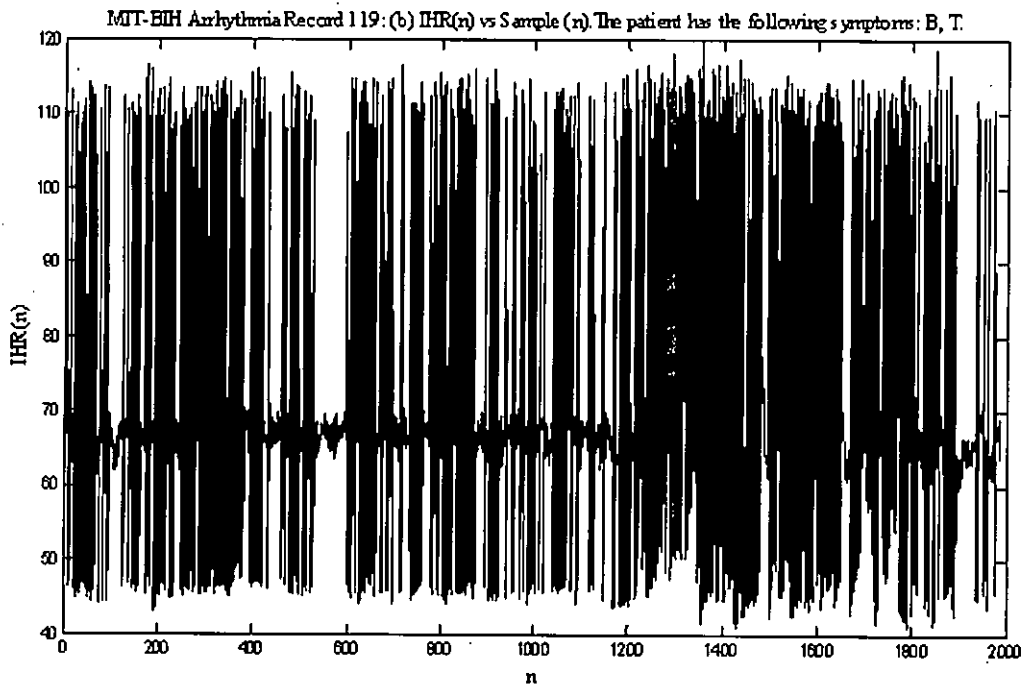
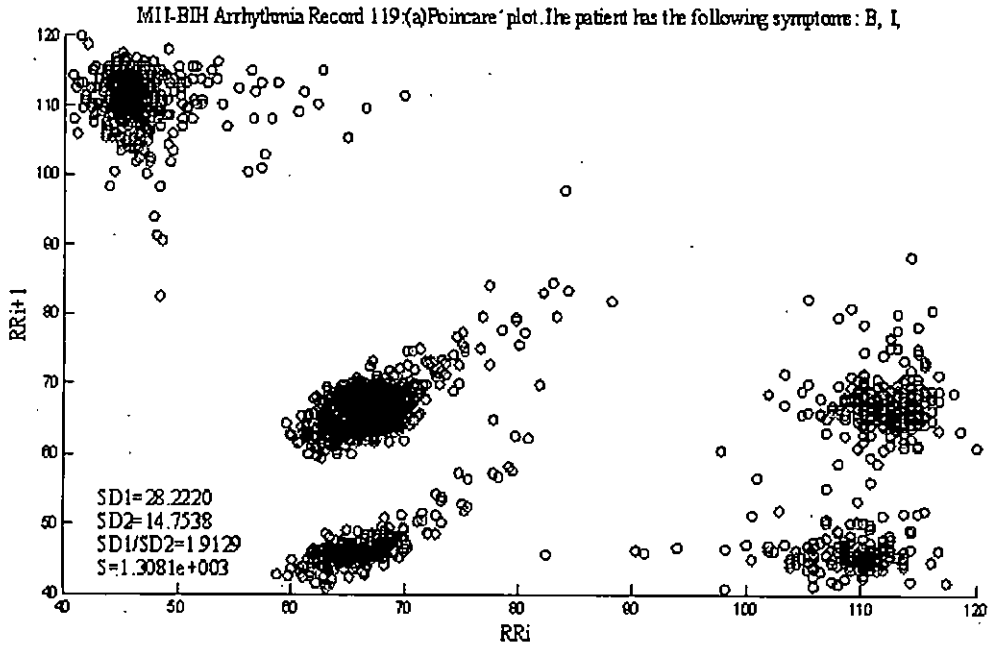
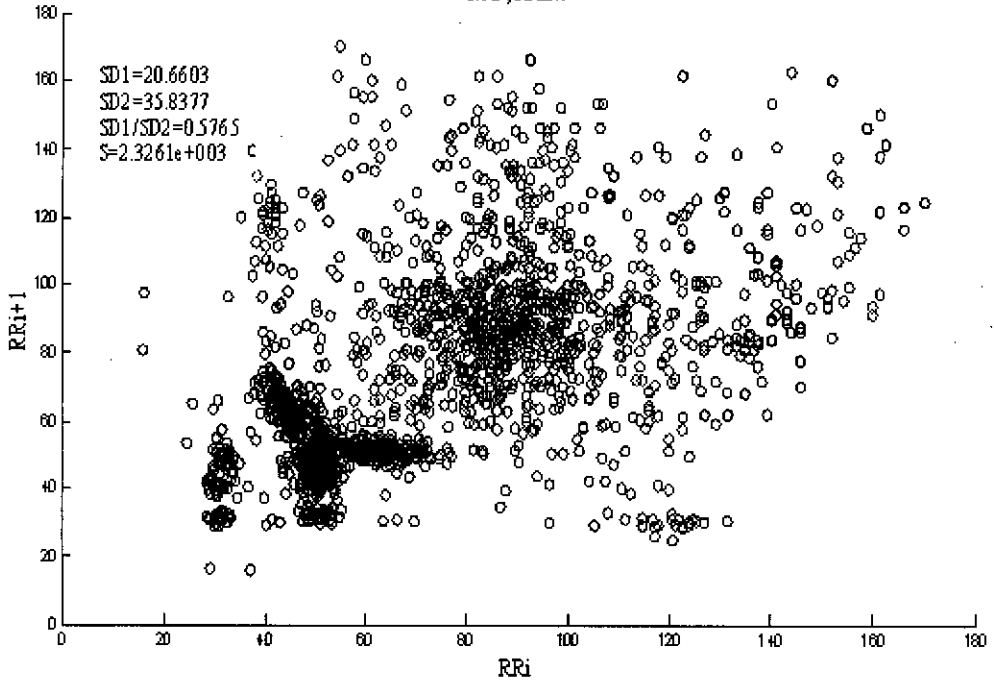


Fig 5.10: MIT-BIH Arrhythmia Record_119: (a) Poincare plot (RR_i vs. RR_{i+1}); (b) IHR (n) vs. Sample (n). The patient has following symptoms: T, B.

MIT-BIH Arrhythmia Record 201: (b) IHR(n) vs Sample (n). The patient has the following symptoms: T, SVTA, NOD, AFIB.



MIT-BIH Arrhythmia Record 201: (b) IHR(n) vs Sample (n). The patient has the following symptoms: T, SVTA, NOD, AFIB.

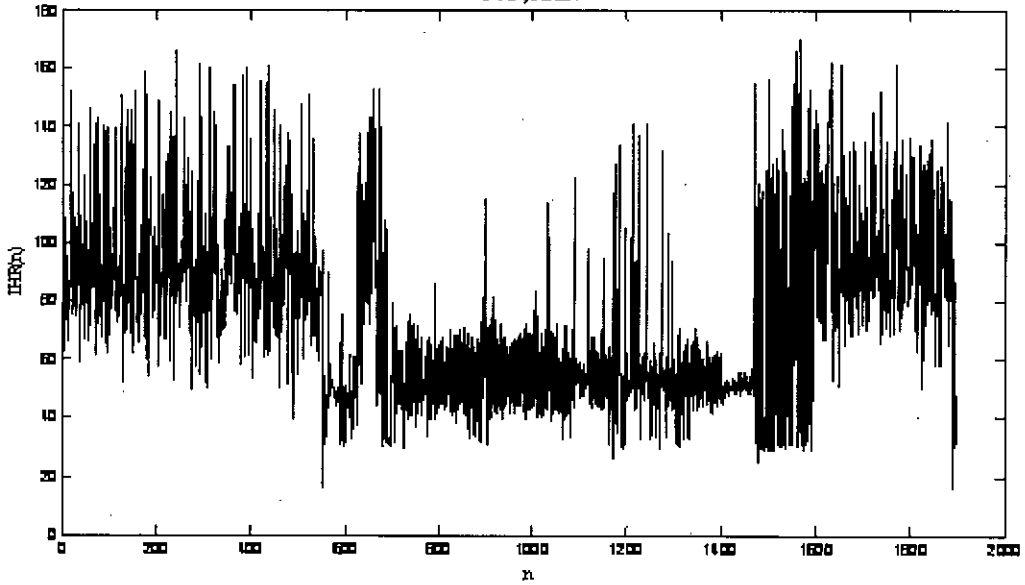


Fig 5.11: MIT-BIH Arrhythmia Record_201: (a) Poincare plot (RR_i vs. RR_{i+1}); (b) IHR (n) vs. Sample (n). The patient has following symptoms: T, SVTA, AFIB, and NOD.

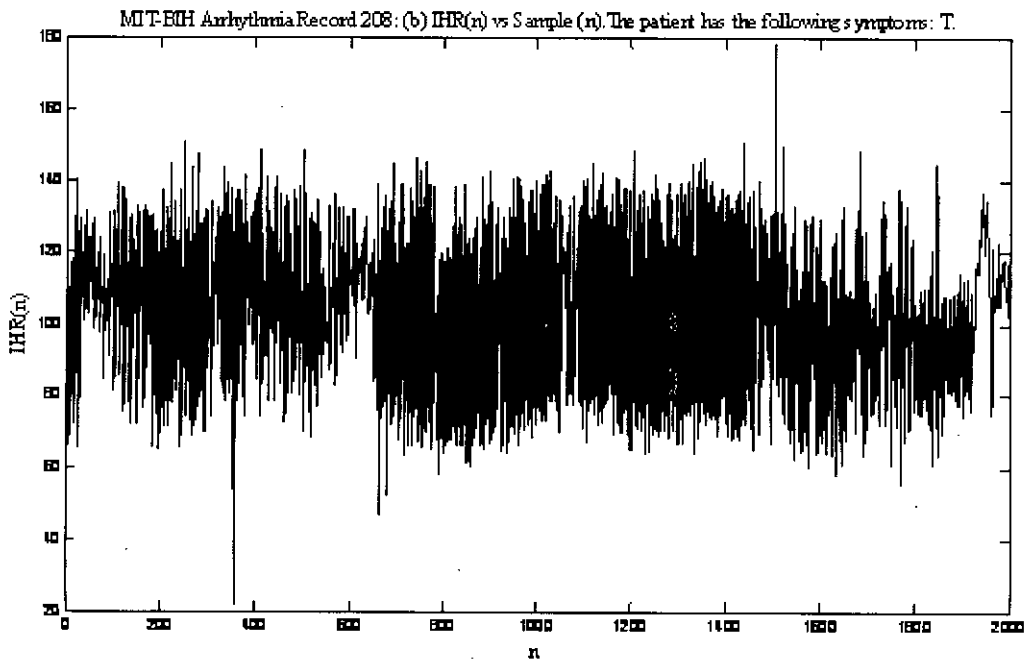
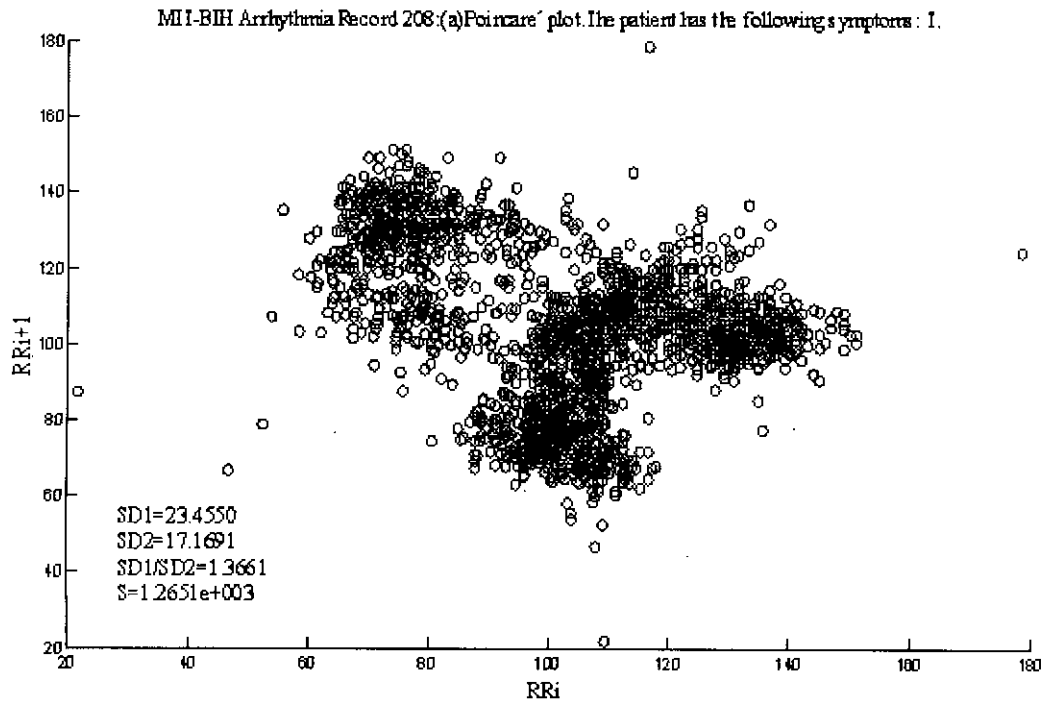


Fig 5.12: MIT-BIH Arrhythmia Record_208: (a) Poincare plot (RR_i vs. RR_{i+1}); (b) IHR (n) vs. Sample (n). The patient has following symptoms: T

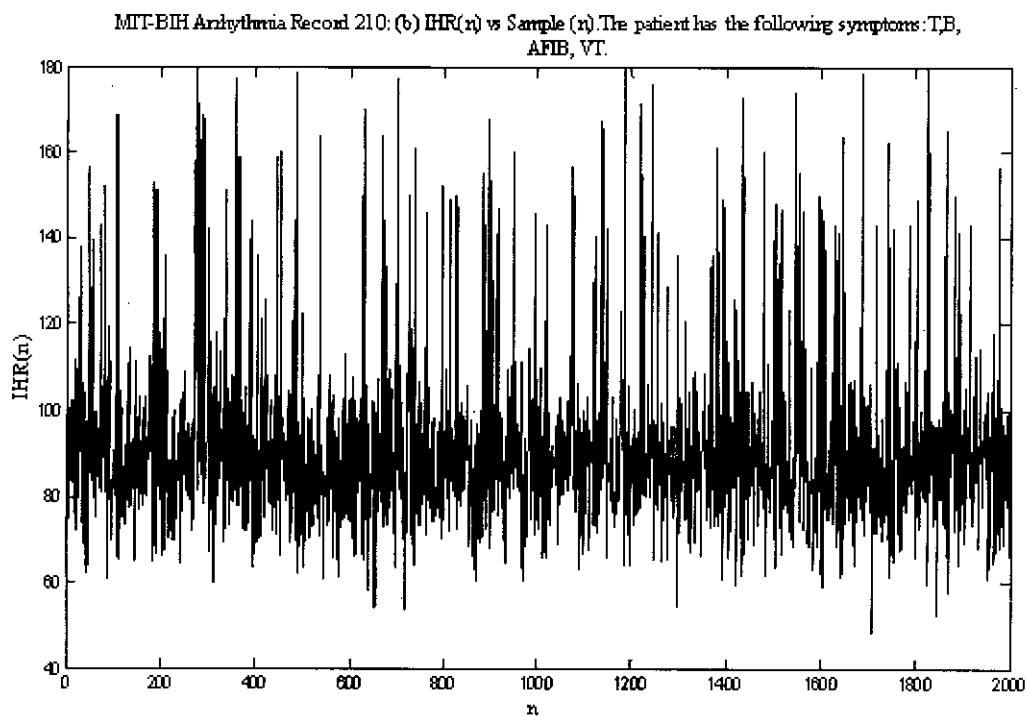
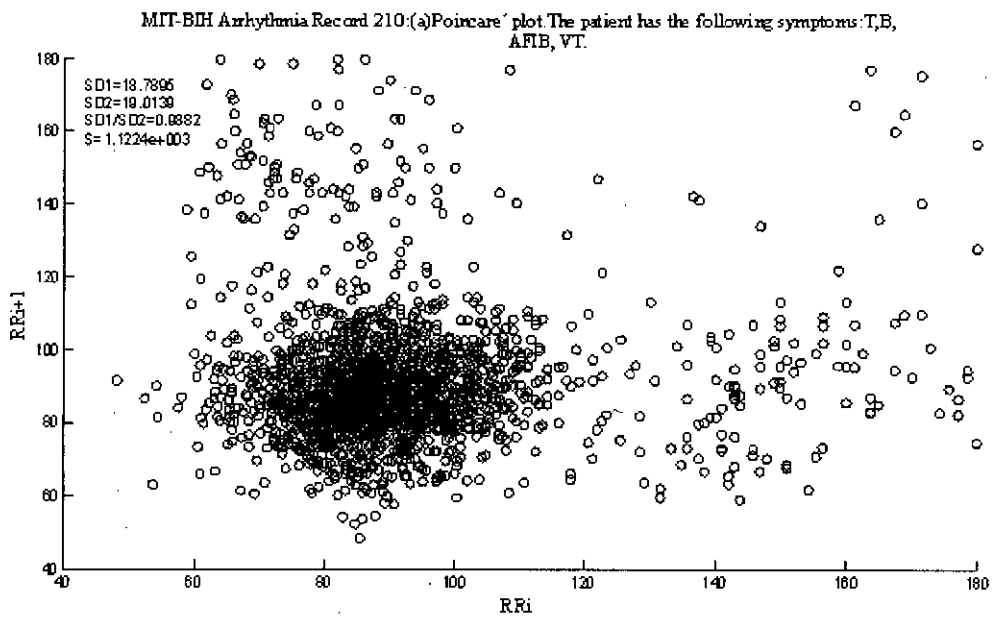


Fig 5.13: MIT-BIH Arrhythmia Record_210: (a) Poincare plot (RR_i vs. RR_{i+1}); (b) IHR (n) vs. Sample (n). The patient has following symptoms: T, B, AFIB, VT.

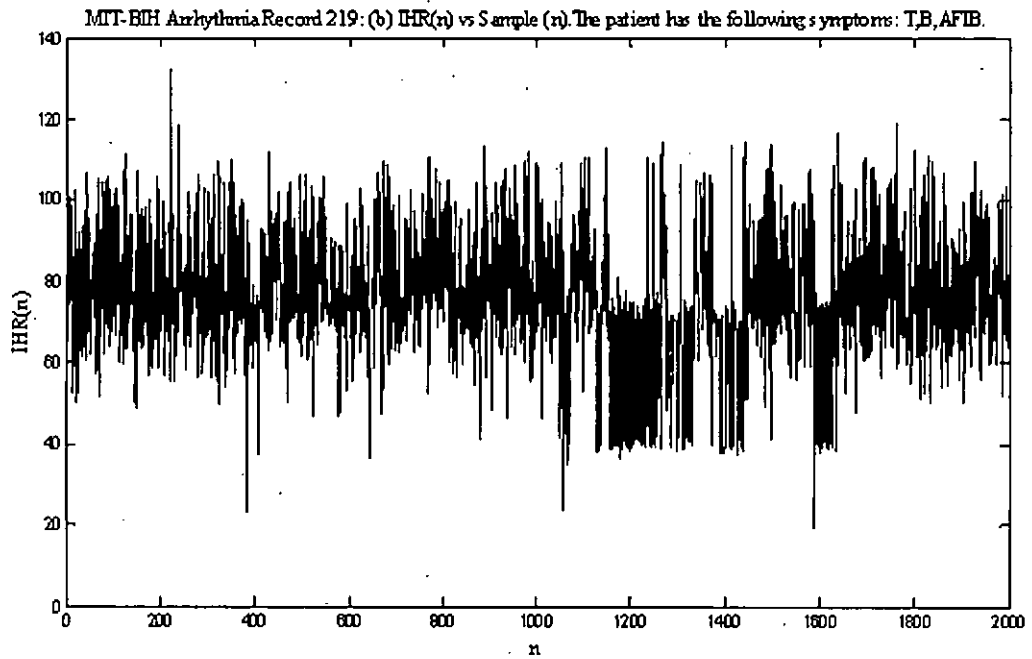
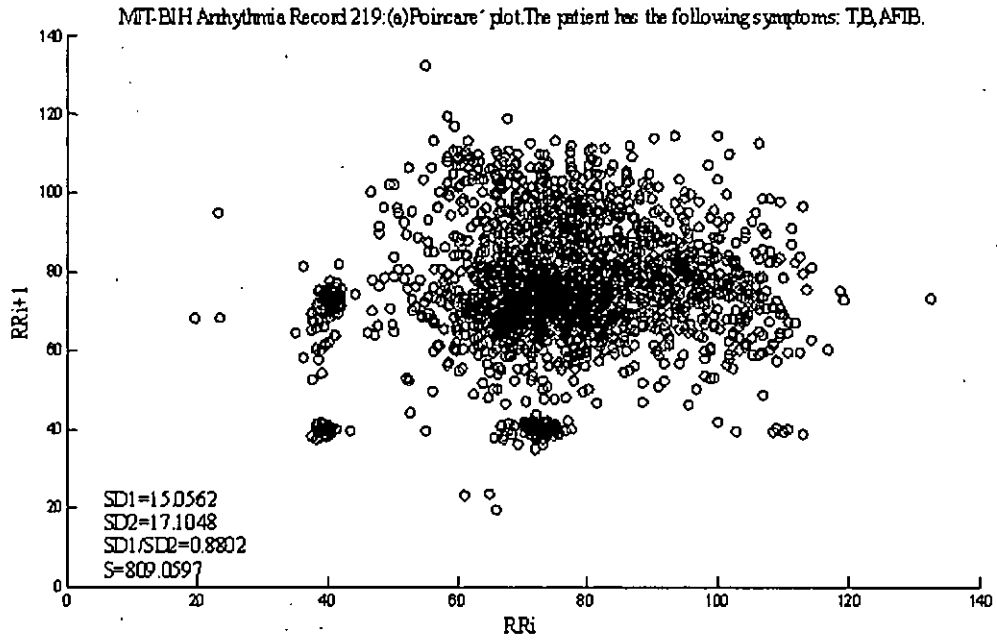


Fig 5.14: MIT-BIH Arrhythmia Record_219: (a) Poincare plot (RR_i vs. RR_{i+1}); (b) IHR (n) vs. Sample (n). The patient has following symptoms: T, B, and AFIB.

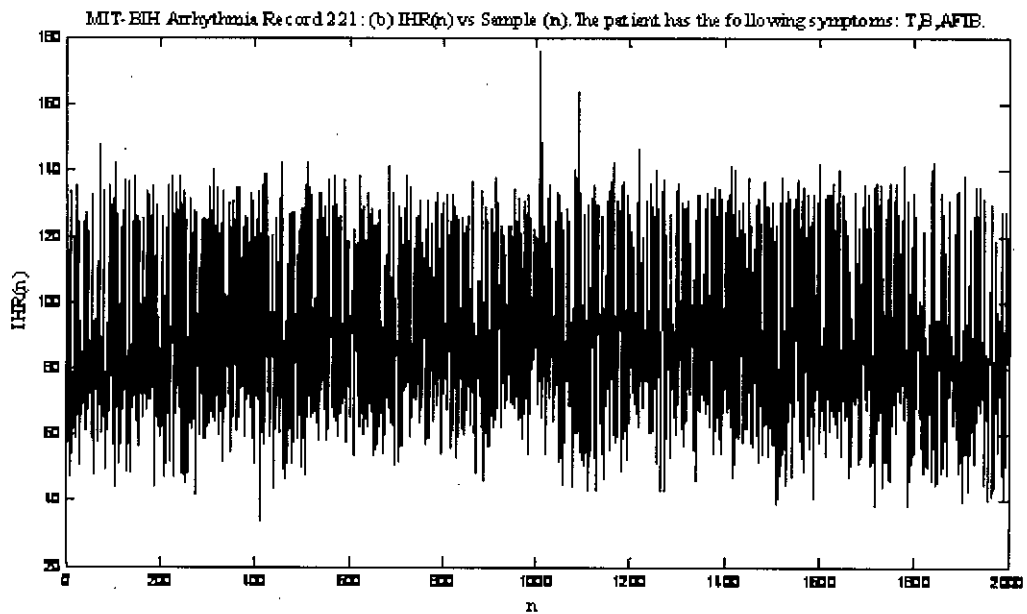
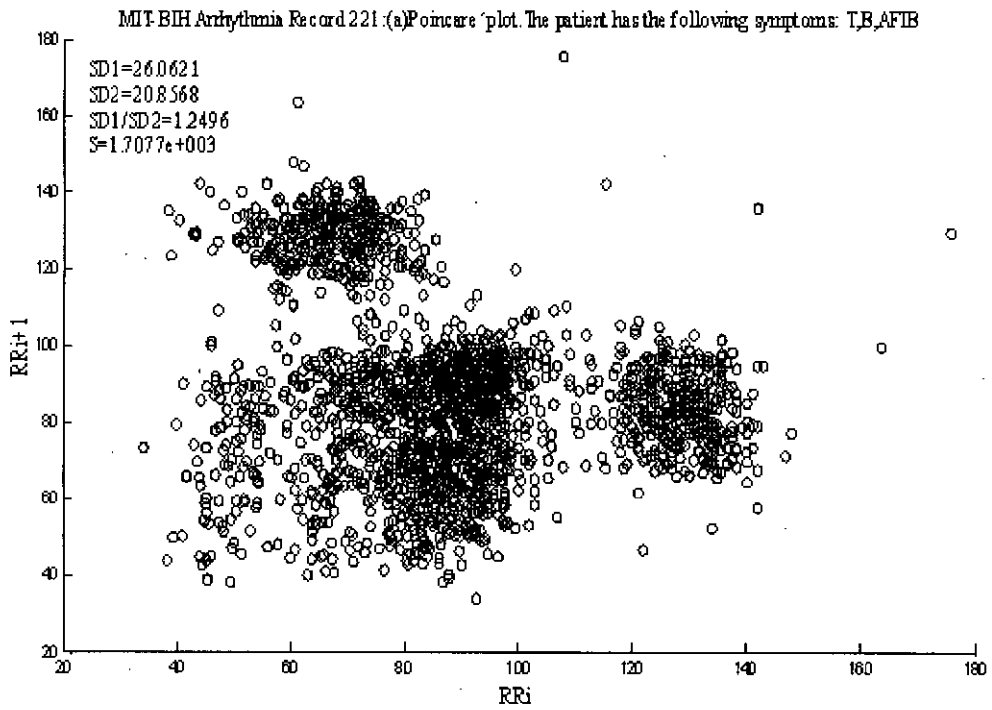


Fig 5.15: MIT-BIH Arrhythmia Record_221: (a) Poincare plot (RR_i vs. RR_{i+1}); (b) IHR (n) vs. Sample (n). The patient has following symptoms: T, B and AFIB

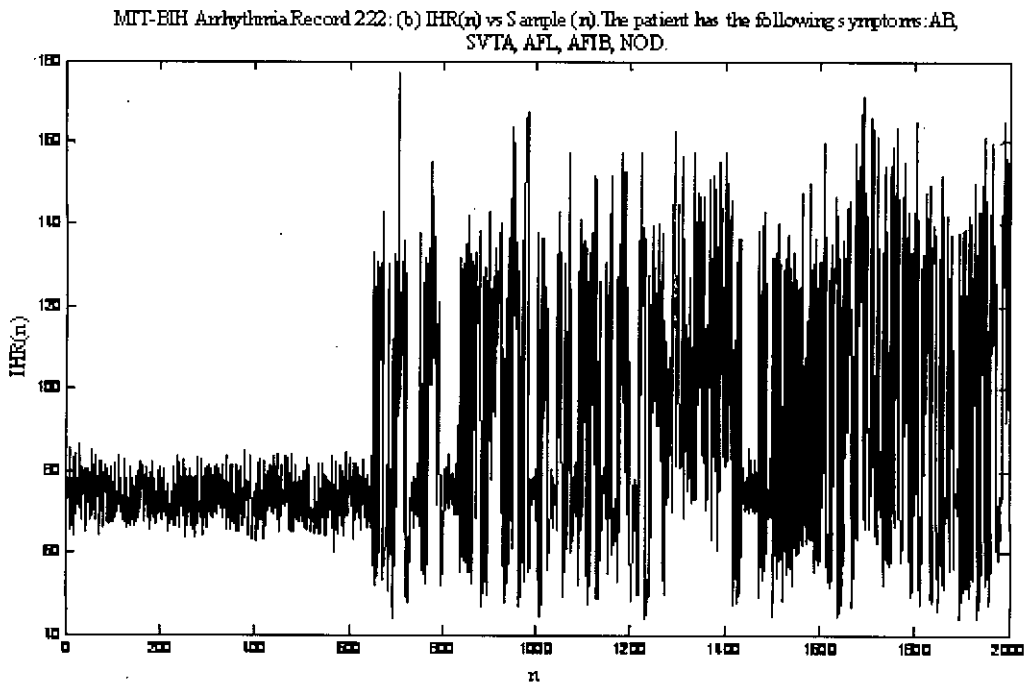
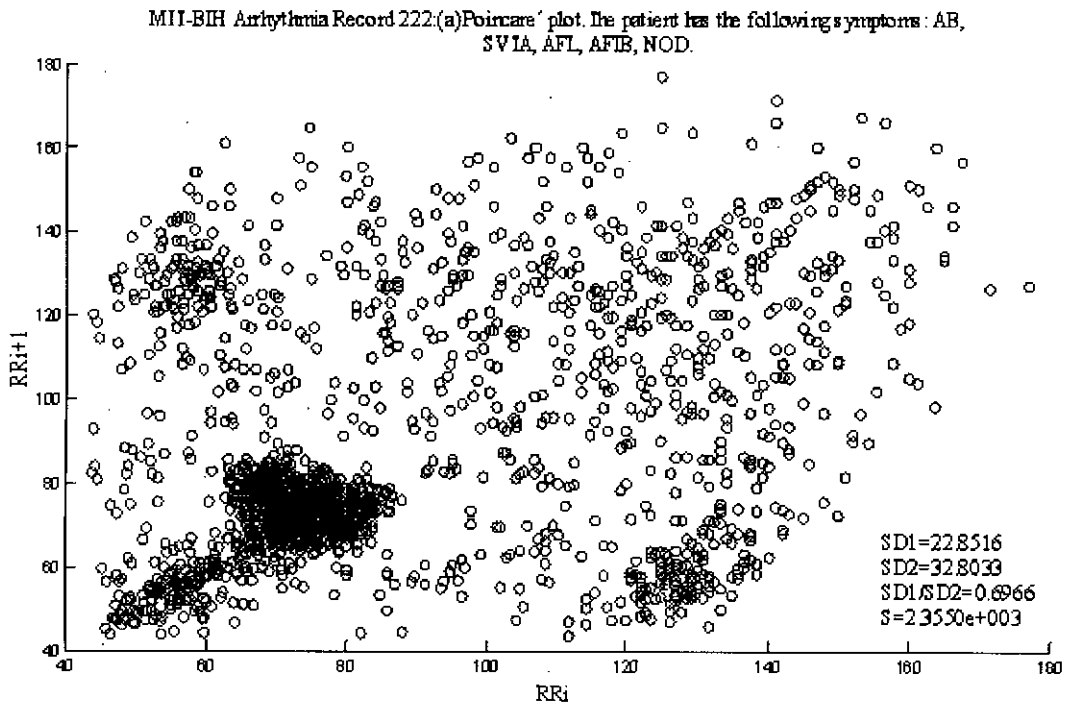


Fig 5.16: MIT-BIH Arrhythmia Record_222: (a) Poincare plot (RR_i vs. RR_{i+1}); (b) IHR (n) vs. Sample (n). The patient has following symptoms: AB, AFIB, SVTA, and NOD.

ApEn between the patient has normal heart beat vs. the patient has following symptoms: T, B, SVTA, AFIB, NOD.

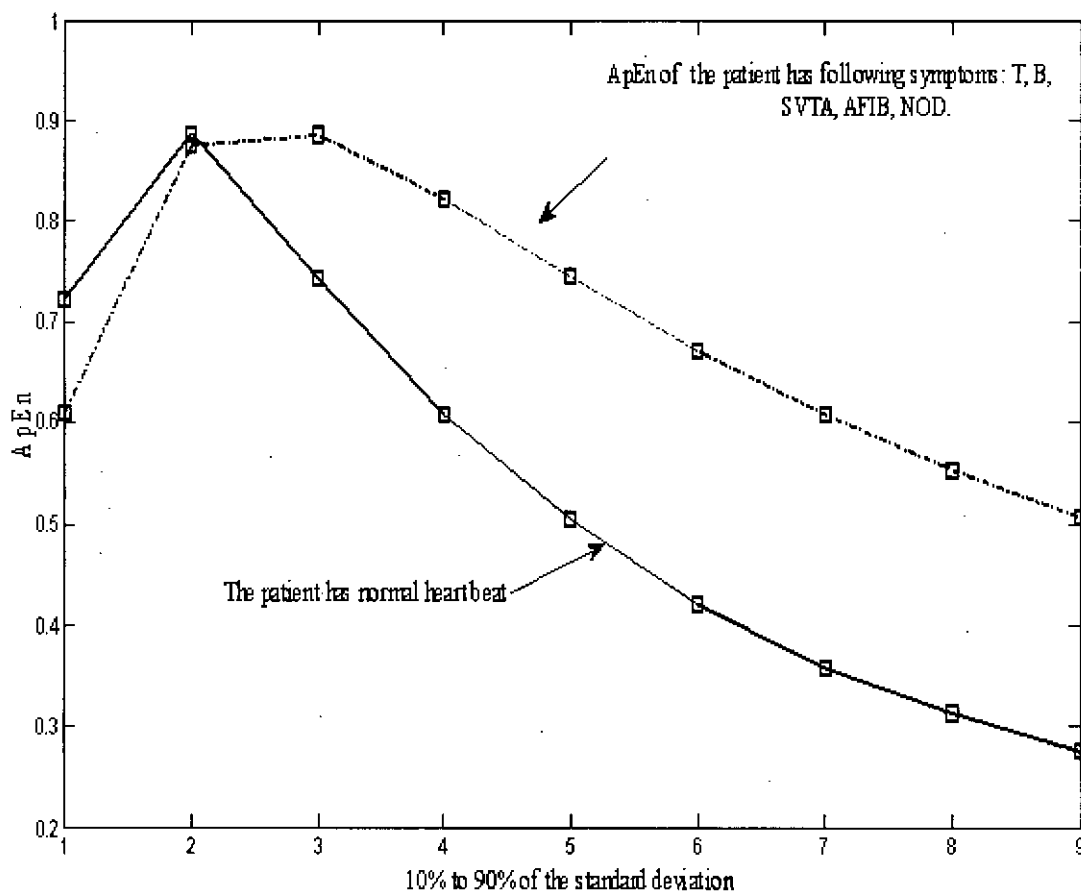


Fig 5.17: ApEn between the patient has normal heart rhythm vs. the patient has following symptoms: T, B, SVTA, AFIB, and NOD.

SampEn between the patient has normal heart beat vs. the patient has following symptoms: T, B, SVTA, AFIB, NOD.

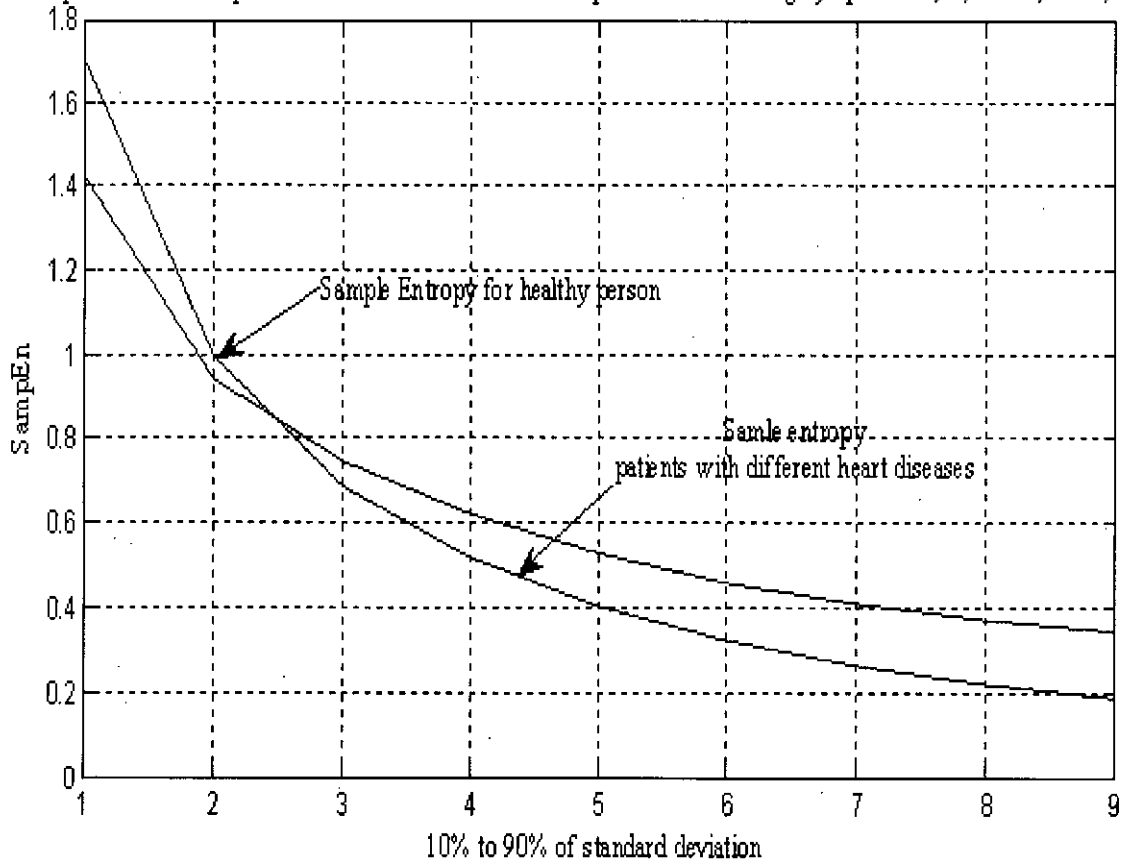


Fig 5.18: SampEn between the patient has normal heart beat vs. the patient has following symptoms: T, B, SVTA, AFIB, and NOD.

Table 5.2 Sample Entropy (SampEn) , Approximate entropy (ApEn) obtained from IHR of normal rhythm (continued)

Record	Symptoms→ duration (in minutes)	ApEn		SampEn	
100	N→ 30:06	10%	0.6702	10%	1.9692
		20%	1.1627	20%	1.3387
		30%	0.9997	30%	0.9769
		40%	0.8191	40%	0.7466
		50%	0.6785	50%	0.5892
		60%	0.5647	60%	0.4787
		70%	0.4807	70%	0.3961
		80%	0.4099	80%	0.3300
		90%	0.3557	90%	0.2757
105	N→ 30:06	10%	1.0073	10%	1.2045
		20%	0.6950	20%	0.6479
		30%	0.4975	30%	0.4079
		40%	0.3883	40%	0.2873
		50%	0.3144	50%	0.2077
		60%	0.2667	60%	0.1497
		70%	0.2324	70%	0.1132
		80%	0.2137	80%	0.0934
		90%	0.2037	90%	0.0824
111	N→ 30:06	10%	0.5019	10%	2.2136
		20%	1.1854	20%	1.5010
		30%	1.0998	30%	1.1528
		40%	0.9289	40%	0.9058
		50%	0.7738	50%	0.7443
		60%	0.6528	60%	0.6188
		70%	0.5598	70%	0.5254
		80%	0.4896	80%	0.4543
		90%	0.4280	90%	0.3929
112	N→ 30:06	10%	0.1326	10%	2.9492
		20%	1.0951	20%	1.7024
		30%	1.1368	30%	1.2053
		40%	1.0249	40%	0.9923
		50%	0.8950	50%	0.8671
		60%	0.7316	60%	0.6960
		70%	0.5947	70%	0.5512
		80%	0.5064	80%	0.4701
		90%	0.4228	90%	0.3954

Table 5.2 Sample Entropy (SampEn) , Approximate entropy (ApEn) obtained from IHR of normal rhythm (continued)

Record	Symptoms→ duration (in minutes)	ApEn		SampEn	
116	N→ 30:06	10%	0.9281	10%	0.9662
		20%	0.5716	20%	0.4550
		30%	0.3953	30%	0.2698
		40%	0.3115	40%	0.1875
		50%	0.2684	50%	0.1486
		60%	0.2440	60%	0.1272
		70%	0.2299	70%	0.1154
		80%	0.2229	80%	0.1101
		90%	0.2185	90%	0.1078
118	N→ 30:06	10%	0.8583	10%	1.1041
		20%	0.6602	20%	0.5498
		30%	0.5117	30%	0.3324
		40%	0.4277	40%	0.2359
		50%	0.3811	50%	0.1905
		60%	0.3487	60%	0.1643
		70%	0.3284	70%	0.1491
		80%	0.3138	80%	0.1392
		90%	0.2989	90%	0.1323
121	N→ 30:06	10%	0.9628	10%	1.1704
		20%	0.6328	20%	0.5484
		30%	0.3701	30%	0.2965
		40%	0.2267	40%	0.1744
		50%	0.1496	50%	0.1110
		60%	0.1053	60%	0.0739
		70%	0.0771	70%	0.0505
		80%	0.0593	80%	0.0360
		90%	0.0477	90%	0.0263
122	N→ 30:06	10%	0.7207	10%	1.9766
		20%	1.1027	20%	1.2785
		30%	0.9294	30%	0.9065
		40%	0.7342	40%	0.6689
		50%	0.5766	50%	0.4957
		60%	0.4499	60%	0.3759
		70%	0.3494	70%	0.2905
		80%	0.2779	80%	0.2308
		90%	0.2250	90%	0.1865

Table 5.3 Sample Entropy (SampEn) , Approximate entropy (ApEn) obtained from IHR of abnormal rhythm (continued)

Record	Symptoms→ duration (in minutes)	ApEn		SampEn	
106	N→ 22:36 B→ 07:15 T→ 00:13 VT→ 0:02	10%	0.8227	10%	0.9352
		20%	0.6777	20%	0.4761
		30%	0.5335	30%	0.2909
		40%	0.4486	40%	0.2097
		50%	0.4065	50%	0.1706
		60%	0.3798	60%	0.1508
		70%	0.3576	70%	0.1395
		80%	0.3442	80%	0.1326
		90%	0.3355	90%	0.1297
119	N→ 22:36 B→ 03:55 T→ 03:34	10%	0.8505	10%	0.7511
		20%	0.6068	20%	0.4527
		30%	0.5104	30%	0.3802
		40%	0.4777	40%	0.3662
		50%	0.4679	50%	0.3673
		60%	0.4611	60%	0.3721
		70%	0.4557	70%	0.3757
		80%	0.4505	80%	0.3765
		90%	0.4475	90%	0.3782
201	N→ 12:57 SVTA→ 0:02 AFIB→ 10:06 NOD→ 0:24 T→ 6:37	10%	0.5399	10%	0.8524
		20%	0.7662	20%	0.5191
		30%	0.8637	30%	0.4729
		40%	0.8492	40%	0.4517
		50%	0.7724	50%	0.4055
		60%	0.6729	60%	0.3381
		70%	0.5915	70%	0.2804
208	N→ 24:43 T→ 5:22	10%	0.5967	10%	1.5636
		20%	0.9365	20%	1.0001
		30%	0.8397	30%	0.7185
		40%	0.7030	40%	0.5570
		50%	0.5889	50%	0.4435
		60%	0.5091	60%	0.3713
		70%	0.4466	70%	0.3240
		80%	0.3979	80%	0.2900
		90%	0.3544	90%	0.2652

Table 5.3 Sample Entropy (SampEn) , Approximate entropy (ApEn) obtained from IHR of abnormal rhythm (continued)

Record	Symptoms→ duration (in minutes)	ApEn		SampEn	
210	AFIB→ 29:30 B→ 0:23 T→ 0:07 VT→ 0:06	10%	0.5712	10%	2.4891
		20%	1.1315	20%	1.7002
		30%	1.1856	30%	1.3140
		40%	1.1094	40%	1.0396
		50%	0.9958	50%	0.8422
		60%	0.8771	60%	0.6927
		70%	0.7692	70%	0.5724
		80%	0.6684	80%	0.4781
		90%	0.5859	90%	0.4030
219	N→ 6:01 AFIB→ 23:47 B→ 0:08 T→ 0:10	10%	0.3705	10%	1.2829
		20%	1.0293	20%	1.1870
		30%	1.2310	30%	1.2266
		40%	1.1858	40%	1.1460
		50%	1.0881	50%	1.0274
		60%	0.9858	60%	0.9049
		70%	0.8830	70%	0.7911
		80%	0.7835	80%	0.6898
		90%	0.7005	90%	0.6025
221	AFIB→ 29:17 B→ 0:03 T→ 0:42 VT→ 0:04	10%	0.4981	10%	2.3384
		20%	1.1579		
		30%	1.2009		
		40%	1.0808		
		50%	0.9538		
		60%	0.8404		
		70%	0.7511		
		80%	0.6826		
		90%	0.6249		
222	N→ 15:57 AB→ 1:28 SVTA→ 0:08 AFL→ 7:03 AFIB→ 1:44 NOD→ 3:45	10%	0.6271	10%	1.1496
		20%	0.6967	20%	0.6116
		30%	0.7250	30%	0.3309
		40%	0.7176	40%	0.1928
		50%	0.6877	50%	0.1295
		60%	0.6490	60%	0.1157
		70%	0.6078	70%	0.1282
		80%	0.5704	80%	0.1476
		90%	0.5315	90%	0.1643

Table 5.4 Sample entropy values of Healthy groups and the groups with abnormal rhythm

Sample entropy	Healthy groups	The patient has the following symptoms: T, AFIB, VT.
SampEn (m=2,r=0.1SD)	1.6979	1.4203
SampEn (m=2,r=0.2SD)	0.9981	0.9417
SampEn (m=2,r=0.3SD)	0.6881	0.7456
SampEn (m=2,r=0.4SD)	0.5171	0.6214
SampEn (m=2,r=0.5SD)	0.4077	0.5299
SampEn (m=2,r=0.6SD)	0.3226	0.4612
SampEn (m=2,r=0.7SD)	0.2611	0.4098
SampEn (m=2,r=0.8SD)	0.2205	0.3718
SampEn (m=2,r=0.9SD)	0.1884	0.3426

Table 5.5 Approximate entropy values of Healthy groups and the groups with abnormal rhythm

Approximate Entropy	Healthy groups	The patient has the following symptoms: T, AFIB, VT.
Apen (m=2,r=0.1SD)	0.7227	0.6096
Apen (m=2,r=0.2SD)	0.8882	0.8753
Apen (m=2,r=0.3SD)	0.7425	0.8862
Apen (m=2,r=0.4SD)	0.6077	0.8215
Apen (m=2,r=0.5SD)	0.5047	0.7451
Apen (m=2,r=0.6SD)	0.4205	0.6719
Apen (m=2,r=0.7SD)	0.3566	0.6078
Apen (m=2,r=0.8SD)	0.3117	0.5530
Apen (m=2,r=0.9SD)	0.2750	0.5064

Table 5.6 HRV parameters in the time domain for Healthy subjects

Record	Symptoms→ duration (in minutes)	Mean	Variance	SD	SDSD	RMSSD
100	N→ 30:06	75.5387	24.7178	4.9705	6.7130	75.7021
105	N→ 30:06	85.8641	122.8857	11.0826	15.7009	86.5764
111	N→ 30:06	70.5141	14.2333	3.7718	3.9680	70.6149
112	N→ 30:06	84.9856	6.4985	2.5486	2.5156	85.0238
116	N→ 30:06	80.5884	113.1342	10.6338	17.4009	81.2869
118	N→ 30:06	75.6381	82.2148	9.0650	11.2915	76.1793
121	N→ 30:06	62.0243	36.2113	6.0159	3.5873	62.3154
122	N→ 30:06	82.4816	21.5514	4.6412	2.2309	82.6121

Table 5.7 HRV parameters in the time domain for the patient has the following symptoms: T, AFIB, VT.

Record	Symptoms→ duration (in minutes)	Mean	Variance	SD	SDSD	RMSSD
106	N→ 22:36,B→ 07:15,T→ 00:13,VT→ 0:02	75.1375	740.2180	27.2001	43.2350	79.9093
119	N→ 22:36,B→ 03:55,T→ 03:34	72.0916	507.2049	22.5155	39.9055	75.5259
201	N→ 12:57,SVTA→ 0:02,AFIB→ 10:06,NOD→ 0:24,T→ 6:37	73.7016	855.6514	29.2438	29.1912	79.2914
208	N→ 24:43,T→ 5:22	103.2344	422.4636	20.5488	33.1580	105.2597

Table 5.7 HRV parameters in the time domain for the patient has the following symptoms: T, AFIB, VT.

Record	Symptoms→ duration (in minutes)	Mean	Variance	SD	SDSD	RMSSD
210	AFIB→ 29:30,B→0:23,T→ 0:07,VT→ 0:06	91.4078	357.2318	18.8959	26.5548	93.3405
219	N→ 6:01,AFIB→23:47,B→ 0:08,T→ 0:10	75.5340	259.6622	16.1100	21.2834	77.2329
221	AFIB→ 29:17,B→0:03,T→ 0:42,VT→ 0:04	87.6909	556.8831	23.5925	36.8294	90.8091
222	N→15:57,AB→1:28,AFL→ 7:03,AFIB→ 1:44 , NOD→3:45	87.4670	799.6668	28.2713	32.3258	91.9225

Table 5.8 Average HRV parameters in the time domain.

Average	Healthy groups	The patient has the following symptoms: T, AFIB, VT.
Variance	52.6808	562.3727
Standard Devision	6.5875	23.2972
SDSD	7.9260	32.8103
RMSSD	77.5388	86.6614

Table 5.9 HRV parameters in the Poincare plot for Healthy subjects.

Record	Symptoms→ duration (in minutes)	SD1	SD2	SD1/SD 2	Area of the ellipse(s)
100	N→ 30:06	4.7468	5.1885	0.9149	77.3745
105	N→ 30:06	11.1128	11.0629	1.0045	386.2278
111	N→ 30:06	2.8096	4.5365	0.6193	40.0414
112	N→ 30:06	1.7803	3.1347	0.5679	17.5326
116	N→ 30:06	12.3134	8.6464	1.4241	334.4725

Table 5.9 HRV parameters in the Poincare plot for Healthy groups

Record	Symptoms → duration (in minutes)	SD1	SD2	SD1/SD2	Area of the ellipse (s)
118	N → 30:06	7.9933	10.0299	0.7969	251.8671
121	N → 30:06	2.5426	8.1208	0.3131	64.8668
122	N → 30:06	1.5712	6.3713	0.2466	31.4484

Table 5.10 HRV parameters in the Poincare plot for the patient has the following symptoms: T, AFIB, VT.

Record	Symptoms → duration (in minutes)	SD1	SD2	SD1/SD2	Area of the ellipse (s)
106	N → 22:36, B → 07:15, T → 00:13 VT → 0:02	30.5780	23.3466	1.3097	2243
119	N → 22:36, B → 03:55, T → 03:34	28.2220	14.7538	1.9129	1308
201	N → 12:57, SVTA → 0:02, AFIB → 10:06, NOD → 0:24, T → 6:37	20.6603	35.8377	0.5765	2326
208	N → 24:43, T → 5:22	23.4550	17.1691	1.3661	1265
210	AFIB → 29:30, B → 0:23, T → 0:07, VT → 0:06	18.7895	19.0139	0.9882	1122
219	N → 6:01, AFIB → 23:47, B → 0:08, T → 0:10	15.0562	17.1048	0.8802	809
221	AFIB → 29:17, B → 0:03, T → 0:42, VT → 0:04	26.0621	20.8568	1.2496	1708
222	N → 15:57, AB → 1:28, SVTA → 0:08, AFL → 7:03, AFIB → 1:44, NOD → 3:45	22.8516	32.8033	0.6966	2355

107622



Table 5.11 Average values of HRV parameters in the Poincare plot

Average	Healthy groups	groups with heart diseases
SD1	5.6087	23.2093
SD2	7.1364	22.6107
Ratio	0.7359	1.1225
Area of the ellipse, (s)	150.4789	1642

5.3 Time domain indexes

All assessed conventional HRV parameters in the time domain (Variance, SDNN, SDDSD, RMSSD) were reduced in Healthy groups (Table. 5, 6, 7). Variance, SDNN, SDDSD, RMSSD of healthy groups are shown in table 5 and groups with abnormal rhythm are shown in table 6. A clear difference was found by analyzing these two groups. The average values of variance, standard deviation, standard deviation of successive difference (SDDSD), root mean square of standard deviation (RMSSD) are shown in table 7. HRV parameters in the time domain were higher in unhealthy groups.

5.4 Poincaré plot indexes

Sixteen ECG recordings were analyzed. Table 8 and table 9 summarize the results from Poincaré indexes of the two groups. A clear reduction of SD1 and SD2 in the healthy group was observed, and we found significant differences between two groups. For healthy group **SD1= 5.6087** and for heart diseases group **SD1=23.2093**. For healthy group **SD2= 7.1364** and for heart diseases group **SD2=22.6107**. Additionally, we may define a parameter which reflects the total variability as measured by the Poincaré plot, $s = \pi SD1SD2$. This is the area of the ellipse. A significant difference was found between two groups for healthy group **Area of the ellipse=150.4789** for abnormal rhythm group **Area of the ellipse=1642** which was much higher than healthy group. Poincaré plot can provide supplementary information about beat to beat HRV structure which cannot be obtained by conventional time and frequency domain analysis [22]. From the Poincaré plot indexes, healthy subjects had all measures reduced (Table 8, 9, 10). The decreased long term HRV (represented by the length, SD2) and decreased beat to beat HRV (represented by the width, SD1) was expected in healthy group. SD1/SD2 represents the ratio of short term and long term variability. The smaller SD1/SD2

ratio for the healthy subjects reflects that a lower percentage of its overall variance is beat-to-beat variance. A significant difference was found between the SD1/SD2 ratio values of the two groups. The average SD1/SD2 for healthy group was **0.7359** and other group was **1.1225**.

5.5 Approximate entropy (ApEn) values of IHR between Healthy groups and the groups with abnormal rhythm

ApEn of a time series $RR(i)$ measures the logarithmic likelihood that runs of patterns of length m that are close to each other will remain close in the next incremental comparisons, $m+1$. A greater probability of remaining close (high regularity) produces smaller ApEn values, and, vice-versa, low regularity produces higher ApEn values. The two parameters, m and r , must be fixed to compute approximate entropy. The values $m = 2$ and r between 10% and 90% of the standard deviation of the data sets $RR(i)$ were used in this project. Figure shows the changes of Apen with $m=2$ and $r=0.1*SD$ to $0.9SD$ of IHR data for healthy groups and unhealthy groups. Lower values of ApEn were found for healthy group. A significant difference was found between healthy group and unhealthy group.

5.6 Sample Entropy (SampEn) values of IHR between Healthy groups and the groups with abnormal rhythm

To investigate the complexity of the heart rate variability, the sample entropy (SampEn) of the IHR signals was calculated. Figure (5.18) demonstrates the change of SampEn with $m=2$ and $r=0.1*SD$ to $0.9*SD$ of IHR data for healthy groups and unhealthy groups. The mean values of SampEn of the healthy group were found to be lower than that of unhealthy groups at all r values except at $0.1*SD$ and $0.2*SD$. Statistically, healthy groups and unhealthy groups were found to be significantly different at $r>0.3*SD$. SampEn values at different r values are summarized in Table 4. Figure (5.18) shows the SampEn ($m=2$, $r= [0.1-0.9]*SD$) values of all subjects in this study. Lower values of SampEn reflect more regular time series while higher values are associated with less predictable (more complex) time series. The lower SampEn values for the healthy group indicates an increase in regularity and a decrease in complexity in the IHR.

5.7 Conclusion:

So, if we compare the three techniques we used namely Poincare plot, time domain analysis and entropy measure, we can come out with the following facts. Poincare plot gives us a visual observation of the ECG signals, whether they are from normal or abnormal rhythms. Best result found by using area of ellipse. On the contrary, by using time domain indexes we can identify the abnormality (in this project reduced value found for normal rhythms), entropy measure quantifies the abnormality levels present in the ECG signals (lower values of ApEn and SampEn found for normal rhythms). Moreover, it roughly gives an idea about the abnormality type as observed in our work.

Chapter 6

CONCLUSIONS

6.1 Discussions

In this work, effective ECG classification techniques are presented based on the Entropy of Chaotic Attractor. This work describes the application of sample entropy (SampEn), approximate entropy (ApEn), Poincaré plot analysis and time domain HRV parameters to differentiate the normal rhythm from the abnormal one. First IHR were calculated from the ECG of eight normal rhythms as well as from that of eight abnormal rhythms, each approximately 30 minutes in duration, and total sixteen data sets were constructed. Then sample entropy (SampEn), approximate entropy (ApEn), Poincaré plot analysis and time domain HRV parameters were determined from those data sets and results are compared for normal and abnormal rhythm data sets. The entropy analysis is able to differentiate between normal and abnormal ECG. This is a crucial step in cardiac signal analysis. The results show that there is a significant difference between the ApEn of normal rhythm data sets and that of abnormal rhythm data sets. A lower ApEn were found for normal rhythm and higher in abnormal rhythm data sets. Approximate entropy (ApEn) counts each sequence as matching itself. In contrast, SampEn abates this self matching. So a better result found by using SampEn. A lower SampEn were found for normal rhythm data sets and higher in abnormal rhythm data sets. Poincaré plot images represent short and long-term variability. The results show that there is a significant difference between the Poincaré plot analysis of normal rhythm data sets and that of abnormal rhythm data sets. A clear reduction of SD1 and SD2 in the healthy groups was observed, and we found a significant difference between two groups. All assessed conventional HRV parameters in the time domain (Variance, SDNN, SDDSD and RMSSD) were reduced in Healthy groups. Hence, we can easily classify normal and abnormal ECG records based on the Entropy of Chaotic Attractor. Moreover, for the BBI data set, we could roughly distinguish some types of abnormalities like Atrial fibrillation (AFIB), Ventricular bigeminy(B), Ventricular trigeminy(T). However, we could not generalize this abnormality detection since we worked only with a limited number of records. Based on the computational simplicity and high levels of accuracy obtained by this technique, it is logical to use this approach to classify the ECG as normal or abnormal.

6.2 Future work

In this work, we have tried to classify ECG using the Entropy of Chaotic Attractor. We have shown that approximate entropy (ApEn), sample entropy (SampEn) and Poincaré plot analysis can distinguish between normal and abnormal cardiac rhythm. In this work, we determined the Poincaré plot indexes and conventional HRV parameters in the time domain (Variance, SD, SDNN, RMSSD) of the whole ECG record which contains both the normal and abnormal beats. Future work may include working with more number of abnormal records to generalize the detection of beat abnormality type.

Appendix-A: QRS Detection Algorithm

An ECG beat is defined as the signal sample from one R-wave to the next. Figure A.1 shows the block diagram of the QRS detection algorithm.

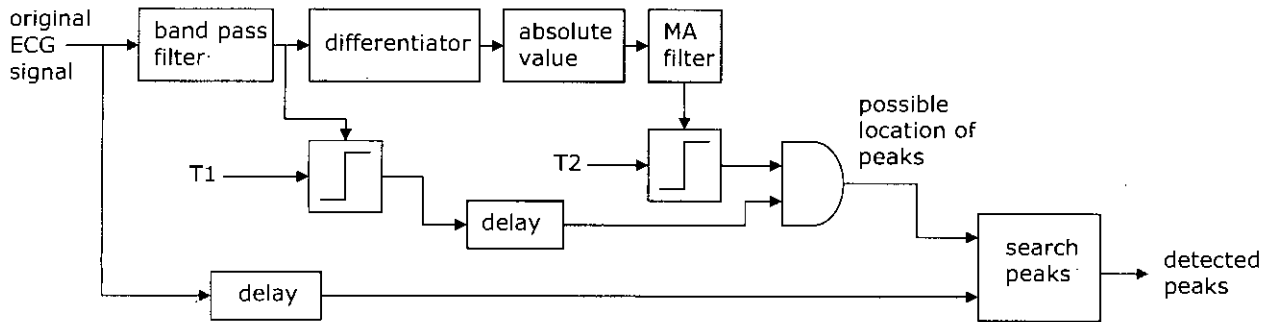


Fig A.1: Block diagram of QRS detection algorithm

At first the ECG signal is passed through a linear phase bandpass filter (4 hz to 40 hz) for smoothing operation and reducing base line shifting. The impulse response, magnitude response and the phase response of the bandpass filter is shown in the figure A.2 and A.3. The group delay of the filter is 150.

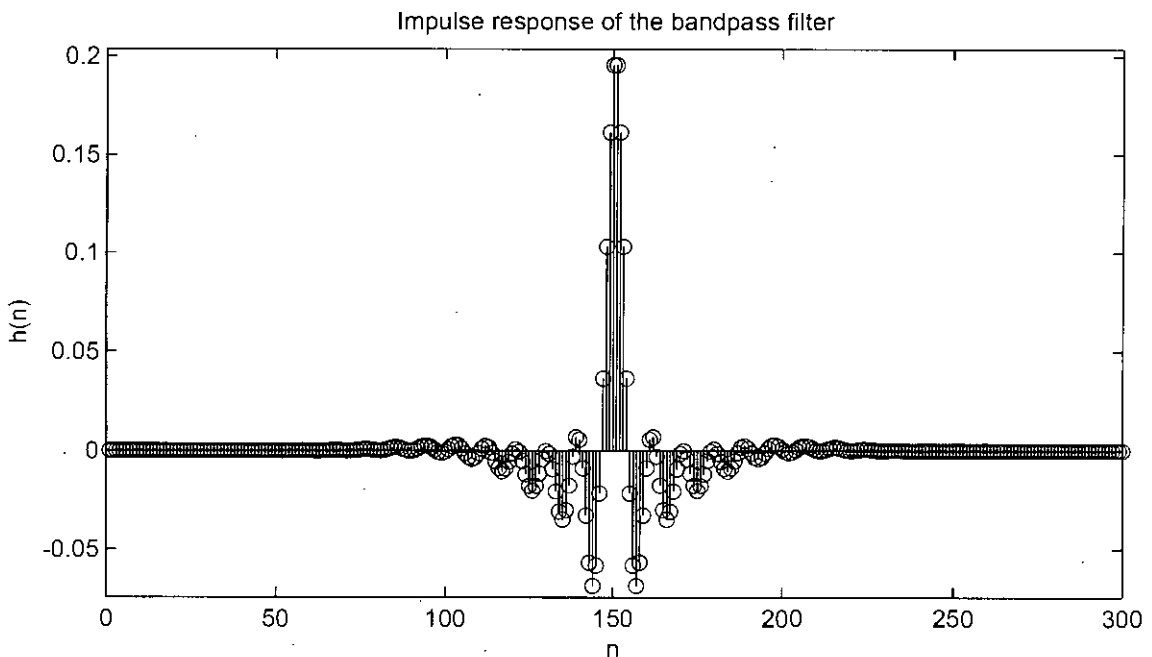


Fig A.2: Impulse response of the bandpass filter

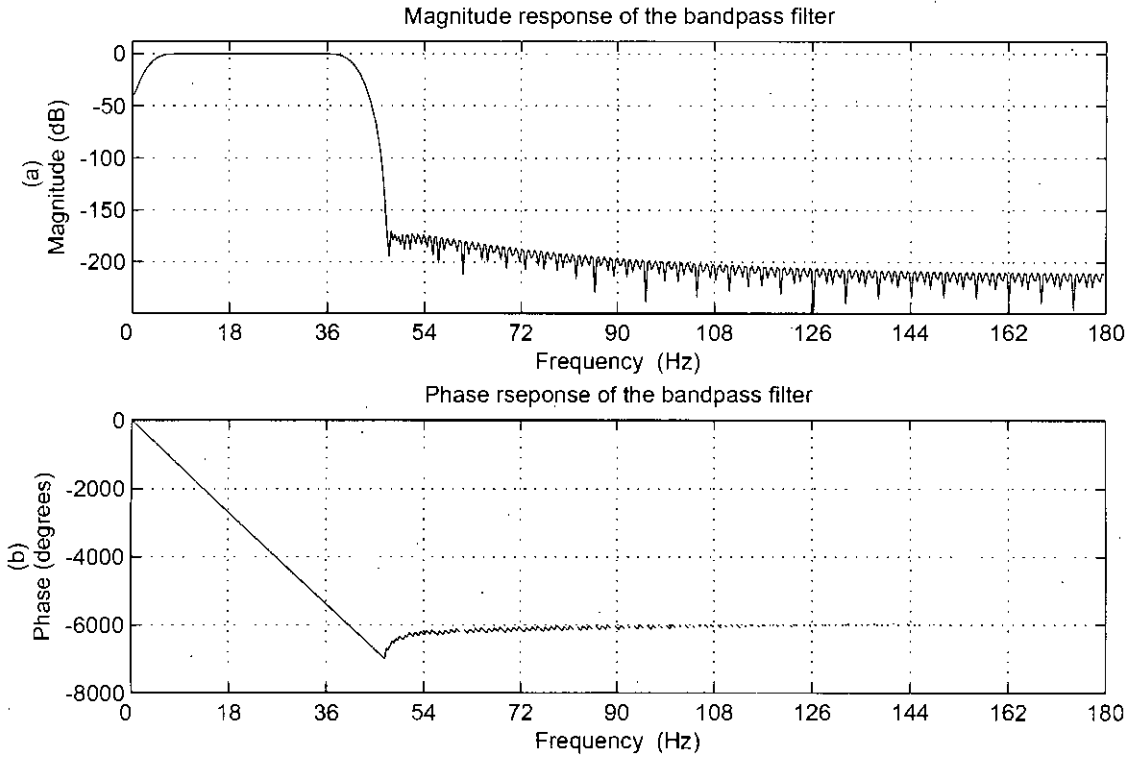


Fig A.3: (a) The magnitude and (b) phase response of the bandpass filter

Differentiation of the filtered signal provides the slope information of QRS complexes. Since there are quick rise and fall times of the QRS complex in the ECG signals, the derivative makes it easier to detect the time of occurrence of the QRS complexes. The transfer function of the five-point differentiation equation is given by

$$H(z) = \left(\frac{1}{8}\right)(-z^{-2} - 2z^{-1} + 2z^1 + z^2) \quad (\text{A.1})$$

The absolute value of the output of derivative filter can be found by the following operation

$$y(n) = \sqrt{x(n)^2} \quad (\text{A.2})$$

This absolute valued signal is then passed through a moving average filter which produces high value at the region of QRS complex. The window size has to be taken properly, neither so wide that merges the QRS complex and T wave together, nor so narrow that produces several peaks in the integration waveform. It is calculated from the equation below:

$$y(n) = \left(\frac{1}{N}\right)[x(n-(N-1)) + x(n-(N-2)) + \dots + x(n)] \quad (\text{A.3})$$

where N is the width of the integration window. This work takes N as 18

As seen in figure A.1, the algorithm sets two thresholds $T1$ and $T2$ to make decisions. $T1$ is set 40 for the filtered ECG, and $T2$ is set 7 for the signals produced by the moving window integration. The thresholded filtered signal is then delayed by 10 samples and a logical 'and' operation is performed with the thresholded moving squared average signal. As a result, possible location of peaks of the original ECG signal is found. These locations are delayed by 160 samples from the original ECG. Searching in these regions, we get the peaks. If there are more than one peak in the vicinity of 50 samples, the highest peak is considered from those.

The sequences of the QRS detection algorithm is shown in figure A.4 where the ECG signal is taken from the record 100/MLII of MIT-BIH arrhythmia database.

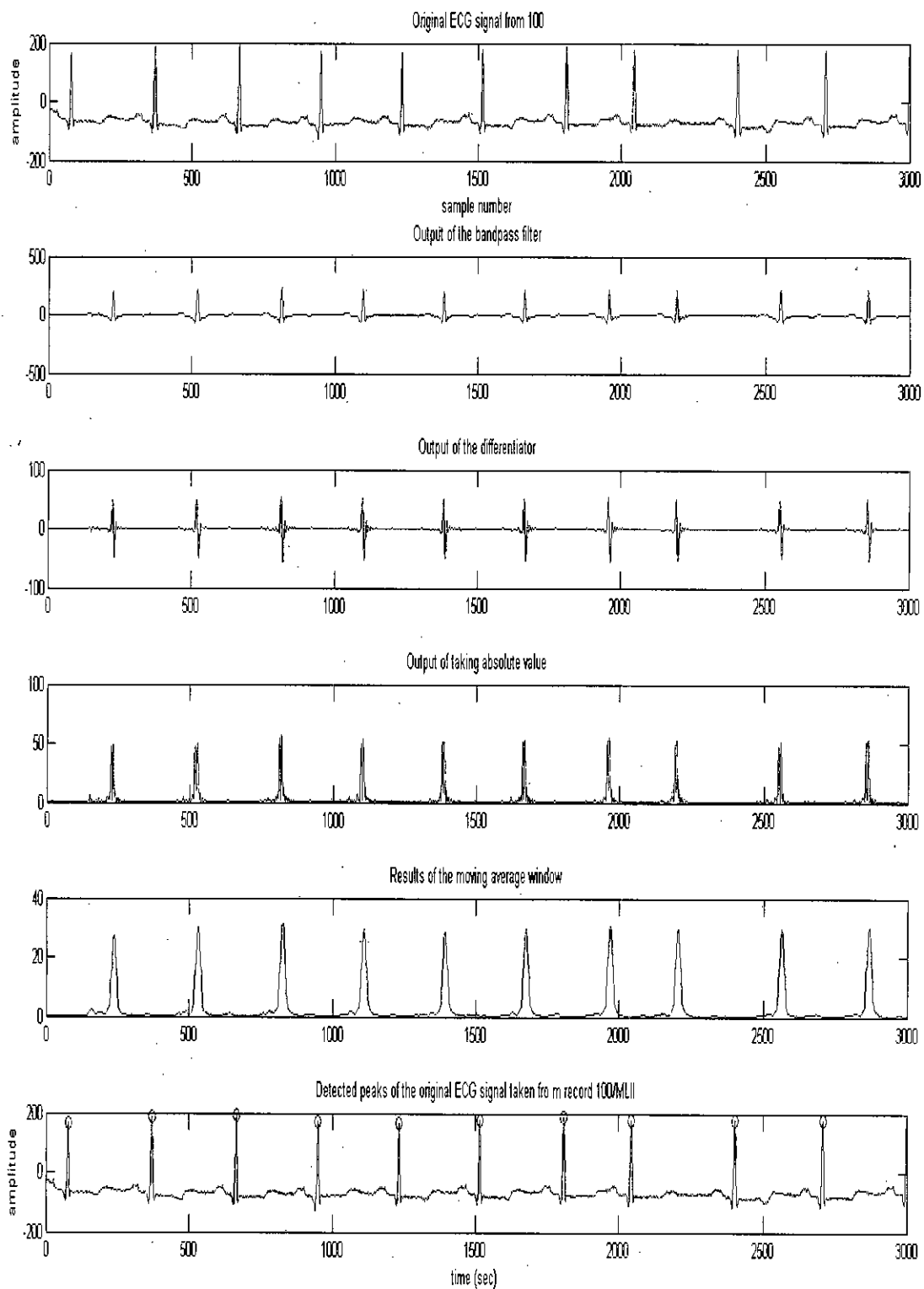


Fig A.4: The sequences of peak detection of the ECG signal taken from the record 100 (MLII) of MIT-BIH arrhythmia database.

References

- [1] K. Minami, H. Nakajima, and T. Toyoshima, "Real-time discrimination of ventricular tachyarrhythmia with Fourier-transform neural network," *IEEE Trans. Biomed. Eng.*, vol. 46, pp. 179–185, Feb. 1999.
- [2] Thakor, N.V.; Pan, K, "Tachycardia and fibrillation detection by automatic implantable cardioverter-defibrillators: sequential testing in time domain," *IEEE Trans. Biomed. Engg.* Vol. 09, pp. 21-24, March.1990.
- [3] Richman JS, Moorman JR,"Physiological time-series analysis using approximate entropy and sample entropy" *Am J Physiol* 2000, Vol. 278: H2039–H2049.
- [4] Thakor, N.V.; Pan, K, "Tachycardia and fibrillation detection by automatic implantable cardioverter- defibrillators: sequential testing in time domain," *IEEE Trans. Biomed. Engg.* Vol. 09, pp. 21-24, March.1990.
- [5] Thakor, N.V.; Pan, K, "Tachycardia and fibrillation detection by automatic implantable cardioverter-defibrillators: sequential testing in time domain," *IEEE Trans. Biomed. Engg.* Vol. 09, pp. 21-24, March.1990
- [6] Jarosław Piskorski, Przemysław Guzik, "Filtering Poincaré plots" *Computational methods in Science and Technology* 11(1), pp. 39-48, June 2005.
- [7] Przemysław Guzik, Jaroslaw Piskorski, Tomasz Krauze, Raphael Schneider, Karel H. Wesseling, Andrzej Wykre, Towicz, and Henryk Wysocki, "Correlations between the Poincaré Plot and Conventional Heart Rate Variability Parameters Assessed during Paced Breathing" *J. Physiol. Sci.* Vol. 57, pp. 63–71, Feb. 2007.
- [8] Volkan Tuzcu, Selman Nas "Sample Entropy Analysis of Heart Rhythm Following Cardiac Transplantation" *2005 IEEE International Conference on Systems, Man and Cybernetics Waikoloa*, Hawaii October 10-12, 2005.
- [9] Ahsan H. Khandoker, Herbert F. Jelinek, Marimuthu Palaniswami, "Heart rate variability and complexity in people with diabetes associated cardiac autonomic neuropathy" *30th Annual International IEEE EMBS Conference Vancouver, British Columbia, Canada*, August 20-24, 2008.

- [10] P Castiglioni, M Di Rienzo, "How the Threshold "R" Influences Approximate Entropy Analysis of Heart-Rate Variability" *Computers in Cardiology* 2008;Vol. 35: pp. 561–564.
- [11] Hudson, D.L.; Cohen, M.E.; Deedwania, P.C, "Chaotic ECG analysis using combined models " Engineering in Medicine and Biology Society, 1998. Proceedings of the 20th Annual International Conference of the IEEE, vol. 3, pp. 1553-1556, 29 Oct-1 Nov.1998.
- [12] Jovic, Alan; Bogunovic, Nikola, "Feature Extraction for ECG Time-Series Mining Based on Chaos Theory," *Information Technology Interfaces, 2007. ITI 2007. 29th International Conference on Information Technology Interfaces, Cavtat, Croatia*, pp. 63-68 June 25-28, 2007.
- [13] N. Srinivasan, M. T. Wong, S. M. Krishnan, "A new Phase Space Analysis Algorithm for Cardiac Arrhythmia Detection", *Proceedings of the 25th Annual International Conference of the IEEE EMBS Cancun, Mexico* September 17-21,2003.
- [14] Kamen PW, Tonkin AM, "Application of the Poincaré' plot to heart rate variability: a new measure of functional status in heart failure" *Aust N Z J Med* 1995, 25:18–26.
- [15] B. Anuradha, K. Suresh Kumar and V. C. Veera Reddy " Classification of Cardiac Signals Using Time Domain Methods" *ARPN Journal of Engineering and Applied Sciences*, VOL. 3, NO. 3, JUNE 2008.
- [16] Huszar RJ. "Basic Dysrhythmias: interpretation & management", 2nd ed. St. Louis, Missouri: Mosby Lifeline; 1994.
- [17] D.A. Litvack, T.F. Oberlander, L.H. Carney, and J.P. Saul. "Time and frequency domain methods for heart rate variability analysis" a methodological comparison. *Psychophys-Vol. 32*,492–504, 1995.
- [18] Jovic, Alan; Bogunovic, Nikola, "Feature Extraction for ECG Time-Series Mining Based on Chaos Theory," *Information Technology Interfaces, 2007. ITI 2007. 29th International Conference on Information Technology Interfaces, Cavtat, Croatia*, pp. 63-68 June 25-28, 2007.

- [19] Brennan M, Palaniswami M, Kamen P, "Do existing measures of Poincaré' plot geometry reflect nonlinear features of heart rate variability?" *IEEE Trans Biomed Eng* 2001, 48:1342–1347.
- [20] Steven M. Pincus "Approximate entropy as a measure of system complexity" *Proc. Nati. Acad. Sci. USA*, Vol. 88, pp. 2297-2301, March 1991.
- [21] R Alcaraz, JJ Rieta, "Robust Prediction of Atrial Fibrillation Termination Using Wavelet Bidomain Entropy Analysis" *Computers in Cardiology* 2007, Vol. 34, pp. 577–580.
- [22] Task force of the European society of cardiology and the North American society of pacing and electrophysiology. Heart rate variability – standards of measurement, physiological interpretation, and clinical use. *Circulation*, 93(5):1043–1065, March 1996.
- [23] Juan Pablo Martínez, Rute Almeida, Salvador Olmos, *Member, IEEE*, Ana Paula Rocha, and Pablo Laguna, *Member, IEEE*, "A Wavelet-Based ECG Delineator: Evaluation on Standard Databases" *IEEE Transactions on Biomedical Engineering*, Vol. 51, NO. 4, APRIL 2004.
- [24] Jiapupan and Willis J. Tompkins," A Real-Time QRS Detection Algorithm" *IEEE Transactions on Biomedical Engineering*, Vol. BME-32, NO. 3, MARCH 1985.
- [25] Claudia Lerma, Oscar Infante, He'ctor Pe'rez-Grovas and Marco V.Jose , "Poincare' plot of heart rate variability capture dynamic adaptations after haemodialysis in chronic renal failure patients" *Clin Physiol & Func Im* (2003) 23, pp. 72–80, October 2002.
- [26] Brennan M, Palaniswami M, Kamen P, "Poincaré plot interpretation using a physiological model of HRV based on a network of oscillators" *Am J Physiol Heart Circ Physiol* 2002, 283, H1873-H1886.
- [27] Kitlas A, Oczeretko E, Kowalewski M, Borowska M, Urban M, "Nonlinear dynamics methods in the analysis of the heart rate variability" *Roczniki Akademii Medycznej w Białymstoku* · Vol. 50, 2005.
- [28] MIT-BIH Arrhythmia Database CD-ROM, 3rd ed. Cambridge, MA: Harvard–MIT Div. Health Sci. Technol., 1997.
- [29] The interpreter of the MATLAB language may be downloaded from <http://www.Mathworks.com/products/statistics/>

